



ICONI



KOREAN SOCIETY FOR INTERNET INFORMATION

The 15th International Conference on Internet (ICONI 2023)

Dec. 17-20, 2023, Lotte Hotel Saigon, Hochiminh, Vietnam

<http://www.iconi.org>

Proceedings of ICONI 2023

| Organized by |

Korean Society for Internet Information (KSII)

| Co-Organized by |

Advanced Institute of Convergence Technology(AICT)

Message from Honorary Chair

We express our warm welcome to all of attendants of the 15th International Conference on Internet (ICONI) that will be held from December 17~20, 2023, in Lotte Hotel Saigon, Hochiminh, Vietnam.

Over the past 15 years, ICONI has increased its academic impact and influence in the area of Information Science, Computer Engineering and Internet Technology by bringing together international researchers from academia, industry, and the government to exchange novel ideas, explore enabling technologies, discuss innovative designs, and share field trial experiences and lessons learned.

This year, ICONI 2023 again covers a broad range of topics which are related with Internet technologies and applications such as wireless and sensor networks, security & privacy, multimedia, image processing, intelligent systems, mobile computing, Internet of things, and wireless communications. It is firmly believed that each paper to be presented will be a basis for further constructive discussion.

We would like to express our heartfelt gratitude to every one who makes this conference successful. First of all, we are so grateful to all contributors who submit their valuable papers, review the papers, and chair the sessions. We also thank many sponsors for their enormous support. Our special thanks must go to the Keynote Speakers Mr. Vo Ngoc Han, Prof. Pham The Bao and Dr. Ton Long Phuoc.

Last but not least, we deeply appreciate limitless devotion of the Organizing Committee members. Without their contribution, this conference would not be realized.

We truly welcome you all to this beautiful country and wish you pleasant and joyous stay in Hochiminh, Vietnam!



Dr. Kwanghoon Pio Kim
Kyonggi Univ.,
Rep. of Korea

Dr. Kwanghoon Pio Kim
Honorary Chair of ICONI 2023

Message from Conference Chair

On behalf of the organizing committee of the 15th International Conference on Internet (ICONI) 2023, I cordially welcome you and sincerely thank you for participating in this international conference to present and discuss valuable state-of-the-art research and development results.

We have been able to successfully plan for ICONI 2023, prioritizing all safety measures. More than 115 papers have been submitted in 17 Internet-related fields, including mobile Internet computing, Internet security, Internet of Things (IoT), wireless and sensor networks and multimedia, image processing and Internet application management. Among these papers, 115 papers will be presented during the conference.

In addition, there are 3 prominent world-famous Keynote Speakers that will be presenting during the ICONI 2023 conference, which include Mr. Vo Ngoc Han, CEO of Vietnam-Korean Trade Investment Consultancy and Prof. Pham The Bao, Dean in Department of Computer Science from Saigon University, and Dr. Ton Long Phuoc, Faculty of Information Technology from Industrial University of Hochiminh.

I owe a great amount of gratitude to the organizing committee members and all contributors for the extraordinary work they have done in organizing this conference. Without their dedicated support, this event would not be possible.

I wish you all a very pleasant time here in Hochiminh, Vietnam and a productive and successful conference.



Dr. Ihnhan Bae
Daegu Catholic Univ.,
Rep. of Korea

Dr. Ihnhan Bae
Conference Chair of ICONI 2023

Message from Program Chairs

It is a pleasure to welcome you to the 15th International Conference on the Internet (ICONI 2023). ICONI 2023 will be held in Hochiminh, Vietnam, from Dec 17-20. The hosting organization (KSII) is one of the largest ICT-related academic societies in South Korea.

Over the past 15 years, ICONI has grown to be a major international conference in the Internet Technology area. The conference is organized into 17 tracks in the Internet Technology field. ICONI continues the endeavor of high-quality, broad international participation in all areas of Internet Technology.

The successful organization of ICONI has required the talents, dedication, and time of many volunteers from Malaysia, USA, Vietnam, China, Nepal and Republic of Korea. Special gratitude and appreciation go to all organizing committee members as they are primarily responsible for the conference. This ICONI conference serves as a unique and exciting gathering at the forefront of the most innovative and captivating realms of computer science. Your presence here provides a valuable opportunity to share novel ideas, research findings, and, importantly, to learn from each other's experiences, fostering mutual growth. I look forward to engaging in deep discussions, sharing fresh ideas, and fostering connections that will lead to newfound perspectives.

We hope that you will find the conference both enjoyable and valuable and also enjoy the architectural, cultural, and natural beauty of Hochiminh, Vietnam. I express my sincere gratitude to all participants for your tremendous support, and I genuinely hope that ICONI 2023 becomes a meaningful and enriching experience for each one of you. Thank you.



Dr. Junho Ahn
Kyonggi Univ.,
Rep. of Korea



Dr. Imran Ghani
Virginia Military Institute, USA

Dr. Junho Ahn, Dr. Imran Ghani
Program Chairs of ICONI 2023

Organizing Committee

| Honorary Chair |

Kwanghoon Pio Kim, Kyonggi University, ROK

| Conference Chair |

Ihnhan Bae, Daegu Catholic University, ROK

| Program Chairs |

Junho Ahn, Kyonggi University, ROK

Imran Ghani, Virginia Military Institute, USA

| Publication Chairs |

Jong-Moon Chung, Yonsei University, ROK

Minho Jo, Korea University, ROK

Moonseong Kim, Seoul Theological University, ROK

| Financial Chairs |

Yoosin Kim, Airdeep Ltd., ROK

Namgi Kim, Kyonggi University, ROK

| Workshop Chairs |

Kiwon Kwon, Korea Electronics Technology Institute, ROK

Woochun Jun, Seoul National University of Education, ROK

Tai-Woo Chang, Kyonggi University, ROK

| Steering Chairs |

DongKyoo Shin, Sejong University, ROK

Dinh-Lam Pham, Dong Nai Technology University, Vietnam

Keon Chul Park, Advanced Institute of Convergence Technology, ROK

Ziyang Liu, Kyonggi University, ROK

Organizing Committee

| Program Committee |

Jang-Hyun Kim, Sungkyunkwan University, ROK
Ji-Hwan Kim, Sogang University, ROK
Woochan Lee, Incheon National University, ROK
WEIJIA LI, Jiaxing University, China
JeongHyeon Chang, Kyonggi University, ROK
Jeongwook Seo, Hanshin University, ROK
Lanh Thanh Le, Dong Nai Technology University, Vietnam
Dohoon Kim, Kyonggi University, ROK
Ilkyeun Ra, University of Colorado, Denver, USA
Rui Liang, Beijing Union University, China
Kyungyong Chung, Kyonggi University, ROK
Mingsheng Wang, Anshan Normal University, China
Namje Park, Jeju National University, ROK
Jongwook Woo, California State University Los Angeles, USA
Ir. Made Sudarma, Udayana University, Indonesia
Hafiz Adnan Hussain, Universiti Kebangsaan Malaysia, Malaysia
Hangbae Chang, Chung-Ang University, ROK
Uzma Jafar, Universiti Kebangsaan Malaysia, Malaysia
Kyung Han You, Jeonbuk National University, ROK
Hoang-Hai Tran, Hanoi Univ. of Science and Technology, Vietnam
Sangmin Lee, CHA University, ROK
Eui-Nam Huh, Kyung Hee University, ROK
Deden Witarsyah, Telkom University, Indonesia
Yong-Suk Park, Korea Electronics Technology Institute, ROK
Chi Gong, Beijing Normal University, China
Yongjoon Lee, Far East University, ROK
Ogan Gurel, L'Institut des Sciences et Arts, Canada
Sudan Prasad Uprety, AICC Solution Lab, Nepal
Sungyoon Cho, Korea Electronics Technology Institute, ROK

Conference Program

| | |
|-------------------------|---------------------------|
| Sun. / Dec. 17, 2023 | Room 1 |
| 18:00-19:30 | VIP Reception & Greetings |

| Mon. / Dec. 18, 2023 | Room 1 (Sapphire) | Room 2 (Topaz) | Room 3 (Emerald I) | Room 4 (Emerald III) |
|-------------------------|--|-------------------|-----------------------|-------------------------|
| 08:30- | Registration (2F) | | | |
| 09:00-10:30 | Session 1 | Session 2 | Session 3 | Session 4 |
| 10:40-11:10 | Opening Ceremony & Keynote Speech (Sapphire Ballroom) | | | |
| 11:20-12:20 | | Session 5 | Session 6 | Session 7 |
| 12:20-14:00 | Lunch (11:30-14:30) / The Canvas GF | | | |
| 14:10-15:40 | Session 8 | Session 9 | Session 10 | Session 11 |
| 15:50-17:20 | | Session 12 | Session 13 | Session 14 |
| 18:00-19:30 | Award Ceremony & Banquet (Sapphire Ballroom) | | | |

Conference Program

| Tue. / Dec. 19, 2023 | Room 1 (Topaz) | Room 2 (Emerald I) | Room 3 (Emerald III) |
|----------------------|---|-------------------------------|-------------------------|
| 08:30- | Registration (2F) | | |
| 09:00-10:30 | Session 15 | participant's waiting room | Session 16 |
| 10:40-12:10 | Session 17 | | |
| 12:10-14:00 | Lunch (11:30-14:30) / The Canvas GF | | |
| 14:10-15:30 | Panel Discussion : Collaborative Research | | |

| Wen. / Dec. 20, 2023 | Room |
|----------------------|---|
| 10:00-12:00 | Visit to Industrial University of Hochiminh |

Conference Program

Sun, Dec. 17, 2023

18:00-19:30

VIP Reception & Greetings

Mon, Dec. 18, 2023

09:00-10:30

Session 1: Interactive Presentation Session I

Session 2: Industrial AI Special Workshop

Session 3 : Artificial Intelligence/Time Series Machine Learning

Session 4: Scientific public security/ crime prevention/intelligent CCTV application

10:40-11:10

Opening Ceremony & Keynote Speech

Dr. Jim Jansen / Prof. Pham The Bao

11:20-12:20

Session 5: Graphics/image processing/video deep learning/video control

Session 6: Security system/information security

Session 7: Defense information technology/high performance computing & Cyber Education/Social Media/Metaverse/UX/SNS

Conference Program

Mon, Dec. 18, 2023

| | |
|-------------|--|
| 12:20-14:00 | Lunch (11:30-14:30) / The Canvas (GF) |
| 14:10-15:40 | Session 8: KETI Workshop I |
| | Session 9: KETI Workshop II |
| | Session 10 : Technology management/ e-commerce/information system/ERP/BPM/DX |
| | Session 11: WIET Education Workshop I |
| 15:50-17:20 | Session 12: KETI Workshop III |
| | Session 13: Network/Internet of Things/ Mobile Computing & Technology management/ e-commerce/information system/ERP/BPM/DX |
| | Session 14: WIET Education Workshop II |
| 18:00-19:30 | Keynote Speech & Award Ceremony & Banquet Dr. Ton Long Phuoc |

Conference Program

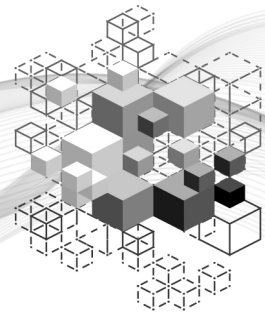
Tue, Dec. 19, 2023

| | |
|-------------|--|
| 09:00-10:30 | Session 15: Interactive Presentation Session II |
| | Session 16: Intelligent ICT Convergence/Smart Building & Big data platform/analysis/application & Cloud Computing/Digital Twin/Smart Factory |
| 10:40-12:10 | Session 17: Interactive Presentation Session III |
| 12:10-14:00 | Lunch (11:30-14:30) / The Canvas (GF) |
| 14:10-15:30 | Panel Discussion : Collaborative Research |

Wen, Dec. 20, 2023

| | |
|-------------|---|
| 10:00-12:00 | Visit to Industrial University of Hochiminh |
|-------------|---|

Contents



| | | |
|-----|--|-------|
| 1-1 | Improving Cybersecurity Education in line with the Defense Innovation 4.0 in Military Departments of Universities Jungho Eom, Kyungil Kim, Sangpil Cheon (Daejeon Univ., ROK) | 1-2 |
| 1-2 | Focus on Green Consumption Exploring Key Influencing Factors in the Intent-Behavior Gap Weijia Li (Jiaxing Univ., China), Ziyang Liu (Kyonggi Univ., ROK) | 3-5 |
| 1-3 | Visualization analysis of performing arts production by SNA method Chang Ryong Kim, Ohung Kwon, Joonho Hyeong, Jooyoung Kim, Donghyeon Seo, Jongryun Roh (Korea Institute of Industrial Technology, ROK) | 6-7 |
| 1-4 | A Fuzzy Ontology based Context-aware Traffic Control Method Eunju Lee, Ihnhan Bae (Daegu Catholic Univ., ROK) | 8-9 |
| 1-5 | Egress Initialization for Graph Convolution Network in Recommendation Systems Sang-min Lee, Kim Heechan, Jeon Janggun, Namgi Kim (KyonggiUniv., ROK) | 10-11 |
| 1-6 | Establishment and Operation of an Online Integration Test Environment for Verifying Heterogeneous Parallel Processor Software Seungwoo Rho, Gibeom Gu (Korea Institute of Science and Technology Information, ROK) | 12-14 |
| 1-7 | Integrated Dataset for 10 Facial Emotion Recognition Hoyoung Yoo, Hyeonjin Kim (Electronics & Telecommunications Research Institute, ROK) | 15-17 |

Contents



| | | |
|------|---|-------|
| 1-8 | Research on Customer Satisfaction in Cruise Ship Tourism Based on Big Data Analytics – An Empirical Analysis from OTA Website Rui Liang, Yu Feng (Kyonggi Univ., ROK) | 18-20 |
| 1-9 | Verification of Terror Factors Based on GTD data: Application of the RTM Model Eun-bi Cho, Sang-Eun Ahn, Jeong-Hyeon Chang (Kyonggi Univ., ROK) | 21-23 |
| 1-10 | Exploratory testing methodology based on MITRE ATT&CK and D3FEND Framework Gaeul Kim, Dohoon Kim (Kyonggi Univ., ROK) | 24-26 |
| 1-11 | Optimal Operation Plan for Unmanned Systems in Future Intelligent Battlefield Jungho Eom, Kyungil Kim (Daejeon Univ., ROK) | 27-30 |
| 2-1 | Experimental Results on Deep Learning-Based Defect Classification in Lithium-ion Battery Component Manufacturing Thi-Thu-Huyen Vu, Haejoong Kim, Taiwoo Chang (Kyonggi Univ., ROK) | 31-33 |
| 2-2 | Developing Human-Following Robot to Collaborate with Railroad Workers YeonSoo Kim, Tai-Woo Chang, Jong-Un Won, Baek-Hyun Kim (Kyonggi Univ., ROK) | 34-36 |
| 2-3 | Assessing Digital Transformation Levels of Partners for Supply Chain Collaboration Yong-Jun Im, Tai-Woo Chang (Kyonggi Univ., ROK) | 37-39 |
| 2-4 | Establishing a digital twin system using logistics simulation and BI tools – Case study in an automotive industry Gyusun Hwang, Jihyeon Park (Univ. of Ulsan, ROK), Haejoong Kim (Kyonggi Univ., ROK) | 40-42 |

Contents



| | | |
|-----|---|-------|
| 2-5 | Classification of fire risk small electric home appliances using deep learning Jun-Seong Kim, Yong Jun Im, Hyun-Soo Kim (Kyonggi Univ., ROK) | 43-45 |
| 2-6 | Real-time heuristics for balancing mixed-model assembly lines Minseok Kim (Korea National Univ. of Transportation, ROK), Hyeongjun Son, Byeong-uk Ha (Kyonggi Univ., ROK), Soyun Kim, Byeonghyun Cha (Korea National Univ. of Transportation, ROK), Haejoong Kim (Kyonggi Univ., ROK) | 46-47 |
| 3-1 | Analyzing Similar Redundancy Among AI Core Courses using Hugging Face Transformer and Hierarchical Analysis Tai Hoon Kim, In Seop Na (Chonnam National Univ., ROK) | 48-53 |
| 3-2 | Pedestrian Collision Risk Assessment Based on Pedestrian's Motion Prediction for Autonomous Vehicles Sunghwan Kim, Jaeseung Kim, Yeongwon Yu (Korea Univ., ROK), Ir. Made Sudarma (Udayana Univ., Indonesia), Deden Witarasyah Jacob (Telkom Univ., Indonesia), Minho Jo (Korea Univ., ROK) | 54-56 |
| 3-3 | A Facial Wrinkle Segmentation Method Based on Unet++ Model Junsuk Lee, Hyeonwoo Kim, Jonghwa Shim, Eunbeen Kim, Eenjun Hwang (Korea Univ., ROK) | 57-59 |
| 3-4 | Comparison of Multi-Person Action Recognition Methods with YOLOv7-Pose Jae hyun Rho, Youngwoo Lee, Hyunbeom Choi, Jeongwook Seo (Hanshin Univ., ROK) | 60-61 |
| 3-5 | Vision-Based Artificial Intelligence System for Detecting Abnormal Customer Behaviors at Unmanned Cafes Yoonseo Kim, Sehun Lee, Sunghyun Kang, Junho Ahn (Kyonggi Univ., ROK) | 62-64 |
| 4-1 | A Study of Automated Bitcoin Transaction Flow Analysis Method SangYong Kang, Hongbi Kim, Daeil Jang, SeungGu Ji (Korea Internet & Security Agency, ROK) | 65-67 |

Contents



| | | |
|-----|--|-------|
| 4-2 | Text Mining Analysis of Crime Prevention Through Environmental Design (CPTED) Project in Seoul: Focusing on LDA topic modeling Junho Park, Sooyun Lee, Euigab Hwang (Kyonggi Univ., ROK) | 68-70 |
| 4-3 | A Yolo based CPTED environmental factor detection model Tsagaantsooj Batzaya, Eunbi Cho, DoHyeon Lee, Jeong-Hyeon Chang, Kwanghoon Pio Kim (Kyonggi Univ., ROK) | 71-73 |
| 4-4 | BLIP-Based Model for Detecting Risky Situations in Video Eunbi Cho, Tsagaantsooj Batzaya, DoHyeon Lee, Jeong-Hyeon Chang, Kwanghoon Pio Kim (Kyonggi Univ., ROK) | 74-76 |
| 4-5 | Vision-Based User Abnormality Detection System for Unmanned Laundry Safety Jiyoung Na, Sunghyun Kang, Junho Ahn (Kyonggi Univ., ROK) | 77-79 |
| 4-6 | A Study on Person Re-identification with Feature Matching Methods Lanh Thanh Le (Dong Nai Technology Univ., Vietnam), Dinh-Lam Pham, Joo-Chang Kim, Sang-eun Ahn, Kyunghee Sun, Kyong-Sook Kim, Kwanghoon Pio Kim (Kyonggi Univ., ROK) | 80-82 |
| 4-7 | Reproduction of CCTV-Videos Using Contextualization Data Generated by Deep Learning Technology Kyunghee Sun, Sangeun Ahn, Kwanghoon Pio Kim (Kyonggi Univ., ROK) | 83-85 |
| 5-1 | Deep Learning-Based Indoor Object Detection and Area Inpainting from RGB-D Camera Wonseop Shin (Chung-Ang Univ.,ROK), Jaeseok Yoo (SweetK, ROK), SangHyun Byun, Sanghyun Seo (Chung-Ang Univ.,ROK) | 86-90 |
| 5-2 | When Pyramid Pooling Transformers Assists Lightweight Decoder for Salient Object Detection in Edge Intelligence Thien-Thu Ngo, Md. Delowar Hossain, Teh-Jen Sun, Yuri Seo (Kyunghee Univ., ROK), Hoang-Hai Tran (Hanoi Univ. of Science & Technology, Vietnam), Eui- Nam Huh (Kyunghee Univ., ROK) | 91-93 |

Contents



| | | |
|-----|---|---------|
| 5-3 | A Study on Problems of the Analysis of AI Medical Images Sun-Kuk Noh (Chosun Univ., ROK) | 94-96 |
| 5-4 | Image encryption based on the maximum-periodic chaotic sequence generator Sung-Won Kang, Kyung-Hyune Rhee (Pukyong National Univ., ROK) | 97-98 |
| 6-1 | Privacy-preserving credential smart contracts using Zokrates Geunyoung Kim, Hongseok Seo, Jaecheol Ryou (Chungnam National Univ., ROK) | 99-101 |
| 6-2 | Digital Vulnerabilities in QR-Based Photo Access: An Analysis and Countermeasures for Self-Photo Studio Systems Hong Seok Seo, Jina Hong, Jaecheol Ryu (Chungnam National Univ., ROK) | 102-104 |
| 6-3 | Comparative Analysis of Blockchain-Based Voting Systems Using Machine Learning Techniques Uzma Jafar, Mohd Juzaidin Ab Aziz, Zarina Shukur, Hafiz Adnan Hussain, Zulkefli Mansor (Universiti Kebangsaan Malaysia, Malaysia) | 105-110 |
| 6-4 | Artifact Analysis of Android Device connected to Hardware Wallet Jina Hong, Hongseok Seo, Jaecheol Ryu (Chungnam National Univ., ROK) | 111-113 |
| 7-1 | HPC Cluster-based Customized Container Image Manager and Builder Gukhua Lee (Korea Institute of Science & Technology Information, ROK) | 114-116 |
| 7-2 | Data-Driven Service Design Education Framework for New Product Development Seohee Kim, Hyemin Park, Yu Cheng, Inseop Lee, Sanghun Sul (Sungkyunkwan Univ., ROK) | 117-119 |

Contents



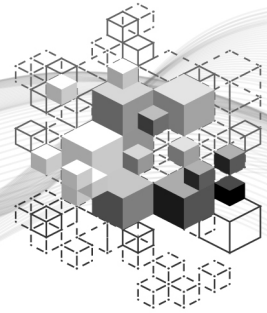
| | | |
|-----|---|---------|
| 7-3 | Derivation of Home Appliances Features with Stretchable Displays based on UX Seung Eun Chung, Youjin Seo, Han Young Ryoo (Ewha Womans Univ., ROK) | 120-123 |
| 7-4 | Analyzing the effects of alcohol-related image types on drinking efficacy and intentions on Instagram: An Eye-Tracking Approach Min hee Lee, Kyung Han You, Du Mengqi, Yifan Wang (Jeonbuk National Univ., ROK), Jeong Kyu Lee (Univ. of Oklahoma, USA) | 124-126 |
| 8-1 | Lightweight Deep Learning-based Product Object Classification Scheme Dongchil Kim, Kyeongeun Seo, Chang Mo Yang (Korea Electronics Technology Institute, ROK) | 127-129 |
| 8-2 | Smart Video Surveillance System using 5G Mobile Communication and Edge Server Chang Mo Yang, Dongchil Kim, Kyeongeun Seo (Korea Electronics Technology Institute, ROK) | 130-133 |
| 8-3 | Autonomous driving service effectiveness verification technology in perspective of laws and regulations Kyungwon Kim, Jongbin Park (Korea Electronics Technology Institute, ROK) | 134-136 |
| 8-4 | A Study for 3D Cross-Modal Retrieval using Self-supervised Learning Yeon-Seung Choo, Hyun-Sik Kim, Yong-Suk Park (Korea Electronics Technology Institute, ROK) | 137-138 |
| 8-5 | Design and Implementation of 3D Visualization Service Architecture for 2D CAD Data Seulki Song, Minju Cho, Jungwook Wee (Korea Electronics Technology Institute, ROK) | 139-141 |

Contents



| | | |
|------|--|---------|
| 8-6 | Deep Learning-Based Video Editing for Privacy Protection Yong-Suk Park, Yeon-Seung Choo, Hyun-Sik Kim (Korea Electronics Technology Institute, ROK) | 142-143 |
| 8-7 | Interaction Recognition in Video for Edge Devices Yong-Suk Park, HyoungSung Kim, Hyun-Sik Kim (Korea Electronics Technology Institute, ROK) | 144-145 |
| 8-8 | Metaverse Asset Creation Tool for Copyright Protection Yong-Suk Park, Yeon-Seung Choo, HyoungSung Kim (Korea Electronics Technology Institute, ROK) | 146-147 |
| 8-9 | Emotion Recognition Enhancement with Gesture Interpretation Yong-Suk Park, HyoungSung Kim, Yeon-Seung Choo (Korea Electronics Technology Institute, ROK) | 148-149 |
| 8-10 | Design and Implementation of a Communication Repeater Operating on UAV Frequencies for Unmanned Vehicles Sung-Chan Choi, Sungwook Jung, Jong-Hong Park, Ilyeop Ahn (Korea Electronics Technology Institute, ROK) | 150-151 |
| 8-11 | Designing a Pipeline to Leverage NFTs for Access Control in Web3.0 HyoungSung Kim, Yong-Suk Park, Hyun-Sik Kim (Korea Electronics Technology Institute, ROK) | 152-153 |
| 8-12 | Efficient Hierarchical Platform Structure Designs for Smart Aids-to-Navigation in Marine Environments Jaeho Lee, Sungyoon Cho, Wongi Jeon, Byoungchul Song, Yangsub Kim, Donghyun Kim, Junsu Kim, Kiwon Kwon (Korea Electronics Technology Institute, ROK) | 154-155 |
| 8-13 | Designing a Blender Add-On for Managing Creative Commons Zero Licensed 3D Assets HyoungSung Kim, Hyun-Sik Kim, Yong-Suk Park (Korea Electronics Technology Institute, ROK) | 156-157 |

Contents



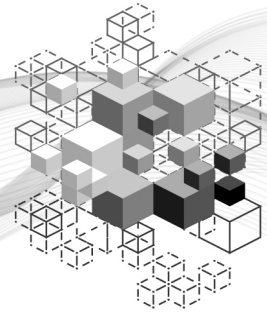
| | | |
|-----|---|---------|
| 9-1 | <p>A Novel Design of Shipment Decision Framework in a Land-based Fish-farm Management with Big-data Based Aquaculture Productivity, Price, and Costs Modeling</p> <p>Jaeho Lee, Juhyoung Sung (Korea Electronics Technology Institute, ROK), Jong Ho Paik (Seoul Women's Univ., ROK), Kiwon Kwon, Yangsub Kim, Kyungwon Park, Sungyoon Cho (Korea Electronics Technology Institute, ROK)</p> | 158-160 |
| 9-2 | <p>High Speed Template Matching System for Wafer Alignment based on FPGA</p> <p>Jae Hyuk So, Dong Hyun Lee, Tae min Hwang, Min joon Kim, Hye Won Kang (Korea Electronics Technology Institute, ROK)</p> | 161-163 |
| 9-3 | <p>AI model deployment management capabilities for oneM2M based AIoT system</p> <p>SeungMyeong Jeong, Chungjae Choe (Korea Electronics Technology Institute, ROK)</p> | 164-166 |
| 9-4 | <p>Sinusoidal Embedding for WiFi CSI-based Human Action Recognition</p> <p>Taehyeon Kim, Seho Park (Korea Electronics Technology Institute, ROK)</p> | 167-169 |
| 9-5 | <p>Melody Extraction for Non-Vocal Music based on Tone-Octave Network</p> <p>Dalwon Jang, Daeun Chae, Jaewon Lee, JongSeol Lee (Korea Electronics Technology Institute, ROK)</p> | 170-172 |
| 9-6 | <p>Study on Video Surveillance System Data Interface for Interoperability with Heterogeneous Physical Security System</p> <p>Sungjoo Park, Dongchil Kim (Korea Electronics Technology Institute, ROK)</p> | 173-175 |
| 9-7 | <p>A Study on Reactive Power Characteristics and Compensation According to Voltage Regulation in Low Voltage Premises System</p> <p>Seunghyeon Park, Kiwoong Kwon, Sanghun Kim, Yongho Kim (Korea Electronics Technology Institute, ROK)</p> | 176-178 |

Contents



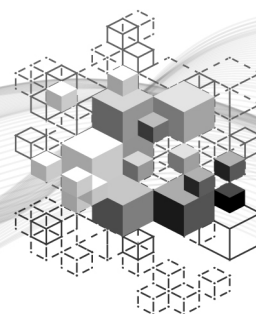
| | | |
|------|---|---------|
| 10-1 | Enhancing Blockchain Platform Security through the Analytics of User-Centric Psychometric Paradigm Jihye Park, Juno Lee, Hangbae Chang (Chung-Ang Univ., ROK) | 179-181 |
| 10-2 | Experimental Verification of Temporal Activity Sequencing Anomaly Detection Algorithm in Business Process Do-Hyeon Lee, Jin-Hyeong Lee, Eunbi Cho, Tsagaantsooj Batzaya, Thanh-hai Nguyen, Kwanghoon Pio Kim (Kyonggi Univ., ROK) | 182-185 |
| 10-3 | A Study on the Utilization of Blockchain for Ensuring Corporate Sustainability Based on the Protection of Personal Information Dabin Lee, Yeji Choi, Jaewook Byun, Hangbae Chang (Chung-Ang Univ., ROK) | 186-188 |
| 10-4 | The effects of users' perceived platform indispensability on their actual usage: A comparative study of Korea and Finland Moonkyoung Jang (Gachon Univ., ROK), Seongcheol Kim (Korea Univ., ROK) | 189-191 |
| 10-5 | A Structural Anomaly Detection Approach in Process Mining Min Jae Park (Daelim College, ROK), Kwanghoon Pio Kim (Kyonggi Univ., ROK) | 192-194 |
| 11-1 | A Case Study on Security Courses for Information Security Personnel Eunsun Choi, Jooyeon Han, Namje Park (Jeju National Univ., ROK) | 195-197 |
| 11-2 | Analysis of Hallyu research trends using text mining Miyoung Kim (Suwan Middle School, ROK), SunJu Park (Gwangju National Univ. of Education, ROK) | 198-200 |
| 11-3 | Comparative Analysis of Chinese Personal Information Protection Textbooks Eunsun Choi, Pureum Kim, Mijin Kim, Yujin Jung, Jinsu Kim, Namje Park (Jeju National Univ., ROK) | 201-203 |

Contents



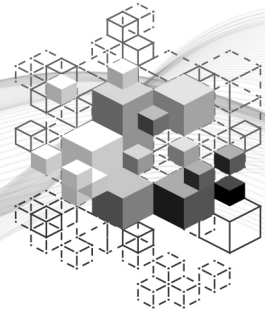
| | | |
|------|---|---------|
| 11-4 | <p>Perception Study on Digital Literacy Teachers Should Possess Through Text Mining Analysis</p> <p>Eunsun Choi, Hyunjeong Lim, Keunbae Choi, Heuipi Kim, Ji-Yun Kim, Hosoo Lee, Yeon-Soo Joo, Namje Park (Jeju National Univ., ROK)</p> | 204-206 |
| 11-5 | <p>Analysis of the Effectiveness of SW•AI Education Camp for Elementary School Students</p> <p>Jaekwoun Shim, Wooyeol Kim, Youngkwon Bae (Daegu National Univ. of Education, ROK)</p> | 207-210 |
| 12-1 | <p>Position Estimation using Stereo Camera on AR Glass and 3D Reconstructed Map</p> <p>Sukwoo Jung, Youn-Sung Lee, Yunju Lee, KyungTaek Lee (Korea Electronics Technology Institute, ROK)</p> | 211-212 |
| 12-2 | <p>An Autonomous Control System Implementation for Maintaining Dissolved Oxygen Level in Aquaculture</p> <p>Juhyoung Sung, Sungyoon Cho, Wongi Jeon, Yangseob Kim, Kiwon Kwon (Korea Electronics Technology Institute, ROK)</p> | 213-215 |
| 12-3 | <p>Homography-based Object Position Tracking Method using Fixed Camera</p> <p>Youn-Sung Lee, Sukwoo Jung, Kyung-Taek Lee (Korea Electronics Technology Institute, ROK)</p> | 216-217 |
| 12-4 | <p>An Analysis of the Imported Ornamental Fish Quarantine Environment for CNN based Recognition of Ornamental Fish Species in Korea</p> <p>Hyun-Hak Song, Sungyoon Cho, Yangsub Kim, Wongi Jeon, Kyungwon Park, Byoungchul Song (Korea Electronics Technology Institute, ROK), Jongho Paik (Seoul Women' s Univ., ROK), Kiwon Kwon (Korea Electronics Technology Institute, ROK)</p> | 218-219 |
| 12-5 | <p>Automatic Calibration of Multiple LIDAR using Plane in Vehicle Intersection Environment</p> <p>Hyunhak Song, Byoungchul Song, Kyungwon Park (Korea Electronics Technology Institute, ROK)</p> | 220-221 |

Contents



| | | |
|-------|--|---------|
| 12-6 | A Reconstruction of Korean Traditional Performance in Virtual Reality through 3D Pose Estimation across Multi-views Taemin Hwang, Myoungjin Kim, Jaehyuk So, Minjoon Kim (Korea Electronics Technology Institute, ROK) | 222-223 |
| 12-7 | Spatial Information Scanning Based on a Mems Mirror for Mobile 3D Scan Device Young Bo Shim, Sung Hun Chae (Korea Electronics Technology Institute, ROK) | 224-225 |
| 12-8 | Range Measurement Technique Based on VCSEL Source and SiPM Receiver Young Bo Shim, Sung Hun Chae (Korea Electronics Technology Institute, ROK) | 226-227 |
| 12-9 | Implementation of a Standard Extension Interface Module for Real-Time Control of Generative Art Content Sunghun Chae, Young Bo Shim (Korea Electronics Technology Institute, ROK) | 228-229 |
| 12-10 | Emotion and Theme Recognition by combining multiple classifiers Dalwon Jang, Hanyi Shin, Jaewon Lee, JongSeol Lee (Korea Electronics Technology Institute, ROK) | 230-231 |
| 12-11 | Comparison of object recognition performance using Embedded AI SoC Seho Park, SaetByul Yu (Korea Electronics Technology Institute, ROK) | 232-234 |
| 12-12 | Metric Learning for Gait Recognition with WiFi Sensing Seho Park, Myeongseop Kim, Taehyeon Kim (Korea Electronics Technology Institute, ROK) | 235-237 |
| 12-13 | Method of Operation and Control of Detection Equipment for Patrol Drone-Cop System Sungwook Jung, Jonh-Hong Park, Sungchan Choi, Wonseok Jung, Il-yeup Ahn (Korea Electronics Technology Institute, ROK) | 238-239 |

Contents



| | | |
|------|--|---------|
| 13-1 | Task Offloading Strategy Based on Different Network Planning Approaches in Edge Computing Networks Md. Delowar Hossain, Tangina Sultana, Sharmen Akhter, Md. Imtiaz Hossain, Md Nosin Ibna Mahbub, Ngo Thien Thu, Ga-Won Lee, Eui-Nam Huh (Kyung Hee Univ., ROK) | 240-245 |
| 13-2 | A Research on the Visualization Services in BIM-Based Virtual Reality Environments Kijung Ryu, Joonyoung Ahn, Dongho Kim (Soongsil Univ., ROK) | 246-248 |
| 13-3 | Dynamic Access Control Policy Based on Blockchain and Machine Learning to Enhance Security in IoT Smart Homes Hafiz Adnan Hussain, Zulkefli Mansor, Zarina Shukur, Uzma Jafar (Universiti Kebangsaan Malaysia, Malaysia) | 249-253 |
| 13-4 | A Study On Adaptation Tasks Between Suitable DBMS For Cloud Migration Joonyoung Ahn, Changik Oh, Sugil Hwang, Dongho Kim (Soongsil Univ., ROK) | 254-256 |
| 13-5 | AI Companion for Children: A Service Architecture to aid Social-Emotional Development Seunghee Han, Soo Jung Chang, Keon Chul Park (Advanced Institute of Convergence Technology, ROK) | 257-259 |
| 13-6 | A study on college students' telemedicine acceptance Semi Han, Sangmin Lee (CHA Univ., ROK) | 260-262 |
| 13-7 | A Study On The Access Devices For The Activation Of Web 3.0 Joonyoung Ahn, Dongho Kim (Soongsil Univ., ROK) | 263-265 |
| 14-1 | Development of Elementary Information Education Program based on SDGs Jaekwoun Shim, Youngkwon Bae, Wooyeol Kim (Daegu National Univ. of Education, ROK) | 266-269 |

Contents



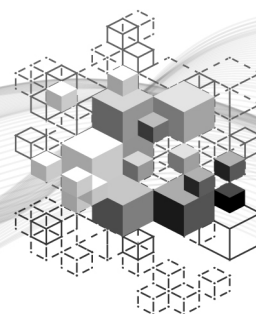
| | | |
|------|--|---------|
| 14-2 | The U.S. Pre-service Teachers' Perspectives and Analysis of NCTM Illuminations: Interactives Yong Joon Park (Indiana State Univ., USA) | 270-272 |
| 14-3 | Development of Automated English Speech Evaluation System Using Machine Learning Woochun Jun (Seoul National Univ. of Education, ROK) | 273-274 |
| 14-4 | A survey on the perception of prospective teachers after learning EPL using the tinkering method Byoung Rae Han (Chinju National Univ. of Education, ROK) | 275-277 |
| 15-1 | Q-Learning-based Access Priority Control in Internet of Medical Things Jung-Hyok Kwon, Jaehoon Park, Eui-Jik Kim (Hallym Univ., ROK) | 278-279 |
| 15-2 | The Spillover Effects of US Economic Policy Uncertainty on China's Economy: Monetary Policy Uncertainty vs. Fiscal Policy Uncertainty Chi Gong (Beijing Normal Univ., China), Shang Wu (Sichuan Univ., China), Hong Jin (China Univ. of Political Science and Law, China), Xiaoye Lu (Beijing Normal Univ., China) | 280-286 |
| 15-3 | A Study on the Improvement of Pet Registration System Using Zero-Knowledge Proof Joonyoung Ahn, Kijung Ryu, Yunhoo Na, Dongho Kim (Soongsil Univ., ROK) | 287-289 |
| 15-4 | Thin Graph Convolution Network in Recommendation Systems Sang-min Lee, Kim Heechan, Jeon Janggun, Namgi Kim (Kyonggi Univ., ROK) | 290-291 |
| 15-5 | A Study On Conversion Vector Between Suitable DBMS Operated As A Cloud Service Joonyoung Ahn, Changik Oh, Sugil Hwang, Dongho Kim (Soongsil Univ., ROK) | 292-293 |
| 15-6 | Egocentric Human Pose Estimation for VR Interactions Using RGB-D Data Seongmin Baek (Electronics and Telecommunications Research Institute, ROK), Yejin Kim (Hongik Univ., ROK) | 294-296 |

Contents



| | | |
|------|---|---------|
| 15-7 | Over-smoothing in Light GCN doesn' t happen much Sangmin Lee, Heechan Kim, Janggun Jeon, Namgi Kim (Kyonggi Univ., ROK) | 297-298 |
| 15-8 | Analyzing the Financial Effects of Smart Factory Support Program using Propensity Score Matching Sangun Park, Tai-Woo Chang (Kyonggi Univ., ROK) | 299-301 |
| 16-1 | Comparing Scalable Multiclass Classification Models to Rate Amazon Books Hsiu-Ping Lin (California State Univ., Los Angeles, USA), Min-Goo Kang (Hanshin Univ., ROK), Jongwook Woo (California State Univ., Los Angeles, USA) | 302-304 |
| 16-2 | ViT-based Classification with Digital Signal Processing and Image Transformation of Time Series Data Gyu-Il Kim, Seo-El Lee, Ji-Won Baek, Hyun Yoo, Kyungyong Chung (Kyonggi Univ., ROK) | 305-307 |
| 16-3 | NLP Models Classifying Helpful Ratings in OpenTable Dataset Sunnyoung Lee (Dongguk Univ., ROK), Hsiu-Ping Lin (California State Univ., LA, USA), Jihyun Park, Eugene Lim (Dongguk Univ., ROK), Jongwook Woo (California State Univ., LA, USA) | 308-310 |
| 17-1 | Study on Classification of mobile application based on mobile log data Kyung Han You, Byung Hoon So, Minhee Lee (Jeonbuk National Univ., ROK), Jaehee Cho (Sogang Univ., ROK) | 311-314 |
| 17-2 | An Operator Task Model for Constructing Human-in-the-loop Simulation in Manufacturing Systems Byeonkuk Kim, Gibaek Kim, Hyeonsu Lee, Dongmin Shin (Hanyang Univ., ROK) | 315-317 |
| 17-3 | A Study of Stochastic Simulation Algorithm for Bond Pricing Model Kisoeb Park, Moonseong Kim (Seoul Theological Univ., ROK) | 318-319 |

Contents



| | | |
|------|--|---------|
| 17-4 | Digital Transformation and its Implications for Technological Innovation in China's Cultural Industry: A Panel Data Analysis Qilong Shi (Anshan Normal Univ., China) | 320-324 |
| 17-5 | Does the Brain's Pre-existing Information Guide Our Explore Gaze? Ji-Chan Yun, Yeong-Woo Lim, Junsik Lee, Kyunghee Kim, Jae min Kim, Sanghoon Lee, Do-Hyung Park (Kookmin Univ., ROK) | 325-327 |
| 17-6 | Effect of Fan Systems for Lowering Leaf Temperature of Korean Ginseng Byungsoon Kim (Andong National Univ., ROK), Myung Hwan Jang (Punggi Ginseng Research Institute, ROK), Ilkyeun Ra (Univ. of Colorado, Denver, USA) | 328-329 |
| 17-7 | A Compound Gateway-Activity Discovery Approach in Process Mining Thanh-Hai Nguyen, Jin-Hyung Lee, Batt Chaya, Eun-Bee Cho, Sang-Eun Ahn, Kyung-Hee Sun, Kyoung-Sook Kim, Dinh-Lam Pham, Kwanghoon Pio Kim (Kyonggi Univ., ROK) | 330-332 |
| 17-8 | Chat-GPT application for AICC of the Bank Sudan Prasad Uprety (AICC Solution Lab, Nepal), Yoosin Kim (Airdeep Ltd., ROK) | 333-334 |
| 17-9 | Application of Block-Chain Technology in Digital Music Copyright Management: A Case Study of VNT Chain Platform Mingsheng Wang (Anshan Normal Univ., China) | 335-339 |
| | | |
| | | |

Improving Cybersecurity Education in line with the Defense Innovation 4.0 in Military Departments of Universities

Jungho Eom, Sangpil Cheon and Kyungil Kim

Department of Security Convergence, Daejeon University

Daejeon, South Korea

[e-mail: eomhun@gmail.com, skyfeel69202@naver.com, kimki@itzone.co.kr]

*Corresponding author: Jungho Eom

Abstract

This paper presents a cybersecurity education plan in the military department of at universities to acquire knowledge and technology in the cybersecurity domain in terms of building AI-based core advanced power among the 5 major strategies for Defense Innovation 4.0. The key strategy to secure AI-based core advanced forces establish an AI-based next-generation command&control(C2) system and manned/unmanned complex combat system, and strengthen operational capabilities in the future space, cyber, and electronic warfare domain. In order to effectively operate manned/unmanned complex combat systems, space, cyber, and electronic weapon systems in the future multi-domain battlefield, it is necessary to have cybersecurity knowledge and technology based on advanced information and communication technology. While military academies that train primary grade officers systematically run a curriculum related to cyber warfare or security, the proportion of courses related to cyber security is not high in the curriculum of military departments at universities. Therefore, we propose a cybersecurity education plan using minor course, nanodegrees, and extracurricular programs.

Keywords: Defense Innovation 4.0, Cyber Warfare, Advanced ICT, Cybersecurity Education

1. Introduction

In the future battlefield that is a multi-domain battlefield including space, cyber, and electronic domains, advanced ICT and defense science technologies based intelligent weapon systems will be operated. In order for these weapon systems to operate effectively, the units system must be protected from errors and cyberattacks. In other words, military junior leaders who will lead the future battlefield must acquire knowledge and skills to protect systems. Therefore, this research proposes curriculum for cybersecurity education in the military department of a university.

2. Related Works

2.1 Defense Innovation4.0

In Defense Innovation 4.0, the promotion strategy proposes constructing AI-based core advanced military power, redesigning of the defense R&D/force build-up system, and innovating military structure and education & training. This means that it will also train manpower capable of operating advanced power based on advanced ICT technology and defense science and technology. Among these, in order for the advanced military power system to operate effectively, it must be protected from

cyber threats occurring internally and externally [1].

2.2 Cybersecurity-related Curriculum in Military Department

Checking at the military department curriculum at universities, it focuses on humanities and social science subjects rather than science and engineering subjects. Military studies is composed of humanities and social sciences subjects such as military theory/strategy/policy and military power construction/support/operation and so on [2]. Although each university is different, cybersecurity-related subjects are organized 1~3 subjects at each university, including cyber information electronic warfare or military security theory. Because future battlefields will take intelligent warfare based on advanced ICT and science & technology, there is a need for subjects that allow primary grade officers to acquire security knowledge and skills for the operation of such intelligent equipment and systems.

3. "1+1 of 3 Ways" Curriculum

We propose a "1+1 of 3 ways" curriculum like Fig. 1 for cybersecurity education. We considered several factors in the proposed method. First, it maintains the existing curriculum, but consider the management capabilities of the university and department. Second, it is considered the future battlefield environment, priority will be given to security subjects related to drones, robots, space, etc. Lastly, teaching method is focused on practice rather than theory education.

Among the methods proposed in this research, the first is that students who want to pursue a career in a branch specialized in information and communications select computer engineering or security-related departments as a double major among sub-majors. If student choose a career in another branch, he can choose information and communications as your second specialty, by taking a minor or nanodegree. Second, if the university or department has sufficient operating budget, it is recommended to develop and operate a cyber security-related extracurricular program. The extracurricular program has the advantage of being able to be organized into essential security subjects for students to protect

advanced systems&equipment, and intelligent weapon systems operated on future battlefields. Finally, after checking the existing curriculum, unnecessary subjects or subjects that can be integrated/converged are adjusted and cyber security-related subjects are added. At this time, related subjects such as wireless security, artificial intelligence security, space security, and electronic security must be added in consideration of the future intelligent battlefield environment.

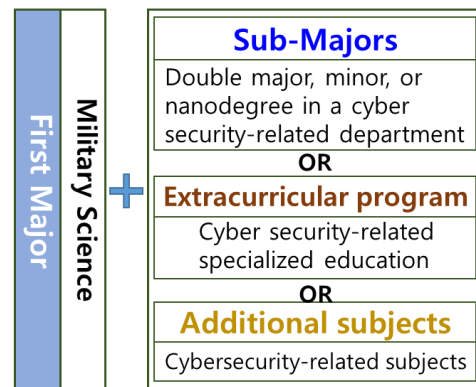


Fig. 1. "1+1 of 3 ways" education course

4. Conclusions

Cybersecurity education in the military department proposed in this paper was approached from the perspective that it is a necessity, not an option, in the future intelligent battlefield environment. If military students choose one of the three methods to acquire cybersecurity, there will be no shortage of advanced intelligent equipment, system operation, and security capabilities in the military.

References

- [1] "Defense redesign at the level of the 2nd Founding forces, fostering a strong AI science and technology force, Ministry of National Defense press release, 2023.
- [2] Jung-Ho Eom, Keun-Seog Park, Sang-Pil Chun, "A Redesign of the Military Education Structure of General Universities based on Defense Innovation 4.0," *Journal of Information and Security*, Vol.22, No.4, pp. 35-44, 2022.

Focus on Green Consumption Exploring Key Influencing Factors in the Intent-Behavior Gap

Weijia Li and Ziyang Liu

Department of Business, Jiaxing University
Jiaxing, China

Department of Global Business, Kyonggi University
Suwon, South Korea

[e-mail: weijia1006@naver.com]

*Corresponding author: Ziyang Liu

Abstract

This paper delves into the psychological pathways of green intention formation and analyzes the underlying causes of the intention-behavior gap. Study One focuses on pivotal factors influencing green consumer intention, while Study Two primarily scrutinizes the factors impacting the gap between intention and actual behavior. Conducting a survey with 732 consumers and employing both questionnaire and grounded theory analysis, we utilized SPSS 27.0 and AMOS 24.0 software for data analysis and hypothesis testing. The findings reveal that subjective norms, personal norms, perceived behavioral control, and attitude exert a direct positive influence on consumer intention. Conversely, perceived costs, ingrained habits, and situational factors emerge as pivotal obstacles hindering the translation of intention into behavior. This study aspires to provide robust support for theoretical research and offer theoretical and empirical guidance for enterprise production and marketing.

Keywords: Green consumption, Intention-behavior gap, Comprehensive Action Determination Model, In-depth interviews, Grounded Theory analysis

1. Introduction

As the global environmental concerns intensify, society's focus on the significance of sustainable consumption behavior is growing. Scholars argue that addressing these environmental issues can be achieved through modifying individuals' consumption behavior[1]. Urgent changes in current consumption habits are deemed necessary for a safer and healthier lifestyle for future generations. Green consumption has emerged as a crucial trend in current and future consumer landscapes, attracting attention across various industries[2]. The essence of green consumption lies in consumers' actions of purchasing and using products and services that

benefit the environment and possess sustainable production capabilities. However, green consumption extends beyond mere purchasing; it encompasses choices related to sustainable lifestyles, such as energy conservation, waste reduction, and support for environmental projects.

In the study of green consumption, understanding consumers' willingness and actual behavior is a crucial aspect. This research defines green consumption as an individual's attitude and intention to willingly adopt a series of environmentally friendly behaviors and purchase eco-friendly products, reflecting a desire for resource conservation, environmental-friendly consumption, green purchasing, and moderate

consumption to promote sustainable living. Factors influencing green consumption are currently focused on two aspects: the content or types of green consumption and the influencing factors of green consumption, including population characteristics, psychological factors, product-related factors, external contextual factors, etc[3].

However, green consumption faces challenges such as high costs, information asymmetry, and the availability of green products[4]. Despite consumers expressing concern for resources and the environment, there is often a disparity between green intentions and actual behavior, termed the attitude-behavior gap. Therefore, studying consumers' motives, attitudes, and behavior in green consumption, as well as addressing this gap, holds significant importance.

In summary, this paper, grounded in the Theory of Planned Behavior and the Value-Belief-Norm Theory, employs a mixed-method approach to deeply explore the reasons influencing consumers' green consumption intentions and the causes of the intention-behavior gap. Understanding the factors influencing green consumption intentions helps reveal consumers' motivations and values. Investigating why intentions fail to translate into behavior provides theoretical support and empirical guidance for green consumer behavior theory and production marketing for green enterprises. Specifically, the research objectives are: (1) constructing a model of factors influencing green consumption intentions based on mature theoretical foundations and consumer characteristics; (2) identifying key factors influencing the conversion of green consumption intentions into behavior through in-depth interviews; (3) proposing feasibility recommendations for promoting green consumption behavior across society based on empirical analysis results.

2. Reconstruction of Volumetric Models

2.1 Study 1

The hypotheses for Study 1 are shown below:

H1: Subjective norm has a significantly positive influence on personal norm

H2: Subjective norm has a significantly positive influence on consumption intention

H3: Perceived behavioral control has a significantly positive influence on personal norm

H4: Perceived behavioral control has a significantly positive influence on consumption intention

H5: Attitude has a significantly positive influence on consumption intention

H6: Personal norm has a significantly positive influence on consumption intention

H7: Biocentric value has a significantly positive influence on awareness of consequence

H8: Awareness of consequence has a significantly positive influence on responsibility attribution

H9: Responsibility attribution has a significantly positive influence on personal norm

H10: Awareness of consequence has a significantly positive influence on subjective norm

H11: Awareness of consequence has a significantly positive influence on attitude

H12: Biosphere value has a significantly positive influence on environmental self-identity

H13: Environmental self-identity has a significantly positive influence on responsibility attribution

H14: Environmental self-identity has a significantly positive influence on subjective norm

H15: Environmental self-identity has a significantly positive influence on perceived behavioral control

H16: Environmental self-identity has a significantly positive influence on attitude

2.2 Study 2

The hypotheses for Study 2 are shown below:

H1: Perceived costs have a significantly positive influence on the intention-behavior gap.

H2: Contextual factors have a significantly positive influence on the intention-behavior gap.

H3: Inherent habits have a significantly positive influence on the intention-behavior gap.

3. Data analysis and results

The method employed for data collection in this research is the questionnaire survey, which serves as a cost-effective and efficient means of gathering a substantial amount of information

and data. Through this method, the perspectives, opinions, and experiences of the respondents can be acquired, ensuring the collection of the necessary information and data. The survey instrument, designed for this research, follows the questionnaire format and is based on well-established scales from previous studies. Modifications and customizations were applied to the selected scales to align with the specific objectives and content of this study. The questionnaire comprises three main sections: the Scale of Factors Influencing Intention toward Green Consumption Behavior, the Scale of Intention-Behavior Discrepancy in Green Consumption, and Personal Demographic Information. Subsequent to the design of the survey questionnaire, a randomized sample of Chinese consumers was solicited and the questionnaires were distributed and collected through online platforms.

Employing the methodology of individual in-depth interviews to gather textual data, followed by a grounded theory analysis, we deeply delve into the underlying factors influencing consumers' transformation of green intentions into actual consumption behavior. This elucidates the gap between green consumption intentions and actions.

Utilizing the SPSS 27.0 software, we conducted a descriptive statistical analysis and examined the reliability and validity of the sample data. This aimed at gaining insights into the respondents' fundamental characteristics while ensuring the dependability and efficacy of the sample data. Subsequently, employing AMOS 24.0, we crafted a structural equation model and scrutinized the model's fit and hypothesis outcomes through an extensive empirical dataset. The findings revealed the validation of assumptions in both Study One and Study Two.

4. Conclusions

In consumer behavior research, purchase intention is regarded as a pivotal determinant of predicting buying behavior. Research one integrates various theoretical frameworks such as the Theory of Planned Behavior (TPB), the Norm Activation Model (NAM), and the Value-Belief-Norm Theory (VBN). It deeply explores the determinants of green consumption intention from different theoretical perspectives, extending existing frameworks, enhancing

theoretical applicability, and emphasizing the intricacies and diversities involved in the formation of green consumption intentions. Research two conducts a thorough analysis of factors influencing the "intention-behavior gap," investigating why people's intentions to engage in green consumption fail to translate into actual behavior. It particularly emphasizes the impact of perceived costs, inherent consumption habits, and situational factors on this gap, holding crucial significance for the practical promotion of sustainable consumption. This comprehensive study aids in filling gaps in existing research, providing a more comprehensive perspective for the theoretical framework of consumer behavior. By delving into factors influencing the "intention-behavior gap," a better understanding of obstacles and challenges consumers face in realizing green consumption is achieved, thereby offering targeted strategies and recommendations for promoting sustainable consumption.

While this study has achieved significant results in exploring the relationship between green consumption intentions and behavior, there are still limitations that require further exploration and refinement in future research.

References

- [1] Li, W. J., Jin, M. X., & Liu, Z. Y., "Green Food Purchase Behavior and Subjective Well-being Factors from the Perspective of Experiential Marketing," *Journal of the Korean Content Society*, Vol 23, No. 9, pp. 356-369, 2023.
- [2] Yan, L., Keh, H. T., and Chen, J., "Assimilating and differentiating: the curvilinear effect of social class on green consumption," *Journal of Consumer Research*, Vol 47, No. 6, pp. 914-936, 2021.
- [3] Shi, J., & Jiang, Z., "Willingness to pay a premium price for green products: does a reference group matter?," *Environment, Development and Sustainability*, Vol 25, No. 8, pp. 8699-8727, 2023.
- [4] Marcon, A., Ribeiro, J. L. D., Dangelico, R. M., de Medeiros, J. F., and Marcon, E., "Exploring green product attributes and their effect on consumer behaviour: A systematic review," *Sustainable Production and Consumption*, Vol 32, pp. 76-91, 2022.

Visualization analysis of performing arts production by SNA method

**Chang Ryong Kim, Ohung Kwon, Joonho Hyeong, Jooyoung Kim, Donghyeon Seo
and Jongryun Roh**

Digital Transformation R&D Department, Korea Institute of Industrial Technology
Gyeonggi, South Korea

[e-mail: ckdfyd52@kitech.re.kr, ohung@kitech.re.kr, freegore@kitech.re.kr,
jjjoshk@kitech.re.kr, dhseo@kitech.re.kr, ssaccn@kitech.re.kr]

*Corresponding author: Jongryun Roh

Abstract

This study is about analyzing work patterns by performing arts production process period for the classification of performing arts platform functions and the design of main functions in stages. In this study, data from communication tools used in the actual K-pop concert production process were collected, and based on this, data coding was performed for each business category using the Open Coding method. After coding the data by business category, we analyzed the communication tools used in the production of two concerts using social network analysis (SNA) techniques. This analysis visualized the key work patterns of each concert, examining the patterns across various production stages, including preparation for performances, rehearsals, and concert days. In this paper, we analyzed the performing arts platform's functionality, focusing on response timings and forms across concert preparation stages. Through data-driven analysis of actual concert productions, we integrated specific functions into our system, improving work efficiency and identifying alternative operational methods.

Keywords: Concert Production Process, Communication Tool, Cooperation, Social Network Analysis, Data Visualization

1. Introduction

In the process of producing performing arts, various communication tools are currently employed for distinct tasks such as communication, scheduling, and archiving. The lack of integration between tools makes effective collaboration difficult. A comprehensive and integrated platform is needed to efficiently manage projects and prevent confusion among various stakeholders in performing arts production, and it is important to understand the process and key tasks throughout the production process to develop these platforms [1].

2. Approach & Method

2.1.1 Open Coding

Open Coding refers to an analytical process, predominantly utilized in qualitative research, notably within the Grounded Theory methodology. This analytical process involves systematically reviewing raw data, such as interview transcripts or observational notes, to identify and label variables, themes, patterns, and categories [2].

2.1.2 Follow-up Task

The concept of "follow-up task" is devised to understand the relationships between tasks. It refers to tasks that are triggered and ensue in response to the completion of a particular task.

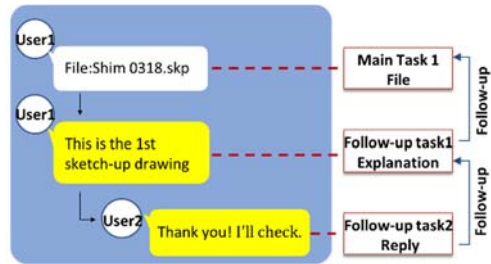


Fig. 1. The conceptualization of a follow-up task

2.2 Social Network Analysis

SNA is used to visualize and analyze structures and patterns within a network, focusing on objects (nodes) and their relationships (links) [3]. Using SNA (Social Network Analysis), it's possible to identify work structures and patterns, such as interactions between team members and information flow within the organization [4]. Specifically, it aids in understanding the importance and role of specific tasks within the network.



Fig. 2. Task process log-based organizational structure based on SNA

3. Visualization & Analysis of Work Patterns in Performing Arts Production (Focused on K-Pop Concerts)

Using Social Network Analysis (SNA) techniques, the communication tools utilized at two concert production sites were examined. This enabled a visualization and analysis of the overarching work patterns for each concert, as well as a breakdown of patterns according to distinct production stages: preparation, rehearsal, and the day of the concert.

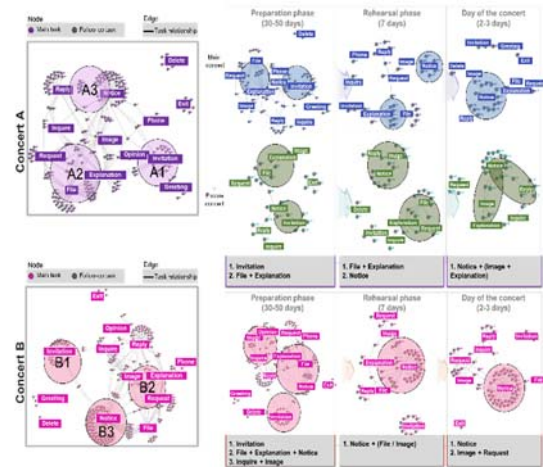


Fig. 3. Visualization of task patterns during performing arts production

4. Conclusions

In this paper, we were able to identify the timing and forms of responses in the functionality of the performing arts platform, through visualizing and analyzing task patterns according to the stages of concert preparation. Furthermore, we enhanced work efficiency in the performing arts domain by incorporating functions into our system and pinpointing alternative functions through a data-driven examination of real concert productions.

References

- [1] Shakeri, Hedieh, and Mohammad Khalilzadeh, "Analysis of factors affecting project communications with a hybrid DEMATEL-ISM approach (A case study in Iran)." *Heliyon* Vol. 6,8 e04430. 4 Aug. 2020.
- [2] Barney Glaser, Anselm Strauss, "Discovery of Grounded Theory: Strategies for Qualitative Research," *Routledge*, 2017.
- [3] Wasserman, S. and Faust, K., "Social Network Analysis: Methods and Applications," *Cambridge University Press*, 1994.
- [4] Seong-Cheol Hong, Yonghyuk Lee, Jinwoo Kim, and Injun Choi, "An methodology for redesigning an organizational structure based on business process models using SNA Techniques," *International journal of innovative computing*, 2012.

A Fuzzy Ontology based Context-aware Traffic Control Method

Eunju Lee and Ihnhan Bae

School of Computer Software, Catholic University of Daegu
Gyeongbuk, South Korea

[e-mail: becom7566@naver.com, ihbae@cu.ac.kr]

*Corresponding author: Ihnhan Bae

Abstract

This paper presents a fuzzy ontology based context-aware traffic control method for autonomous vehicles. The proposed a FOCATC4K provides the traffic control model for the intersection with the most complex traffic situation and the underpass with causing the greatest damage by fuzzy ontologies and generates the traffic control rules for those. The FOCATC4K is a traffic control model that is only suitable for Korea because the model is built according to the intersection layouts and traffic regulations of Korea.

Keywords: Autonomous vehicle, driving context information, fuzzy logic, fuzzy ontology, traffic regulation, web ontology language.

1. Introduction

The complexity of autonomous vehicles and the huge variety of possible scenarios, lead to a significant increase of required development and testing efforts. One of the requirements to an autonomous vehicle is that its behavior has to be in line with current traffic regulations at all times [1].

Ontologies have been used in the field of autonomous vehicles. M. A. Mohammad et al [2] proposed a novel ontology design for intelligent controlling of traffic signal. M. I. Alipo et al [3] developed an ArRoad system that monitors and analyzes vehicle traffic and flooded area then predicts possible reroute areas.

In this paper, we propose a Fuzzy Ontology based Context-Aware Traffic Control for Korea (FOCATC4K). The proposed FOCATC4K is a framework to allow modular fuzzy ontology building for the development of automated vehicles.

2. Related Works

2.1 Fuzzy Ontology

Fuzzy ontologies are an extension of the domain of crisp ontologies for solving the uncertainty

problems. To represent formally the fuzzy knowledge, the concepts in the ontology with a degree of fuzzy defined based on fuzzy logic [4, 5]. Fig. 1 presents an illustrative example of a fuzzy concept and its set of fuzzy property.

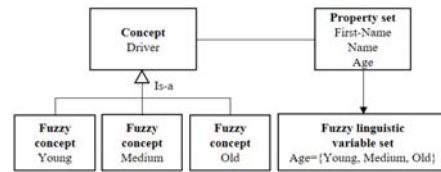


Fig. 1. Fuzzy concept of driver

2.2 Driving Context Information

In context-aware vehicular networks, any information that describes the driving situation is called driving context information. The context information has two components: the static context and the dynamic context. The elements of static context are not updated in real-time. But, the elements of dynamic context require to be able to update the context dynamically [6].

3. FOCATC4K

We design the FOCATC4K that is a fuzzy ontology model of a traffic flow management system for intersections and underpasses. Fig. 2

shows the system structure of FOCATC4K, where context-aware service management provides a wide range of transportation applications: collision warning, road hazard condition notification, traffic condition notification, etc.

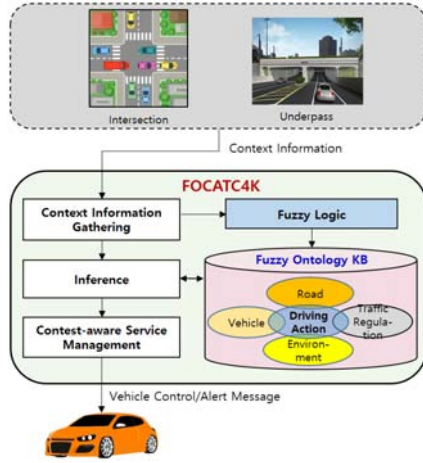


Fig. 2. System structure of FOCATC4K

The proposed FOCATC4K is built with fuzzy ontology and shown in Fig. 3, where Visibility, Weather, Pedestrian, Road, Tailgating and so on have fuzzy ontology subclasses.

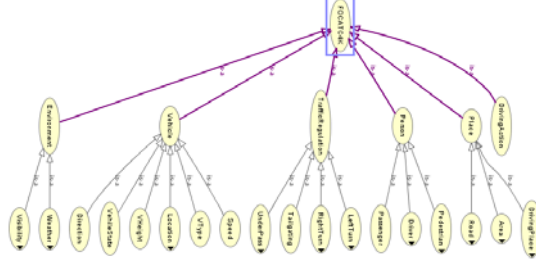


Fig. 3. OWL based fuzzy ontology for FOCATC4K

Fig. 4(a) shows the fuzzy membership function for tailgating according to the driving distance (DD_x) in an intersection, where L_x is the length of the intersection. DD_x can be computed by Eq.(1), where v_{ic} , T_{gs} and t_c represent the vehicle speed entering the intersection, green signal period, the elapsed time within green signal period, respectively.

$$DD_x = v_{ic} \times (T_{gs} - t_c) \quad (1)$$

Fuzzy rules for traffic congestions (TC) according to road occupancy (CR) and vehicle speed (VS) ratios are illustrated in Fig. 4(b), where CR, $VS \in \{\text{Low, Medium, High}\}$ and $TC \in \{\text{Smooth, Mild-Congested, Congested, Heavy-Congested}\}$ are fuzzy linguistic variables.

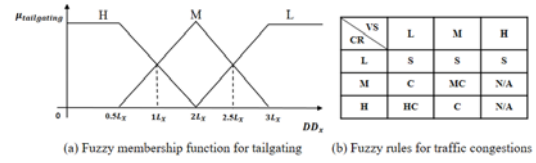


Fig. 4. Fuzzy logics for FOCATC4K

Table 1 shows the FOCATC4K based JessTab rule for an underpass entry in the roads near to Catholic University of Daegu.

Table 1. Example of FOCATC4K based JessTab rule

```
(defrule TCRule1 (Driver LastName ?x) (hasVehicle ?x Location)
(hasPlace Location ?y) (hasArea ?y HaYang) (hasWeather HaYang
LightRain) (hasRoad ?y Underpass) (hasSection ?y ExternalStraightLane)
(TrafficCongestion smooth) (< VHeight (HasTR_Underpass ?y
GeumrakTR_Height)) (and (UWaterLevel TR_WaterLevel))) => (assert
(Entry ?x ?y))
```

4. Conclusions

We proposed the FOCATC4K that can be used as a framework for autonomous vehicles. Future work includes mapping digital maps into FOCATC4K and integrating it with navigation.

References

- [1] M. Buechel et al., "Ontology-Based Traffic Scene Modeling, Traffic Regulation Dependent Situational Awareness and Decision-Making for Automated Vehicles," in *Symposium on Intelligent Vehicles*, pp. 1471-1476, 2017.
- [2] M. A. Mohammad et al., "New Ontology structure for intelligent controlling of traffic signal," *Procedia Computer Science*, Vol. 207, pp. 1201-1211, 2022.
- [3] M. I. Alipio et al., "Vehicle Traffic and Flood Monitoring with Reroute System Using Bayesian Networks Analysis," in *Global Conf. on Consumer Electronics*, pp. 1-5, 2017.
- [4] -S Theory Based Fuzzy Ontology Context Modeling and Similarity Based Reasoning," in *Proc. of the CIS*, pp. 707-711, 2013.
- [5] T. S. Nguyen, H. K. Nguyen, "Fuzzy Ontology Based Model for Supporting Safe Driving," *Int. Journal of Computer Science and Network Security*, Vol.19, No.4, pp. 111-115, 2019.
- [6] S. C. Moussaoui et al., "Dynamic Context Awareness in Autonomous Navigation," in *Proc. of Int. Conf. on SMC*, pp. 2808-2813, 2022.

Egress Initialization for Graph Convolution Network in Recommendation Systems

Sang-min Lee, Kim Heechan, Jeon Janggun and Namgi Kim

Department of Computer Science, Kyonggi University
Suwon, South Korea

[e-mail: d9249@kyonggi.ac.kr, train1212@kyonggi.ac.kr, 201812285@kyonggi.ac.kr,
ngkim@kyonggi.ac.kr]

*Corresponding author: Namgi Kim

Abstract

Recommendation Systems (RS) help users quickly find the content they want on the Internet, where there is a lot of information. Graph Convolution Network (GCN) algorithms are widely used to efficiently implement such RIS. Xavier Uniform Initialization has been mainly used as a weight initialization in the learning process of GCN algorithms. However, to optimize the learning of GCN, it is necessary to apply a learning method that fits the graph structure. Therefore, we propose Egress Initialization, which initializes the distribution according to the graph embedding size, rather than Xavier Uniform Initialization, which has been traditionally used.

Keywords: Graph Convolution Network, learning method, Recommendation Systems

1. Introduction

Recommendation systems are a key tool for presenting information or content that users want in a world of information overload, and they effectively present the most relevant information or products by analyzing users' past behavior, preferences, and interests. Mainly utilized in various fields such as e-commerce, entertainment, and social media, recommendation systems are used on platforms such as YouTube, Netflix, and Amazon to present users with the content they want. Recent research on recommendation systems utilizes Graph Convolution Network (GCN) to improve performance, and many studies have been conducted to improve the learning process of graphs. Therefore, in this paper, we propose an optimized initialization function for graph learning.

2. Proposal methods

2.1 Egress Initialization

Equation 1 is the popular Xavier Uniform Initialization, where N_{in} is the number of input units and N_{out} is the number of output units (embedding size). The Xavier Initialization is designed to ensure that the variance of the output at each layer of the neural network remains the same as the variance of the input, which prevents the gradient from becoming too small or too large during training.

$$W \sim U \left(-\sqrt{\frac{6}{N_{in} + N_{out}}}, \sqrt{\frac{6}{N_{in} + N_{out}}} \right) \quad (1)$$

Xavier Initialization is especially effective when the activation function is linear or when using a symmetric activation function such as tanh. Equation 2 is the proposed Egress Initialization, which has an initialization distribution that only considers the output of the layer.

$$W \sim U\left(-\frac{1}{\sqrt{N_{out}}}, \frac{1}{\sqrt{N_{out}}}\right) \quad (2)$$

If the two initialization functions have 10,000 input units and 64 output units, the Xavier Initialization will initialize in the range (-0.024, 0.024), while the proposed Egress Initialization will initialize in the range (-0.125, 0.125). This is a 5.2x higher range. This can lead to problems such as weight divergence/explosion, but in the case of Xavier Initialization, the embedding value of the graph neural network is converging too small, so having a wider range will improve the learning process of the GNN.

3. Experiments

For the experimental data, we used the Yelp2018 dataset, which is often used in previous studies. The configuration of the dataset is shown in Table 1, and the experimental results of the existing method and the proposed initialization method are shown in Table 2. The experimental results show that by applying the egress initialization method, the learning epoch can be reduced, and the accuracy is slightly improved or degraded within the error range. This confirms that having a wider initialization distribution is superior in terms of learning speed.

Table 1. Experiments data

| | Yelp2018 | |
|-------------|-----------|-----------|
| | Train | Train |
| User | 31,668 | 31,668 |
| Item | 38,048 | 38,048 |
| Interaction | 1,237,259 | 1,237,259 |
| Sparsity | 99.90% | 99.90% |

Table 2. Experiments result

| | yelp2018 | | |
|------------------|----------|---------|---------|
| | Epoch | recall | NDCG |
| BUIR(Xavier) | 290 | 0.04371 | 0.03557 |
| BUIR(Egress) | 79 | 0.04578 | 0.03718 |
| LightGCN(Xavier) | 231 | 0.05949 | 0.04887 |
| LightGCN(Egress) | 186 | 0.06059 | 0.04993 |
| SGL(Xavier) | 27 | 0.06791 | 0.05585 |
| SGL(Egress) | 20 | 0.06480 | 0.05304 |
| SimGCL(Xavier) | 18 | 0.07248 | 0.05976 |
| SimGCL(Egress) | 18 | 0.07217 | 0.05932 |
| XSimGCL(Xavier) | 14 | 0.07277 | 0.06002 |
| XSimGCL(Egress) | 13 | 0.07241 | 0.05964 |

4. Conclusions

This paper proposed an egress initialization optimized for the learning process of GNN, and showed an increase in performance and a reduction in the learning process through experiments, and we plan to conduct future research on initialization distributions that consider the number of users and items instead of only considering the embedding size during the initialization process.

Acknowledgments

This research was supported by the National Research Foundation of Korea (NRF) Basic Science Research Program funded by the Ministry of Education (No. 2020R1A6A1A03040583) and this research was supported by the MSIT(Ministry of Science and ICT), Korea, under the National Program for Excellence in SW supervised by the IITP(Institute of Information & communications Technology Planning & Evaluation) (2021-0-01393) and this work was supported by Korea Institute for Advancement of Technology(KIAT) grant funded by the Korea Government(MOTIE) (P0008691, HRD Program for Industrial Innovation)

References

- [1] X. He, K. Deng, X. Wang, Y. Li, Y. Zhang *et al.*, “LightGCN: Simplifying and Powering Graph Convolution Network for Recommendation,” in *Proc. Special Interest Group in Information Retrieval (SIGIR)*, China, pp. 639–648, 2020.
- [2] J. Wu, X. Wang, F. Feng, X. He, L. Chen *et al.*, “Self-supervised Graph Learning for Recommendation,” in *Proc. Special Interest Group in Information Retrieval (SIGIR)*, Canada, pp. 726–735, 2021.
- [3] D. Lee, S. Kang, H. ju, C. Park and H. Yu “Bootstrapping User and Item Representations for One-Class Collaborative Filtering,” in *Proc. Special Interest Group in Information Retrieval (SIGIR)*, Cabada, pp. 317-326, 2021.

Establishment and Operation of an Online Integration Test Environment for Verifying Heterogeneous Parallel Processor Software

Seungwoo Rho and Gibeom Gu

Korea Institute of Science and Technology Information
Seoul, South Korea

[email: seungwoo0926@kisti.re.kr, gibeom.gu@kisti.re.kr]

*Corresponding author: Gibeom Gu

Abstract

Advancements in technology necessitate the incorporation of more advanced heterogeneous parallel processors in modern computing systems. Korea has recently developed a heterogeneous parallel processor called AB21. Given the limited research on its performance, this study examines the software stack for this processor and elucidates the process of building and operating an online integrated test to verify its performance. We established configuration, execution, and issue-handling scenarios for Open Computing Language–Conformance Test Suite (OpenCL-CTS)-based integration testing. The OpenCL-CTS used to verify the AB21 software stack function consisted of thousands of large unit tests. For these large-scale integrated tests, we employed the commonly used GitLab CI/CD pipeline to automatically build and distribute an integrated online environment. Using the integrated testing environment, we efficiently discovered, verified, and resolved software issues in real time, thereby reducing the integrated testing time. The testing environment developed in this study serves as a valuable reference for building other testing environments to verify the performance of other recently developed heterogeneous parallel processors.

Keywords: Heterogeneous parallel processor, AB21, integrated testing environment, software stack

1. Introduction

Looking at the recently announced TOP500 supercomputers in June 2023, the proportion of systems adopting graphical processing units has increased, and approximately 37% (185/500) of the systems use accelerators. This paradigm shift is due to the increasing demand for artificial intelligence computing and accelerators with improved actual measurement performance ratios and energy efficiencies [1]. Following this trend, Korea has continued to carry out research and development projects for heterogeneous parallel processors such as AB21 through

matrix/vector calculations and developed software stacks to utilize these processors [2].

Regarding AB21, it is a RISC-V core-based heterogeneous processor with a computational performance (FP64) of up to 8 TFLOPS. Its software stack comprises an AB21 emulator, device drivers, LLVM compilers, Open Computing Language (OpenCL) runtime, and the OpenCL Conformity Test Suite (OpenCL-CTS) provided by the Khronos Group for functional verification. The integration of these components and large-scale unit tests is important for the performance verification and effective utilization of the self-developed AB21.

In the OpenCL-CTS test, approximately 1,700 unit tests, excluding the graphics, OpenCL 2.0 and 2.1, and extension categories, were used. In the AB21 function verification test, the sequential configuration and installation process of each software stack was completed, following which the subunit tests of OpenCL-CTS were executed. OpenCL-CTS is configured to run sequentially on a single node, and parallel processing is only possible for large-scale conversion tests and mathematical tests with long execution times. To parallelize the entire unit test, the code was modified to be set up, installed, and executed for each unit test, and abstracted installation and execution templates were applied to enable distributed parallel processing. Additionally, to build and operate an integrated test environment using these methodologies, a commonly used GitLab CI/CD tool was employed. These DevOps platforms facilitated the automation of the overall recurring workflow, including software building, installation, issue management, and pipeline management. To the best of our knowledge, this is the first study to parallelly verify the performance of AB21 using OpenCL-CTS in a cluster environment.

2. Software Stack for Verifying the Heterogeneous Parallel Processor Function

The software stack used for verifying the AB21 function comprised four target software packages and one functional verification software package. Regarding these software, AXPUSIM is an emulator developed for AB21 processor simulation. It is a pre-built binary-type library file and supports the architectures X86_64 and AArch64. XDE is a device driver emulator developed for AB21 support [3]; the compiler adds features to support the AB21 processor based on the open-source LLVM [4]. OpenCL Runtime was modified and developed based on the OpenCL Portable version, the Portable Computing Language (PoCL) [5] project; it is defined as SoCL [4]. We performed the functional testing of AB21 using OpenCL-CTS. Specifically, we utilized LLVM v12.0, PoCL v4.0, and OpenCL-CTS v2023-05-16.

3. Establishment of an Integrated Testing Environment

The steps in constructing and executing the large-scale integrated test environment for verifying AB21 is as follows.

- 1) The software for integration and functional verification were selected; AXPUSIM, XDE, LLVM, and SoCL were used for integration, and OpenCL-CTS was used for functional verification.
- 2) The code of OpenCL-CTS, namely the main code of the test and the CMake environment configuration file, was modified to select only the categories and tests to be tested in the entire OpenCL-CTS.
- 3) Unit tests were selected for the parallel and distributed execution of OpenCL-CTS, and an installation and execution setup file (.yml) was generated. This setup file, which was a subunit of the gitlab-ci setup file, separated the build and execution parts independently for each unit test, thereby allowing each test to be run individually.
- 4) Using the Python application programming interface (API) provided by GitLab, issue management was performed in two ways: using a script that generated an issue when a task would fail (issue_open.py) and one (issue_close.py) that would close the issue when the operation was successful.
- 5) A Python script utilizing the same API was written to send a test summary email when the entire test was completed.
- 6) By leveraging the subtree functionality provided by Git, a virtual integrated directory was built for testing.
- 7) A configuration file (gitlab-ci) was created for the GitLab CI/CD. This file was at the root of the virtual integration project and enabled the GitLab-CI/CD pipeline to run. It consisted of stages including variables, default, install-xde, install-axpusim, install-compiler, install-socl, install-opencl-cts, and email-summary; each stage included the installation of each software and a subtest part of the OpenCL-CTS. Include a subsetting part of GitLab-ci and include the unit test installation and executable files created in step 7. The variables represented the installation and log path of each software; default used the afterscript keyword to execute an issue management script depending on whether the unit test was successful, install_* represented

each applicable software installation script, and the email summary executed the Python script file for task completion after all the tasks had been completed. Additionally, dependence and execution trigger sources were specified for each stage.

8) The default settings for the GitLab CI/CD pipeline execution were changed. The GitLab runner could be registered, released, and changed, and timeout and token management tasks could be performed. Tokens needed to be issued and registered manually using Python API GitLab. Regarding the GitLab runner, it is an agent running on a real computational node and the subject of the unit testing of the OpenCL-CTS in GitLab. In this study, a GitLab runner was used in a cluster consisting of four computational nodes. The current variables in its configuration file (`/etc/gitlab-runner/config.toml`) were set to 40—this value is generally appropriate if set to 80% or less of the total number of CPUs—for each parallel execution thread. Timeout is the maximum time that a GitLab runner can execute per unit operation; its default is 1 h. In this study, it was 24 h; if the execution time per task exceeded this, it indicated task failure.

An AArch64-based cluster of four nodes was set up to run and verify the integrated tests based on OpenCL-CTS. The GitLab runner handled the task in parallel using 40 threads per node. The initial pipeline was executed manually to run the entire OpenCL-CTS unit test, and the pipeline was partially executed in accordance with the resolution of the registered issues. When it was executed, the initial state changed to "running" and then to either "passed" or "failed" depending on the result. The status, execution results, and total execution time could be checked in the CI/CD pipeline menu. When the pipeline ends, the email_summary task is executed, and a pipeline summary email is automatically sent to the administrator.

4. Conclusions

This paper introduces a software stack for AB21 and a real-time large-scale integrated test pipeline construction and execution method. For large-scale integration testing, GitLab CI/CD tools were used to build 160 distributed/parallel task test environments across four nodes, and the code was modified to handle OpenCL-CTS unit

tests in parallel. Additionally, by applying automatic issue processing according to unit test execution, automatic pipeline execution based on code updates, and a notification system that sends summary emails based on the completion status of the test, a real-time large-scale integrated test environment was established and operated. The results of this study can aid future research on verification approaches and testing environments for other types of heterogeneous parallel processor software.

Acknowledgements

This research was supported by the Korean government (Ministry of Science and ICT), as well as the National Research Foundation of Korea's Supercomputer Development Leading Project (No.: 2020M3H6 A1084853)

References

- [1] D. Reed, D. Gannon, and J. Dongarra, "Reinventing high performance computing: challenges and opportunities," *arXiv preprint arXiv*, 2203.02544, 2022.
- [2] K. Gu, S. Rho et al., "Supercomputing S/W technology for supercomputing SoC," *The Korean Society of Electronics and Engineering Conference*, Vol 45. No.1, pp. 2635-2636, 2022.
- [3] E. J. Lim, S. Y. Ahn, Y. Kim, and Y. M. Park, "Device driver emulator for parallel computing accelerator simulator supporting OpenCL," in *2022 13th International Conference on Information and Communication Technology Convergence (ICTC)*, pp. 2352-2355, 2022.
- [4] Y. Shin, J. Ryu, J. Park, and H. Sung, "Supreme-K supercomputer oriented OpenCL compiler and runtime," *Academic Conference of the Korean Society of Electronic Engineering*, pp. 2643-2647, 2022.
- [5] P. Jääskeläinen, C. S. de La Lama, E. Schnetter et al., "pocl: A performance-portable OpenCL implementation," *International Journal of Parallel Programming*, Vol 43, pp. 752-785, Issue 5, 2015.

Integrated Dataset for 10 Facial Emotion Recognition

Hoyoung Yoo and Hyeonjin Kim

Electronics and Telecommunications Research Institute
Daejeon, South Korea
[e-mail: yoohy@etri.re.kr, jini@etri.re.kr]

Abstract

Facial expression recognition classify is one of the tasks in computer vision where given a facial image, it is classified into predefined emotion label. There are numerous datasets available for facial expression recognition. However, the labels of facial expressions included in each dataset vary from one dataset to another and the data types are also different. To expand the range of recognizable facial expressions, we have tried integrating and relabeling various datasets for effective use in practical applications. The integrated dataset consists of a total 93543 images with 10 facial expressions including 'Angry', 'Calm', 'Disgust', 'Dislike', 'Fear', 'Happy', 'Neutral', 'Sad', 'Sleepy', 'Surprise'. We utilized Resnet50 as the classification model for integrated dataset training. The best accuracy of integrated dataset was 69.32%.

Keywords: Facial emotion recognition, Facial expression, 10 emotions dataset

1. Introduction

Facial expression recognition is one of the tasks in computer vision. Recently, its performance has improved dramatically with the advancement of deep learning[8], there are many efforts to apply facial expression recognition in a wide range of human computer interaction(HCI) applications such as medical care, education, gaming, etc[9]. There are various datasets for facial expression recognition, but the data type and expression labels are different for each dataset. Since it is better to recognize various facial expressions in the application, we integrated and relabeled the existing datasets to form a dataset for 10 facial expressions.

2. Datasets for Facial Emotion Recognition

2.1 Pre-existing datasets

The dataset used for integration is shown in Table 1. The CK dataset[1] has been used for facial expression recognition even before the advent of deep learning models. It was created in a laboratory setting and consists of 7 different facial expressions. The Ryerson dataset was produced in a studio environment where professional actors performed and underwent a inspection process[2]. Ryerson dataset contains 8 facial expressions. In contrast, the RAF dataset was collected online, with images tagged by 40 annotators using 7 facial expressions[3]. The YawDD dataset, specifically designed for drowsiness detection, comprises two labels: drowsy and non-drowsy[4].

This work was supported by an internal fund/grant of the Electronics and Telecommunications Research Institute(ETRI). [23ZT1100, ICT convergence technology support and development based on local industry in the metropolitan area]. This research was also supported by the Korea Evaluation Institute of Industrial Technology(KEIT) grant funded by the Korea government (MOTIE), (No.20022137).

2.2 ETRI emotion dataset

ETRI emotion dataset is a dataset constructed for 22 detailed facial expressions (Tense, Nervous, Worried, Anxious, Relaxed, Fatigued, Sleepy, Board, Confused, Surprised, Envious, Apologetic, Calm, Happy, Suffering, Irritated, Angry, Furious, Disgusted, Disappointed, Dislike). It consists of video data where actors performed various scripts for each facial expression.

The dataset comprises 200 Korean actors, considering gender and age. The recordings were conducted in a studio equipped with HD cameras (1920 by 1080 resolution, 30 frames per second), taking into account lighting, background, head position, and noise conditions. Three version of videos were recorded for each actor, corresponding to the 22 facial expressions resulting in a total of 13,200 mp4 videos in the dataset.

Table 1. Datasets for facial emotion recognition

| Dataset | # of actors | Race | # of emotions | Data Type | Characteristic |
|---------|-------------|----------------|---------------|-----------|------------------------------------|
| CK+ | 210 | North American | 7 | Images | Gray-scale |
| Ryerson | 24 | North American | 8 | Videos | Professional North American actors |
| RAF | 29672 | World | 7 | Images | Internet images, 40 annotators |
| YawDD | 107 | World | 2 | Video | Drowsiness detection |
| ETRI | 200 | Korea | 8 | Video | Studio camera env. |
| Total | 30142 | World | 10 | Images | Integration of above |

2.3 Integration of datasets

The ETRI emotion dataset, consisting of emotional acting performed by non-professionals, could not be directly used as training data, because there are differences between the facial expressions played in acting by actors and the facial expressions that are actually perceived by annotators. We selected 8 overlapped facial expressions with other pre-existing datasets: 'Angry', 'Calm', 'Disgust', 'Dislike', 'Happy', 'Sad', 'Sleepy', and 'Surprise' from the ETRI emotion dataset, and performed video sampling. ETRI dataset used in the paper involved capturing video only for the middle parts where the actors' performances were at their peak. Face detection was performed from images captured, and detected faces were

provided to three different individuals of Korean nationality. Participants were asked to select one of the 8 facial expressions. Only data where all three individuals agreed were retained and refined.

The Ryerson dataset and YawDD dataset were also captured from videos and converted into facial images. Three Korean individuals performed the annotation of these datasets. Subsequently, only the data on which all three individuals agreed was retained and refined. The total number of integrated data samples was 93543, consisting of 81227 training data and 12266 validation data. Samples for each facial expression are shown in Fig. 1.



Fig. 1. Samples of ETRI emotion dataset for each facial expression

3. Benchmark Implementation

We implemented the vanilla deep learning model consisting of ResNet50 and classifier to assess the performance of facial expression recognition. We also utilized algorithms from FAN [6] and DMUE [7] to verify the performance of the integrated dataset. [6] is an algorithm that utilizes an attention network of features extracted by passing sequential frames through a CNN. We applied this algorithm to sampled sequential frames from a video, and in the case of a single image, we replicated it to use attention algorithm.

To improve the performance of all algorithms, data augmentation processes such as normalization, color jitter and adding Gaussian noise were applied to the training data. Pretrained Resnet50 was applied to all algorithms. The vanilla CNN model and FAN are trained with the same hyperparameters such as the Adam optimizer, learning rate of 0.0001, batch size of 80. Table 2 shows the highest accuracy of each algorithm. The best performing model was FAN, which achieved an accuracy of 69.32%. Compared to the state of the art

accuracy of the Ryerson dataset in [10], 74.92%(visual only) for 8 expressions, the accuracy of integrated dataset for 10 facial expression was 5.6% lower. Due to the insufficient experimentation with various hyperparameter conditions for the DMUE model, there is a possibility that it may not be properly finetuned for a new dataset.

Table 2. The number of samples for each expression

| Algorithm | Vanilla CNN classifier | FAN [6] | DMUE [7] |
|-----------|------------------------|---------|----------|
| Accuracy | 65.71 % | 69.32 % | 65.93 % |

4. Conclusions

We integrated existing datasets for facial expression recognition, including the ETRI emotion dataset, and constructed a dataset for 10 facial expressions. In the bench-mark performance test, the best accuracy for the 10 facial expressions was 69.32%. We believe that using more sophisticated deep learning models will improve the performance.

References

- [1] Patrick Lucey, Jeffrey F Cohn, Takeo Kanade, Jason Saragih, Zara Ambadar, and Iain Matthews, "The extended cohnkanade dataset (ck+): A complete dataset for action unit and emotion-specified expression," in CVPRW, 2010.
- [2] S. R. Livingstone and F. A. Russo, "The ryerson audio-visual database of emotional speech and song (ravdess): A dynamic, multimodal set of facial and vocal expressions in north american english," PloS one, vol. 13, no. 5, 2018.
- [3] Shan Li, Weihong Deng, and JunPing Du. Reliable crowdsourcing and deep locality-preserving learning for expression recognition in the wild. in CVPR, pages 2852–2861, 2017.
- [4] Shabnam Abtahi, Mona Omidyeganeh, Shervin Shirmohammadi, and Behnoosh Hariri, "YawDD: A Yawning Detection Dataset", In Proceedings of the 5th ACM Multimedia Systems Conference, pages 24-28, 2014.
- [5] Ching-Hua Weng, Ying-Hsiu Lai, Shang-Hong Lai, "Driver Drowsiness Detection via a Hierarchical Temporal Deep Belief Network", in Asian Conference on Computer Vision Workshop on Driver Drowsiness Detection from Video, Taipei, Taiwan, Nov., 2016.
- [6] Debin Meng, Xiaojiang Peng, Kai Wang, Yu Qiao, "Frame Attention Networks for Fa-cial Expression Recognition in Videos", in ICIP, 2019.
- [7] Jiahui She, Yibo Hu, Hailin Shi, Jun Wang, Qiu Shen, Tao Mei, "Dive into Ambiguity: Latent Distribution Mining and Pairwise Uncertainty Estimation for Facial Expression Recognition", in CVPR, 2021
- [8] Huang, Yunxin and Chen, Fei and Lv, Shaohe and Wang, Xiaodong, " Facial Expres-sion Recognition: A Survey", in Symmetry, 2019.
- [9] Anil Audumbar Pise, Mejdal A. Alqahtani, Priti Verma, Purushothama K, Dimitrios A. Karras, Prathibha S, and Awal Halifa, "Methods for Facial Expression Recognition with Applications in Challenging Situations", in Computational Intelligence and Neurosci-ence, May, 2022
- [10] Kateryna Chumachenko, Alexandros Iosifidis and Moncef Gabbouj, "Self-attention fu-sion for audiovisual emotion recognition with incomplete data", arXiv:2201.11095v1, Jan., 2022.

Research on Customer Satisfaction in Cruise Ship Tourism Based on Big Data Analytics - An Empirical Analysis from OTA Website

Rui Liang and Yu Feng

Professor, Beijing Union University

Ph. D. Student, Department of Global Business, Kyonggi University

[e-mail: liangruilr@qq.com, Yukihihi@qq.com]

*Corresponding author: Yu Feng

Abstract

This study utilized Python technology to scrape data from tourists' reviews on Online Travel Agencies (OTAs) and conducted an in-depth analysis of tourists' satisfaction with international cruise ships in the Asian market. Research methodologies including text analysis, word frequency analysis, and sentiment analysis were employed. The subjects of the study were Carnival Cruise Group's Majestic Princess, Royal Caribbean Cruise Group's Spectrum of the Seas, and Genting Hong Kong Limited's Norwegian Joy. The primary objective was to acquire a comprehensive understanding of the perceived preferences and distinct needs of Asian travelers. Delving deeply into big data, the study scrutinizes the challenges and opportunities for the enhancement of cruise groups in the Asian market. Ultimately, by integrating consumer behavior research with tourism market analysis, the study offers recommendations and strategic directions for international cruise groups to enhance customer satisfaction in the Asian market. The study elucidates the principal factors influencing tourist satisfaction among the three major cruise groups across four dimensions: cruise facilities, services, family travel services, and tourist perceptions, potentially contributing to the sustainable development of cruise tourism in Asia.

Keywords: Tourist satisfaction, Cruise Ship, Online reviews, Big data, Network Text Analysis

1. Introduction

In the era of digitalization and big data, the tourism industry, particularly the cruise tourism sector, has undergone significant transformations. The development of OTA platforms and big data analysis has paved new ways to explore customer satisfaction. OTA platforms, as a novel mode of interaction between tourists and the tourism industry, have not only simplified the booking process but also become repositories for customer reviews, preferences, and behaviors. Utilizing this data, cruise operators can gain in-depth insights into customers' needs and expectations. The cruise

industry offers a wide range of services and experiences, with numerous factors influencing tourist satisfaction. Compared to traditional surveys, big data analysis can more effectively process and analyze large datasets, aiding in identifying key factors affecting customer satisfaction, analyzing feedback, discerning trends, and predicting future customer behavior. [1] The Asian cruise market is highly competitive and possesses tremendous potential, with projections indicating that by 2025, Asia will account for 30% of the global cruise customer market. [2] Consequently, this study contributes to delineating the current challenges faced by cruise enterprises in the Asian market and can aid

international cruise groups in better integrating into the Asian market in the future.

2. Subjects and Methods

2.1 Overview of Research Subjects

In this study, the selected research subjects were Centurion Princess (a subsidiary of Carnival Cruise Group), Spectrum of the Seas (a subsidiary of Royal Caribbean Cruise Group), and Norwegian Joy (a subsidiary of Genting Hong Kong Limited). The selection of these research subjects was primarily based on the following considerations: Firstly, these three cruise ships are affiliated with the world's leading cruise groups and exhibit minimal differences in tonnage. Secondly, these cruise ships were among the first to enter the Asian market during 2017-2019, a period marked initially by rapid growth followed by a gradual decline due to various incompatibilities in the Asian cruise market. [3] These three cruise ships are thus more representative for examining the development trends and characteristics of the Asian cruise market.

2.2 Data Collection

This study employed Python technology to collect a substantial amount of tourist review data. The data were sourced from Online Travel Agency (OTA) websites, with the collection period extending from October 2017 to April 2023. Necessary preprocessing was conducted on the data, including text cleaning, removal of irrelevant information, and elimination of duplicate comments, resulting in 12,190 valid comments. Based on the principle of random sampling, 2,000 comments from each subject were selected as the final research sample.

3. Manipulation of Virtual Model

Initially, a comprehensive text analysis was conducted, which involved extracting the top 200 high-frequency words to generate a visual "word cloud" map (see [Fig. 1](#)) representing tourists' perceptions. This process was followed by further analysis of the top 50 high-frequency words. Among the analyzed groups of high-frequency words, 27 words were found to be repeated.



Fig. 1. Word Cloud Map

This repetition suggests a significant consistency in tourists' perceptions of the needs of different cruise ships. The high-frequency words were subsequently categorized into four dimensions: cruise facilities, cruise services, family traveling services, and tourist perceptions. Numerous linguistic comments on online travel platforms exhibit emotional tones or tendencies expressed by users. These can be analyzed to understand tourists' overall sentiments towards the three cruise tourism products. The results indicate that, overall, tourists exhibit positive emotions towards the three cruise ship tours. Specifically, the Royal Caribbean cruise ship received the highest percentage of positive comments from tourists, while Princess had the lowest percentage of negative comments. Conversely, Norwegian Joy received the lowest percentage of positive comments and the highest percentage of negative comments in the reviews. As depicted in **Fig. 2**, in terms of average sentiment score, Norwegian Joy recorded the highest score of 12.42, followed by Royal Caribbean at 9.15 and Princess with the lowest at 8.65.

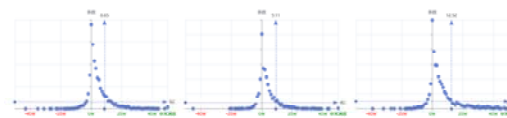


Fig. 2. Tourist Emotional Score

4. Conclusions

The results of the aforementioned analyses indicate that tourists generally exhibit positive emotional willingness and high satisfaction with the three cruise ships, viewing cruise tourism as a leisurely, comfortable, and enjoyable mode of travel. The perceived image of tourists aligns fundamentally with the international image of cruise tourism, suggesting that international cruise ships have largely succeeded in establishing their product images in the Asian market, although numerous detailed issues persist. In conjunction with the study's findings,

the following management recommendations are proposed.

Optimize the core image shaping of cruise products to create distinctiveness. The Asian market is intensely competitive, necessitating that cruise lines enhance their competitiveness through distinctive brand characteristics. The study discovered that although the cruise companies conveyed their core product image to tourists, there was no significant difference in tourists' perception of various cruise brand images, with the different brands being less distinguishable. Tourist reviews primarily focused on the cruise product itself, with minimal emphasis on brand positioning. Therefore, cruise lines should strengthen their brand image in the Asian market by leveraging their history and heritage from Europe and the United States

Enhancing the service capabilities of tour guides and leaders is essential. It was observed in the study that evaluations pertaining to leaders and guides were frequent. While tourists generally expressed satisfaction with the guide services, issues such as lack of enthusiasm, inadequate explanations, and poor problem-solving skills were noted. The inconsistent professionalism of service personnel led to suboptimal service experiences for some customers. Consequently, it is imperative to intensify training in relevant skills for leaders to ensure that every employee at customer touchpoints demonstrates the company's commitment to service quality. The study also revealed a significant number of tourists are families, suggesting a focus on child and elderly services in the training. This approach will firmly grasp the considerable potential of the family travel market.

Enhancing the efficiency of collective tourist activities is crucial. In evaluations related to collective activities like boarding, disembarking, and safety drills, tourists' overall feedback was negative, with disorganized activities significantly dampening their enthusiasm. Cruise companies have not adequately considered the needs and experiences of different age groups. Clear guidance and expedited customs queues are a universal desire among tourists. Therefore, implementing phased and streamlined approaches to improve crowd management should be considered. Organizing collective activities based on the diverse needs of different groups is a more rational approach.

Appropriately providing free services, such as WiFi and complimentary drinking water services, is crucial. These are both common pain points for tourists and attractive needs.^[4] Once such expectations are met, satisfaction levels are likely to rise sharply. Fulfilling these needs can create a sense of delight, thereby enhancing brand loyalty among tourists. Cruise companies could reduce costs or offer free periods, and establish complimentary water stations onboard. This approach can attract and retain customers, thereby enhancing the brand's competitiveness.

References

- [1] Kim.S, & Lee. W. S, "Network text analysis of medical tourism in newspapers using text mining: The South Korea case," *Tourism Management Perspectives*, No.31, pp.332–339, 2019.
- [2] Qi-Yun.Wu, "Liu Zinan, Chairman of Royal Caribbean Group Asia: 'Seeding Interest' + Sprinting to Deploy Multiple Large Ships in China," *Beijing Business Today*, No. 006, 2023.
- [3] Sun. X, Feng. X, Gauri. D. K. "The cruise industry in China: Efforts, progress and challenges," *International Journal of Hospitality Management*, Vol. 42, No. 1, pp. 71-84, 2014.
- [4] Chunhua. Jin, Junshan. Gao, "A Study on the Quality Factors of B2C E-Commerce Websites Based on the Kano Model," *China Business and Market*, Vol. 32, No. 04, pp. 77-84, 2018.

Verification of Terror Factors Based on GTD data: Application of the RTM Model

Eun-bi Cho¹, Sang-Eun Ahn¹, and Jeong-Hyeon Chang^{1*}

¹Contents Convergence Software Research Center, Kyonggi University
[e-mail:crime_tiger564@kyonggi.ac.kr]

*Corresponding author: Jeong-Hyeon Chang

Abstract

The increasing threats of terrorism worldwide have emphasized the importance of identifying and understanding the underlying factors that contribute to such activities. Utilizing the comprehensive Global Terrorism Database (GTD), this study aimed to verify the predominant terror-related factors in recent years. We applied the Risk Terrain Modeling (RTM) method, a spatial risk analysis technique, to provide a more granulated understanding of the areas prone to terrorist activities and the reasons behind them. Our findings highlighted certain hotspots where the convergence of various risk factors has led to heightened terrorist activities. Additionally, the RTM model proved effective in not only predicting potential terror-stricken regions but also in offering policy recommendations to mitigate risks. This study underscores the need for an integrated approach, combining rich datasets like GTD with advanced modeling techniques like RTM, to develop robust strategies against terror threats.

Keywords: Global Terror Data(GTD), World Victim Survey(WVS), Risk Terrain Model(RTM), Terror

1. Introduction

The escalating frequency and severity of terrorist incidents globally necessitate a deeper understanding of the factors driving such events. This study leverages the extensive data from the Global Terrorism Database (GTD) to examine the key factors contributing to terrorism in recent years. By employing Risk Terrain Modeling (RTM), a sophisticated spatial risk analysis technique, we aim to identify and analyze the geographic hotspots of terrorist activities and understand the confluence of risk factors leading to these incidents [1].

The significance of this research lies in its potential to enhance current counter-terrorism strategies. By pinpointing high-risk areas and understanding the underlying causes, our findings can guide more effective and targeted

interventions. This study also demonstrates the value of combining comprehensive datasets with advanced analytical methods to predict and mitigate terrorism risks, providing valuable insights for policymakers and security professionals [2].

The paper is organized to first contextualize the importance of data-driven approaches in addressing terrorism, followed by a detailed examination of our methodology and findings, culminating in a discussion of policy implications and recommendations for future research.

2. Theoretical Background

"Spatial Analysis of Terrorism: Insights from the Global Terrorism Database and Risk Terrain Modeling, we address the evolving nature of

terrorism, emphasizing its shift from regional conflicts to a global threat. This section underscores the complexity and varied motivations behind terrorist activities, illustrating the necessity of a nuanced understanding in today's context. The critical role of data in analyzing terrorism is highlighted, focusing on how diverse data types help unravel terrorist patterns, methods, and motivations [3]. The Global Terrorism Database (GTD) is introduced as an essential resource, providing comprehensive data for global terrorism analysis, along with its development, application, and limitations. Risk Terrain Modeling (RTM) is briefly overviewed as a key method in spatial risk analysis, particularly relevant to terrorism studies [4].

The section underlines RTM's methodology and its prior applications. Finally, the integration of GTD and RTM is discussed, outlining how this combination can offer deeper insights into the spatial dynamics of terrorism, thereby aiding in the development of more effective counter-terrorism strategies. This summarized background provides a foundational understanding of the study's approach, blending comprehensive terrorism data with advanced spatial analysis [5].

3. Research Model

The proposed model for the study on "Spatial Analysis of Terrorism" integrates various components to understand and predict terrorism activities effectively. At its core, the model focuses on "Terrorism Analysis," which is fed by two primary data sources: the Global Terrorism Database (GTD) and Risk Terrain Modeling (RTM).

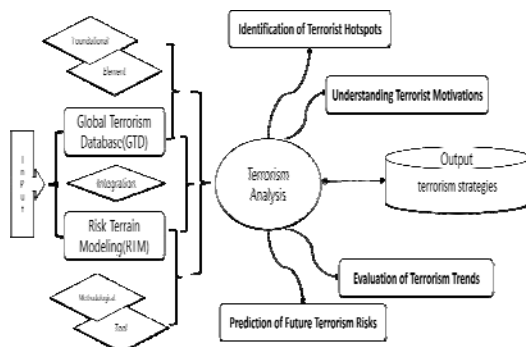


Fig. 1 . Research Model

4. Methodology

This research proposes a methodology to develop a terrorism prediction model by applying the AutoRegressive Integrated Moving Average (ARIMA) analysis to data from the Global Terrorism Database (GTD). The process begins with the collection and preparation of GTD data, focusing on essential variables and converting them into a time series format. Following this, the data undergoes a series of analyses, including decomposition to understand its underlying patterns, and stationarity testing, essential for the ARIMA model. The optimal parameters for the ARIMA model are then determined. Once the model is built, it is rigorously validated through diagnostic checks and performance assessments using standard statistical metrics. Finally, the validated model is used for forecasting future terrorism incidents, with the predictions aimed at informing actionable policy recommendations for counter-terrorism. This approach aims to provide a systematic and data-driven methodology for predicting terrorism trends, thereby aiding in the development of informed and proactive counter-terrorism strategies.

5. Results

In this study, we have conducted an analysis to identify the nations most impacted by incidents over a half-century, from 1970 to 2020. Our analysis reveals a distinct concentration of incidents in specific geographical regions. Following Afghanistan ranks as the first affected nation, underscoring the ongoing challenges it faces. Pakistan emerges as the second on this list, reflecting its significant share of incidents. India and Colombia, ranking third and fifth respectively, complete this list of the top five nations most affected by incidents from 1970 to 2020. This detailed breakdown paves the way for a more granular examination of the annual incident trends within these countries, offering insights into the temporal dynamics of these occurrences.

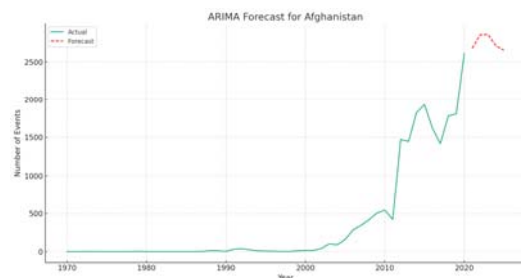


Fig. 2 . ARIMA for 1st

Based on this data, it is anticipated that the ARIMA model for Colombia, having the lowest AIC (Akaike Information Criterion) and BIC (Bayesian Information Criterion) values, will exhibit relatively better predictive performance compared to other models. Conversely, the model for Pakistan, which has the highest AIC and BIC values, is expected to have comparatively lower predictive accuracy.

Table 1. Evaluation of Model

| Country | AIC | BIC |
|-------------|--------|--------|
| Afghanistan | 670.50 | 681.85 |
| Pakistan | 675.48 | 686.83 |
| India | 596.95 | 608.31 |
| Colombia | 592.23 | 603.58 |

6. Conclusions

This research serves as an initial step in combining the Global Terrorism Database (GTD) with Risk Terrain Modeling (RTM) for predicting terrorism. The primary aim was to assess the feasibility of using GTD data for this purpose. Our findings, particularly through the application of the ARIMA model, show promising results, especially in the case of Colombia, where low AIC and BIC values suggest a higher predictive accuracy. However, the higher AIC and BIC values for Pakistan's model indicate some limitations and the need for further refinement.

The study highlights the value of integrating comprehensive datasets with advanced analytical methods to improve our understanding of terrorism's spatial and temporal aspects. It suggests that further research should include more variables and diverse modeling techniques to enhance the predictive accuracy. The research underscores the importance of continuous model updating to keep pace with the evolving nature of

terrorism.

In summary, this study demonstrates the potential of using GTD data in terrorism prediction and sets the groundwork for future, more sophisticated research in this field, contributing to more effective global counter-terrorism strategies.

References

- [1] LaFree, G., "The global terrorism database (GTD) accomplishments and challenges," *Perspectives on Terrorism*, 4(1), 24-46, 2010.
- [2] Apel, R., & Hsu, H. Y., "Interrupted time series analysis in the study of terrorism," *The handbook of the criminology of terrorism*, 276-293, 2016.
- [3] Drawve, G., Thomas, S. A., & Walker, J. T., "Bringing the physical environment back into neighborhood research: The utility of RTM for developing an aggregate neighborhood risk of crime measure," *Journal of Criminal Justice*, 44, 21-29, 2016.
- [4] Kocher, M., & Leitner, M., "Forecasting of crime events applying risk terrain modeling," *GI_Forum-Journal of Geographic Information*, 1, 30-40, 2015.
- [5] Onat, I. "An analysis of spatial correlates of terrorism using risk terrain modeling," *Terrorism and Political Violence*, 31(2), 277-298, 2019.

Exploratory testing methodology based on MITRE ATT&CK and D3FEND Framework

Gaeul Kim and Dohoon Kim*

Department of Computer Science, Kyonggi University
Suwon, South Korea

[e-mail: zxx52539@kyonggi.ac.kr, karmy01@kyonggi.ac.kr]

*Corresponding author: Dohoon Kim

Abstract

Evaluating the quality of a specific target system necessitates security testing based on stability and reliability. Thus, this paper proposes an exploratory test-based assessment method for conducting objective and intuitive security evaluations. Particularly, we utilize the MITRE ATT&CK and D3FEND frameworks to craft test scenarios (cyberattacks and defenses) from the perspectives of attack trees and defense trees, aiming to propose a systematic testing methodology. Ultimately, our research develops an intuitive table that maps and manages various vulnerabilities identified during the creation of test scenarios to security controls compliant with the NIST 800-53 standard, thereby guiding the development of a more structured security framework.

Keywords: Exploratory Test, MITRE Framework, Attack Tree(AT), Defend Tree(DT), NIST 800-53

1. Introduction

In recent years, various modern weapon systems have become highly network-dependent and exposed to various network-based cyber threats due to the increasing importance of software, and various security guidelines have been proposed to improve them. Therefore, security testing based on stability and reliability is necessary to evaluate and improve the quality of weapon systems.

This research aims to develop a more structured security framework by proposing an exploratory test-based evaluation method to perform objective and intuitive security evaluation. Furthermore, the exploratory tests performed based on TTPs* utilize MITRE ATT&CK [1] and D3FEND Framework [2] to create AT/DT scenarios. An intuitive table that maps and manages the various vulnerabilities identified during the creation of test scenarios to security controls that comply with the NIST 800-53 standard is developed to understand the

relationship between the two frameworks, which is useful for assessing the effectiveness of security control items and suggesting ways to counter real-world threats. In the end, the proposed research will allow for efficient assurance that the cyber capability of a weapon system is an important factor affecting the quality of the weapon system.

2. Solution Approach

The procedure for performing exploratory testing by scenarios is as follows. Basically, when performing exploratory testing based on possible cases, attack/defense scenarios are satisfied simultaneously step by step.

Fig. 1 introduces a step-by-step procedure based on AT/DT to derive cyber-attack and defense scenarios that can be identified in a specific weapon system, i.e., an unmanned autonomous vehicle, and presents a proposal for conducting exploratory tests based on it.

*TTPs : Tactics, Techniques, and Procedures

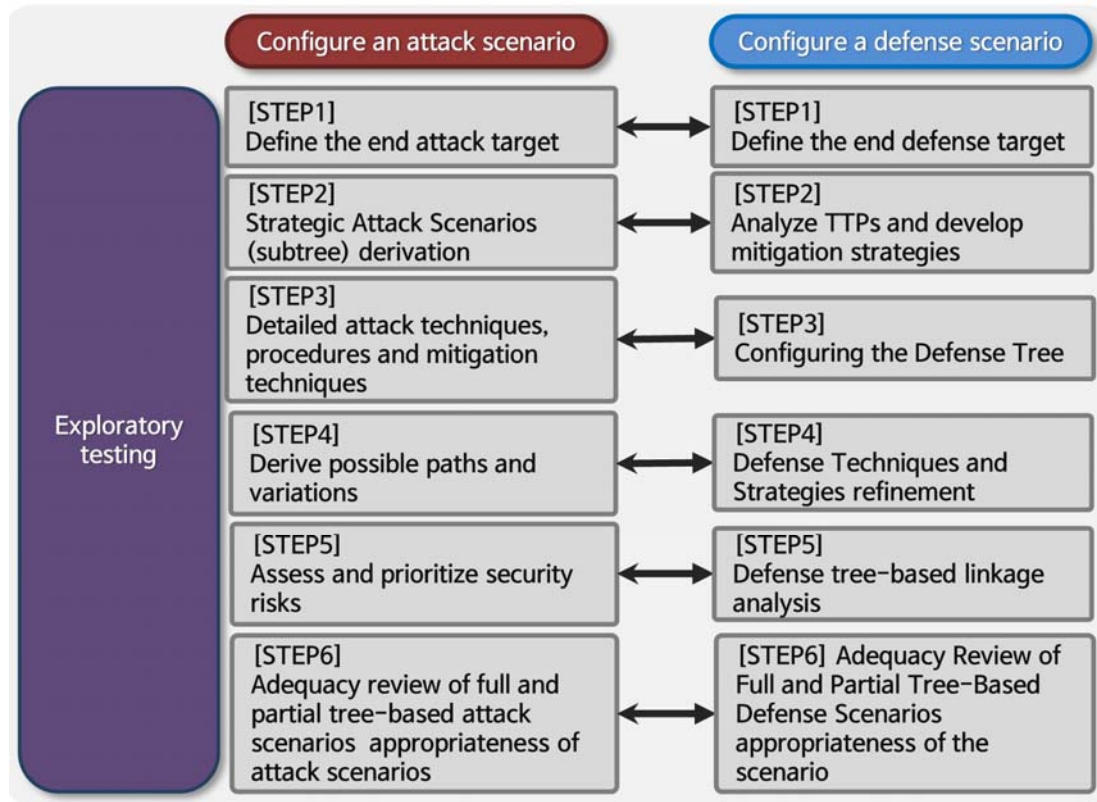


Fig. 1. Configure offensive/defensive scenarios for exploratory testing

In [STEP1], we identify the AT/DT tree components and define the final offensive and defensive objectives, such as "intentional disabling of a free-running tactical vehicle based on malicious behavior".

In [STEP2], a strategic attack/defense scenario based on TTPs is presented based on the goal defined in step 1, and each component is described. It is important to systematically analyze and elaborate the scenarios and attack/defense paths during the subtree derivation process.

In [STEP 3~4], we analyze detailed attack techniques for each scenario and derive procedures and mitigation techniques to understand the scenarios and develop effective response strategies through the above procedures. In addition, we derive possible paths and variations in attack scenarios based on TTPs. Through this, defense techniques and strategies can be detailed according to the attack path.

In [STEP 5], the security risk of the target system is analyzed through AT/DT and NIST 800-53 controls are mapped to perform risk prioritization based on the CIA model and vulnerability assessment through CVE [3].

In [STEP 6], we analyze and review the appropriateness of the scenario.

The above procedures provide a systematic and objective process for efficiently assessing and assuring the impact of a weapon system's cyber capabilities on system quality.

3. Experiment

First, if we consider only the AT, based on the scenario mentioned above, the entire AT can be configured as shown in Fig. 2. After selecting the partial scenario (red box) in Fig. 2, we will perform exploratory tests that can determine the success of the attack.

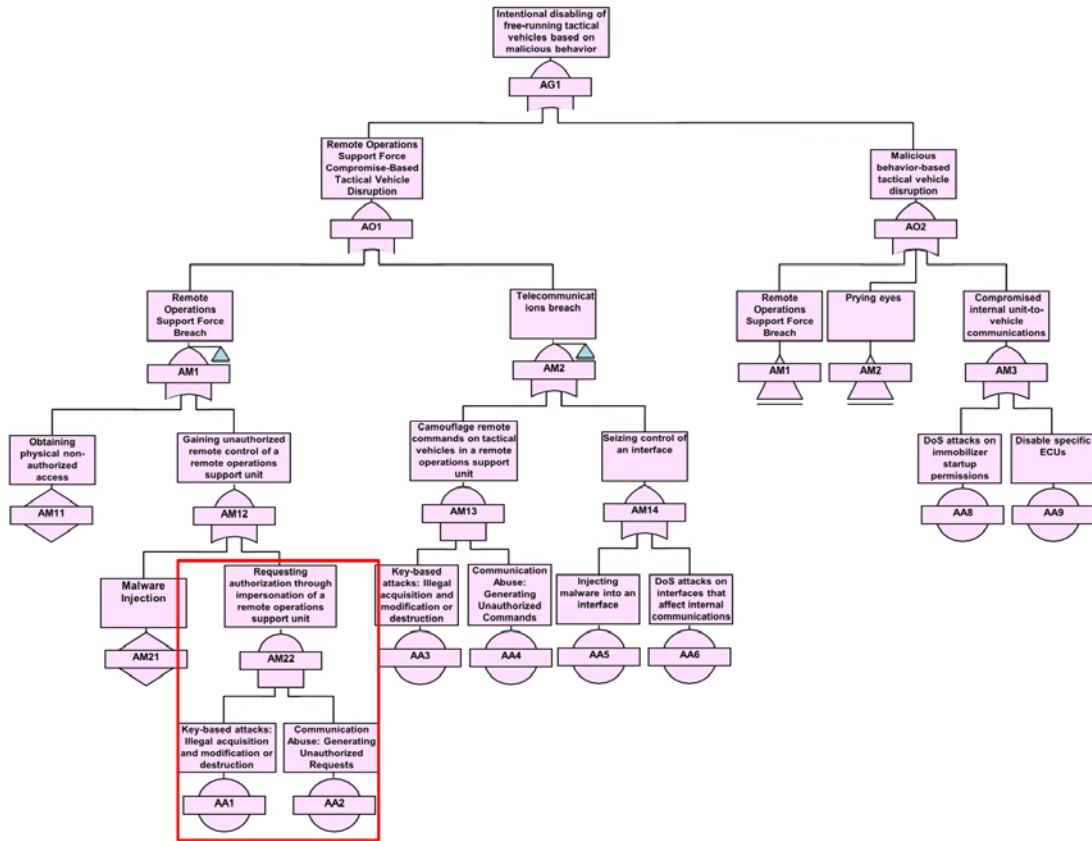


Fig. 2. Partial scenario for intentional disabling of autonomous tactical vehicles based on malicious behavior.

In order to seize control of the normal operation mode (driving/mission), it must be done through the OR gate in the state of system infection due to malware insertion into the control setting communication described above, or in the state of DoS attack on the normal operation mode control setting communication. A DoS attack on the communication system related to establishing control rights for remote driving of unmanned vehicles will reduce the availability of remote driving and autonomous driving communication networks. As a result, it is disconnected and driving, mission execution becomes impossible.

4. Conclusions

In an era where the cyber capabilities of a weapon system have a significant impact on its quality, we present a system development process based on exploratory testing to assess and improve the quality of weapon systems. Used in conjunction with the MITRE ATT&CK and D3FEND Framework, exploratory testing provides an intuitive way to map and manage

vulnerabilities to security controls that comply with the NIST 800-53 standard. This can help you strengthen the security of your systems and improve their quality.

4. Acknowledgments

This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the National Program for Excellence in SW supervised by the IITP(Institute of Information & communications Technology Planning & Evaluation)(2021-0-01393)

References

- [1] MITRE ATT&CK official site (accessed on Nov. 1) : <https://attack.mitre.org/>
- [2] MITRE D3FEND official site (accessed on Nov. 1) : <https://d3fend.mitre.org/>
- [3] CVE official site (accessed on Nov. 1) : <https://www.cve.org/>

Optimal Operation Plan for Unmanned Systems in Future Intelligent Battlefield

Jungho Eom and Kyungil Kim

Department of Security Convergence, Daejeon University
Daejeon, South Korea

[e-mail: eomhun@gmail.com, kimki@itzone.co.kr]

*Corresponding author: Jungho Eom

Abstract

The military operates small unmanned aerial vehicles and IoT systems and plans to operate intelligent robots or drones in the future, but they are operated using different communication ways and networks. In particular, when operating various unmanned systems, different communication systems are used depending on the manufacturer or distribution (development) company, so a separate system is established to ensure interoperability. In this research, we propose a plan to build a base communication environment that can be integrated and operated by unmanned systems such as drones, robots, and autonomous vehicles operated at military bases. The military is preparing basic communication environment standards for unmanned system operation and conducting research to ensure interoperability. We propose an optimized construction plan in terms of data link and network among the basic communication environments for efficient operation and management of unmanned systems.

Keywords: Unmanned System, Interoperability, Communication Environments, Operation Standards

1. Introduction

The military seeks to develop, introduce, and utilize unmanned systems according to each military's battlefield and combat characteristics. The Army seeks to maximize combat efficiency by complementing the combatant-oriented combat system, as integrating unmanned systems such as attack drones and combat robots with existing manned combat systems. The Navy plans to operate water drones, underwater drones, and aerial drones as a power support system, focusing on unmanned weapon systems such as unmanned aerial vehicles, unmanned surface vehicles, and unmanned submarines. The Air Force seeks to construct a multi-communication system, including satellite communication, so that wired and wireless communication can be connected at all times while operating mainly on armed unmanned aerial vehicles with low

observability that prevent exposure to enemy wiretapping and anti-radiation equipment. In addition, the regional network between unmanned fighter aircraft seeks to interoperate with external system networks such as tactical data links for cooperative operations with manned aircraft and connection with the intelligence/surveillance/reconnaissance system and command&control system[1].

Future unmanned systems must have operated manned/unmanned complex combat systems in accordance with development policies and strategies such as convenience of operation, automation, effectiveness, and autonomous operation. Therefore, standardized communication and control technologies such as networks, wireless communications, frequencies, etc. are needed.

In this research, we propose a plan to optimize

the basic communication environment for effective operation of unmanned systems. We explain the current status of the military's unmanned system operation in chapter 2 and explain the current basic communication environment in chapter 3. And we present the optimization plan in chapter 4, and finally conclude in chapter 5.

2. Status of Military Unmanned System Operation

2.1 Military unmanned systems[2,3]

Unmanned systems refer to mobile vehicles, including ground, sea, and air, that recognize the external environment on their own, move according to situational judgment, and perform necessary tasks without human intervention. Depending on the application environment, unmanned systems can be classified as follows.

- Aerial unmanned vehicle: A powered drone that takes off and flies autonomously or is controlled by remote control based on flight dynamics principles without a pilot.

- Ground unmanned vehicle: A mobile vehicle or operation management system that can move and perform operational missions through autonomous control and remote control by automated computing systems inside and outside the vehicle.

- Marine unmanned vehicle: An unmanned vehicle operated on the water or underwater without ground control or piloting personnel.

Ground unmanned systems perform missions such as aerial reconnaissance to collect intelligence in dangerous areas where troops cannot approach, surveillance/reconnaissance in valleys/narrows areas, environmental monitoring in combat areas, supply support, patient rescue, and equipment transportation.

The Unmanned Maritime Systems perform missions such as maritime and underwater surveillance, attacks on enemy surface and underwater forces, mine search and removal, and surveillance/reconnaissance in dangerous sea areas that are difficult for manned systems to approach.

Aerial unmanned systems are mainly used for reconnaissance and attack, and perform missions such as surveillance/reconnaissance, target acquisition, combat damage assessment support, and counter-power warfare.

2.2 The development trends of domestic /international unmanned systems[4-6]

The Ministry of National Defense is presenting unmanned system research goals for the effective operation of various unmanned systems for each military service. According to the Future Defense 2030 Technology Strategy, the development goals are commonization, automation, swarming, modularization, and miniaturization.

- Commonization: Technologies commonly used in platforms such as communication, materials, and batteries for mutual compatibility of each unmanned system component and interoperability of communication, etc.

- Automation: Automation technology such as autonomous flight/driving, autonomous decision-making, and autonomous mission performance combined with artificial intelligence technology

- Swarming: Development of communication and control technology for swarm unmanned systems to efficiently perform various missions simultaneously

- Modularization: Common mission equipment and software technology for major parts and functions to enable multi-purpose use of unmanned system missions

- Miniaturization: Technology related to ultra-miniaturization so that it is not exposed to the enemy and can be easily moved/transported in order to increase the effectiveness of operational missions

The U.S. military seeks to strengthen the core functions of the unmanned combat system by judging it to be the optimal means of combat that can complement the shortcomings of the manned combat system&combatants and maximize the effectiveness of combat missions on the battlefield. The U.S. military is proposing seven long-term development elements to strengthen the functions of unmanned combat systems.

- Interoperability: The ability to share data and information between unmanned combat systems in real time and complete assigned missions through cooperation

- Autonomy: The ability to autonomously choose decisions and actions based on collected and processed data and information

- Airspace Integration : Ensuring sufficient airspace for testing, training, evaluation, and operational flights of aerial unmanned combat systems

- Communication: Technologies such as data links, frequency bands, and network infrastructure for sending and receiving data between unmanned combat systems or between unmanned systems and ground control centers
- Training: Skills and knowledge for operating unmanned combat systems, continuous training for effective operation of autonomous unmanned combat systems
- Propulsion and Power: Ensuring batteries that can maximize operational mission time, and the ability to minimize refueling and power recharging
- Manned-Unmanned Teaming: Integration and construction of cooperative relationships between manned and unmanned combat systems that can maximize combat mission effectiveness

3. Limitations of Unmanned System Operation

Each military forces strengthens operational capabilities using unmanned systems such as combat ground/sea/air unmanned systems, base perimeter surveillance unmanned systems, IoT-based equipment/facility remote monitoring systems, and robot detection systems using unmanned, based on operational characteristics and bases of each military forces. However, there are still problems that need to be resolved to fully perform combat missions. For example, they are lack of integration due to different manufacturers, delays in standardizing operations, and lack of application of security/encryption technologies to protect systems and data.

First of all, the operational unmanned systems currently in the military or those scheduled to be introduced are made by different manufacturers or developers, so they use different communication methods and networks. This is a problem that must be firstly resolved in the integrated operation of unmanned systems.

Second, it is needed measure for integrated operation or interoperability between unmanned systems. In order to effectively achieve combat missions, collected and processed information must be shared in real time between unmanned systems or between manned and unmanned systems, and the mission performance of all unmanned combat systems must be confirmed at the ground command and control center.

Third, security technologies related to system operation and protection of data transmitted

between unmanned systems, between manned and unmanned systems, and between unmanned systems and ground control stations are needed. Recently, gaining right of drone control or leaking transmitted data continue to happen, so the military also needs security measures to prepare for such cyber attacks.

4. Optimization Plan for operating Military Unmanned System

In this chapter, we present a plan to optimize the operation of unmanned systems in terms of the limitations presented in Chapter 3 in order to maximize the combat mission effectiveness of unmanned systems operated by the military[7-9]. First of all, the standardization of data links must be prepared and applied in consideration of the future introduction and expansion of unmanned systems. Since different data links are applied depending on the platform and operating concept of the unmanned system, policy research and technological standardization on communication standards are necessary. Because commercial mobile communication technology is continuously developing, a standardized communication infrastructure system must be established in order to apply communication technology to the unmanned systems that the military plans to operate. For example, since mobile communication technology is moving from 5G to 6G, standardized communication regulations must be established in order to apply this communication technology to military unmanned systems, and a communication infrastructure environment for unmanned systems must be established based on this.

Second, for interconnection and data integration between unmanned systems, an exclusive network must be established between the ground systems and control equipment of all unmanned systems through the defense broadband transmission network (wired), and the collected data and operational data of unmanned systems must be integrated[7]. For example, the aerial unmanned system transmits the collected data and unmanned system operation data to the ground control system through the downlink, and the ground control system analyzes and stores the received data. Data integration is necessary to manage received data efficiently and provide higher quality services. Various types of data transmitted from the aerial unmanned system

must be converted into standardized rules and formats for data management, and the converted data must be stored in the integrated database related to aerial unmanned system.

Third, standardized communication protocols that take into account interoperability between heterogeneous unmanned systems must be established. The North Atlantic Treaty Organization (NATO), focusing on the United States and Europe, has established the STANAG standard, which defines standards for terminology and communication technologies and methods to ensure interoperability between military aerial unmanned systems among NATO countries[8]. The Korean military also needs to develop standards for interoperability of various unmanned systems to be operated by the military based on these standard specifications. Of course, standard communication technology is currently applied introduced unmanned systems to communicate, but physical communication devices and system integration technology are dependent on vendors, making it difficult to expect interoperability and integration effects between heterogeneous systems.

Lastly, the data collected by the unmanned system must be managed by assigning security levels according to importance, and related infrastructure and security systems must be provided so that it can be linked/integrated with the defense AI platform. When data collected from various unmanned systems is refined and converted to extract data necessary for military operations, a secret level can be assigned according to data importance[9]. For this reason, the extracted data must be managed according to its confidentiality level. In addition, security software and systems must be in place to protect the hardware and OS that make up the unmanned system from cyber-attacks. In particular, security technologies such as encryption communication technology, firewalls, and intrusion detection systems must be applied to prevent communicated data from being leaked by spoofing or sniffing attacks.

5. Conclusions

In the future battlefield, a complex manned/unmanned combat system will be operated. In order for these systems to effectively achieve combat missions, they must have operation integration, interoperability,

operational standardization, and security & encryption technologies. In this paper, an optimization plan to meet above conditions is presented.

References

- [1] Jung-Eup Gye, "Trends and Acquisition Strategies on Defense Unmanned Robot Core Technology," *Electronics and telecommunications trends*, Vol.29 No.3, pp. 118-130, 2014.
- [2] Sung-Jin Kim, "National defense unmanned system concept definition and classification plan," *KIDA Defense Issues & Analyses*, Vol.1919, pp.1-12, 2022.
- [3] Gyeong-Su Kim and Yon-Gun Lee, "Classification of unmanned weapon systems and human roles and manned/unmanned complex systems," *Defense & Technology*, Vol.483, pp.130-139, 2019.
- [4] "Future Defense 2030 Technology Strategy," Defense Technology Advancement Research Institute, 2022.
- [5] "'21-'35 Core Technology Plan," Defense Industry Technology Promotion Research Institute, 2021.
- [6] Dong-Seon Park and Kyung-Won Oh, "Study on Development of Korean Unmanned Systems through Analysis of U.S. Unmanned Systems Policy," *Journal of Aerospace System Engineering*, Vol.15, No.3, pp.65-70, 2021.
- [7] Gil-Nam Kim, "Direction for efficient construction of a network for aerial unmanned systems," *KIDA Defense Issues & Analyses*, Vol.1935, pp.1-10, 2023.
- [8] Gyeongrae Nam et al, "A Study on the Analysis and Improvement of STANAG 4586 / MAVLink Protocol for Interoperability Improvement of UAS," *Journal of KIMS Technology*, Vol.23, No.6, pp.618-638, 2020.
- [9] National Security Agency. CSI_CNSA_2.0_FAQ, https://media.defense.gov/2022/Sep/07/2003071836/-1/-1/0/CSI_CNSA_2.0_FAQ_PDF

Experimental Results on Deep Learning-Based Defect Classification in Lithium-ion Battery Component Manufacturing

Thi-Thu-Huyen Vu¹, Haejoong Kim¹, and Taiwoo Chang^{2,*}

¹Department of Industrial and Systems Engineering

²Intelligence and Manufacturing Research Center

Kyonggi University

154-42 Gwanggyosan-ro, Yeongtong-gu, Suwon, Gyeonggi 16227, Korea.

[thuhuyen, haejoong.kim] @kyonggi.ac.kr

*Corresponding author: keenbee@kgu.ac.kr

Abstract

In manufacturing, ensuring the quality of products is imperative prior to their introduction to the market. Detecting and classifying product defects during the production process is one of the top priorities in smart factories. Introducing defective products to the market put consumers at risk and significantly affect the company's brand and value. Hence, before introducing and delivering products to clients, we must strictly control defective on products before taking them to market. Product errors can occur at many stages during the production process, and the most reasonable control is to identify errors and defects right on the production line for each product component. In this study, we present experimental results using deep learning-based methods for classifying defect on micro-fasteners in Lithium-ion battery manufacturing. Exploring the defect detection of tiny objects from Lithium-ion battery component helps to eliminate errors in the final production and reduce costs for the company.

Keywords: Defect detection, Micro-fasteners defect, Deep learning-based defect classification, Defect in manufacturing, smart factories.

1. Introduction

In smart factories, the production line is automated using connected devices, machines, and manufacturing systems to collect and share production data continuously. The data is then used to inform process improvement decisions and determine issues that may arise during manufacturing operations. There are many smart production methods used in a smart factory, such as applying deep learning algorithms, data analysis methods, cloud computing, internet of things (IoT) to operate production. Managing

product quality is indispensable in operating the production line in smart factories, and one of them is detecting product errors (defects) during operations [1]. We can apply and use technologies and different methods to detect errors, defects, or flaws in products to achieve the goal of zero-defect manufacturing [2]. The application of defect classification on products has many important implications and brings significant benefits to manufacturers.

In this paper, we applied different deep learning-based methods for classifying defective products. To be more precise, we have studied

and implemented YOLOv8 [3], ResNet [4], and EfficientNet [5] into micro-fasteners' images to classify not-good (NG) and good (OK) products. The experimental results indicate that using the EfficientNet method achieved the best F1 scores compared to the YOLOv8 and ResNet50 models. The results of this study provide useful information for manufacturers in choosing deep learning models and implementing the defect classification system in their factories for detecting defects on micro-fasteners in Lithium-ion battery component manufacturing.

2. Dataset

The training dataset contains 209 images of micro-fasteners (NG: 137, OK: 72), a part in the Lithium-ion battery and was captured from the production line in a real factory in Korea. Fig. 1 illustrates the micro-fasteners with defects (NG) and no defects (OK). For enriching the dataset, we conduct augmentation for the dataset with the "albumentations" library [6] and have 836 images of micro-fasteners (NG: 548, OK: 288). We use the test dataset with 28 images (NG:14, OK: 14) to evaluate and compare the effectiveness of different proposed models.

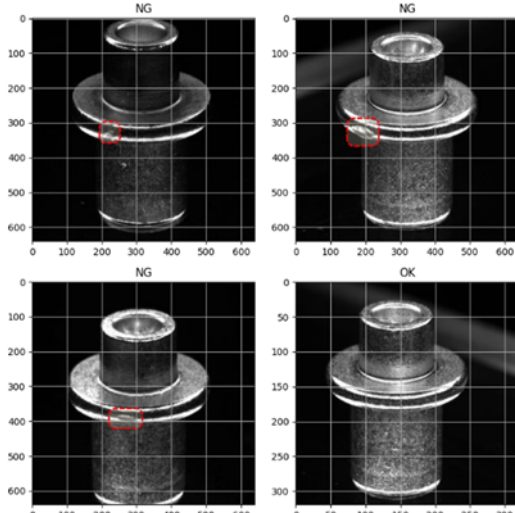


Fig. 1. Example micro-fasteners in the dataset

3. Experimental Results

After preparing and preprocessing the dataset, we conducted training the dataset with different methods. All the models were trained and

converged through 100 epochs.

With the YOLOv8 network, we annotated the NG images with the bounding box surrounding the defects inside the images. After that, we evaluate the model by predicting the defects on test images. Images containing predicted defects are marked as NG images. As we can see in Fig. 2, the trained model consistently improved accuracy, reaching a convergence point after 100 epochs. The training accuracy curve represents a continuous decrease in the loss function, indicating effective learning. The mAP50 is achieved at 0.871, however, mAP50-95 is only achieved at 0.431.

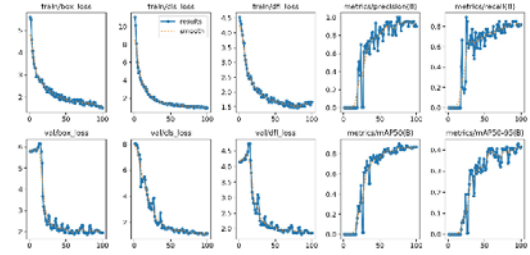


Fig. 2. Training results with the YOLOv8 on the dataset

With the Resnet model, we used the Resnet50 network for training the model. Fig. 3 represents this model's training and validation accuracy on the dataset. As shown in the Figure, the training and validation accuracy were converted to 1 through 100 epochs.

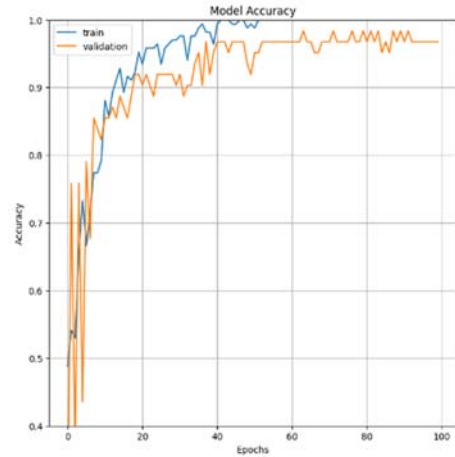


Fig. 3. Training results with the Resnet50 on the dataset

Similar to Resnet50, using EfficientNetB5 resulted in the convergence of training and validation accuracy to 1 after iterating 100 epochs. These values highlight the performance and convergence observed through different networks.

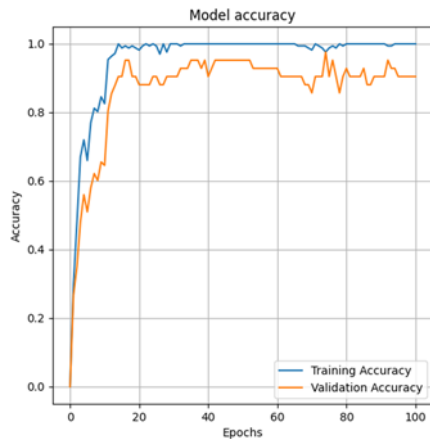


Fig. 4. Training results with the EfficientNetB5 on the dataset

For evaluating the trained models, we calculate the F1 scores with the predicted label of images from the test dataset. The results of evaluating F1 scores are represented in **Table 1** and calculated based on the equation (1) below.

$$F_1 = \frac{2*TP}{2*TP+FP+FN} \quad (1)$$

Where:

- TP: True positive
- FP: False positive
- FN: False negative

Table 1. Comparison of F1 scores between different models

| | YOLOv8 (Conf: 0.6) | Resnet50 | EfficientNetB5 |
|----------------|-----------------------|----------|----------------|
| F ₁ | 0.833 | 0.933 | 0.9655 |

As shown in **Table 1**, EfficientNetB5 outperforms the remaining models in classifying the defect images from the test dataset, achieving F1 score of 0.9655.

4. Conclusions

In this paper, we conducted evaluating different deep learning methods for classifying defective products in manufacturing. In

particular, the YOLOv8, Resnet50, and EfficientNetB5 were used with a real dataset of micro-fasteners for classifying not good and good products. The experimental results showed that EfficientNetB5 outperforms the remaining models in classifying defects on micro-fasteners.

References

- [1] P. Osterrieder, L. Budde, and T. Friedli, "The smart factory as a key construct of industry 4.0: A systematic literature review," *Int. J. Prod. Econ.*, vol. 221, no. November 2017, p. 107476, 2020, doi: 10.1016/j.ijpe.2019.08.011.
- [2] F. Psarommatis, G. May, P. A. Dreyfus, and D. Kiritsis, "Zero defect manufacturing: state-of-the-art review, shortcomings and future directions in research," <https://doi.org/10.1080/00207543.2019.1605228>, vol. 58, no. 1, pp. 1–17, Jan. 2019, doi: 10.1080/00207543.2019.1605228.
- [3] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2016, vol. 2016-Decem, pp. 779–788. doi: 10.1109/CVPR.2016.91.
- [4] K. H. X. Z. S. R. Jian, "Deep Residual Learning for Image Recognition arXiv:1512.03385v1," *Enzyme Microb. Technol.*, vol. 19, no. 2, pp. 107–117, 2015, [Online]. Available: <http://image-net.org/challenges/LSVRC/2015/>
- [5] Q. V. Le Mingxing Tan, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks Mingxing," *Can. J. Emerg. Med.*, vol. 15, no. 3, p. 190, 2013.
- [6] A. Buslaev, V. I. Iglovikov, E. Khvedchenya, A. Parinov, M. Druzhinin, and A. A. Kalinin, "Albumentations: Fast and flexible image augmentations," *Inf.*, vol. 11, no. 2, Feb. 2020, doi: 10.3390/INFO11020125.

Developing Human-Following Robot to Collaborate with Railroad Workers

YeonSoo Kim, Tai-Woo Chang, Jong-Un Won and Baek-Hyun Kim

Department of industrial and Systems Engineering, Kyonggi University
Korea Railroad Research Institute Artificial Intelligence Laboratory
Suwon, South Korea

[e-mail: yeonsoo98@kyonggi.ac.kr, keenbee@kgu.ac.kr, juwon@krri.re.kr, bhkim@krri.re.kr]

*Corresponding author: Tai-Woo Chang

Abstract

This study aims to recognize the importance of safety and accident prevention for railway workers, as well as to improve the railway work environment. Considering the characteristics of the railway work environment, the goal is to develop a transport robot capable of carrying heavy loads. The transport robot is designed to autonomously follow railway workers to ensure collaboration. If it detects workers or obstacles within a certain distance, it will automatically stop moving, taking the worker's safety into account. Configured based on the ROS 2 platform for seamless integration of hardware and software, the transport robot employs YOLOv5 for real-time object detection. By implementing these technologies, it is expected to establish a secure and automated work environment, thereby enhancing the convenience and efficiency of railway workers.

Keywords: Robotics development, Human-following robot, ROS 2, Railroad workers

1. Introduction

In the railway industry, railroad safety accidents occur continuously. There is a need to enhance safety and accident prevention measures for railroad workers and improve the work environment. The railway work environment involves a variety of heavy equipment and tools, leading to frequent tasks involving the transportation of heavy loads [1,2,3]. To mitigate these challenges that adversely impact workers' safety and operational efficiency, we propose the development of a transport robot. The robot collaborating with workers in material handling environments has two functions of autonomous driving for self-navigation and human-following for collaboration. This can improve convenience

and increase work efficiency while ensuring worker safety. To ensure seamless integration between the hardware and software of the transport robot, the development is planned based on the ROS (Robot Operating System) 2 platform [4]. Considering the ability to detect workers and obstacles in real-time as the robot moves and aiming for relatively high accuracy, the YOLOv5 model has been employed [5].

2. Robot Design

2.1 Robot Hardware Design

The transport robot was designed with a lightweight frame to improve energy efficiency. Structures and connectors were built for flexible loading and modularization of assembly and

This work was supported by the Technology Innovation Program (20019178, Development of AI-based manufacturing process logistics optimization technology that supports operation optimization for each process situation) funded by the Ministry of Trade, Industry & Energy (MOTIE, Korea). This work was also supported by the Korea Institute for Advancement of Technology (KIAT) grant funded by the Korea government(MOTIE) (P0008691, HRD Program for Industrial Innovation).

disassembly to facilitate collaboration with the operator. NVIDIA Jetson Orin NX was used as a single board computer to control the robot, and a depth camera and LiDAR were used for near and long-range object detection.

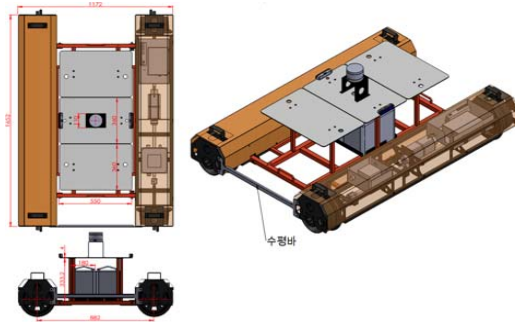


Fig. 1. Robot Design Configuration

2.2 Robot Software Design

The transport robot system has been set up utilizing the ROS 2 platform. Through the integration of hardware and software, the transport robot collaborates with railroad workers. The operating system is Linux Ubuntu 20.04, and the system includes communication code written in Python, keyboard and joystick control codes, motor connectivity code, among others.

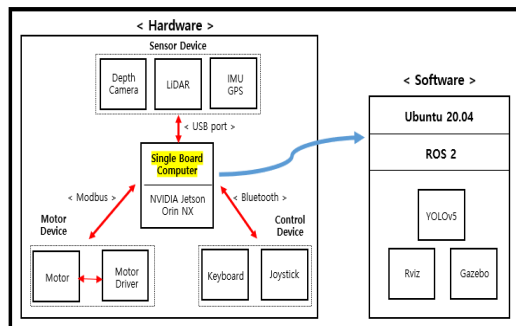


Fig. 2. Configuration of the Following Robot System

3. Robot Development

3.1 Human-Following & Safety Driving

Before the robot begins its tasks, it automatically initiates camera node connection and object detection using YOLOv5. It collects ROS 2 based real-time vision sensor data and measures

distances. When workers are detected in front, the robot drives towards the closest person at a consistent speed and follows. If workers or obstacles are detected within 1 meter in front of the robot, it immediately halts. Through this human-following and safe driving, it can efficiently support the tasks of the worker and ensure safety.

3.2 Hardware Development

The required functions were realized according to the design plan presented earlier. The transport robot can carry 20kg to 30kg at a time and is equipped with a safe driving function that stops when it detects workers or obstacles within a certain distance, 1m.



Fig. 3. Human-Following Robot

4. Conclusions

This study describes the process and results of developing a transport robot that collaborates with a railroad worker to carry heavy loads, aiming to improve the working environment of railroad workers. The robot enhances workers' convenience and work efficiency by following them and implements a safe driving function considering their safety. In the future, we plan to improve the robot's performance, to add autonomous mobile driving technology, to develop robot data collection and interfaces and to create a simulation environment similar to the real one to build a digital twin.

References

- [1] Seoung-Ji Kim, Lee, Choul-Ki, Byeong Soo Yum and Won-Mo Gal, "A Study on the Cause Analysis of Railway Worker Accidents Using Heinrich and Bird's

- Domino Theory,” *Journal of The Korean Institute of Plant Engineering*, Vol 27, No.1, pp.21-32, 2022.
- [2] You-Ho Kim, Soo-Hwan Lee, Soo-Joo Lee and Wang-Seok Oh. “Case Study of Protecting Maintenance Worker from Accidents and Study of Train Detection Technology,” *Journal of The Korean Society For Urban Railway*, Vol 4, No.4, pp.593-602, 2016.
- [3] Jong-Gyu Hwang and Hyun-Jeong Jo. “A Study on Bidirectional Detection Safety Equipment Mechanism for Casualty Accidents Protection of Railroad Workers and Motorcars,” *The Transactions of the Korean Institute of Electrical Engineers P*, Vol 59, No.4, pp.384-389, 2022.
- [4] J.-M. Kim, J.H. Jo, and C.-b. Moon, "People Tracking using the Grid and CAD Map in the Large Scale Logistics Environments," *Korean Society for Precision Engineering*, pp.291-292. 2017
- [5] Hyun-sik Son, Deok-keun Kim and Seung-hwan Yang, “Development of Following Algorithm for Human-Following Robots using Deep-Learning Technology and Depth Camera,” *Journal of Institute of Control, Robotics and Systems*, Vol 28, No.2, pp.95-101, 2022.

Assessing Digital Transformation Levels of Partners for Supply Chain Collaboration

Yong-Jun Im and Tai-Woo Chang*

Department of Industrial and Systems Engineering, Kyonggi University
Suwon, South Korea

[e-mail: imyj@kyonggi.ac.kr, keenbee@kgu.ac.kr]

*Corresponding author: Tai-Woo Chang

Abstract

Although information and communication technology is advancing rapidly, digital transformation (DX) within the manufacturing and associated sectors remains insufficient. Notably, challenges like recurring communication expenses and supply chain inventory shortages across multiple companies persist. This research seeks to enhance collaboration by assessing the level of DX among supply chain partners of a company. In this study, a survey was conducted with the company's current dealers and parts suppliers, and the responses from 85 companies were analyzed. Leveraging an established digital industrial innovation maturity diagnostic model, we incorporated diagnostic elements across seven domains within the survey, conducted an analysis to validate and make practical use of the digital transformation level diagnostic model system through frequency and correlation analyses, and derived insights to enhance supply chain collaboration.

Keywords: Digital transformation, Maturity assessment, Correlation analysis, Clustering

1. Introduction

The recent enactment of Korea's Industrial Digital Transformation Promotion Law has garnered considerable attention from numerous companies in the field of industrial digital transformation (DX). Concurrently, initiatives are in progress to establish digital collaborative factories [1].

In the machinery industry, the absence of necessary parts during production poses challenges, leading to frequent communication through calls, emails, and visits among supply chain partner to ascertain supplier availability. In response to this, companies are making significant efforts to resolve supply chain issues; nevertheless, challenges persist. Therefore, with the objective of constructing a digital

collaborative factory, the company aims to gather real-time data spanning the reception of parts, product assembly, inspection, and delivery. This will be achieved by leveraging data acquired from MES/APS, fostering collaboration through shared information among supply chain participants. So, it is necessary to assess the level digital transformation of partners related to the company and formulate a strategy to enhance their DX capabilities.

2. Related Works

Industrial digital transformation assessment models focused on manufacturing industry include Germany's Industrie 4.0 Readiness Model, Austria's Industry 4.0 Maturity Model, and Singapore's Smart Industry Readiness Index [2,3,4]. In Korea, there also exists a system for

verifying smart factory levels, similar to these [5].

Based on this, Chang *et al.* Developed a digital industrial innovation maturity diagnostic model consisting of seven assessment areas and 28 criteria [6]. Maturity diagnostic models are also employed in other service industries and beyond [7]. This study applies an existing manufacturing-based maturity diagnostic model [6] to assess the DX maturity of a company and its suppliers in the supply chain, and draws strategic implications for building a supply chain DX cooperation system.

3. Research Process

The survey asked for basic information such as the company's industry, revenue, and number of employees, and used 28 evaluation items in 7 areas by applying an existing model to diagnose the DX level.

The survey was conducted among suppliers and distribution dealerships that do business with the agricultural machinery manufacturer T. Out of around four hundred partners, 85 responded, including 46 suppliers and 39 dealerships.

As both online and offline surveys were conducted, data such as revenue and number of workers were not displayed as categories, and some companies entered exact values, so they were grouped into categories and preprocessed. In addition, if there were duplicate responses, such as the industry and the respondent's department, we changed them to a single value. Therefore, we preprocessed the data for 2 cases of company size, 10 cases of industry, 6 cases of annual revenue, 5 cases of number of workers, and 2 cases of respondents' departments.

4. Analysis Results

For the statistical analysis, we analyzed variables such as company classification (supplier, dealer), revenue, and company size. In this paper, we focus on the suppliers.

As shown in Fig. 1, the average score for each assessment area was compared by the percentage of revenue generated by the supplier to the manufacturer T. As a result, companies that have a high percentage of sales to the company T had relatively low scores in the remaining areas, except for the data area. It is necessary to

consider companies with a relatively high proportion of sales to construct a digital collaboration factory.

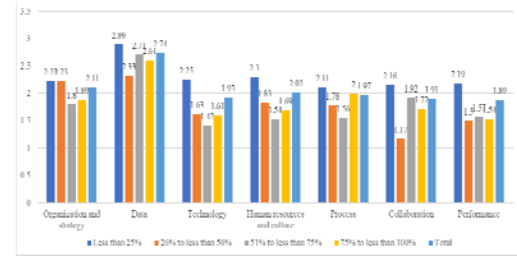


Fig. 1. Partner level assessment by sales ratio

Utilizing survey results of the partners, we conducted correlation analysis and observed that data utilization exhibited lower correlations with other evaluation items, as shown in Fig. 2. On the other hand, we observed a correlation between different evaluation items in the section of collaborating with partners. This could mean that they lack the information collaboration to share their data with their partners. In the supplier data, 'Cooperation with suppliers' and 'Digitalization education and training' have a high score of 0.82, which confirms the need for education and training.

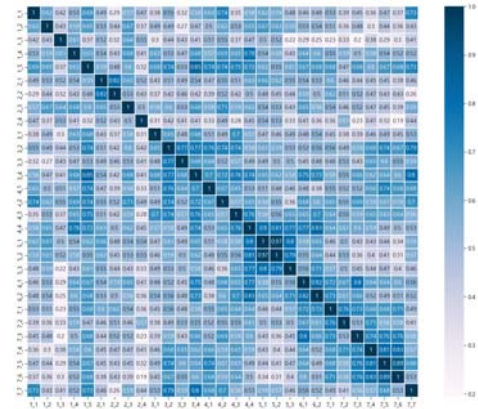


Fig. 2. Correlation analysis among evaluation items

Clustering was performed using data on revenue, revenue share, workforce size, and survey results for each item. While the data can be categorized into three primary groups, we conducted the analysis by segregating and labeling the data into four clusters, as illustrated in Fig. 3, to prevent excessive focus on any one group.

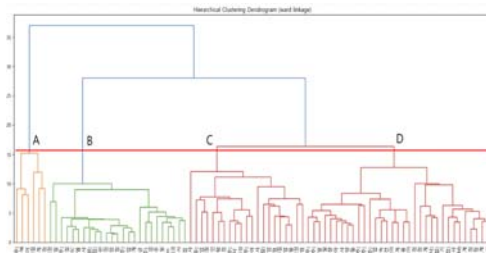


Fig. 3. Clustering Result Dendrogram

Within Group A, suppliers account for 40% of the total, and while dealerships exhibited high scores in the organizational and strategic aspects, their scores were lower in the collaboration category. On the other hand, within Group B, where suppliers constitute 44%, this group falls under the category of companies with the lowest scores. Particularly, the process scores were notably low, indicating a need for activities aimed at process improvement. In the case of Group C, it mainly consists of companies with slightly above-average scores, primarily composed of suppliers (67%). Notably, high scores in the data category are observed within this group. In Group D, suppliers (55%) score low on process and technology categories, and performance scores are also low. So, technology and process categories need to be reviewed. In Group D, the revenue is higher than the other partners. However, the percentage of the revenue is low, and the company needs to be reviewed.

5. Conclusion

This study applied a DX maturity diagnostic model to identify the current level of companies that have a supply chain to construct a digital collaboration factory.

Basic statistical analysis was used to compare the mean of the overall scores of suppliers and dealerships. Correlation analysis was used to determine the validity of the survey items, which revealed a lack of cooperation among suppliers in exchanging data. Four clusters were formed based on the characteristics of the companies and the survey results, and the analysis revealed the current status of digital maturity diagnosis for each group and drew implications for building a supply chain collaboration process.

Based on the analysis, strategies for strengthening DX capabilities in supply chain

collaboration management should be established by utilizing techniques such as AHP and quantitative company information.

References

- [1] Korea industry intelligentization association, "Digital Transformation", Vol. 32, 2023.
- [2] <https://www.industrie40-readiness.de>
- [3] Andreas Schumacher et al., "A Maturity Model for Assessing Industry 4.0 Readiness and Maturity of Manufacturing Enterprises", *Procedia CIRP*, Vol. 52, pp. 161-166, 2016.
- [4] <https://siri.incit.org/>
- [5] Lee et al., "Developing a framework for assessing maturity of future manufacturing system and case study", *Journal of Society for e-Business Studies*, Vol.24, No.2, pp 165-178, 2019.
- [6] Chang et al., "A Study on Digital Maturity Assessment for Digital Industrial Innovation", *Journal of Society for e-Business Studies*, Vol.27, No.3, pp95-111, 2022.
- [7] Choi et al., "Development and Study of Digital Literacy Indicators (Checklist) for Micro Business Owners for Continuous Digital Transformation: Focusing on the Tertiary Industry", *Journal of Information Technology Applications & Management*, Vol.30, No.1, pp81-95, 2023.

Establishing a digital twin system using logistics simulation and BI tools - Case study in an automotive industry

Gyusun Hwang, Jihyeon Park and Haejoong Kim*

School of Industrial Engineering, University of Ulsan, Korea

Department of Industrial and Management Engineering, Kyonggi University, Korea

[e-mail: haejoong.kim@kyonggi.ac.kr]

*Corresponding author: Haejoong Kim

Abstract

The current automotive industry is in the process of transitioning from Internal Combustion Engine vehicles to Electric vehicles. As a result of the shift towards mixed production, the need has arisen to proactively identify and respond to potential production management issues. This research aims to address these challenges by conducting logistics simulation in a digital twin environment. In order to perform a simulation, this research tries to define the required data, establish the system architecture and derive the Key Performance Indicators necessary for result analysis. Additionally, this study tries to define an architecture for connecting Business Intelligence(BI) tools with the digital twin environment to establish a real time data-driven decision making. This study demonstrated that the integration of predictive digital twins and analytical methodologies for decision-making can provide an environment that allows for a rapid response to the changing needs of the evolving factory.

Keywords: Digital Twin; Logistics simulation; Business intelligence; Data driven decision making

1. Introduction

With the increasing focus on sustainable development and carbon emissions across industries, the global automotive industry is also introducing Electric Vehicle(EV) to replace Internal Combustion Engines(ICE), along with carbon neutrality policies. The transition from ICE to EV vehicles has led to major changes in the key components that comprise a car and how they are produced[1].

Nowadays, logistics simulation techniques are becoming increasingly important as a technique for constructing a digital twin environment of a factory. Logistics simulation is a technique for modeling physical or abstract

systems and experimenting with them. It is used to imitate and mimic real-world situations or systems to predict various situations and find or optimize alternative solutions.

With the application of Digital Transformation in the factory environment, a multitude of IoT(Internet of Things) sensors have been deployed in equipment. As a result, it has become possible to establish a more sophisticated Digital Twin environment, allowing for data-driven decision making in the factory using real-time monitoring and predictive logistics simulation results. The systematic computational analysis of data will lead to more informed decisions, which will in turn enhance the effectiveness of smart manufacturing [2].

This study implement a simulation model to predict problem situations and find solutions to situations that cannot be executed in the real world. This study tries to explain how data-driven decisions are possible in a digital twin environment.

2. Problem definition and Research objectives

In this study, a simulation is conducted in a digital twin environment to derive a solution to the material warehousing logistics congestion problem that occurs in the process of mixed ICE and EV production, and then the results analyzed. The problem situation in this study is as follows. 1) it is impossible to predict the logistics flow congestion of material receiving vehicles due to the mixed production of ICE and EV vehicles, 2) As EV parts become larger than ICE parts and the frequency of delivery increases, it becomes difficult to estimate the capacity of logistics resources and the layout of loading and unloading yards suitable to produce EV parts. 3) It is essential to visually inspect simulation scenarios and establish a system framework for utilizing BI tools with the result data.

A logistics simulator is developed using Tecnomatix Plant Simulation, a manufacturing simulation software from Siemens, to predict the logistics flow of materials and products in and out of a mixed production process. The goal of this study is to conduct what-if simulations by configuring various logistics scenarios due to electrification, and to identify potential logistics problems through KPI derivation and data analysis using BI tools from Tibco Spotfire.

3. Literature review

Table 1 summarizes the research contents of the related literature, divided into developed algorithms, whether simulation is used, and whether it presented system architecture.

Table 1. Related research summary

| Related research | Algorithm | What-if Simulation | System Architecture |
|------------------|----------------------------|--------------------|---------------------|
| [3] | Information entropy theory | x | o |
| [4] | Simulation, Optimization | o | x |
| [5] | Simulation | o | o |

4. Development of digital twin environment

Digital twin simulations are conducted in five steps: collecting and pre-processing data, modelling the system, performing simulations, validating simulation accuracy, and making predictions from scenarios (**Fig. 1**).



Fig. 1. Simulation steps

This study define each step according to the simulation development steps and describe the actual adoption real case study. The simulation problem in this paper is a tier 1 vendor supplying parts to a Korean automotive company that produces a mix of ICE and EV parts.

The first step is the simulation data collection and preliminary preparation phase, where the requirements for setting up the system are analyzed, and input data is processed into the required format. This process involves structuring input and output data to reflect the characteristics of the target system for modeling and is a crucial foundational step in modeling. IEC, EV parts production data, Part storage capacity data, Truck and Forklift information data and Work calendar data are collected. The changes in supply due to the increasing demand for electric vehicles have been incorporated into the data for both ICE and EV components.

The second step involves defining the relationships between various entities and resources and incorporating scheduling rules and

various constraints to establish the simulation environment.

The third step involves the actual execution of the simulation. As the scale and complexity of the modeling increase, the execution time also tends to become longer. The digital twin environment developed in this study depicted in Fig. 2.



Fig. 2. Digital Twin display snapshot

The data architecture organization of the simulator is shown in Fig. 3.

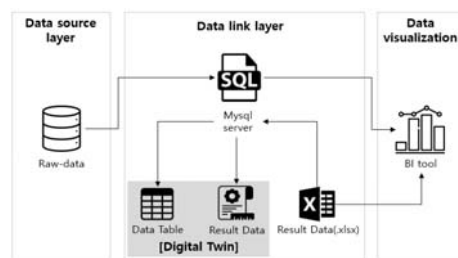


Fig. 3. Connection to data and BI tools

The step of validation and result analysis is dedicated to ensuring the accuracy of the simulation by verifying and confirming whether the constructed simulation environment accurately represents the intended scenarios. This study verified that the algorithm is functioning correctly by cross-referencing the dispatch execution results data table on a day by day basis with the work schedule to ensure that the number of trucks dispatched matches those entering, handling components, and leaving the factory.

This study have developed three KPIs for use in data-driven decision-making. The first is the 'Number of Unloaded Trucks KPI,' which represents the number of trucks. The second is the 'Distribution Compliance Rate,' which is the ratio of actual trucks dispatched to the planned

trucks dispatched. Finally, the 'Distribution Load Factor' indicates the time spent on dispatching during the workday.

5. Conclusion and future works

This study aims to establish a digital twin environment for predicting changes in the manufacturing landscape and to build a system for real-time decision-making. To achieve this, this study defined the electric vehicle hybrid production environment and outlined the simulation procedure for development. Additionally, we designed a system architecture for implementation, thus establishing a real-time decision-making framework.

References

- [1] Poornesh, Kavuri, Kuzhivila Pannickottu Nivya, and K Sireesha. "A Comparative Study on Electric Vehicle and Internal Combustion Engine Vehicles." Paper presented at the 2020 International Conference on Smart Electronics and Communication (ICOSEC), 2020.
- [2] Tao, Fei, Qinglin Qi, Ang Liu, and Andrew Kusiak. "Data-Driven Smart Manufacturing." *Journal of Manufacturing Systems* 48: 157-69, 2018
- [3] Liu, Haijiang, Kaixiang Xu, and Zhenhua Pan. "Modeling and Application of Mixed Model Assembly System Complexity Introduced by Auto-Body Personalization." *The International Journal of Advanced Manufacturing Technology* 93: 43-54, 2017
- [4] Zhou, Binghai, and Zhexin Zhu. "A Dynamic Scheduling Mechanism of Part Feeding for Mixed-Model Assembly Lines Based on the Modified Neural Network and Knowledge Base." *Soft Computing* 25, No. 1: 291-319, 2021
- [5] Pereira, Teresa, Mariana Xavier, Fernanda A Ferreira, and Marisa Oliveira. "Improvement in Planning and Resource Management for an Automotive Company's Parts Feeding System." Paper presented at the Proceedings of the International Conference on Industrial Engineering and Operations Management, 2022.

Classification of fire risk small electric home appliances using deep learning

Jun-Seong Kim, Yong Jun Im and Hyun-Soo Kim*

Department of Industrial and Systems Engineering, Kyonggi University
Suwon, South Korea

[e-mail: jsjs9444@kyonggi.ac.kr, imyj@kyonggi.ac.kr, hskim@kyonggi.ac.kr]

*Corresponding author: Hyun-Soo Kim

Abstract

Recently, the emergence of small electric home appliances with built-in batteries, such as portable small appliances for user convenience, has led to explosion accidents during waste disposal. If a waste lithium battery explodes, it is difficult to put out other fires, and the damage to human life and company assets is enormous. As a result, it is necessary to prioritize the classification of products containing batteries in the recycling process of waste household appliances. In this study, in order to test the possibility of classifying typical hazardous small electric home appliances through image learning, 969 images of small household appliances were applied to a YOLO model capable of real-time object detection, and a fairly high detection rate was obtained.

Keywords: Recycling, Deep learning, YOLOv5, Home appliance, Lithium battery

1. Introduction

The recent surge in the popularity of portable and convenient small electric home appliances has made our lives much easier. However, the recycling process required to dispose of these products has become increasingly risky.

Since most portable small electric home appliances are powered by lithium batteries, the risk of fire and explosion during the recycling process is very high, unlike conventional plug-in appliances. As a result, there have been cases of explosions in the field [1][2].

In the event of a lithium battery fire, it is not only difficult to extinguish the fire, but most recycling companies are facing difficulties due to the high risk of dust explosions and chain fires in the environment of waste electronics recycling centers.

To address these emerging risks, this study

developed a system to distinguish and classify lithium battery products from non-lithium battery products by using the YOLOv5 model, a deep learning-based object detection algorithm.

The results of this study may provide a solution to the typical risk by enabling small-scale e-waste recyclers to effectively categorize hazardous lithium battery products and handle them separately.

2. Related Works

Recently, the recycling industry has been interested in automated sorting using robots, and various studies have been conducted.

In particular, the field of deep learning using YOLO is actively being researched in the field of household waste, and the results of these efforts are beginning to show [3][4][5][6].

Waste Management, a US recycling company, has created software and hardware that

automatically sorts household waste and is now commercializing it, and the products are so good that they can replace human labor [7].

Unlike other recycling fields, existing studies on consumer electronics recycling have focused on product value and process methods, such as "Status and Prospects of Waste Small Appliances Recycling" by In-sang Jung and Jin-hwan Park [8].

While these processes and methods were practical before the widespread use of portable small appliances, further research is needed, such as sorting batteries from small waste appliances for safe recycling.

3. Research Flow

3.1 Collecting Data

To collect the data needed for the analysis, images of various small electric home appliances that are actually being discarded were taken with a smartphone (Fig. 1).



Fig. 1. Images of Waste Home Appliances

We used a total of 15 classes, and Table 1 shows the composition of the dataset for each class. The total data consists of 969 images. Of these, laptops, robot vacuum cleaners, necklace fans, stand fans, and Bluetooth speakers are among the products that contain explosive lithium batteries (35 percent), while the remaining classes consist of small appliances that do not have explosive potential (65 percent).

Table. 1. Data collection status of 15 classes

| Classes | Images |
|---------|--------|
| Laptop | 233 |

| | |
|----------------------|-----|
| Radio | 41 |
| Robot Vacuum Cleaner | 74 |
| Vacuum Cleaner | 16 |
| Necklace Fan | 10 |
| Stand Fan | 5 |
| Blender | 95 |
| Rice Cooker | 167 |
| Bluetooth Speaker | 18 |
| Stand Speaker | 9 |
| Coffee Machine | 20 |
| Coffee Pot | 44 |
| Keyboard | 47 |
| Toaster | 82 |
| Hair Dryer | 108 |
| TOTAL | 969 |

3.2 YOLO Model

YOLO (You Only Look Once) is a deep learning algorithm that performs real-time object detection, dividing the image into a grid and detecting objects in each grid cell in a single forward propagation, making it fast and suitable for real-time applications. In fact, for use in the recycling process, we adopted a YOLO algorithm that works as a one-stage detector to detect objects passing by at high speed. There are various versions of the YOLO model, but compared to previous deep learning models, this study uses YOLOv5, which has demonstrated high performance in a fast and accurate environment [9].

4. Experimental results

In this study, we conducted an experiment to recognize and classify the types of small electric home appliances using YOLOv5 and checked its performance. Model training was performed on an Intel i9-13900K CPU, 24GB RTX 4090 GPU, and 128GB Ram.

The evaluation of the trained model used the mAP0.5 and mAP0.5_0.95 metrics along with precision and recall. mAP0.5 indicates the percentage of object detections and correct answers that match 0.5 or higher, and mAP0.5_0.95 means the average value obtained by measuring the mAP every 0.05 units from 0.5 to 0.95. The training parameters for training are Image size: 320, batch size 16, epochs 300.

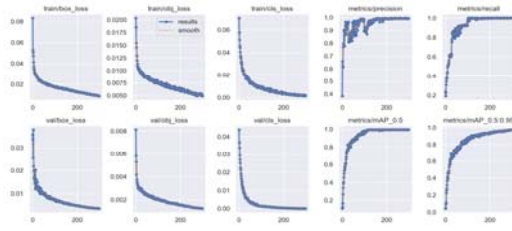


Fig. 2. YOLO Learning Curve

After training the YOLOv5s model, it took 4.3 hours to train 300 epochs and achieved Precision 0.9927, Recall 0.9997, mAP@0.5 0.9948, and mAP@0.5:0.95 0.9736. As shown in Fig. 2, which summarizing the experimental results, the learning curve was unstable at first, but then the learning stabilized progressively.



Fig. 3. Test results using a webcam

Based on the trained dataset, OpenCV was used to recognize small electric home appliances with a webcam, as shown in Fig. 3, and the performance of the test was decreased depending on the angle applied.

5. Conclusions

In this paper, we have defined explosive product groups by utilizing actual field data to classify certain products with fire potential. After collecting the image data, we trained the YOLOv5s model using 969 images, and the result mAP@0.5:0.95 showed an accuracy of 0.9736. Using this, we tested a real webcam and found that it performed well when the angle for the recognition of the image is similar to the image which was captured data we used for training but we obtained poor performance when the applied angles are varied.

Therefore, we suggest the plan to organize an experimental environment similar to the actual recycling process to adjust the camera angles for classifying small electric home appliances, and to acquire additional image data to solve the class imbalance problem and apply it to training. It is expected that it can be applied to replace manual labor in the actual work environment to provide work convenience and solve the problem of fire and explosion hazards for workers and the work environment.

References

- [1] Yonhap News, "Let the uncomfortable wire go"...'portable appliances' are on the rise, Yonhap News, June 17, 2015
- [2] Hyun Il Shin "Household recycling plant fire in Gimcheon, Gyeongbuk...Waste battery storage", Daily News, January 30, 2023
- [3] Kim et al., (2021). "Automatic Garbage Classification System based on YOLOv4 and Raspberry Pi". Journal of Digital Contents Society, vol.22, no.12, pp.2111-2119, 2021
- [4] Kim et al., "YOLO Based automatic Sorting System for Plastic Recycling", Proceedings of the Korean Institute of Information and Commucation Sciences Conference, 2021.10a, pp382-384, 2021.
- [5] Kim et al., "Deep learning algorithms and application". Communications of the Korean Institute of Information Scientists and Enginners, vol.33, no.8, pp 25-31, 2015
- [6] Kim et al., (2022). "A study on the application of deep learning-based object recognition algorithm in automatic sorting system of recyclable garbage", Proceedings of the Korean Institute of Control Robotics Systems Conference, 2022.06a, pp718-719. 2022
- [7] Hyun Jin Jung "AI to Sort Waste by Shape, Size: U.S. Recycler Adopts It", Asia Economic, November 8, 2023
- [8] Jung et al., "Overview and Recent Development of Recycling Small Waste Eletical and Electronic Equipment", Resources Recycling (JKIRR), vol.24, no.4, 38-49, 2015
- [9] roboflow, <https://blog.roboflow.com/yolov4-versus-yolov5>

Real-time heuristics for balancing mixed-model assembly lines

Minseok Kim¹, Hyeongjun Son², Byeong-uk Ha², Soyun Kim¹, Byeonghyun Cha¹, Haejoong Kim²

¹Department of Industrial Management Engineering, Korea National University of Transportation
Chungju-si, South Korea

²Department of Industrial and Systems Engineering Kyonggi University
Suwon-si, South Korea

[e-mail: millenniumd2@naver.com, shjmj187@kyonggi.ac.kr, hbu1370@naver.com,
sally5157@kyonggi.ac.kr, malong0313@naver.com, haejoong.kim@kyonggi.ac.kr]

*Corresponding author: Haejoong Kim

Abstract

In a mixed production process, a bottleneck occurs when products that are heavily loaded on a specific process are continuously inputted. You can prevent bottlenecks by leveling the types of products that are inputted through scheduling in advance to maintain the balance of production facilities. However, in the case of static scheduling, it is not possible to cope with new unplanned factors such as machine breakdowns, so bottlenecks occur in the field, resulting in lower production efficiency. We propose a real-time sequencing heuristic to cope with new unplanned bottlenecks in mixed-model assembly lines. Experiments were conducted on a line producing several types of agricultural tractors to verify the performance of the input ordering method.

Keywords: Bottleneck detection, Dispatching algorithm, Mixed production, Simulation

1. Introduction

Many companies, including those that produce automobiles or auto parts, manufacture similar products of various specifications. At this time, manufacturing products of each standard in each dedicated process requires a lot of production lines and cost investment. To solve this problem, companies are using mixed production, which produces multiple products in one production line. Simulation experiments confirmed that when a product to be input in mixed flow production is arbitrarily determined, production decreases as the types of products vary and the working time of each product varies [1]. In order to increase production efficiency in the mixed flow production process, studies were conducted to plan the inventory input schedule using meta heuristic techniques such as linear planning [2], heuristic techniques [3], and

genetic algorithms [4].

However, planning the entire product input schedule in advance is challenging when dealing with real-world situations such as equipment failure, rework due to defects, and changes in production plans. This makes it difficult to adapt to unexpected problems. In this study, we propose a real-time sequencing approach that can adapt to unpredictable situations that may occur in the field.

2. Bottleneck Determination

2.1 Bottleneck Determination Method

In this study, the flow shop process of a total of N processes is considered. The final bottleneck process is identified by analyzing the performance of n process times with the longest working time of the input product. The process

with the longest working time is determined as a bottleneck by adding data from n processes with a long working time of previously injected products to the working time of all the currently injected products.

2.2 Proposed Method

Our idea is to alleviate bottlenecks by injecting products with the shortest processing time among those waiting in the buffer after the bottleneck process has been determined. Fig. 1 shows the flow chart of the proposed method.

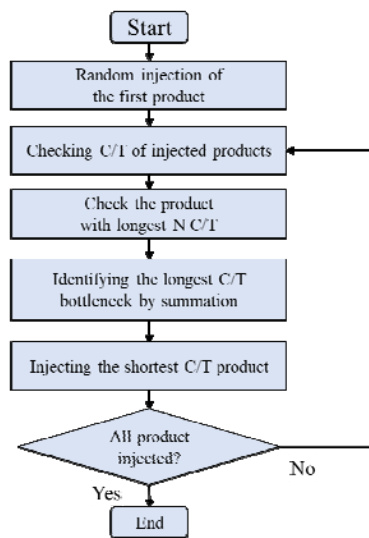


Fig. 1. flow chart of the proposed Method

3. Experiments

This study conducted an experiment by producing a simulation model of one production line that produces 10 types of products with a 28-step process. The model did not reflect variables such as device failure, setup time, and logistics transportation time. The experiment was conducted on the assumption that 10 types of products were waiting for 550 units of 10 and 100 types of products, respectively, in the buffer. The experiment was conducted by changing the n used when considering the bottleneck process from 1 to 10, and the total production time was compared with the method of sequentially injecting the product (Table 1). An additional experiment was conducted on the assumption that 100 units were waiting to see if the ratio of products waiting in the buffer was different, regardless of the total number of products

waiting.

It was confirmed that the time required for overall production was reduced when the proposed method was used compared to when the products were sequentially injected. At this time, the more products waiting in the buffer, the more time decreased.

Table 1. Summary of experiments

| n | Previous | Proposed | Improvement |
|--------|----------|----------|--------------|
| 10*100 | 2.91% | 3.09% | 6.19% |
| 10*550 | 3.40% | 3.56% | 4.71% |

4. Conclusions

The proposed method had the effect of shortening the production time in the mixed flow production process. It was confirmed that the larger the number of products waiting in the buffer, the more effective the proposed method is, and that the n value and total production that minimize production time vary depending on the amount and ratio of products waiting in the buffer.

Further experiments will be conducted by reflecting the uncertainty and complexity of the actual process.

References

- [1] D. Moon, Y. Lee, and Y. Shin, "The Effect of Mixed-Model Production in Automotive Body Shops Considering Assembly Methods and Part Transfer Policies," *Journal of the Korean Institute of Industrial Engineers*, Vol 44. No.5, pp. 391-403, 2018.
- [2] Y. Jang, D. Son, B. and Jeong, "A Study on Production Planning of Multi-Products Mixed Production Lines: Focused on Automobile Camera Production Lines," *Journal of the Korean Society of Supply Chain Management*, Vol 20. No.1, pp. 99-109, 2020.
- [3] Caskey, Kevin, and Richard Lee Storch, "Heterogeneous dispatching rules in job and flow shops," *Production Planning & Control*, Vol 7. No.4, pp. 351-361, 1996.
- [4] C. Hyun, Y. Kim, and Y. Kim "A Genetic Algorithm for Multiple Objective Sequencing Problems in Mixed Model Assembly Lines," *Applied Intelligence*, Vol 13. No.3, pp. 247-258, 1998.

Analyzing Similar Redundancy Among AI Core Courses using Hugging Face Transformer and Hierarchical Analysis

Tai Hoon KIM and In Seop NA*

{School of Elec. and Comp. Engineering, Division of Culture Contents}, Chonnam National University
Yeosu, South Korea

[e-mail: taihoonn@jnu.ac.kr, ypencil@hanmail.net]

*Corresponding author: In Seop NA

Abstract

This paper presents an approach for identifying and analyzing similar redundancy among Artificial Intelligence (AI) core courses using the Hugging Face Transformer model and hierarchical analysis. We outline five key steps taken in this process. In the initial stage, we collected the syllabi of 18 target courses. During the preprocessing stage, we encoded course names and categorized items, such as weekly learning objectives, course content, course name, course code, and week number. For analysis, we employed hierarchical clustering with the Ward method to assess the similarity among courses. Finally, we visualized the results using a dendrogram. This analysis is expected to provide valuable insights for optimizing curriculum efficiency and reducing redundancy.

Keywords: Similar redundancy, Hugging face transformer, Hierarchical analysis, Ward method, Visualization

1. Introduction

With the rapid advancement in the field of Artificial Intelligence (AI), universities and educational institutions are developing AI-related curricula and offering a variety of AI courses. These diverse courses play a crucial role in providing students with the knowledge and skills required in the AI domain. However, it is essential to consider the efficiency and redundancy of AI education [1].

Upon examining 18 AI core courses, it is evident that a wide range of topics and areas are covered. This diversity can offer rich learning experiences for students but may also lead to content

redundancy and academic burdens. For instance, when a course titled "AI Mathematics" is divided into "AI Mathematics Fundamentals" and "Advanced AI Mathematics," it raises questions about the potential overlap in content.

Moreover, redundancy can occur within "AI Fundamentals" and "AI Core" courses, with topics like "Machine Learning" being covered in both. Such redundancy can diminish the efficiency of the curriculum.

This paper introduces a study on identifying and optimizing the redundancy among the 18 AI core courses. To achieve this, we utilize the Hugging Face Transformer model and hierarchical analysis to assess the similarity among courses

"Following are results of a study on the "Convergence and Open Sharing System" Project, supported by the Ministry of Education and National Research Foundation of Korea" AND "Korea Institute of Planning and Evaluation for Technology in Food, Agriculture and Forestry (IPET) and Smart Farm R&D Foundation (KosFarm) through the Smart Farm Innovation Technology Development Program, funded by the Ministry of Agriculture, Food and Rural Affairs (MAFRA) and Ministry of Science and ICT (MSIT), and Rural Development Administration (RDA) (421017-04)."

and provide insights for curriculum enhancement.

This paper is structured on five key steps (data collection, preprocessing, model selection, similarity analysis, visualization) as shown in Fig. 1. These five key steps aim to identify and optimize redundancy among AI core courses, contributing to the efficiency and effectiveness of AI education.

2. Proposed System

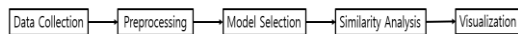


Fig. 1. Flow-chart of proposed system

2.1 Data Collection Phase

The data collection phase marks the initiation of our research, during which we gather the syllabi of 18 AI core courses. This phase comprises the following detailed procedures:

- a. **Selection of AI Common Courses:** Firstly, we meticulously select the AI common courses to be included in our study. This selection is made with the aim of representing the AI curriculum and encompassing various subfields within AI. The consideration of a variety of courses ensures the comprehensiveness of our research.
- b. **Data Collection:** The syllabi include comprehensive information about the courses, such as course objectives, course outlines, weekly learning objectives, weekly course content, schedules, assignments, and essential materials.
- c. **Data Storage:** Data is collected in various formats, including Korean (HWP), Word (DOC), PDF, and web pages. These documents are securely stored and preserved.
- d. **Quality Assurance:** During the data collection process, we verify the accuracy and completeness of the collected syllabi. We check for any missing information or unclear content and make the necessary corrections.

This modified data collection phase ensures the careful collection and secure storage of data from the very beginning of our research, enhancing the reliability of our study.

2.2 Preprocessing Phase

Details about data preprocessing, including encoding course names and categorizing syllabus items into weekly learning objectives, course content, course name, course code, and week number.

- a. **Subject Code Assignment:** Each course name is encoded according to the coding scheme that categorizes the 18 courses into 4 domains as shown in Fig. 2. For instance, "Artificial Intelligence Mathematics" belongs to the "A" domain, and "Artificial Intelligence Mathematics Fundamentals" is encoded as "A01," while "Artificial Intelligence Mathematics Advanced" is encoded as "A02."

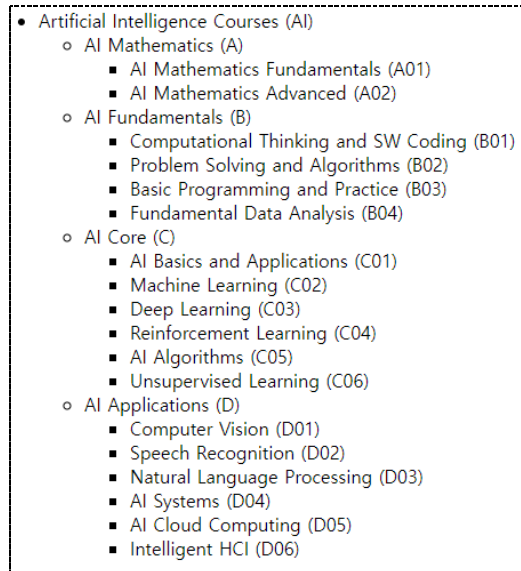


Fig. 2. Subject code assignment

- b. **Item Segmentation:** The syllabi are divided into specific items, including weeks, course objectives, course content, course name, course code, and weeks.
- c. **Language Processing:** Text within the data is cleansed and standardized. This step involves correcting typos or grammatical errors and ensuring consistency in the formatting of the text.
- d. **Data Format Integration:** Since data may exist in various formats, it is integrated into the new coding system.
- e. **Duplicate Item Handling:** Duplicate information or redundant items are

identified and resolved, ensuring the accuracy of the analysis.

- f. **Data Organization:** Data is restructured into a format suitable for analysis. Database or spreadsheet tools are used to organize and clean the data.
- g. **Data Validation:** A validation step is carried out to verify the consistency and completeness of the data. Data errors or omissions are identified and rectified to enhance data quality.
- h. **Data Storage:** Processed data is securely stored. Systems or storage media are set up to ensure data is accessible for future use.

This process, adapted to the coding scheme, effectively prepares the data for analysis, maintaining accuracy and reliability.

2.3 Model Selection

Selecting a Hugging Face Transformer model is a crucial decision in the context of natural language processing (NLP) tasks. Here's an explanation of why choosing a Hugging Face Transformer model is a suitable choice:

- a. **Efficiency of Pretrained Models:** Hugging Face offers a wide range of pretrained Transformer models that have been trained on extensive text data. This means that these models already possess a general understanding of language, making it easier to adapt them to new NLP tasks with less data and training time, ultimately improving performance.
- b. **Multilingual Support:** Hugging Face Transformer models support various languages, making them valuable for international projects or multilingual environments where NLP tasks need to be performed.
- c. **Diverse Model Portfolio:** Hugging Face provides a variety of Transformer [2] models, such as GPT-2 [3], BERT [4], RoBERTa [5], T5, XLNet [6], each tailored for different types of NLP tasks. This diversity allows you to select the most suitable model for your specific task.
- d. **User Community and Resources:** Hugging Face has an active user community and offers rich resources, providing support and solutions for

potential issues that may arise when using the models.

- e. **Ease of Model Fine-Tuning:** Hugging Face Transformer models come with user-friendly APIs and libraries for fine-tuning. This enables you to adapt the models to specific tasks and improve their performance.
- f. **Transfer Learning and Generalization:** Hugging Face Transformer models excel at transfer learning, allowing knowledge learned from different NLP tasks to be transferred to new tasks. This capability results in strong performance even with small datasets.
- g. **Real-time Conversations and Interactive Interfaces:** Hugging Face offers real-time conversation and interactive interfaces that allow testing and interaction with the models in practical settings.

For these reasons, Hugging Face Transformer models are often chosen as a suitable option for NLP tasks and serve as strong candidates during the model selection process.

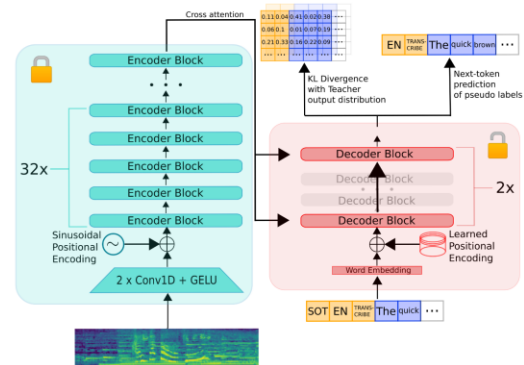


Fig. 3. Distil-Whisper Transformer Model [7]

2.4 Similarity Analysis Phase

An exploration of the hierarchical analysis process, using the Ward method to evaluate redundancy among AI courses and a discussion of the results. The similarity analysis stage in the research paper, guided by an exploration of the hierarchical analysis process using the Ward method, aims to assess redundancy among AI courses and subsequently discuss the outcomes. Here's an explanation of this stage:

- a. **Hierarchical Analysis Process:** The hierarchical analysis process involves

organizing the AI courses into a hierarchical structure based on their similarities. This typically begins by representing the courses as data points and using clustering techniques to group them into clusters. The Ward method is a hierarchical clustering algorithm that helps identify patterns and relationships among these courses.

- b. **Ward Method:** The Ward method, a widely used hierarchical clustering technique, is employed to evaluate the similarity and redundancy among AI courses. It calculates the distances or dissimilarities between courses and iteratively merges clusters to minimize the total within-cluster variance. This method is particularly effective at identifying clusters of courses with similar content or objectives.

The Ward method (or Ward's Method) is one of the clustering algorithms used in hierarchical cluster analysis. This method is based on the principle of minimizing the increase in variance between clusters. Below is an explanation of the Ward method with mathematical formulas:

- a. **Measuring Distance Between Clusters:** To calculate the distance (or increase in variance) between two clusters A and B, Euclidean distance is commonly used. The distance between two clusters can be calculated as follows:

$$d(A, B) = \sqrt{\frac{1}{N_A + N_B} \sum_{i=1}^{N_A+N_B} \sum_{j=1}^p (x_{ij} - \bar{x}_j)^2}$$

Where:

- . $d(A, B)$ is the distance between clusters A and B.
- . N_A is the number of data points in cluster A.
- . N_B is the number of data points in cluster B.
- . x_{ij} is the j^{th} feature of data point i .
- . \bar{x}_j is the mean of the j^{th} feature.

- b. **Calculating the Increase in Variance:** The Ward method considers the increase in variance for each cluster when merging clusters. This indicates how much the variance increases when a new cluster is formed. The formula to calculate the increase in variance is as follows:

$$\Delta V = V_{A+B} - (V_A + V_B)$$

Where:

- . ΔV is the increase in variance when clusters A and B are merged.
- . V_{AB} is the variance of the new cluster formed by merging A and B.
- . V_A is the variance of cluster A.
- . V_B is the variance of cluster B.

- c. **Merging Criterion:** The Ward method proceeds by merging the cluster pair with the smallest increase in variance at each step. This means that it consistently merges clusters that result in the least increase in variance.

2.5 Visualization

An explanation of how the results are visualized using a dendrogram and the interpretation of the visualized data. To gain a deeper understanding of the similarities and redundancies among the 18 AI core courses, we utilized a dendrogram for visualization. This stage allowed us to create a hierarchical structure and represent it visually. The dendrogram is a graphical representation of the clustering results, showing how each course is hierarchically combined. It consists of branches representing clusters and illustrates the order of mergers, with the height indicating the distance between clusters. The dendrogram visually captured the relationships among the courses, highlighting courses with shared topics and overlapping content. Through the dendrogram, we analyzed the similarity and redundancy of the 18 AI core courses in our research. By examining the distances and merger order of grouped courses, we identified clusters of similar courses. This insight informed decisions related to course restructuring and redundancy reduction, contributing to the improvement of the curriculum. The dendrogram visualization played a pivotal role in our research, providing a clear overview of the relationships and redundancies among the AI core courses.

3. Experimental Results

The experiment involved creating a dendrogram through hierarchical cluster analysis using the Ward method to assess the similarity of course titles, course objectives, and course content in the collected course syllabi.

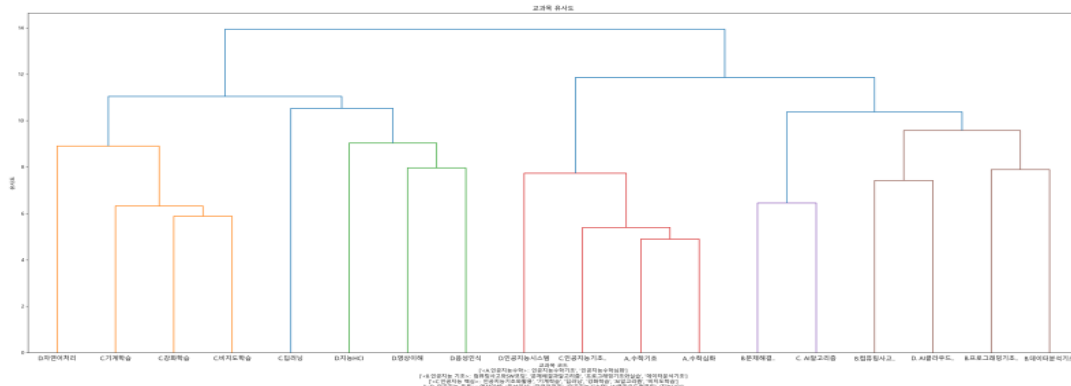


Fig. 4. Subject Similarity

The experimental results revealed distinct clusters in the dendrogram. As shown in Fig. 4, In the title category, "Unsupervised Learning," "Reinforcement Learning," "Machine Learning," and "Natural Language Processing" were closely clustered. In another cluster, "Speech Recognition," "Computer Vision," and "Intelligent HCI" exhibited proximity. Additionally, "Mathematics Fundamentals," "Advanced Mathematics," "AI Fundamentals," and "AI Applications" formed a separate cluster, while "Problem Solving and Algorithms" and "AI Algorithms" were also found to be closely clustered. These findings provide insights into the grouping of related AI core course topics based on the similarity of course titles, objectives, and content. Fig. 5 represents the similarity of course objectives in the experimental results.

As shown in Fig. 5, "Unsupervised Learning," "Reinforcement Learning," "Machine Learning," and "Natural Language Processing" cluster closely together.

Similarly, "Speech Recognition," "Computer Vision," "Intelligent HCI," and "Deep Learning" are observed in proximity. Additionally, "Advanced Mathematics," "Mathematics Fundamentals," "AI Fundamentals," and "AI Applications" form another tight cluster, while "AI Algorithms" and "Problem Solving and Algorithms" are closely grouped together.

These results visually demonstrate the grouping of AI core courses based on the similarity of their course objectives. Fig. 6 illustrates the similarity of course content in the experimental results. As depicted in Fig. 6, "Intelligent HCI" and "Reinforcement Learning" cluster closely together, while "Speech Recognition" and "Computer Vision" as well as "AI Systems" are observed in close proximity.

Furthermore, "Advanced Mathematics" and "Mathematics Fundamentals" form a tight cluster, and "Natural Language Processing," "AI Algorithms," and "Unsupervised Learning" cluster closely together.

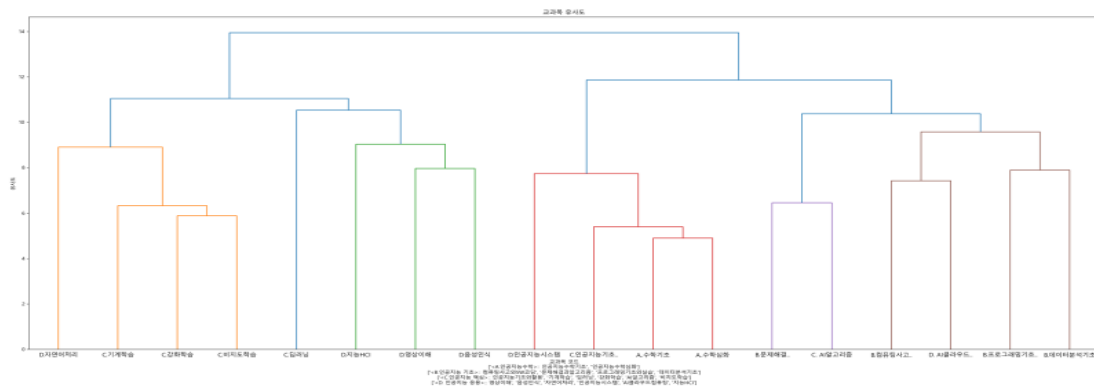


Fig. 5. Objective Similarity

These results visually represent the grouping of AI core courses based on the similarity of their course content.

4. Conclusions

Based on the analysis of course titles, course objectives, and course content among the 18 AI core courses using the hierarchical similarity comparison through the Ward method, our study has successfully explored and compared the relationships and redundancies among these courses. Our findings reveal the distinct clustering patterns among these courses, shedding light on how they are related and can be optimized within the AI curriculum.

In conclusion, this research contributes to a deeper understanding of the interrelatedness and thematic similarities among AI core courses. By utilizing the Ward method and dendrogram visualization, we have provided valuable insights for curriculum designers and educators to make informed decisions about optimizing the AI core course offerings. This work can serve as a foundation for curriculum enhancement and streamlining, ultimately benefiting students and the field of artificial intelligence education.

References

- [1] Y. Noh, I. Ahn, S. Choi, "A Study on Developing the Curriculum for Introduction to Library and Information Science Courses," *Journal of the Korean Society for Library and Information Science*, Vol 47, No. 1, pp. 287-325, 2013.
- [2] A. Vaswani et.al, "attention is all you need," *Advances in Neural Information Processing Systems 30 (NIPS 2017)*.
- [3] A. Radford, et. al., "Language Models are Unsupervised Multitask Learners," 2019.
- [4] J. Devlin, et. Al., "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, vol. 1, pp. 4171-4186, 2019.
- [5] Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., & Stoyanov, V. (2019). RoBERTa: A Robustly Optimized BERT Pretraining Approach. *ArXiv, abs/1907.11692*.
- [6] Z. Yang, et. Al, "XLNet: generalized autoregressive pretraining for language understanding," *Nips'19*, pp. 5753-5763, 2019.
- [7] Distil-Whisper Transformer, <https://github.com/huggingface/transformer/releases>, Access: 2023.11.13.

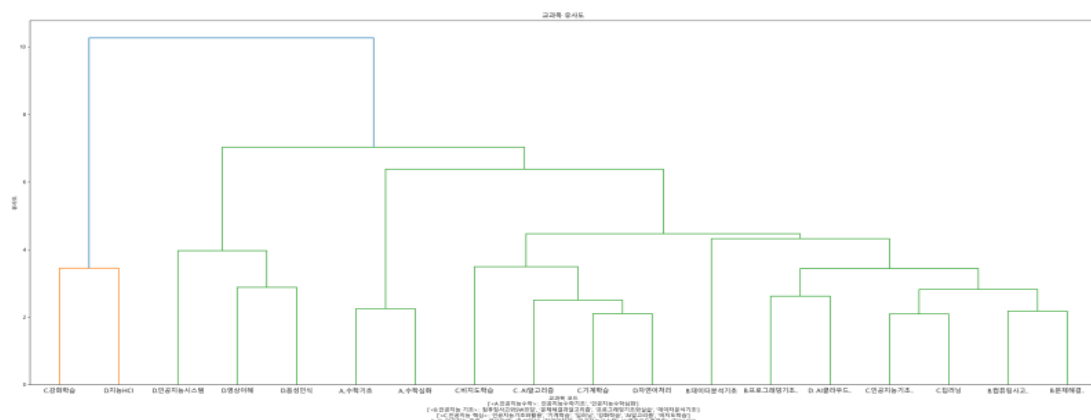


Fig. 6. Course Content Similarity

Pedestrian Collision Risk Assessment Based on Pedestrian's Motion Prediction for Autonomous Vehicles

Sunghwan Kim¹, Jaeseung Kim¹, Yeongwon Yu¹, Ir. Made Sudarma², Deden Witarsyah Jacob³,
and Minho Jo¹

¹ Department of Computer Convergence Software, Korea University
Sejong, South Korea

² IT Track, Department of Electrical Engineering, Udayana University,
Bali, Indonesia

³ School of Industrial and Systems Engineering, Telkom University,
Jawa Barat, Indonesia

[e-mail: hwann@korea.ac.kr, js6088@korea.ac.kr, won3667@naver.com, msudarma@unud.ac.id,
dedenw@telkomuniversity.ac.id, minhojo@korea.ac.kr]

*Corresponding author: Minho Jo

Abstract

Pedestrian collision avoidance is one of the key points for safe autonomous driving. In order to prevent pedestrian collisions in complex urban areas, it is important to consider the risk of different pedestrian types. Different targets of pedestrians, children, the disabled, and personal ride carriers such as bikes, and motor kickboards, have different risk levels. This paper proposes a pedestrian collision risk assessment algorithm based on motion prediction of different targets of pedestrians. We prioritize collision risk depending on different targets of pedestrians by predicting time-to-collision (TTC) with an autonomous vehicle. The experimental results show that the proposed risk priority algorithm proves effectiveness.

Keywords: Autonomous Driving, Collision Risk Assessment, Deep Learning, Motion Prediction, Time-to-Collision, Pedestrian Behavior Analysis

1. Introduction

In the evolving environment of autonomous vehicles (AVs), ensuring pedestrian safety remains a critical challenge. In complex urban areas, different targets of pedestrians, for example, children, the disabled, and personal ride carriers such as bikes, and motor kickboards, have different risk levels. Existing methods do not distinguish between various pedestrian targets. We consider different risk levels

depending on speed of pedestrian targets in [Table 1 \[1\]\[2\]](#). The collision risk is not low even if objects are far away from a car, because motor kickboards and bikes have a higher risk of collision depending on their speed. Disabled people have a higher risk of collision than other pedestrians in the same location due to their slower reaction time. This paper proposes a multi-pedestrian collision risk assessment algorithm based on motion prediction of different targets of pedestrians. We prioritize collision risk

depending on different targets of pedestrian by predicting time-to-collision (TTC) by an autonomous vehicle.

Table 1. Speed of pedestrian targets

| Pedestrian Targets | Speed (km/h) |
|---------------------------|--------------|
| Adults | 5 |
| Children | 3 |
| The disabled (Wheelchair) | 10 |
| Bikes | 15 |
| Motor Kickboards | 20 |

2. Proposed Methodology

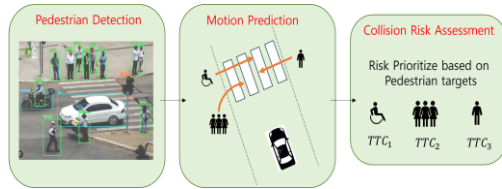


Fig. 1. Architecture of the proposed pedestrian collision risk assessment model

2.1 Pedestrian Detection

This research employs CityPersons dataset of urban traffic scenarios including various pedestrian densities and backgrounds [3]. YOLOv7, a state-of-the-art object detection model, was used for detecting pedestrians due to high accuracy and real-time processing capabilities [4]. We utilize a pre-trained YOLOv7 model and fine-tuned it with our target pedestrians including adults, children, people in wheelchairs, bikes, and motor kickboards.

2.2 Pedestrian Motion Prediction

Deep SORT is an advanced algorithm for object tracking, particularly effective in handling scenarios with moving cameras or multiple moving objects, like pedestrians in urban environments [5]. Pedestrians are tracked with Kalman Filter for predicting the future state (e.g., position and velocity) of pedestrians based on their past states. Deep SORT employs the Hungarian algorithm for data association, which solves the problem of matching detected objects with existing tracks.

2.3 Collision Risk Assessment

Targets of pedestrian have different risk levels depending on their speed and reaction time. A

motor kickboard that is farther away from an AV has a higher level of risk than an adult that is nearby because the reaction time of motor kickboard is slow due to braking distance depending on the speed. Likewise, unexpected situations can occur in children or people in wheelchairs because they have poor attention and slow reaction time. Therefore, we propose a method using risk priorities according to the speed and reaction time of pedestrians to prevent accidents. The proposed risk priorities are as follow:

$$\begin{aligned} \text{The disabled} &= \text{Children} > \text{Motor Kickboard} \\ &= \text{Bike} > \text{Adults}. \end{aligned}$$

The future trajectory of an AV is denoted as y and pedestrians as z_i with T steps, where $i \in [1, \dots, N]$ is the number of pedestrians and T is a timestep in seconds. We first calculate the closest distance between the AV and pedestrians to determine if they have possibility to be in collision:

$$D_i = \min_{t=1}^T \|y^t - z_i^t\|_2. \quad (1)$$

D_i compares the distance between pedestrians and the AV present in a situation. If $D_i > \varepsilon$, where ε is a threshold of safe distance between two objects, the AV and the pedestrian do not interfere with each other. Otherwise, we calculate the time steps between the AV and pedestrians when a collision occurs:

$$C_i = \operatorname{argmin}_{t=1}^T \|y^t - z_i^t\|_2. \quad (2)$$

C_i defines the time step whenever each pedestrian meets with the AV. The pedestrian with smallest C_i meets the AV earliest while the largest C_i does the AV last.

When there are pedestrians with the same C_i value, the AV responds differently according to risk priority level. If a person in a wheelchair and an adult have the same C_i value, then the AV should reduce the speed for the disabled much faster than the adult, due to the disabled person's slower responds.

3. Experimental Results

The speed of an AV and pedestrians are assumed to be constant. The speed of the AV is

50 km/h, and the fixed speeds of the pedestrians are shown in Table 1. The AV at 50km/h needs a braking distance of 16m and takes 2 seconds to stop [6]. The experiment was conducted assuming that a collision will occur if the AV cannot stop within 2 seconds after detecting pedestrians and predicting the motion. When the risk priority algorithm is implemented, experimental results showed that the AV starts braking 2 seconds in advance for disabled people and children, and 1 second in advance for motor kickboards and bikes.

Table 2. Comparison of collision rates with the proposed risk priority algorithm

| Targets of Pedestrian | Collision Rate | |
|---------------------------|-----------------------|--------------------|
| | Without Risk Priority | With Risk Priority |
| Adults | 25% | 25% |
| Children | 30% | 22% |
| The disabled (Wheelchair) | 23% | 14% |
| Bikes | 23% | 17% |
| Motor Kickboards | 35% | 28% |

Table 2 shows comparison of collision rates with proposed risk priority algorithm. Adults are not considered a priority, there was no effect on the collision rates in the experimental results. However, children, disabled people, bikes, and motor kickboards that take risk priority have the effect of reducing the collision rates by an average of 7%. The results prove that our proposed risk priority algorithm based on targets of pedestrian is effective.

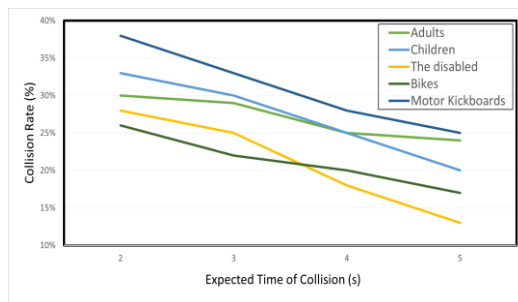


Fig. 2. Comparison of collision rates according to expected time of collision

Fig. 2 shows the collision rates when each pedestrian target has the same collision time. The results show that as the expected time of collision is given sufficiently, the AV is able to brake to

prevent a collision. Moreover, it shows that children, disabled people, bikes, and motor kickboards with risk priority tended to see a greater reduction in collision rates compared to adults without risk priority.

4. Conclusion

This paper proposed a pedestrian collision risk assessment algorithm that considers various targets of pedestrian, including children, the disabled, bikes, and motor kickboards. The proposed algorithm predicts the time-to-collision (TTC) with different pedestrian targets, allowing AV to brake in time for avoiding collisions. This advancement makes a significant step towards improving the interaction between autonomous vehicles and pedestrians, ensuring safety in urban traffic environments. For the future work, we plan to implement the proposed risk priority algorithm in a real vehicle on the road and to check if it works properly in the road test.

References

- [1] Preferred walking speed, Wikipedia. Accessed: Oct. 17, 2023. Available: https://en.wikipedia.org/wiki/Preferred_walking_speed
- [2] Wheelchair, Wikipedia. Accessed: Oct. 17, 2023. Available: <https://en.wikipedia.org/wiki/Wheelchair>
- [3] Shanshan Zhang, Rodrigo Benenson, Bernt Schiele, "CityPersons: A Diverse Dataset for Pedestrian Detection," *arXiv preprint arXiv:1702.05693*, 2017.
- [4] C.-Y. Wang, A. Bochkovskiy, and H.-Y. M. Liao, "Yolov7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors," *arXiv preprint arXiv:2207.02696*, 2022.
- [5] N. Wojke, A. Bewley and D. Paulus, "Simple online and realtime tracking with a deep association metric," in *Proc. of 2017 IEEE International Conference on Image Processing (ICIP), Beijing, China, 2017*, pp. 3645-3649, doi:10.1109/ICIP.2017.8296962.
- [6] "Braking distance," Wikipedia. Accessed: Oct. 17, 2023. Available: https://en.wikipedia.org/wiki/Braking_distance

A Facial Wrinkle Segmentation Method Based on Unet++ Model

Junsuk Lee, Hyeonwoo Kim, Jonghwa Shim, Eunbeen Kim, and Eenjun Hwang*

School of Electrical Engineering, Korea University
Seoul, Republic of Korea

[e-mail: {lilemoest48, guihon12, indexlibrorum3822, gichanac, ehwang04}@korea.ac.kr]

*Corresponding author: Eenjun Hwang

Abstract

Facial wrinkle segmentation is the process of dividing facial image into wrinkle segments and representing them by a mask. Existing deep learning-based methods have difficulty segmenting fine wrinkles due to the insufficient amount of wrinkle data available and the imbalance between wrinkle and non-wrinkle data. To address these problems, in this paper, we propose an efficient facial wrinkle segmentation method based on Unet++ model. For segmentation, we first enhance the wrinkle dataset by manually annotating wrinkles on the entire face. Secondly, we perform facial feature extraction to extract only the skin region from the facial image. Lastly, we train Unet++ model using both dice loss and focal loss to alleviate the class imbalance problem. Through several comparative experiments, we demonstrate that the proposed method outperforms previous facial wrinkle segmentation method.

Keywords: Facial wrinkle segmentation, image segmentation, wrinkle dataset, class imbalance, dice loss, focal loss

1. Introduction

Facial wrinkle segmentation divides wrinkle regions in a face image and represents them as a mask. Recently, deep learning-based methods have been proposed to segment wrinkles. In [1], the authors proposed a wrinkle inpainting model with a facial wrinkle segmentation component. Despite its performance, this method had difficulty accurately segmenting wrinkles due to the insufficient amount of wrinkles dataset available and class imbalance of wrinkle elements. With the imbalance of wrinkle elements, the model tends to focus on the non-wrinkle elements that contain the majority of facial images. As a result, the class imbalance can lead the model to falsely predict wrinkles as non-wrinkles.

In this paper, we propose a new facial wrinkle segmentation scheme to overcome these problems. To do that, we constructed an enhanced wrinkle dataset with both prominent and fine wrinkles. Then, we specify region of interest where wrinkles appear on facial images. Finally, to alleviate the imbalance of wrinkles and non-wrinkle elements, we train the Unet++ model [2] using both dice loss [3] and focal loss [4]. To show the effectiveness of the proposed scheme, we perform comparative experiments using our enhanced wrinkle dataset.

The structure of this paper is as follows. In Section 2, we describe our proposed scheme in detail. Section 3 evaluates the performance of the proposed scheme through comparative experiments. Finally, Section 4 concludes the paper.

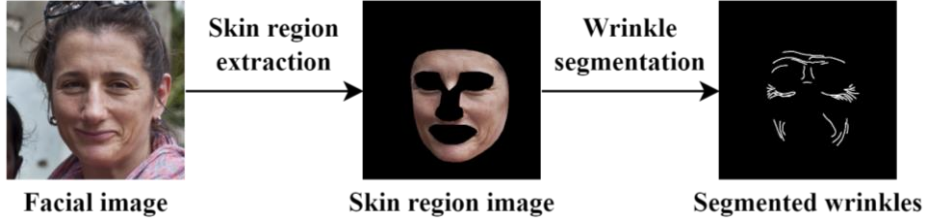


Fig. 1. Overview of the proposed facial wrinkle segmentation method.

2. Methods

Fig. 1 depicts the overview of our proposed scheme for accurate facial wrinkle segmentation, which is composed of three stages: (1) creating an enhanced wrinkle dataset, (2) extracting the skin region of the face image, (3) and segmenting facial wrinkles using Unet++ model with dice loss and focal loss. In the following subsections, we describe more details on each component.

2.1 Enhanced Wrinkle Dataset

For more accurate wrinkle segmentation, we created an enhanced wrinkle dataset from the public dataset, Flickr-Faces-HQ (FFHQ) [5]. From the dataset, we selected a thousand images of people of different ages, genders, and ethnicities. Then, we annotated various shapes, sizes, and complexities of facial wrinkles.

2.2 Skin Region Extraction

In this stage, we extract only skin regions from facial images to focus on wrinkle segmentation of facial skin. To specify region of interest where wrinkles appear, we removed facial landmarks and image background of facial images, isolating the skin region from the images. By using skin region images as input, the facial wrinkle segmentation model can focus more on learning wrinkle patterns without noise from non-wrinkle elements.

2.3 Facial Wrinkle Segmentation Model

Model architecture. We implemented a segmentation model based on Unet++. Unet++ model consists of encoder, decoder, and sub-networks interconnected through densely nested skip pathways. The encoder down-samples skin region images to extract wrinkle feature maps with contextual information. The down-sampled wrinkle feature maps are then propagated to the corresponding layer of sub-networks through

densely nested skip pathways. Then, the decoder up-samples features from the corresponding layer of sub-networks to generate wrinkle masks.

Loss function. We use two loss functions in our model: dice loss and focal loss. We expect that the dice loss can penalize wrong predictions, while the focal loss will help the model to down-weight non-wrinkle elements and concentrate on wrinkle elements.

Dice Loss is a region-based loss that quantifies the similarity between the predicted segmentation mask and the ground truth mask, highlighting the overlapping regions. This loss function effectively mitigates the problem of class imbalance as it punishes false predictions. The dice loss can be defined by Eq. (1).

$$L_D(p, q) = 1 - \frac{2 \sum_{i=1}^{H \times W} p_i q_i}{\sum_{i=1}^{H \times W} p_i^2 + \sum_{i=1}^{H \times W} q_i^2} \quad (1)$$

Here, p and q denote the predicted mask and the ground truth, respectively. In addition, H and W represent the height and width of the input image, respectively.

On the other hand, focal loss reshapes the classical distribution-based loss function to overcome class imbalance. In facial wrinkle segmentation, non-wrinkle facial elements are easy to distinguish, but wrinkles are difficult to identify. During training, focal loss assigns low weights on non-wrinkle elements and high weights on the wrinkle elements. Focal loss can be defined as Eq. (2).

$$L_F(p_t) = \begin{cases} -\alpha(1 - p_t)^\gamma \log(p_t), & \gamma = 1 \\ -(1 - \alpha)p_t^\gamma \log(1 - p_t), & \text{otherwise} \end{cases} \quad (2)$$

Here, α stands for weighting factor, γ denotes focusing parameter, and p_t represents the predicted probability of correct class. In this study, we set the weighting factor to 0.75 and the focusing parameter to 2.

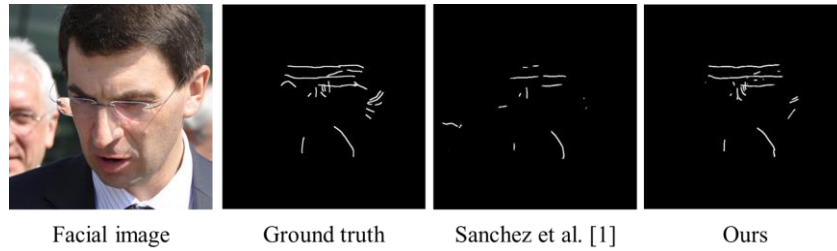


Fig. 2. Qualitative comparisons of facial wrinkle segmentation methods.

3. Experiments

3.1 Qualitative Evaluation

Fig. 2 depicts the qualitative comparison of the previous deep learning method [1] and our proposed method. In the figure, we can observe that the proposed model outperforms the deep learning-based method in the wrinkle extraction. That is, our method reproduces the wrinkles in the ground truth while avoiding non-wrinkle elements. This indicates that our model is capable of learning wrinkle patterns more accurately and robustly.

3.2 Quantitative Evaluation

In this section, we quantitatively compare the performance of wrinkle segmentation methods, including the proposed method, in terms of F1 score, Intersection over Union (IoU), and pixel accuracy. For all metrics, scores closer to 1 indicate higher quality.

Table 1 shows the comparison results. As shown in the table, the proposed method shows better performance than other method in all metrics. This indicates that our method can accurately segment diverse wrinkles.

Table 1. Quantitative comparison of wrinkle segmentation methods.

| Segmentation methods | F1 score \uparrow | IoU \uparrow | Pixel acc. \uparrow |
|----------------------|---------------------|----------------|-----------------------|
| Sanchez et al. [1] | 0.4925 | 0.3317 | 0.9919 |
| Ours | 0.5949 | 0.4275 | 0.9929 |

4. Conclusions

In this paper, we proposed a facial wrinkle segmentation method that can effectively segment wrinkles on the entire face. To do that, we constructed an enhanced wrinkle dataset with deep and shallow wrinkles across different ages,

racess, and genders. We also extracted skin regions from facial images to focus on the wrinkle patterns of facial skin. To alleviate the class imbalance problem of wrinkles and non-wrinkle elements, we used both dice loss and focal loss to train Unet++ model. Through comparative experiments on our wrinkle dataset, we showed that our proposed method outperforms existing facial wrinkle segmentation method.

References

- [1] M. Sanchez, G. Triginer, C. Ballester, L. Raad, and E. Ramon, "Photorealistic Facial Wrinkles Removal," *Asian Conference on Computer Vision*, pp. 117–133, 2022.
- [2] Z. Zhou, M. M. Rahman Siddiquee, N. Tajbakhsh, and J. Liang, "Unet++: A nested u-net architecture for medical image segmentation," *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, pp. 3–11, 2018.
- [3] W. R. Crum, O. Camara, and D. L. G. Hill, "Generalized overlap measures for evaluation and validation in medical image analysis," *IEEE Trans Med Imaging*, vol. 25, no. 11, pp. 1451–1461, 2006.
- [4] T. Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollar, "Focal Loss for Dense Object Detection," *IEEE International Conference on Computer Vision*, pp. 2980–2988, 2017.
- [5] T. Karras, S. Laine, and T. Aila, "A Style-Based Generator Architecture for Generative Adversarial Networks," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, vol. 43, no. 12, pp. 4401–4410, 2021.
- [6] S. Xie, R. Girshick, P. Dollár, Z. Tu, and K. He, "Aggregated residual transformations for deep neural networks," *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1492–1500, 2017.

Comparison of Multi-Person Action Recognition Methods with YOLOv7-Pose

Jae hyun Rho, Youngwoo Lee, Hyunbeom Choi and Jeongwook Seo

School of Computing and Artificial Intelligence, Hanshin University
Osan, South Korea

[e-mail: shwogus@hs.ac.kr, duddn1124@hs.ac.kr, chl9717@hs.ac.kr, jwseo@hs.ac.kr]

*Corresponding author: Jeongwook Seo

Abstract

Multi-person action recognition refers to the technology of identifying and analyzing the actions of multiple individuals within a single scene simultaneously. As pose estimation models in multi-person action recognition, OpenPose and PoseNet are usually considered. However, their low pose estimation speeds may cause reducing real-time data processing speed and energy efficiency. Therefore, in this paper, we compare and analyze the multi-person action recognition methods with YOLOv7-Pose with high pose estimation speed and action classification models such as long-short term memory (LSTM), gated recurrent unit (GRU), and spatial temporal-graph convolution network (ST-GCN). Experimental results showed that the combination model of YOLOv7-Pose and ST-GCN could achieve the highest accuracy of about 91%, and the combination model of YOLOv7-Pose and LSTM combination model could achieve the fastest action recognition speed of about 1.2 ms.

Keywords: Action recognition, LSTM, Pose estimation, Time series data, YOLOv7

1. Introduction

Multi-person action recognition is a technology that identifies human actions within a scene involving multiple active participants. This technology has previously been implemented by utilizing pose estimation models such as OpenPose and PoseNet. These models, however, they exhibited slower inference speeds than their contemporary counterparts, such as YOLOv7-Pose[1]. In this paper, we propose an action recognition method that combines deep learning models that perform action classification and YOLOv7-Pose and compare these methods considering testing time and accuracy as important indicators.

2. Action Recognition Method

We implemented an multi-person action recognition method that combines YOLOv7 for pose estimation and three deep learning models, LSTM, ST-GCN, and GRU, for action classification[2].

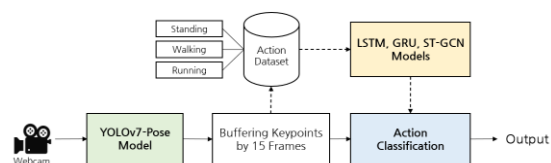


Fig. 1. Action recognition model with YOLOv7-Pose

As shown in Fig. 1, Human skeleton data is extracted while buffering 15 frames of keypoints based on the YOLOv7-Pose model, and action dataset is classified for recognizing three actions of standing, walking, and running based on the LSTM, ST-GCN, and GRU models.

3. Experimental Results

This section explains about the experiment procedures and results. We use 80% of the dataset for training, and holdout 20% for testing. The operating system is Ubuntu 20.04 and GPU model is CUDA 11.4. We construct neural network models using Pytorch 1.9.1[3]. We built video data sets of three actions: walking, running, and standing, for pre-training of the action classification model.

Table 1. Action recognition models performance evaluation

| Metric | Y-LSTM | Y-GRU | Y-ST-GCN |
|--------------|---------|--------|-----------|
| Accuracy | 0.87 | 0.81 | 0.91 |
| F1-Score | 0.82 | 0.75 | 0.90 |
| Testing time | 1.2ms | 1.5ms | 7.2ms |
| Parameters | 352,775 | 80,452 | 7,922,301 |

As shown in Table 1, the accuracy of Y-ST-GCN was the highest at 91%, but the testing time of Y-LSTM was the fastest at 1.2 ms and the accuracy was also relatively high at 87%. Therefore, we used Y-LSTM in demonstrations, hoping to combine it with edge devices in the future[4].



Fig. 2. Multi-Person action recognition model example

As shown in Fig. 2, the demonstrations of this experiment was conducted with more than 3 people.

4. Conclusion

In this paper, we implemented an advanced multi-person action recognition method and compare Y-LSTM, Y-GRU, Y-ST-GCN methods. As comparison, The accuracy of Y-ST-GCN is the highest at 91%, and the testing time of Y-LSTM is the fastest at 1.2ms. We expected that these model will be able to achieve better performance by combining it with edge devices such as Raspberry Pi and Jetson series in the future.

References

- [1] D. Dao, T. Le, T. Tran, Y. Nguyen, and T. Duy, "The combination of face identification and action recognition for fall detection," *Tạp chí Khoa học và Công nghệ-Đại học Đà Nẵng*, Vol. 20, No. 12, pp. 37-44, 2022.
- [2] S. Yan, Y. Xiong, and D. Lin, "Spatial temporal graph convolutional networks for Skeleton-Based Action Recognition," in *Proc. of the Association for the Advancement of Artificial Intelligence(AAAI)*, 2018.
- [3] S. Chheda, A. Curtis, E. Siegmund, and B. Chapman, "Performance study on cpu-based machine learning with pytorch," in *Proc. of the High Performance Computing in Asia-Pacific Region(HPC Asia)*, 2023
- [4] S. Wang, I. Li and W. Wang, "Human Action Recognition of Autonomous Mobile Robot Using Edge-AI," *IEEE Sensors Journal*, Vol. 23, No. 2, pp. 1671-1682, 2023

Vision-Based Artificial Intelligence System for Detecting Abnormal Customer Behaviors at Unmanned Cafes

Yoonseo Kim, Sehun Lee, Sunghyun Kang, Junho Ahn

Department of AI Computer Science and Engineering, Kyonggi University
Suwon, South Korea

[e-mail: mikooyoon@kyonggi.ac.kr, redkye2@naver.com, tsugu00@kyonggi.ac.kr,
jha@kyonggi.ac.kr]

*Corresponding author: Junho Ahn

Abstract

This study designs a real-time video recognition algorithm for crime prevention in unmanned cafes. As the number of unmanned cafes has increased, crimes targeting them have also on the rise. Even though evidence of crimes was obtained from CCTV, it took a lot of time. Thus, it was difficult to respond to these crimes in real time. Therefore, this study investigates the research trends of two algorithms used for real-time video recognition. YOLO is an object detection algorithm, and it recognizes objects and identifies their locations. Hence, it can quickly assess whether there has been damage to objects. MediaPipe, which is a pose detection technology, employs the method of recognizing movements from a human perspective. MediaPipe is implemented using CNN-based feature extraction and classification algorithms. Algorithms to be applied to the anomalous sign detection system are each selected based on the analysis results, these algorithms are utilized to design an artificial intelligence video algorithm.

Keywords: Unmanned cafes, Artificial intelligence, Abnormality, Object recognition, Pose detection

1. Introduction

With advances in the Internet of Things (IoT), there has been an increase in the number of unmanned cafes, which are managed and operated without human staff. However, there has also been an increase in the number of vandalism crimes that exploit the fact that unmanned cafes are not staffed by humans [1]. In 2022, there were 6,018 cases of theft at unmanned stores, but only 3,609 criminals were apprehended. This outcome can be associated with the characteristics of unmanned cafes. Although the evidence of a crime can be obtained from CCTV, the owner of the cafe can view the recorded video footage after assessing the extent

of damage to the store. Therefore, this study analyzes algorithms used in video recognition technology and designs a real-time video recognition algorithm for crime prevention. Additionally, this study employs the You Only Look Once (YOLO) algorithm to perform object recognition to assess the performance of the model.

2. Proposed Algorithm and System

2.1 Research Trends in Object Detection

The object detection algorithm detects the main objects the user wants the algorithm to

☆“This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the National Program for Excellence in SW supervised by the IITP(Institute of Information & communications Technology Planning & Evaluation)”(2021-0-01393)

detect and distinguishes these objects by drawing bounding boxes around them. Early object detection models employed two techniques for locating and segmenting objects. They utilized convolutional neural networks (CNNs) to extract bottom-up regions and used a paradigm of training large CNNs when there was an insufficient amount of labeled training data [2]. However, early algorithms encountered difficulties in separating the objects from the background. As a result, they were unsuitable for real-time data processing. Joseph Redmon and his research staff addressed this issue and developed the YOLO object detection algorithm [3]. The YOLO algorithm is fast because it only needs to obtain the final detection results after inputting images into the neural network. Moreover, since it encodes the global information by using the entire image, it has the characteristic of reducing the error of detecting the background. Since it was first introduced, there have been many updates to the YOLO algorithm, including the development of a model that can be used even in an environment without graphics processing units (GPUs) [4].

2.2 Research Trends in Pose Detection

The most intuitive way to detect a person's pose is to use the person's joints as the key points and learn the person's poses based on the movements of these joints. The early DeepPose model presented a method of using a deep neural networks (DNNs) regression model to accurately predict the positions of a person's joints [5]. However, DNNs are more specialized for processing sequential data than for visual learning. Hence, it is difficult to train DNNs with a small amount of data. To compensate for this issue, Google developed an open-source library called MediaPipe, which mainly uses CNNs [6]. MediaPipe uses reusable components to quickly prototype recognition pipelines

2.3 Design of a System for Applying an Artificial Intelligence Video Algorithm to Unmanned Cafes

This section designs the algorithm that can be applied to unmanned cafes based on the research trends of the two algorithms analyzed earlier. As shown in Fig. 1, the main objects of the cafe, such as the kiosk and coffee vending machines,

are identified using the YOLO algorithm. To determine whether the identified objects have been damaged, the data collected from CCTV cameras are fed into the YOLO model, and the boundaries of the damaged objects are identified.



Fig. 1. Recognition of the kiosk and coffee machine objects

Table 1. Object Recognition Rate

| | Coffee machine | Kiosk |
|-----------|----------------|-------|
| Recall | 0.92 | 0.93 |
| Precision | 0.98 | 0.98 |
| Accuracy | 0.90 | 0.90 |
| F1 Score | 0.95 | 0.95 |

The following is a table of results compiled based on the results derived by collecting data for recognizing the devices inside the cafe and training the model using the YOLO object detection algorithm to assess the extent to which the model recognizes the coffee machines and kiosk.

Most of the metrics are 0.9 or greater in the table above. This result indicates that the YOLO object detection algorithm accurately detected the kiosk and coffee machines in most cases. Therefore, this model can reliably detect objects in an unmanned cafe environment.

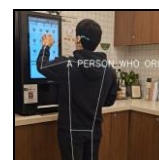


Fig. 2. Recognition of behaviors for using the kiosk

If damage to an object is detected, the system generates an alert. In addition, the system utilizes MediaPipe's movement detection to monitor the behaviors of a customer from the moment the customer enters the cafe and identifies normal and abnormal behaviors, as shown in Fig. 2 and Fig. 3.



Fig. 3. Recognition of behaviors for drinking a beverage in the café

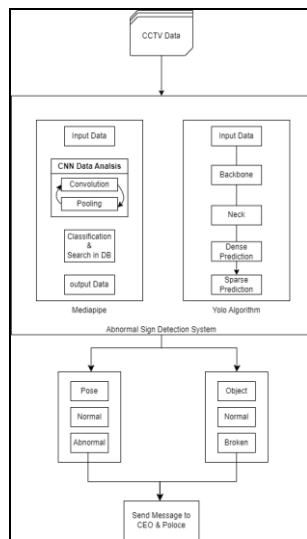


Fig. 4. Architecture of Anomaly Detection System

If abnormal behaviors are detected, the system generates an alert. If both the MediaPipe and YOLO algorithms generate alerts, the system combines the alert data and sends an alert message to the cafe owner and public institution. The architecture of the system is shown in Fig. 4. A crime prevention system like this one is expected to enhance the stability of the operation of unmanned cafes and contribute to crime prevention.

3. Conclusions

In this study, we designed an AI video algorithm system that will be applied to unmanned cafes based on the research trends of the YOLO object detection and MediaPipe pose detection algorithms. This system detects potential crimes and abnormal signs that can occur in unmanned cafes and respond to them in real time. It employs the YOLO object detection algorithm to detect damages inside a cafe, quickly assess whether objects inside the cafe have been damaged, and take necessary actions. The characteristics suitable for fast and real-time data

processing were utilized to reduce background detection errors and build a system suitable for real-time data processing. Moreover, the system utilizes MediaPipe to monitor the behaviors of customers in real time, and it was designed such that it can take actions if anomalous signs or abnormal behaviors by a customer are detected. The system monitors the situation in an integrated manner by combining the data from the two algorithms. Furthermore, the system was built to send an alert message to the cafe owner and public institution when alerts from the two algorithms are sent to the system. It is expected that this system will improve the safety and efficiency of unmanned cafes and contribute to improving crime prevention and response.

References

- [1] https://www.kicj.re.kr/board.es?mid=a20201000000&bid=0029&tag=&act=view&list_no=13375
- [2] Girshick, Ross, et al. "Rich feature hierarchies for accurate object detection and semantic segmentation." Proceedings of the IEEE conference on computer vision and pattern recognition, p. 580-587, 2014.
- [3] Redmon, Joseph, et al. "You only look once: Unified, real-time object detection." Proceedings of the IEEE conference on computer vision and pattern recognition, p. 779-788, 2016.
- [4] Huang, Rachel, Jonathan Pedoeem, and Cuixian Chen. "YOLO-LITE: a real-time object detection algorithm optimized for non-GPU computers." 2018 IEEE international conference on big data (big data). IEEE, p. 2503-2510, 2018.
- [5] Toshev, Alexander, and Christian Szegedy. "Deeppose: Human pose estimation via deep neural networks." Proceedings of the IEEE conference on computer vision and pattern recognition, p. 1653-1660, 2014.
- [6] Liguarsi, Camillo, et al. "Mediapipe: A framework for building perception pipelines." arXiv preprint arXiv:1906.08172 (2019).

A Study of Automated Bitcoin Transaction Flow Analysis Method

SangYong Kang, Hongbi Kim, Daeil Jang and SeungGu Ji*

Korea Internet & Security Agency
Naju, South Korea

[e-mail: scytalezz@gmail.com, khb@kisa.or.kr, dale@kisa.or.kr, eleca@kisa.or.kr]

*Corresponding author: SeungGu Ji

Abstract

In recent years, crimes using virtual assets, such as ransomware distribution, exchange hacking, drug purchases, have caused a steady increase in economic losses to the public. In addition, existing crimes such as the concealment of criminal funds are evolving into cybercrime using virtual assets. However, the analysis technology for tracking illegal virtual assets is still insufficient, resulting in a lot of time and manpower being spent on estimating and apprehending criminals.

In this paper, we propose a graph-based automatic transaction analysis system for tracking illegal transaction funds, which takes advantage of the decentralization and anonymization features of blockchains. A directed graph is constructed by setting a suspicious specific wallet as the root node, and adding new nodes and edges through transaction analysis, etc., to extract the destination of the specific funds. The proposed method is expected to contribute to cybercrime investigation by automatically analyzing the flow of illegal funds, and tracking the exact destination effectively.

Keywords: Blockchain, Transactions, Cryptocurrency, cybercrime, Bitcoin

1. Introduction

The bitcoin network is a peer-to-peer payment network that allows users to send and receive cryptocurrency by sending digitally signed messages to the network. The bitcoin transaction ledger is recorded transparently so that it can be verified by anyone, but it is not possible to know who owns the wallet. As a result of this anonymity feature of cryptocurrencies, they have recently been exploited for crimes such as spreading ransomware, hacking exchanges and buying illegal goods such as drugs. This has also led to a steady increase in economic losses to the public.

In this paper, we propose a graph-based automatic transaction analysis system to track

the funds of illegal transactions that abuse the decentralization and anonymization characteristics of blockchains. By setting a suspicious specific bitcoin wallet as the root node, and creating directed graphs by adding new nodes and edges through transaction analysis, we can quickly analyze the flow of funds.

2. Automated Bitcoin Transaction Flow Analysis Method

The automated bitcoin tracking method proposed in this paper consists of four main steps:

(1) creating the root node; (2) performing the tracking; (3) checking the tracking termination

condition; (4) terminating the tracking and refining the graph.

Through these steps, the flow of the bitcoin to be tracked is automatically analyzed and the final wallet it reaches is identified.

2.1 Preparation

Algorithm 1: Analysis Transaction Flow

Data: $Node_{root}$
Result: $G \leftarrow$ bitcoin flow graph
 $G.Node_{depth=0} \leftarrow Node_{root}$
 $depth_{current} \leftarrow depth=0$
while $depth_{current} < depth_{max}$ **do**
 $NodeList \leftarrow G.Node_{depth=current}$
 for $node \leftarrow NodeList$
 if $node.balance < dustAmount$ **then**
 continue
 else if $node.cluster \neq \text{'unknown'}$ **then**
 continue
 else
 $G.Node \leftarrow \text{Find Next Node}(node)$
 end
 end
 $depth_{current} = depth_{current} + 1$
end

First, a directed multi-graph is created for transaction flow tracking. The graph consists of nodes containing wallet information and edges containing transaction information. Each node contains information such as wallet address hash, transaction time, tracking balance and cluster name and can have directed multi-edges. The edge contains information such as the current node, the next node and the weight. The weight represents the size of the transfer currency amount in the transaction.

To analyze the transaction flow, the wallet address of the bitcoin to be tracked and the suspicious transaction or reference time and tracking amount are specified. Based on this information, a root node similar to Fig. 1 is generated.



```
'36EA-QviN': {
  'time': '2018-12-04 00:00:00',
  'balance': 1900,
  'cluster': 'unknown'
}
```

Fig. 1. Root Node Creating Example

2.2 Perform a Track

Algorithm 2: Find Next Node

Data: $node$
Result: $G.Node \leftarrow node_{new}$
 $TxsOutList \leftarrow \text{Get TxsOut List}(node)$
for $txs \leftarrow TxsOutList$
 if $txs.time < node.time$ **then**
 continue
 else
 $node_{new} \leftarrow txs.receivedAddress$
 $node_{new}.time \leftarrow txs.time$
 $node_{new}.balance \leftarrow \text{calc}(txs.amount)$
 $node.balance \leftarrow node.balance - node_{new}.balance$
 end
 if $node.balance < dustAmount$ **then**
 break
 end
end

Next, we query the outgoing transactions that occurred from the corresponding node based on the transaction time. We record the information of the transaction, such as the received wallet address and the sent amount, in the graph(nodes, edges). As shown in Fig. 2, the edge records the names of the current node and the next node(wallet address hash), and the transfer currency amount is stored as the weight. The balance of the current node is reduced by the transfer amount, and the next node stores the amount received and the time the transaction occurred. In this case, dusting transactions (less than 0.1% of the tracking amount) and change (the currency amounts sent to oneself, such as Bitcoin's UTXO) are excluded from the tracking target.

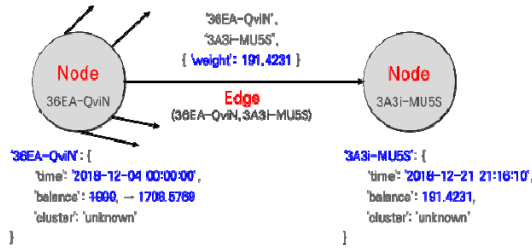


Fig. 2. Next Node and Edge Creating Examples

2.3 Check the tracking termination conditions

When a new node is reached, the tracking termination condition is checked. The tracking termination conditions are defined as follows:

(1) The balance reaches the threshold value(dusting value); (2) The search reaches the specified depth; (3) The wallet address is identified as an exchange

The size of the dusting value or the specified depth can vary depending on the number of transactions, so they need to be set flexibly. In this paper, the default values are set to 5 for the specified depth and 0.1% of the tracking amount for the dusting value.

When a new node is created, the above three conditions are checked. If one or more of these conditions are met, the node is set as a termination node and tracking is stopped. If the termination condition is not met, the same process is repeated for the nodes at the next depth. In other words, all nodes are searched by this process using BFS(Breadth First Search).

2.4 Terminate tracking & Refine a Graph

Once all the nodes have been searched, tracking is completely terminated and the graph is refined. If the size of the graph is too complex (in the case where the number of nodes and edges is very large), edges with small weights are deleted first. For example, if the number of nodes is more than 1,000, edges with $weight < (0.05 * node_{root}.balance)$ are deleted. In this case, nodes close to the end node or nodes with a large depth are deleted first, and nodes that satisfy the end condition (3) are excluded from the deletion target.

2.5 Exception: multi-input and output of bitcoin

In the case of bitcoin, multi-input and multi-output are common. Therefore, it is necessary to deal with them in order to analyze bitcoin transactions. When multi-input occurs in a transaction, we only consider the sent amount of the wallet corresponding to the current node. When multi-output occurs, we extract the receiving ratio of the issuing wallets and distribute it by the ratio, and then track it. If change(UTXO) is included in the output, the amount is excluded from the tracking amount.

$$node.balance = node.balance - \sum weight(node_{depth}, node_{depth+1}) \quad (1)$$

$$weight = \left(\begin{array}{l} txs.input_{node}.sentAmt \\ - txs.output_{node}.recvAmt \end{array} \right) \times \frac{txs.output_{node}.recvAmt}{\sum txs.output.recvAmt} \quad (2)$$

3. Conclusions

Through the proposed method, we were able to automatically analyze the flow of bitcoin and effectively and accurately track the destination of criminal proceeds. This is expected to dramatically reduce the amount of time and manpower required to estimate and apprehend criminals.

This paper does not deal with transaction obfuscation techniques that make tracking difficult, such as commingling and token swapping. In the future, it will be necessary to develop and apply techniques to undo these obfuscation techniques in order to increase the practicality of the tracking technology.

References

- [1] Oggier, Frédérique, Silivanxay Phetsouvanh, and Anwitaman Datta. "BiVA: Bitcoin network visualization & analysis." *2018 IEEE international conference on Data Mining Workshops (ICDMW)*. IEEE, 2018.
- [2] Di Battista, Giuseppe, et al. "Bitcoveview: visualization of flows in the bitcoin transaction graph." *2015 IEEE Symposium on Visualization for Cyber Security (VizSec)*. IEEE, 2015.

Text Mining Analysis of Crime Prevention Through Environmental Design(CPTED) Project in Seoul: Focusing on LDA topic modeling

Junho Park, Sooyun Lee and Euigab Hwang
Department of Criminology, Kyonggi University
Suwon, South Korea

[e-mail: junho95@kyonggi.ac.kr, lee_soosoo@kyonggi.ac.kr, flyhwang@gmail.com]

*Corresponding author: Euigab Hwang

Abstract

The Seoul Metropolitan Government of Korea has named CPTED as a “Life Safety Design” and they focus on building infrastructure to prevent crime by installing emergency bells, reflectors, and CCTV around old alleys. This research attempts to identify the trends of Crime Prevention Through Environmental Design (CPTED) project in Seoul from 2014 to 2021. In this study, topic modeling analysis was conducted by using unstructured data from 221 cases related to Crime Prevention Environment Design (CPTED). These CPTED cases were extracted from the life safety map provided by Ministry of the Interior and Safety. We visualized the trends of the CPTED projects in Seoul and regional differences in project type. Practical strategies for efficient CPTED project were suggested based on the findings.

Keywords: Text Mining, LDA, The Seoul Metropolitan Government, CPTED, Crime Prevention

1. Introduction

Empirical studies in domestic and foreign contexts have shown that effective police activities in the community cannot be implemented only by post-processing responses. Against this backdrop, the concept of CPTED (Crime Prevention Through Environmental Design) proposed by Oscar Newman[1], an architect, began to draw attention as an innovation of crime prevention strategies. In the case of Korea, Compared to the precedents of CPTED in Western contexts, the crime prevention design has not yet been fully established. Therefore, this research aims to identify trends in the CPTED project conducted in Seoul from 2014 to 2021. Latent Dirichlet Allocation (LDA) analysis will be

conducted by using unstructured data related to the Seoul Metropolitan Government CPTED projects extracted from the Ministry of the Interior and Safety’s Life Safety Map. Through the findings derived from our analysis, we identified the main topics of the CPTED project in the Seoul Metropolitan Government and suggested the elements of proper CPTED project types to seek policy strategies for efficient CPTED project implementation.

2. Theoretical Background & Review of Previous Studies

2.1 Types and Characteristics of Domestic CPTED Business

In the late 2000s, CPTED project in Korea was carried out on the back of the

establishment of an institutional foundation. Fig. 1 is a schematic diagram of the current status of the domestic CPTED projects.

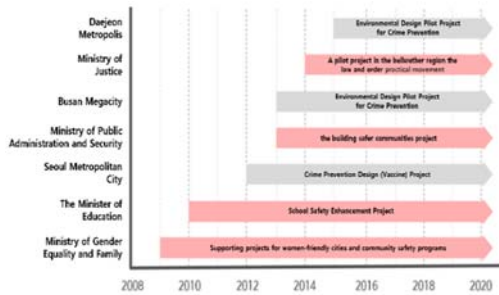


Fig. 1. Current status of CPTED in Korea[2,3]

After the 1990s, second-generation CPTED emerged in Western society, reflecting the concept of community reinforcement along with improvements in the physical environment. This statement emphasizes the importance of enhancing the social environment in which local residents actively engage.

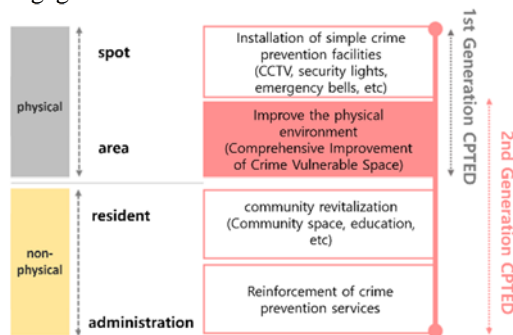


Fig. 2. Scope of CPTED business[4]

2.2 CPTED Project in Seoul Metropolitan Government

The Seoul Metropolitan Government enacted three ordinances in 2021: the Urban Environmental Design Ordinance, the Public Design Promotion Ordinance, and the Universal Design Urban Creation Ordinance to revitalize the CPTED project[5,6]. But there are also limitations of CPTED projects. Typically, CPTED structures installed in the past should be maintained and managed with interest based on the participation of local authorities, police, and residents, but their continuity is not guaranteed due to urban redevelopment[6].

3. Research Method

3.1 Data collection method

In this study, the Life Safety Map provided by Ministry of the Interior and Safety was used to identify the main topics of the CPTED projects in Seoul Metropolitan Government. Total 221 cases were used for analysis.

3.2 Method of Analysis

In this study, R 4.3.1 software was used for analysis. First, after extracting unstructured data based on text, text pre-processing was preceded for LDA Topic Modelling analysis. And then, we looked into the frequency of CPTED projects and major frequent words by period in Seoul. Finally, the main topics of CPTED project in Seoul from 2014 to 2021 were classified through the LDA-based topic modeling analysis.

4. Analysis result

4.1 Calculating the optimal number of topics

Fig. 3 and Fig. 4 are the analysis results conducted to calculate the optimal number of topics before conducting the topic modeling analysis. In order to calculate the optimal number of topics, both general calculation methods and calculation methods using the harmonic mean were used. In this study, the eight topics judged to be the most suitable were finally selected through analysis.

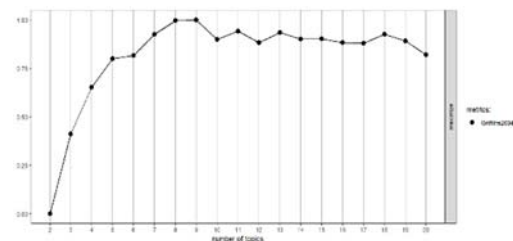


Fig. 3. Graph of the typical number of topic

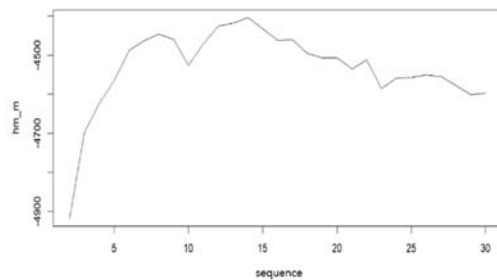


Fig. 4. The most optimal number of topics using the harmonic mean

4.2 LDA(Latent Dirichlet allocation)

To sum up, the types of CPTED projects in Seoul derived from this study could be classified into eight categories: “Crime Vulnerable Zone Improvement Project,” “Relief Staircase Creation Project,” “Park Toilet Improvement Project,” “Residential Environment Improvement Project,” “Women’s Safe Zone Creation Project,” “Safety Village Project,” “Lighting Improvement Project,” and “Women’s Safe Return Home Project.”

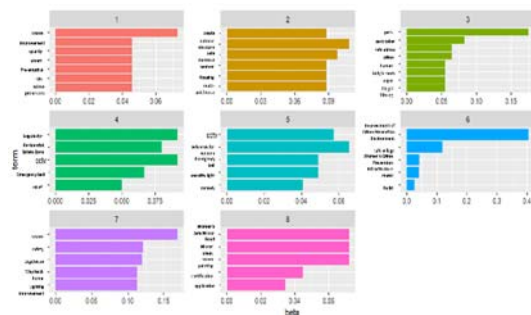


Fig. 5. The result of LDA topic modeling analysis

5. Conclusions

In this research, eight major business topics were finally derived from LDA topic modeling analysis using unstructured data on CPTED project contents in Seoul. These results show that the scope of the CPTED project in Seoul is limited to first-generation CPTED strategy. Of course, the original purpose of the CPTED project is to “prevent crime through physical environmental design,” but if the active participation of local members and an appropriate administrative system are not

combined, the effect can only be very temporary. Therefore, to improve the limitations of project, it is necessary to seek ways to revitalize the local community. For example, in the case of the “Won-teo Park and Wan-teo Ou-maru” project in Nowon-gu, Seoul, they are striving to revitalize the local community by setting up a community space in the park, running local event facilities, programs for residents, and residents’ meet-up. If the CPTED project is carried out with these non-physical methods, it is expected to be last longer effects than now.

References

- [1] Newman, O, “Defensible space: Crime prevention through urban design,” *New York: collier Books*, 1973.
- [2] San-kyu, Cho, “The 2014 Ministry of Justice CPTED Projects: Status and Monitoring Report,” *Architecture & Urban Research Institute*, 2015
- [3] Mi-jin, Seo and Seok-Jin, Kang, “An Analysis on the Characteristics of CPTED Pilot Projects in Korea,” *Architectural Institute of Korea*, Vol 140. No.2, pp. 72-75, 2020.
- [4] Hyung-bok, Lee and Hye-ran, Kim, “A Study on the Effective Implementation of CPTED Projects – In the case of Daejeon Metropolitan City,” *Korean Association of Public Safety and Criminal Justice*, Vol 29. No. 2, pp. 131-162, 2020.
- [5] Su-jin, Lee and Eung-Ryul, Choi, “A Review on the Issues of CPTED Ordinances by Metropolitan Government and Improvement,” *Korean Journal of Criminal Justice*, Vol 5. No.2, pp. 7-32, 2022.
- [6] Jun-ho, Park and Hyun-seok, Jang, “An Exploratory Study on the Spatial Correlation between the Crime Prevention Through Environmental Design (CPTED) Project Implementation Area in Seoul and the Crime dangerous grade at the Administrative dong Level,” *Journal of Community Safety and Security by environmental Design*, Vol 13. No. 3, pp. 185-222, 2022.

A Yolo based CPTED environmental factor detection model

Tsagaantsooj Batzaya¹, Eunbi Cho¹, DoHyeon Lee², Jeong-Hyeon Chang³,
Kwanghoon Pio Kim³

¹Department of Public Safety Big-data Analytics, Kyonggi University Seoul, South Korea

²Department of Computer Science, Kyonggi University

³Contents Convergence Software Research Institute, Kyonggi University

[e-mail: tsbaachka95@kyonggi.ac.kr, eunbee0508@gmail.com, ehgus8@kyonggi.ac.kr,
tiger564@nate.com, kwang@kgu.ac.kr]

*Corresponding author: Kwanghoon Pio Kim

Abstract

Citizens should live in a safe and comfortable environment, free from crime. Crime Prevention Through Environmental Design (CPTED) explains how to achieve such a secure and comfortable environment. The environmental factors described in CPTED have a significant impact on crime rates. With technological advancements, researchers in various countries are already conducting extensive studies using big data. Consequently, research on algorithms for object detection in videos is also actively underway. The purpose of this study is to construct a YOLOv8 model focusing on three environmental factors from CPTED—green spaces, streetlights, and CCTV cameras. Google Street View image segmentation data will be used as training data to predict the model. Furthermore, by utilizing this model, the study aims to understand the correlation between local crime rates and environmental factors in the area. This is expected to provide a basis for applying crime prevention strategies and other diverse fields.

Keywords: (CPTED) Crime Prevention Through Environmental Design, Google Street view, Crime prevention, YOLOV8 model

1. Introduction

Citizens should live in environments that are safe and comfortable, free from crime. The necessity of such secure and comfortable surroundings can be traced back to theories like Oscar Newman's "Defensible Space" theory from 1970 and C. Ray Jeffery's Crime Prevention Through Environmental Design (CPTED) introduced in 1971. Many researchers worldwide, focusing on CPTED theories, have demonstrated the significant impact of environmental factors on crime rates. With technological advancements, researchers across the globe are actively utilizing big data for their studies.[5] Consequently, research on algorithms for object detection in videos is also

thriving. It is now possible to create a deep learning object detection model that can simultaneously recognize environmental factors influencing crime rates, factors previously identified through surveys on crime perception.

In this study, we will focus on key CPTED factors, specifically green spaces, streetlights, and CCTV cameras, selecting them to construct a YOLOv8 model. Google Street View image segmentation data were used in training the model. This model can be applied to recognize the urban landscape level, understand the extent to which environmental factors influence crime rates in specific areas, and can be employed in various fields, including crime prevention strategies and beyond.

2. Related work

2.1 Related work

CPTED is a crime prevention method through environmental design [3]. One of the five basic principles of CPTED, "natural surveillance," is a fundamental principle in preventing crime by effectively designing the environment for monitoring. Exposing buildings or facilities to public view, especially, enables neighbors to monitor each other's surroundings, provides a sense of safety to local residents, reduces fear of crime, and has a positive impact on lowering crime rates. To achieve this, it is important not to place obstacles or trees in the vicinity, provide proper lighting at suitable distances, and install CCTV cameras, all of which significantly impact crime rates. CPTED has evolved beyond its 1st, 2nd, and 3rd generations, with numerous studies on its impact on crime rates emerging with each generation. Looking forward, advancements in technology may lead to further research on the impact of environmental factors on crime rates based on big data platforms and deep learning. Our research also has the same meaning.

YOLO [4] is a deep learning-based model that detects multiple objects in an image simultaneously in a single forward pass. YOLO provides faster processing speeds and real-time prediction capabilities compared to traditional object detection algorithms. The YOLO series includes versions such as YOLOv3, YOLOv4, YOLOv5, and the recent YOLOv8. The latest version, YOLOv8, incorporates various features such as instance segmentation, pose and key point estimation, and classification. Compared to previous versions, YOLOv8 showed significant performance improvements in terms of accuracy and speed, so the YOLOv8 version was selected in this study.

Google Street View is one of Google's online map and image services. This service provides 360-degree panoramic images of roads and city streets worldwide, allowing users to explore areas virtually [2]. Through Google Street View, users can move along streets or examine the surrounding environment from specific points. Integrated with Google Maps, this service helps users locate specific positions on the map and visually inspect their surroundings [1]. Street View images are captured by cameras mounted on vehicles or pedestrians,

providing images of publicly accessible areas. Google Street View images, based on segmentation, contribute to extensive research in various fields related to urban environments.

3. Research method and Analysis results

3.1 Research method

In this study, we aim to automatically verify three representative elements of CPTED (Crime Prevention Through Environmental Design): streetlights, green spaces, and CCTV. To achieve this, we created a YOLOv8 model using Roboflow, utilizing 8 publicly available datasets for streetlights, green spaces, and CCTV, each labeled accordingly. We performed preprocessing to combine these datasets into one, resulting in a total of 13,715 images. The images were categorized into 2,664 for streetlights, 5,293 for green spaces, and 5,758 for surveillance cameras.

Using this preprocessed dataset, we trained the YOLOv8 model through deep learning.

To collect test image data for the trained model, we utilized Google Street View along the 'Enhtainii urgun chuluu' street in Ulaanbaatar, Mongolia. We captured images at 50m intervals, covering a total distance of 5500m from the "Sapporo" bus stop to the "Zuun 4 zam" bus stop, resulting in a set of 110 modified images.

3.2 Analysis results

The model structure created in this study was organized in the sequence depicted in Fig. 1. The model successfully detected various objects such as streetlights, trees, and CCTV in a single image. The image detection results of the model are illustrated in Fig. 2.

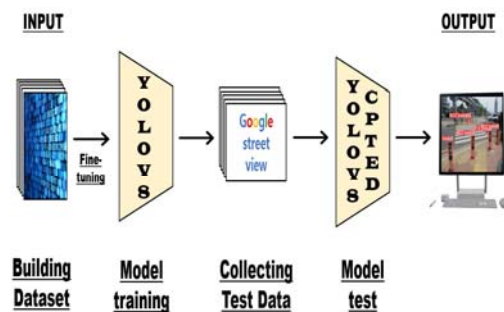


Fig. 1. Model image

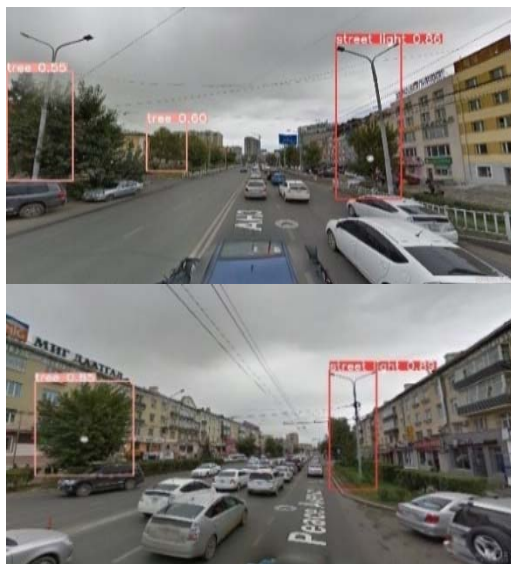


Fig. 2. Object detection image of streetlights and trees

4. Conclusions

This research was conducted with the aim of developing a deep learning based YOLOv8 model to simultaneously detect three environmental elements streetlights, trees, and CCTV emphasized in crime prevention methods through environmental design, as described in Crime Prevention Through Environmental Design (CPTED). The goal was to gain rapid awareness of the local environment. Three different datasets were integrated to create a unified dataset, which was then used to train the YOLOv8 model and develop a new model.

The deep learning based YOLOv8 object detection model, utilizing big data, is expected

to be valuable for swift information acquisition, making it useful for research on crime prevention and the collection of information regarding urban environmental factors. Additionally, it is anticipated to contribute to city planning by providing rapid awareness information about essential environmental elements required for ensuring the safety and comfort of the public.

References

- [1] Kim, Sunjae, and Sugie Lee. "Analysis of the Nonlinear Relationship and Interaction Effect between Urban Environments and Crime Incidence: Using Urban Big Data and Interpretable Machine Learning." *Available at SSRN 4136568*, 2022.
- [2] Amiruzzaman, Md, et al. "Classifying crime places by neighborhood visual appearance and police geonarratives: A machine learning approach." *Journal of computational social science*, 1-25, 2021.
- [3] Park Young Ju. "A study on the effects of the Crime Prevention Through Environmental Design (CPTED)," *Korean Police Research*, 9(1), 55-82, 2010)
- [4] Redmon, Joseph, et al. "You only look once: Unified, real-time object detection." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.
- [5] Lee, Seol; Lee, Yeeun; Hwang, Uigab. "The Effect of CPTED Elements and Crime Prediction by Using Big Data Analysis in Seoul". *Journal of the Korea CPTED Society*, 14.2: 89-116, 2023.

BLIP-Based Model for Detecting Risky Situations in Video

Eunbi Cho¹, Tsagaantsooj Batzaya¹, DoHyeon Lee², Jeong-Hyeon Chang³, Kwanghoon Pio Kim²³

¹Department of Public Safety Big-data Analytics, Kyonggi University

²Department of Computer Science, Kyonggi University

³Contents Convergence Software Research Institute, Kyonggi University

[e-mail: eunbee0508@gmail.com, tsbaachka95@kyonggi.ac.kr, ehgus8@kyonggi.ac.kr ,
tiger564@nate.com, kwang@kgu.ac.kr]

*Corresponding author: Kwanghoon Pio Kim

Abstract

This study introduces a deep learning approach to detect and analyze risk in video footage. By using a captioning model, we convert multiple scenes in the video to text, making it easier to comprehend the content in the frames and quickly identify hazardous elements. We use the 'Abnormal Behavior CCTV Video' dataset, which focuses on assault and fight scenarios, to produce descriptions for each frame, and then compare these with the original frame images to confirm the effectiveness of the model. Our findings offer the benefit of saving storage space through the efficient handling of large-scale video data and the ability to store video as text, facilitating easy processing, analysis, and data retrieval. Moreover, it is anticipated to facilitate effective prevention, control, and expeditious response to both criminal activity and natural disasters.

Keywords: BLIP model, Image Captioning, risky situation detection,, deep Learning, caption generation, video analysis,

1. Introduction

CCTV is becoming the eyes and ears of modern society, and the video data it provides provides deep insights into human behavior, interactions, and the physical environment around them. In particular, modern society is exposed to a variety of threats, such as crime, disasters, and accidents, and there is a growing need for technologies that can detect and respond to these threats in real time. Recent technological advances in object detection, motion tracking, and deep learning-based video analysis algorithms have made it possible to recognize and react quickly to complex situations in CCTV videos. In particular, automatic detection of violent scenes in videos plays a key role in crime

prevention, disaster response, and accident prevention. However, while the number of CCTVs is steadily increasing, the human capacity to monitor them remains severely limited. Furthermore, most CCTV monitoring systems are passive and under the control of human operators. This leads to increased fatigue and diminished concentration, hindering the ability to detect and respond to dangerous situations promptly. These passive systems also require significant labor costs, presenting an economic challenge as staffing levels increase. Consequently, this study is dedicated to leveraging captioning models to automate the detection and analysis of hazardous situations in CCTV footage. By transforming the varied scenarios captured in video into descriptive natural language, the process aims to enhance the identification and timely response to dangers.

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(No. 2020R1A6A1A03040583)

2. Related work

2.1 Vision-Language Pre-training

Vision-Language Pre-training(VLP) is a combination of computer vision and natural language processing(NLP). VLP focuses on comprehending the information present in images and text, which enables models to make intricate inferences and analyses by utilizing visual elements like the shapes and colors of objects and linguistic elements such as emotions and descriptions in text. Pre-trained VLP models are utilized in various fields, including image captioning for generating captions; visual question answering (VQA) to provide answers to questions about images; and image-to-text and text-to-image searches. These models are pre-trained through image and text pairs. Existing VLP models like CLIP[1] and ALBEF[2] use image-text datasets from the internet to enhance their performance. This dataset comprises web text data with extraneous images and irrelevant content. Moreover, VLP model structures tend to have either an encoder-based or an encoder-decoder architecture. The encoder-based architecture is disadvantageous in the text generation task due to weaker performance and the encoder-decoder architecture is less effective in the image-text retrieval task.

2.2 BLIP

Bootstrapping Language-Image Pre-training for Unified Vision-Language Understanding and Generation(BLIP)[3] overcomes the shortcomings of existing VLP models in terms of data and model perspectives by training both understanding and generation tasks together. BLIP reduces the noise of noisy web data through bootstrapping. It utilizes the decoder used in the generation task and the encoder used in the understanding task as captioner and filter, respectively. The captioner generates synthetic captions, and the filter removes noisy elements. BLIP's two main contributions to the VLP field include the Multimodal mixture of Encoder-Decoder (MED) and Captioning and Filtering (CapFilt) techniques. The first, the MED structure enables BLIP to engage in flexible transfer learning, distinguishing it from traditional encoder-based or encoder-decoder models by incorporating three distinct

vision-language objectives: a Unimodal encoder, an Image-grounded text encoder, and an Image-grounded text decoder. The second contribution, CapFilt, is a novel bootstrapping method that refines learning from noisy image-text pairs by generating web-sourced image captions through the captioner and subsequently filtering out the noise.

3. Method

In this study, we use the 'Abnormal Behavior CCTV Video' dataset provided by AI-HUB to convert various situations in CCTV videos into natural language through a captioning model and detect the presence of dangerous situations. The dataset contains 12 types of abnormal behavior (assault, fight, burglary, vandalism, swoon, wander, trespass, dump, robbery, datefight, kidnap, drunken). In this study, we use the assault and fight datasets. The process of the BLIB model-based risk situation detection model is as follows.

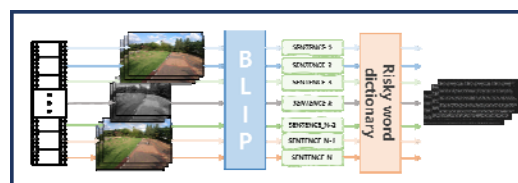


Fig. 1. The process of the BLIB model-based risk situation detection

Extract an image every 5 frames from the 'Abnormal Behavior CCTV Video'. Then, generate a sentence key that describes the contents of the frame using BLIP, a bootstrapping-based image captioning model. The generated sentences are passed through the risky-word dictionary and saved as a JSON file with the risky frame and the sentence describing the frame. The risky-word dictionary consists of nouns and verbs related to dangerous situations related to criminal activities and accidents in order to automatically detect risks in CCTV images. The sentences generated by the BLIP model are reviewed against the words in this dictionary. If a sentence contains at least one word from this dictionary, the corresponding frame is identified as depicting a risky situation.

Table 1. Examples from the Risky-Word Dictionary

| | |
|--------------------|---|
| Risky Nouns | "assault", "axe", "baseball bat", "blood", "broken bottle", "death", "driver", "explosion", "gas", "gun", "hammer", "handcuff", "injury", "kidnap", "knife", "murder", "raid", "scissors", "terror", "violence", "war", "weapon", "wire", "wrench", ... |
| Risky Verbs | "attack", "bleed", "collapse", "drop", "faint", "fall", "fight", "hurt", "kick", "lay", "lie", "resist", "slip", "stumble", "swing", "throw", "topple", "trip", "tumble", ... |

Fig. 2 shows the frames detected as dangerous by the proposed model and their corresponding captions. The frames are of people attacking another person, people fighting with each other, and someone falling down, and the generated captions describe the scene fairly accurately. These examples demonstrate the effectiveness of the model in detecting dangerous situations and providing contextual information through captions.

**Fig. 2.** Video frame risk detection and captioning illustrative examples

4. Conclusions

In this study, we introduce an approach to detect and analyze hazardous situations in CCTV images. We utilize a deep learning-based captioning model to convert dangerous scenes in the video into natural language and effectively capture the risky situations. Moreover, analyzing the riskiness of the captions generated using the risky-word dictionary is crucial in capturing specific dangerous scenarios. Based on the findings of this research, future investigations will implement the technique of cropping objects within the frame to enhance the model's precision. By doing so, we can capture each object's conduct and the correlation between them with greater accuracy, ultimately generating more complex captions. A system like this would enable improved detection and swift response to critical situations in various domains, including crime prevention and disaster relief. In addition, image captioning can automatically detect signs of both physical and social disorder in CCTV footage. This capability aids in reducing crime rates and preventing crimes through preventive methods, including the improvement of the area's environment and facilitating rapid responses.

References

- [1] A. Radford et al., "Learning Transferable Visual Models From Natural Language Supervision," *International Conference on Machine Learning*, arXiv, pp. 8748-8763, 2021.
- [2] J. Li, R. R. Selvaraju, A. D. Gotmare, S. Joty, C. Xiong, and S. Hoi, "Align before Fuse: Vision and Language Representation Learning with Momentum Distillation," *Advances in neural information processing systems*, 34, pp. 9694-9705, 2021.
- [3] J. Li, D. Li, C. Xiong, and S. Hoi, "BLIP: Bootstrapping Language-Image Pre-training for Unified Vision-Language Understanding and Generation," *International Conference on Machine Learning*, arXiv, pp. 12888-12900, 2022.

Vision-Based User Abnormality Detection System for Unmanned Laundry Safety

Jiyoung Na, Sunghyun Kang and Junho Ahn

Department of AI Computer Science and Engineering, Kyonggi University
Suwon, South Korea

[e-mail: jiyoung@kyonggi.ac.kr, tsugu00@kyonggi.ac.kr, jha@kyonggi.ac.kr]

*Corresponding author: Junho Ahn

Abstract

With COVID-19 and the drastic increase in prices, as well as rising labor costs, there is a growing preference for non-face-to-face operations and unmanned stores in various aspects. Among them, the number of newly opened unmanned laundromats is increasing because it is convenient to manage the store, and more customers are using unmanned laundromats. However, unmanned laundromats have a drawback in that machines like washing machines and dryers need to be maintained. This study proposes a system that recognizes abnormal behaviors using the K-nearest neighbors (KNN) algorithm and detects fire using the YOLO.

Keywords: Unmanned laundromats, Unusual behaviors, Object detection, Vision, Safty

1. Introduction

Unmanned laundromats have attracted attention as a start-up business opportunity because, unlike other stores, there are few goods or items to manage, and more people are utilizing unmanned laundromats as they feel burdened by the costs of dry cleaners. However, because unmanned stores are operated without human staff, it is difficult to respond to incidents when they occur. Furthermore, unmanned stores are vulnerable security-wise due to the characteristic of being operated without human staff.

Unmanned laundromats are not difficult to manage, and many people go to unmanned laundromats to use large washing machines and dryers. The number of franchise stores for the top six unmanned laundromats has increased from 3,086 in 2016 to about 6,800 in 2021 [1], and this number is expected to increase even more in the future. Since the machines in unmanned

laundromats are expensive, they need to be maintained thoroughly, and security is needed for them.

In one of the unmanned laundromats in Spain, there was an incident in which a lighter got mixed in with the laundry, and it exploded inside the dryer. Fires involving washing machines and dryers are steadily occurring incidents, and extra caution is required for unmanned stores because there may be no one at the store who can respond to such incidents initially. In an unmanned laundromat in Korea, an incident occurred where an intoxicated female middle school student entered a dryer and caused a disturbance. The operation of unmanned stores generally relies on CCTV, and there were 6,018 cases of theft in 2022. Of these cases, the criminals were apprehended in 4,127 cases; the apprehension rate is about 70% [2]. However, a large portion of these cases involve juvenile offenders, making it difficult to impose punishments. By using this

☆ “This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the National Program for Excellence in SW supervised by the IITP(Institute of Information & communications Technology Planning & Evaluation)”(2021-0-01393)

system, it will be possible to monitor crime scenes in real time and promptly respond to incidents.

2. Proposed Algorithm and System

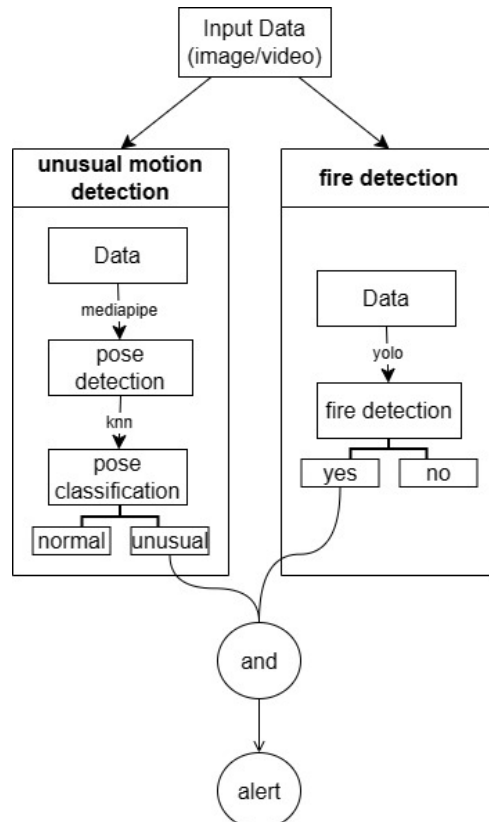


Fig. 1. Flowchart of the proposed algorithm

2.1 Proposed System

A system with the flowchart shown in Fig. 1 is proposed that uses the two systems mentioned below to send a notification to the administrator when abnormal user behaviors or anomalous signs are detected in the store.

When data is input, MediaPipe extracts the joint points and converts them into vector values. Next, the KNN algorithm classifies abnormal and normal behaviors.

2.2 Detections of Fire

YOLO [3] stands for 'You Only Look Once.' It is a system for object detection, that segments the input image or video into a grid and passes it through the network. It classifies objects according to the confidence level based on the

grid. YOLO is used to detect humans in an image or video.

Unlike conventional image processing methods, YOLO processes images in a single pass, which enables fast real-time processing of images and videos. When an image is divided into a grid and then passed through the network, it is possible to obtain overlapping bounding box data that is drawn based on the location data where objects are estimated to be located and the confidence level. YOLO uses these two results to finally detect objects.

Images are collected and used after they are labeled using Roboflow. YOLO is trained with the labeled images to detect matching objects, which are flames. If the confidence level of the trained data and input data is 50% or higher, bounding boxes are drawn.

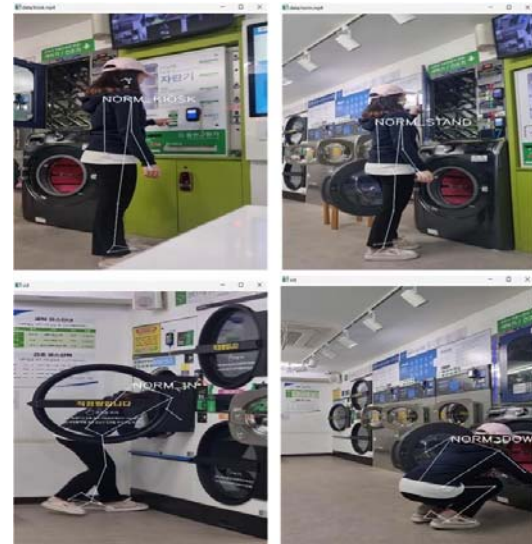


Fig. 2. User normal behavior detections

2.3 Detections of Abnormal Behavior

Both the MediaPipe and K-nearest neighbors (KNN) algorithms are needed to detect individuals who are behaving abnormally. Because it is difficult to extract poses from two-dimensional images or videos, behaviors are recognized by extracting the joint points of human objects through MediaPipe. In addition, the KNN algorithm is used to distinguish whether or not the observed behavior is abnormal.

MediaPipe [4] is Google's AI framework that represents objects in a pipeline format. MediaPipe extracts the joint points from the

images classified as humans by YOLO and represents objects in a pipeline format.

The KNN [5] algorithm is a classification algorithm. It identifies k data points from the existing trained data that are nearest to the new input data and then classifies the new input data as the closest data point among the k data points. If a person behaves abnormally in a laundromat, the behavior is classified using the KNN algorithm, and a bounding box is displayed on the CCTV screen. Fig. 2 uses mediapipe to classify the actions of laundromat users through the knn algorithm.

3. Conclusions

This study proposes a system for detecting anomalous signs and abnormal user behaviors and a service that uses the MediaPipe, YOLO, and KNN algorithms to detect abnormal user behaviors and anomalous signs in an unmanned laundromat. The system monitors video feeds in real time, which allows the system administrator to check the CCTV screens to promptly assess situations involving anomalous signs, like fires, or instances of abnormal user behavior. Most unmanned laundromats are operated 24 hours a day because they are operated without human staff. This system can detect anomalous signs and abnormal behaviors in situations where human staff cannot monitor the store continuously. Moreover, anomalous signs and abnormal behaviors can apply to other types of unmanned stores. In the future, this system can be applied to different types of businesses by collecting data on abnormal behaviors that may occur in different types of stores.

References

- [1] Joon-mu Beak, "Power-on straighteners, 24-hour washing and drying machines... A new 'fire' that no one knows is coming," the Segye Daily, 2023. <https://m.segye.com/view/20230103513204>
- [2] Min-ju Yoo, "The owner of the store who was robbed of an unmanned store, damages 2 million won, but only 30,000 won in compensation?," News1, 2023. <https://www.news1.kr/articles/?5124915>
- [3] Redmon, Joseph, et al. "You only look once: Unified, real-time object detection," *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016.
- [4] Bazarevsky Valentin, et al. "Blazepose: On-device real-time body pose tracking," *arXiv preprint arXiv:2006.10204*, 2020.
- [5] Seung-Joon O, "A K-Nearest Neighbor Algorithm for Categorical Sequence Data," *Journal of the Korea Society of Computer and Information*, 10.2, 215-221, 2005.

A Study on Person Re-identification with Feature Matching Methods

Lanh Thanh Le¹, Dinh-Lam Pham^{1,2}, Joo-Chang Kim², Sang-eun Ahn², Kyunghye Sun²,
Kyong-Sook Kim², Kwanghoon Pio Kim^{2,*}

¹ Faculty of Information Technology, Dong Nai Technology University, Bien Hoa City, Vietnam.
[e-mail: lethanhlanh@dnut.edu.vn, phamlam@dnut.edu.vn,]

² Contents Convergence Software Research Institute, Kyonggi University
154-42 Gwangkyosan-ro, Yongsong-gu, Gyeonggi-do, 16227, Republic of Korea
[e-mail: phamdinhlanh@kgu.ac.kr, kjc2232@naver.com, sangeun1804@kgu.ac.kr,
sunkh0507@gmail.com, khmjmc@kgu.ac.kr, kwang@kgu.ac.kr]

*Corresponding author: Kwanghoon Pio Kim

Abstract

In recent years, analyzing and using video content collected from closed-circuit television (CCTV) cameras have become popular and is highly beneficial in finding individuals, suspects, and criminals through extracted video data from CCTV systems. Using analyzed data helps reduce search time, labor and costs in managing and operating CCTV camera systems, especially for security and police agencies. Various methods to support person re-identification have been researched and introduced to address this issue. Especially in recent years, due to the strong development of deep learning algorithms, the results have been increasingly improved. In this paper, we study and present different feature-matching methods combined with the YOLOv8 algorithm to support searching persons through CCTV effectively. The results promise that the proposed method can be installed and used in practice to support CCTV camera management departments in finding persons in the CCTV systems.

Keywords: Person Re-identification, Feature matching methods, Person searching, Public security, CCTV cameras

1. Introduction

CCTV systems play an important role in monitoring and protecting public and private security. These systems provide the ability to record and monitor activities through cameras, support security assurance, and ensure order and safety in schools, shopping malls, public areas, etc. Using CCTV also supports locating missing people, finding and identifying suspects and criminals, thereby helping to prevent and reduce crime [1][2].

One of the challenging and difficult issues when managing CCTV systems is the handling of the vast amount of video data. Hundreds of gigabytes of video data are generated daily by each CCTV system, especially those installed for monitoring large areas. Managing, storing, and analyzing this data presents numerous challenges. Additionally, resolution and image quality are also concerns for these systems. Fog, rain, snow, or adverse weather conditions can reduce the image resolution from CCTV cameras. Simultaneously, CCTV cameras often struggle in poor lighting conditions or at night, reducing their ability to track and identify individuals

[3][4]. With the advancement of deep learning algorithms, the performance of content searching systems from CCTV videos has significantly improved. This enhancement strengthens the system's capabilities in detecting and re-identification people through CCTV [5][6]. In this paper, we describe and compare different feature matching methods for searching person from CCTV cameras. In particular, we utilize YOLOv8 with ORB and SIFT algorithms to identify and search people within CCTV footage. The next section discusses some of recent related works to person re-identification from CCTV cameras.

2. Related Works

In [5], Tong Xiao et al. introduced an End-to-End deep learning method for person search through video content. In this study, they introduced a convolutional neural network to extract feature maps of persons from an image. After that, a convolution layer with 512 3x3 filters was used and followed by anchor classification and regression layers to predict ROIs of persons. Then, these ROIs were used to pool the feature vector for each box on the feature maps. Finally, three fully connected layers were used to generate the final feature vector and compute distances. The proposed method was evaluated with a dataset containing 18,184 images and 8,432 persons, achieving an mAP of 55.7 and a top-1 accuracy of 62.7.

In [6], Jiale Cao et al. proposed a method using Transformers for searching persons and called it PSTR (End-to-End one-step person search with transformers). In this proposal, they comprised a person search-specialized module with a re-id decoder for person re-identification. The method was evaluated with two popular benchmarks: CUHK-SYSU and PRW, and outperformed many other methods with an mAP of 95.2 (CUHK-SYSU) and an mAP of 56.5.

3. Person Re-identification with YOLOv8 and Feature Matching Methods

In our study, to identify persons from CCTV cameras with the input search image, we first utilize the YOLOv8 algorithm to locate the positions of persons in the CCTV camera footage.

After obtaining the bounding boxes of these persons, we employ different image processing methods to compare the matching features and scores between the input search image and the persons in each frame of the camera feed.

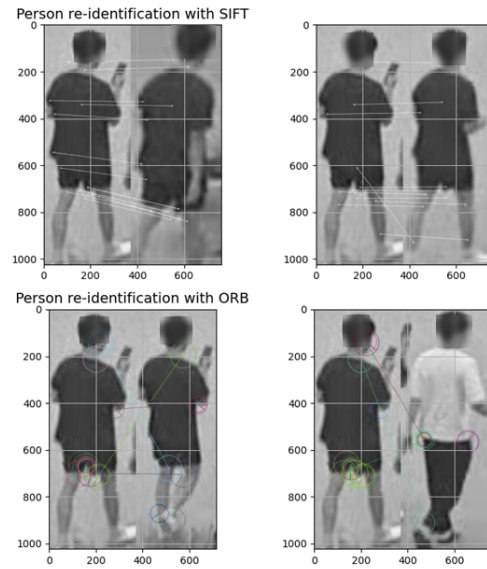


Fig. 1. Person re-identification with SIFT and ORB algorithms.

3.1 Person Re-identification with YOLOv8 and SIFT

The Scale-Invariant Feature Transform (SIFT) [7] is an algorithm used for detecting and description in images. The algorithm's idea is identifying key points that are invariant to scale changes, rotation, and illumination variations. These key points are represented by their local image gradients.

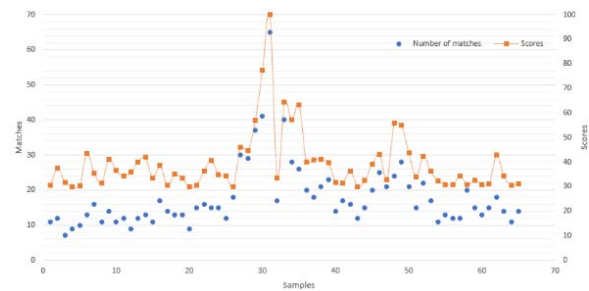


Fig. 2. Value of matches and scores using the SIFT algorithm with search image in Fig. 1.

After using SIFT algorithm to detect key points of input search person and candidate

persons, we evaluate the matching scores with the following equation. $\alpha = \frac{\mu}{\min(\lambda, \sigma)}$ (1) Where:

- α : Score of matching.
- μ : Total number of matches points
- λ : Total number of key points of input image.
- λ : Total number of key points of candidate image.

3.2 Person Re-identification with YOLOv8 and ORB

The Oriented FAST and Rotated BRIEF (ORB) [8] algorithm is used to detect features and description similar to the SIFT algorithm. However, ORB is designed combining the FAST (Features from Accelerated Segment Test) key point detector with the BRIEF (Binary Robust Independent Elementary Features) descriptor. The ORB algorithm uses a binary descriptor to identify key points that are invariant to rotation and scale changes. The scores of matching points are also calculated using the equation (1) described in the previous section.

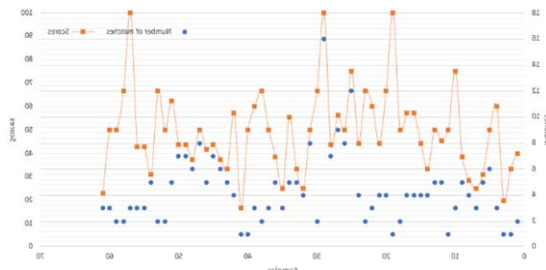


Fig. 3. Values of matches and scores using the ORB algorithm with search image in Fig. 1.

4. Conclusions

This paper presented a study on person re-identification utilizing YOLOv8, ORB and SIFT algorithms. In particular, at the first step, we detect a person through frames in video using YOLOv8. After that, key points detection methods are used with ORB and SIFT to get the list of unique key points of the person and compare the matching points with the input search image. We have identified the candidate persons corresponding to the input through the number of matched points and scores. The results of this method are promised to be applied and

used in practice supporting CCTV management systems.

References

- [1] E. L. Piza, B. C. Welsh, D. P. Farrington, and A. L. Thomas, "CCTV surveillance for crime prevention," *Criminol. Public Policy*, vol. 18, no. 1, pp. 135–159, Feb. 2019, doi: 10.1111/1745-9133.12419.
- [2] M. E. Hollis, "Security or surveillance?: Examination of CCTV camera usage in the 21st century," *Criminol. Public Policy*, vol. 18, no. 1, pp. 131–134, Feb. 2019, doi: 10.1111/1745-9133.12427.
- [3] C. V. Amrutha, C. Jyotsna, and J. Amudha, "Deep Learning Approach for Suspicious Activity Detection from Surveillance Video," *2nd Int. Conf. Innov. Mech. Ind. Appl. ICIMIA 2020 - Conf. Proc.*, pp. 335–339, Mar. 2020, doi: 10.1109/ICIMIA48430.2020.9074920.
- [4] M. P. J. Ashby, "The Value of CCTV Surveillance Cameras as an Investigative Tool: An Empirical Analysis," *Eur. J. Crim. Policy Res.*, vol. 23, no. 3, pp. 441–459, Sep. 2017, doi: 10.1007/S10610-017-9341-6/TABLES/3.
- [5] P. C. Dataset, "End-to-End Deep Learning for Person Search," pp. 1–4, 2017, [Online]. Available: <http://www.ee.cuhk.edu.hk/~xgwang/PS/paper.pdf>
- [6] J. Cao et al., "PSTR: End-to-End One-Step Person Search With Transformers," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2022-June, pp. 9448–9457, 2022, doi: 10.1109/CVPR52688.2022.00924.
- [7] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, Nov. 2004, doi: 10.1023/B:VISI.0000029664.99615.94/METRICS.
- [8] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, "ORB: An efficient alternative to SIFT or SURF," *Proc. IEEE Int. Conf. Comput. Vis.*, pp. 2564–2571, 2011, doi: 10.1109/ICCV.2011.6126544.

Reproduction of CCTV-Videos Using Contextualization Data Generated by Deep Learning Technology

Kyunghee Sun, Sangeun Ahn and Kwanghoon Pio Kim

Contents Convergence Software Research Institute, Kyonggi University

Gyeonggi-do, South Korea

[e-mail: kh_sun@kyonggi.ac.kr, sangeun1804@kyonggi.ac.kr, kwanghoon@kyonggi.ac.kr]

*Corresponding author: Kwanghoon Pio Kim

Abstract

This paper proposes a method to reproduce videos using background image and contextualization data generated with Deep learning technology from CCTV videos. CCTV cameras are installed in a wide area, regardless of city or region, and are typically fixed in place. By analyzing CCTV videos using deep learning, information such as object details, object behavior, and situational information can be stored as contextualization data describing the video. This allows for the preservation and utilization of information even for videos that have expired and been deleted due to the expiration of their storage period. The proposed method reproduces the original video using the background image of a fixed CCTV system, captured object images and the contextualization data generated from the CCTV video using deep learning technology. By applying the proposed technology, it will be possible to determine the past situation of videos deleted due to expiration of the retention period, identify the causes of safety accidents at production sites, and identify situations such as crime prevention and traffic accidents.

Keywords: video reproduction , CCTV video, contextualization data, deep learning, crime prevention

1. Introduction

CCTV cameras are installed in a wide area, regardless of city or region, and CCTV systems are operated by different entities for various purposes. In order to store CCTV videos, a storage device with a large capacity is required. And because CCTV videos contain personal information, the export and use of related information is extremely limited, and the retention period of CCTV videos is often limited to protect personal information. By analyzing CCTV video using deep learning and storing it as context data that explains the video, such as

object information, behavior, and situation information, storage device usage can be significantly reduced, and information on videos whose retention period has expired can also be stored and utilized.

We developed a video-objects contextualization system by using the unified and real-time video-object detection approach of YOLO (You Only Look Once) based upon the deep learning neural network models [1,3]. The system we developed can contextualize and store information on 80 objects, including people and cars, using the YOLO and COCO datasets [4,5]. Additionally, it includes techniques to protect

This research was supported by Industrial Job Advancement Technology Development Program through Korea Planning & Evaluation Institute of Industrial Technology(KEIT) funded by the Ministry of Trade, Industry and Energy(No. 20026458)

everyone's personal information and minimize exposure of sensitive information [2].

However, intelligent CCTV cameras and other CCTV video-related systems, including the system we developed, are systems that manage videos stored in storage devices or detect specific situations. There is no way to utilize information about videos whose retention period has expired or to reproduce the original video.

This paper proposes a method to reproduce videos using background image and contextualization data generated with Deep learning technology from CCTV videos.

2. Reproduction of CCTV-Videos using Contextualization Data by Deep Learning

2.1 Detecting and Contextualizing the Video-Objects

The YOLO system detects 80 different objects from CCTV videos and stores the type of each object and its location information in the video. Additionally, using deep learning technology, information such as a person's gender or clothing color can also be detected. The information such as video frame number, video creation date and time, and weather can also be stored as metadata. In the system we developed, this information is saved as a text file in JSON format [1].

2.2 Background Image of CCTV video

CCTV cameras are installed in a wide area, regardless of city or region, and are typically fixed in place. Therefore, the background image of a CCTV video is fixed and may vary depending on the season, weather, and time. From CCTV videos, we can select and save background images that are appropriate for each situation and that do not contain objects such as people or cars.

2.3 Object Images of CCTV video

Using deep learning technology, we can detect the type, location, and size information of objects included in CCTV videos. Only object images detected for each video frame can be cropped and saved.

3. Implementation of CCTV-video Reproduction using Contextualization Data by Deep Learning

By combining the background image, object images, and contextualization data, the original CCTV video can be reproduced. To reproduce the video, we used the DaVinci Resolution program, one of the free video production programs. This program creates videos by overlaying background images and object images.



```
{
  "Metadata": {
    "key_example": "some property",
    "obj_to_detect": [
      "person",
      "bicycle",
      "car",
      "motorcycle",
      "knife",
      "bus",
      "train",
      "truck"
    ],
    "Rescale": 100,
    "width": 1280.0,
    "height": 720.0,
    "Frame": 1701,
    "fps": 24.0
  },
  "Frame_1": {
    "ID": 1,
    "DetectedObjsDict": {
      "person": 1
    },
    "Detail": {
      "object_1": {
        "obj_id": "person_None",
        "class": "person",
        "location": [
          712,
          256,
          731,
          311
        ],
        "confidence": 0.33,
        "color": "Undetected"
      }
    }
  },
  "Frame_2": {
    "ID": 2,
    "DetectedObjsDict": {
      "person": 1
    },
    "Detail": {
      "object_1": {
        "obj_id": "person_None",
        "class": "person",
        "location": [
          707,
          245,
          732,
          303
        ],
        "confidence": 0.4,
        "color": "Undetected"
      }
    }
  },
  "Frame_3": {
    "ID": 3,
    "DetectedObjsDict": {
      "person": 1
    },
    "Detail": {
      "object_1": {
        "obj_id": "person_None",
        "class": "person",
        "location": [
          706,
          244,
          730,
          300
        ],
        "confidence": 0.42,
        "color": "Undetected"
      }
    }
  },
  "Frame_4": {
    "ID": 4,
    "DetectedObjsDict": {
      "person": 1
    },
    "Detail": {
      "object_1": {
        "obj_id": "person_None",
        "class": "person",
        "location": [
          706,
          244,
          730,
          300
        ],
        "confidence": 0.42,
        "color": "Undetected"
      }
    }
  },
  "Frame_5": {
    "ID": 5,
    "DetectedObjsDict": {
      "person": 1
    },
    "Detail": {
      "object_1": {
        "obj_id": "person_None",
        "class": "person",
        "location": [
          706,
          244,
          730,
          300
        ],
        "confidence": 0.42,
        "color": "Undetected"
      }
    }
  }
}
```

Fig. 1. Example: a background image, captured images of person object and contextualization data.

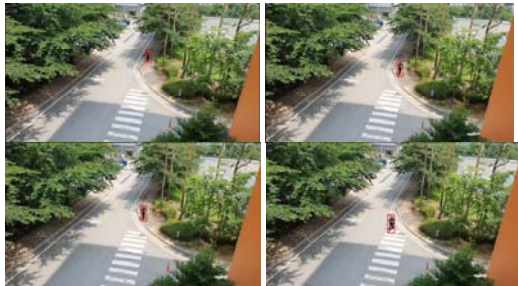


Fig. 2. Example: sample images of reproduced video using a background image, captured images of person object and contextualization data by deep learning.

In this paper, we propose a method to reproduce videos using background image and contextualization data generated with Deep learning technology from CCTV videos. The proposed method reproduces the original video using the background image of a fixed CCTV system, captured object images and the contextualization data generated from the CCTV video using deep learning technology.

Fig. 1 shows sample background image, person object images, and contextualization data used to reproduce the original video. **Fig. 2** shows a sample frame image captured from the result of reproducing the original video using sample data. As a result of the implementation, it was confirmed that the video could be reproduced similarly to the CCTV video.

4. Conclusions

In this paper, we propose a method to reproduce videos using background image and contextualization data generated with Deep learning technology from CCTV videos. The proposed method reproduces the original video using the background image of a fixed CCTV system, captured object images and the contextualization data generated from the CCTV video using deep learning technology.

We plan to develop a system that automatically reproduces original CCTV footage. This system includes a function to automatically save background images classified by weather, season, and time from CCTV videos. We plan to develop a system that selects an appropriate background image using contextualized data generated by deep learning technology and automatically reproduces original CCTV footage by combining object images or alternative images of objects.

By applying the proposed technology, it will be possible to determine the past situation of videos deleted due to expiration of the retention period, identify the causes of safety accidents at production sites, and identify situations such as crime prevention and traffic accidents.

References

- [1] Y. I. Park, S. E. Ahn, D. L. Pham, K. H. Sun, K. S. Kim, and K. H. Kim, "A Deep-Learning-Driven Video-Objects Contextualization System", *The 13th International Conference on Internet*, 2021.
- [2] D. L. Pham, J. C. Kim; S. E. Ahn, Y. E. Lee, J. H. Chang, H. Yoo, K. H. Sun, K. S. Kim, and K. H. Kim, "A Deep Learning-based Video Management System for Privacy-Preserving CCTV Video Data over The Web Environment", *the 17th Asia-Pacific International Conference on Information Science Technology (APIC-IST) 2022*, pp. 296-298, 2022.
- [3] K. H. Sun, H. Yoo, J. H. Chang, D. L. Pham, J. C. Kim, K. S. Kim, S. E. Ahn, and K. H. Kim, "An Active and Contextual Video-Object Bigdata Acquisition Framework", *The 17th Asia-Pacific International Conference on Information Science Technology (APIC-IST) 2022*, pp. 299-301, 2022.
- [4] Website: COCO Dataset in 2021
<https://viso.ai/computer-vision/coco-dataset/>
- [5] Redmon, Joseph and Farhadi, Ali, "YOLOv3: An Incremental Improvement," arXiv, 2018.

Deep Learning-Based Indoor Object Detection and Area Inpainting from RGB-D Camera

Wonseop Shin¹, Jaeseok Yoo², Sanghyun Byun¹, and *Sanghyun Seo³

¹Graduated School of Advanced Imaging Science, Multimedia & Film, Chung-Ang University
Seoul, South Korea

²SweetK, Gwacheon, South Korea

³College of Art and Technology, Chung-Ang University,
Anseong, South Korea

[e-mail: wonseop218@cau.ac.kr, jaeseokyoo9599@gmail.com, egoist12276@cau.ac.kr,
sanghyun@cau.ac.kr]

*Corresponding author: Sanghyun Seo

Abstract

With the development of various 3D-related contents, such as VR and Metaverse, the demand for reconstructing virtual spaces has significantly increased. However, the process of constructing indoor virtual spaces is often time-consuming and costly. In response to this challenge, we propose a method to facilitate the creation of indoor virtual spaces. We employ RGB-D cameras to capture indoor environments, classify and remove indoor objects, and restore occluded areas. Deep learning models based on Convolutional Neural Networks (CNN) are employed for indoor object Detection and occlusion restoration. We qualitatively demonstrate the results of object Detection using a 3D indoor object classification model and area restoration using an image restoration model. This research can be applied the efficiency of indoor virtual space development in the future, offering a more efficient for reconstructing immersive virtual environments.

Keywords: Deep learning, Object detection, Image restoration, 3D Reconstruction

1. Introduction

The proliferation of devices, network infrastructure, and metaverse technologies has propelled content development. Consequently, there is a growing emphasis on designing avatars and creating virtual spaces. Effectively immersing users in VR/AR content requires authentic representations of the real world in virtual environments. Therefore, constructing virtual spaces is an essential endeavor.

To build virtual environments, designers often rely on specialized tools for manual 3D modeling. However, this process is time and cost-intensive and may yield disparities in the quality of models created by different individuals [1].

To address these challenges, we propose a pipeline for reconstructing virtual indoor spaces. This pipeline involves capturing interior environments using RGB-D cameras and creating point clouds with the assistance of Open3D [2]. A deep learning model based on

CNN is used to identify and remove indoor objects within the point cloud while simultaneously restoring obscured regions. Segmentation methods are employed to identify and remove indoor objects, while the restoration of occluded regions is accomplished through the use of DIP (Deep Image Prior) [3].

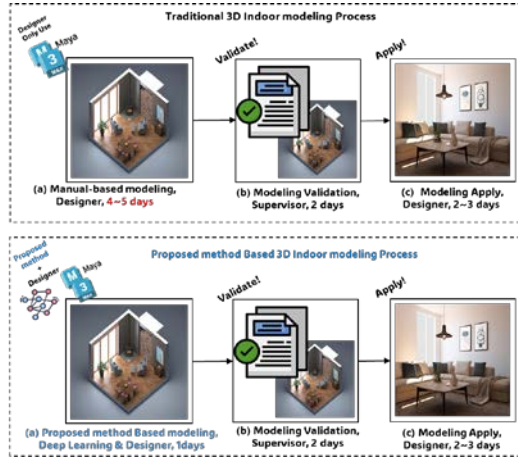


Fig. 1. Comparison between the traditional and proposed method-based modeling. The traditional 3D indoor modeling process is depicted above, while the proposed method-based 3D indoor modeling process is illustrated below.

Fig. 1 demonstrates the time required for manual modeling by a designer and modeling using the proposed method in this study. In the case of manual modeling, the initial modeling process using only 3D authoring tools lasts approximately 4-5 days. Subsequently, the model undergoes review by a supervisor, and the application of the modeling results adds another

5 days. Therefore, manual modeling can take up to 10 days. If modeling is based on the approach proposed in this paper, it can be completed in 1-2 days by extracting a pure 3D space using the method suggested in the initial modeling task and subsequently modifying it by employing a 3D authoring tool. The overall time required can be further reduced if the initial work is significantly decreased. This research presents a method that utilizes deep learning models to expediently create three-dimensional virtual interiors and apply them to modeling tasks.

2. 3D virtual indoor space reconstruction

To achieve the structure of a purely 3D virtual space, we capture color and depth images using an RGB-D camera, which are then used to create the virtual space. The structure of the created space is assessed, enabling the recognition of 3D objects within it, and missing parts are restored through the removal of identified objects. The pipeline proposed in this paper is presented in **Fig. 2**.

2.1 Create a virtual indoor space point cloud

In this paper, we used the Azure Kinect RGB-D camera to capture images of an indoor environment. RGB and depth images were extracted, and point clouds were derived using the Open3D [2] library. The point clouds were obtained by estimating feature points between the extracted images and estimating the relationship for converting between RGB and

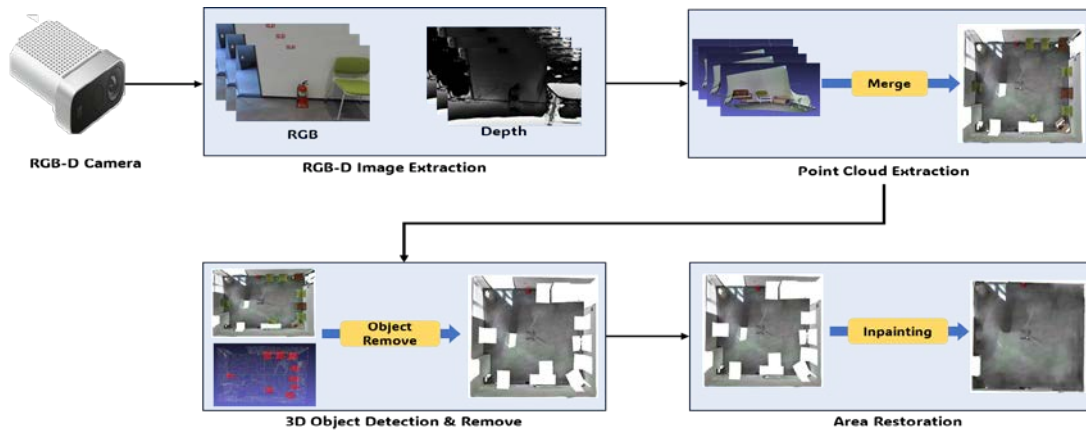


Fig. 2. Deep learning-based 3D modelling pipeline

Depth image pairs. The ORB algorithm [4] was utilized to match feature points between two image pairs, while 5-Point RANSAC[5] was employed to eliminate mismatched points and calculate the camera trajectory.

However, capturing images using an RGB-D camera can introduce noise into the image. As the error accumulates, it becomes difficult to accurately estimate the camera's position. To solve this problem, we need to find a correlation between the camera's position registered in the current frame and the previous pose graph to reduce the accumulated error. We minimize the error by enriching the camera's position recorded in the pose graph. We obtained 5,300 color and depth images and divided them into 100 frames to extract 53 partial point clouds. The Colored-ICP [6] algorithm was then used to detect connectivity in the partial point clouds. Afterwards, we aligned the partial point clouds to enable the extraction of the complete point cloud of the indoor space.

2.2 Object detection in indoor spaces

For the 3D classification of objects, we employed VoteNET [7]. To train our model, data captured exclusively with Kinect V2 from the SUN RGB-D [8] dataset was extracted and utilized.

The SUN RGB-D dataset comprises 10,335 RGB-D images taken with the Intel RealSense, ASUS Xtion, Kinect v1, and Kinect v2, containing 146,617 2D polygons and 64,595 3D bounding boxes. However, the point cloud we obtained was based on the image data acquired using Azure Kinect, which does not align with the point cloud coordinate system used in the model's training, leading to inaccurate object detection. Therefore, the object detection outcomes are imprecise.

To solve this problem, we matched the coordinate system to detect objects. Since the direction of the point cloud we extracted is upward on the y-axis, and the direction of the point cloud used as training data is upward on the z-axis, we rotated the point cloud extracted by 90° in the x-axis direction to match them. After matching the coordinate system, we used the trained model to perform object classification again. After object detection, the bounding box was applied to the point cloud to remove the object. Fig. 3 shows the result of removing the detected objects.

2.2 Inpainting missing areas in indoor spaces

When no point cloud is accessible for the object's location, and there are obstructed or unphotographed areas that require restoration, we implement an image-based reconstruction technique to restore the object. The plane belonging to the entire indoor space's point cloud was converted into a two-dimensional image employing orthographic projection. Restoration is achieved using the DIP [3] approach. A binary image was generated of the damaged area to assist in digital image processing. Vertical vectors were obtained from all planes within the indoor point cloud and ortho-projected, enabling the demarcation of the damaged region in need of restoration. By learning the correlation of non-zero pixels and proceeding with inpainting, the lost area was successfully restored. Figure 4 shows the restored area.

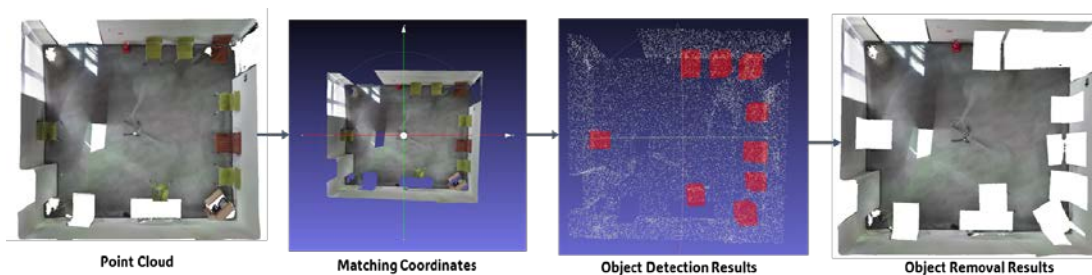


Fig. 3. Results of Object detection & removal



Fig. 4. Area inpainting with DIP Neural neural networks

3. Results

The virtual space created through the proposed pipeline was compared to the virtual space obtained using the image from the RGB-D camera. While constructing the virtual space with the RGB-D camera, the wall surfaces were affected by noise and vanishing areas caused by obstructing objects. However, the wall surfaces of the virtual space created using the proposed pipeline show that the noise and vanishing area issues have been resolved.

3.1 Extension for Application

The virtual spaces produced by the proposed pipeline can be conveniently adjusted and repurposed using the virtual spaces generated by 3D modeling tools. To generate the ultimate indoor virtual space, the restored vanishing area was integrated into a point cloud, resulting in the creation of a virtual space. The objects were subsequently modeled using Autodesk Maya and merged into the virtual space. **Fig. 5** shows the process of constructing a virtual space with Maya software and presents the final interior environment.

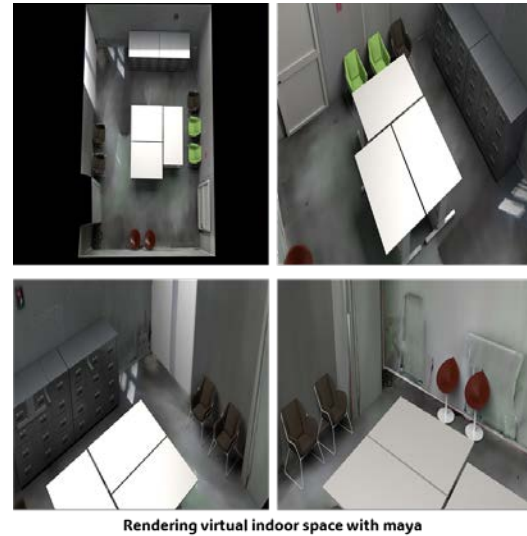


Fig. 5. Virtual indoor space rendering result, incorporating the constructed virtual space and objects modeled in Autodesk Maya

4. Conclusions

In this paper, we proposed a space structure-based modeling pipeline to efficiently create virtual indoor spaces while minimizing manual efforts and overcoming issues with traditional virtual space creation methods. The pipeline utilizes point clouds extracted from real-world indoor spaces acquired through RGB-D cameras as the base dataset. The proposed pipeline operates through indoor virtual space construction, 3D object detection, texture inpainting.

We reconstructed the virtual indoor space using 3D reconstruction techniques based on the Open3D library and addressed the noise in the extracted point clouds through space structure-based modeling. Additionally, we employed the VoteNet [6] for object classification and used the DIP neural network to restore the missing areas after removing the classified objects. As a result, we successfully constructed a virtual indoor space without objects, demonstrating its usability in 3D authoring tools and validating its practicality in VR, AR, and metaverse content creation. The proposed method allows users to construct and edit virtual spaces easily, significantly contributing to the entertainment industry.

Acknowledgements

This research was supported by Culture, Sports and Tourism R&D Program through the Korea Creative Content Agency grant funded by the Ministry of Culture, Sports and Tourism in 2023 (Project Name: Development of digital abusing detection and management technology for safe metaverse service, Project Number:RS-2023-00227686, Contribution Rate: 100%)

References

- [1] C.W. Chu, J.Y.Park, H.W.Kim, J.C.Park, S.J. Lim and B.K. Koo, "Recent Trends of 3D Reconstruction Technology", *In Electronics and Telecommunications Research Institute on Electronics and Telecommunications Trends*, 22(4), 1-0, 2007.
- [2] QY.Zhou, J.Park, and V.Koltun, "Open3D: A Modern Library for 3D Data Processing," *arXiv:1801.09847 [cs.CV]*, 2018.
- [3] D.Ulyanov, A.Vedaldi, and V.Lempitsky, "Deep image prior." *in Proc. of the IEEE conference on computer vision and pattern recognition*, pp. 9446-9454. 2018
- [4] E. Rublee, V.Rabaud, K.Konolige, and G.Bradschi, "ORB: An efficient alternative to SIFT or SURF", *In 2011 International conference on computer vision*, pp. 2564-2571, 2011.
- [5] D. Nistér, "An efficient solution to the five-point relative pose problem." *IEEE transactions on pattern analysis and machine intelligence*, 26(6), 756-770, 2004
- [6] J.Park, Y.Zhou and V.Koltun, "Colored point cloud registration revisited," *in Proc. of the IEEE international conference on computer vision*, 143-152, 2017
- [7] R.Qi, O.Litany, K.He, K and J.Guibas, "Deep hough voting for 3d object detection in point clouds," *in Proc. of the IEEE/CVF International Conference on Computer Vision*, 9277-9286, 2019.
- [8] S. Song, S. P. Lichtenberg and J. Xiao, "SUN RGB-D: A RGB-D scene understanding benchmark suite." *In Proc. of the IEEE conference on computer vision and pattern recognition*, pp. 567-576. 2015.

When Pyramid Pooling Transformers Assists Lightweight Decoder for Salient Object Detection in Edge Intelligence

Thien-Thu Ngo¹, Md Delowar Hossain¹, Teh-Jen Sun², Yuri Seo², Hoang-Hai Tran³,
and Eui-Nam Huh^{1*}

¹ Department of Computer Science and Engineering, Kyung Hee University

² Department of Artificial Intelligence, Kyung Hee University
Suwon, South Korea

³ School of Information and Communication Technology

Hanoi University of Science and Technology, Hanoi, Vietnam

[e-mail: thu.ngo@khu.ac.kr, delowar@khu.ac.kr, dlwmznzl@khu.ac.kr, yuri0329@khu.ac.kr,
haith@soict.hust.edu.vn, johnhuh@khu.ac.kr]

*Corresponding author: Eui-Nam Huh

Abstract

Cutting-edge deep learning models for salient object detection (SOD) excel at pinpointing visually prominent objects but face challenges when deployed on resource-constrained devices. To address this, we present an innovative SOD architecture, featuring pyramid pooling transformers and a lightweight decoder. Our decoder incorporates Wavelet Transform-Centric Aggregation for boundary detection in shallow layers and introduces the Parallel Partial Semantic Aggregation Mechanism to efficiently extract multi-scale semantic cues in deeper layers. Empirical evaluations across five benchmark datasets highlight our model's efficiency, competing with bulkier SOD architectures while reducing computational operations by around 98% and parameters by 77%.

Keywords: transformers, salient object detection, edge intelligence, deep learning

1. Introduction

Salient object detection (SOD) powered by deep learning is vital in various vision applications, isolating important objects with pixel-wise saliency. Despite its relevance in fields like autonomous vehicles and content retrieval, integrating SOD into resource-constrained edge devices poses challenges due to computational demands. While state-of-the-art SOD models excel [1], they hinder edge deployment with

complex architectures. Researchers have proposed lightweight SOD models trading precision for efficiency, but this can be risky in safety-critical settings. Transformers, known for NLP, now impact computer vision. They excel in dense prediction, capturing long-range pixel dependencies and holistic context. Transformers also handle hierarchical feature representations, vital for precise predictions.

This work was supported in part by the Institute of Information and Communications Technology Planning and Evaluation (IITP) grant funded by the Korea Government (MSIT) (Artificial Intelligence Innovation Hub) under Grant 2021-0-02068 and MSIT(Ministry of Science and ICT), Korea, under the ITRC(Information Technology Research Center) support program(IITP-2023-RS-2023-00258649) supervised by the IITP(Institute for Information Communications Technology Planning Evaluation)

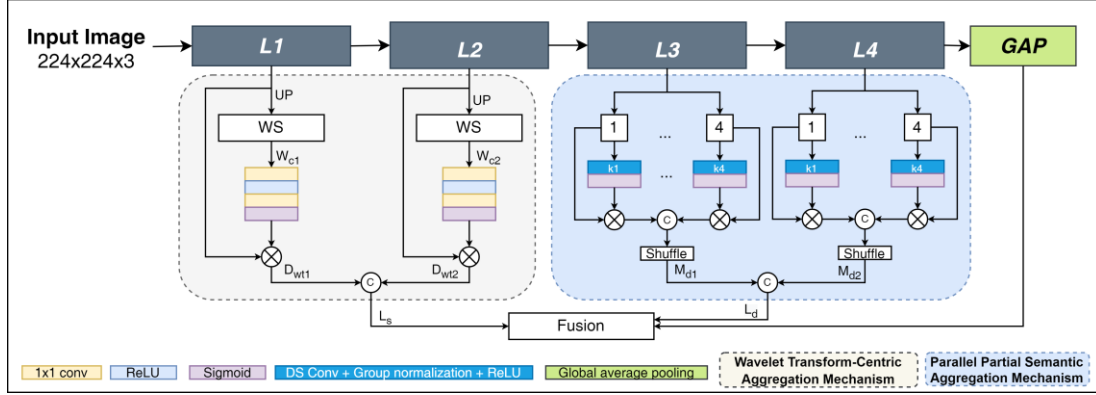


Fig. 1. The proposed framework consists of pyramid pooling transformers and lightweight decoder

To balance efficiency and precision in SOD, we combined pyramid pooling transformers (p2T) with a lightweight decoder, consisting of wavelet-based boundary detection and multi-scale feature recalibration. This streamlining maintains precise object boundary detection with reduced computational load, making it suitable for edge deployment.

2. Related Works

SOD has evolved with the rise of Convolutional Neural Networks and Transformers. Current SOD models integrate robust backbones, often combining CNNs and Transformers, using techniques like knowledge distillation and adversarial training. While powerful, these models can be resource-intensive. Lightweight SOD models prioritize efficiency but may sacrifice accuracy. In 2017, Transformers, like Vision Transformer (ViT)[2], made a mark in natural language processing and image classification. The Pyramid Pooling Transformer (P2T)[3] introduced efficient multi-scale feature extraction, enhancing contextual understanding for SOD. P2T's fusion in SOD sets it apart from ResNet50, offering improved global context awareness.

3. The Proposed architecture

3.1 Wavelet Transform-Centric Aggregation Mechanism

Our approach utilizes wavelet transforms to capture spatial and spectral details in shallow

layers, effectively identifying object edges and reducing noise, without adding computational overhead. As depicted in Fig. 1, we use 2D Discrete Wavelet Transform to decompose upsampled features into four distinct sub-bands L_l, L_h, H_l, H_h , where L_l sub-band captures low-frequency components, and the L_h, H_l , and H_h sub-bands capture high-frequency components. Thus, we combine them into two branches in the equation:

$$W_c = \text{Concat}(L_l, L_h + H_l + H_h)$$

where W_c represents the coefficient matrix. Then, we derive the frequency-aware attention feature vector by channeling it through a sequence of convolutions and activations, and multiply it with input features to yield D_{wti} , then obtain the final vector through the two outputs.:

$$D_{wti} = \text{Sig}(\text{Conv1}(\text{ReLU}(\text{Conv1}(W_c)))) \times L_i$$

$$L_s = \text{Concat}(D_{wt1}, D_{wt2})$$

where Conv1 is 1x1 convolution operation, the DW layer represents a depth-wise separable convolution, group normalization, ReLU. L_s denotes the output of the given module.

3.3 Parallel Partial Semantic Aggregation Mechanism

Our module harnesses high-level semantic cues in deep layers through multi-scale feature augmentation. To reduce computation, we partition features into channel-wise sub-features (H_j) and employ 3x3 dilated convolutions with different dilation rates ($r = 1, 3, 5, 7$) to expand the receptive field for each sub-branch:

$$W_j = \text{Sigmoid}(DW_r(H_j))$$

| Models | Flops (G) | Params (M) | ECSSD | | | PASCAL-S | | | DUTS-TE | | | DUT-OMRON | | | HKU-IS | | |
|--------------------------|--------------|---------------|-------------|-------------|---------------|-------------|-------------|---------------|-------------|-------------|---------------|-------------|-------------|---------------|-------------|-------------|---------------|
| | | | $F\uparrow$ | $S\uparrow$ | $M\downarrow$ | $F\uparrow$ | $S\uparrow$ | $M\downarrow$ | $F\uparrow$ | $S\uparrow$ | $M\downarrow$ | $F\uparrow$ | $S\uparrow$ | $M\downarrow$ | $F\uparrow$ | $S\uparrow$ | $M\downarrow$ |
| RFCN | 102.8 | 134.69 | .898 | .859 | .095 | .837 | .808 | .118 | .782 | .792 | .089 | .738 | .773 | .094 | .894 | .858 | .079 |
| NLDF | 263.9 | 35.49 | .905 | .875 | .063 | .831 | .803 | .099 | .812 | .815 | .065 | .753 | .770 | .079 | .902 | .879 | .048 |
| DSS | 114.6 | 62.23 | .921 | .882 | .052 | .831 | .797 | .093 | .830 | .822 | .052 | .781 | .788 | .063 | .916 | .879 | .040 |
| PicaNet | 37.1 | 32.85 | .935 | .917 | .046 | .857 | .853 | .076 | .860 | .868 | .050 | .803 | .831 | .065 | .919 | .904 | .044 |
| DGRL | 24 | 126.35 | .922 | .903 | .041 | .854 | .836 | .072 | .829 | .841 | .050 | .774 | .806 | .062 | .910 | .895 | .036 |
| PAGR | 100.4 | 23.63 | .927 | .889 | .061 | .856 | .818 | .093 | .855 | .837 | .056 | .771 | .775 | .071 | .918 | .887 | .048 |
| C2S | 20.5 | 137.03 | .896 | .881 | .059 | .829 | .826 | .087 | .790 | .817 | .066 | .733 | .779 | .079 | .883 | .872 | .051 |
| R3Net | 56.16 | 60.24 | .934 | .910 | .040 | .835 | .806 | .092 | - | - | - | .795 | .816 | .063 | .915 | .895 | .035 |
| EGNet | 270.8 | 108.07 | .947 | .924 | .037 | .865 | .852 | .074 | .889 | .887 | .039 | .815 | .840 | .053 | .935 | .917 | .031 |
| BASNet | 121.6 | 55.90 | .942 | .916 | .037 | .854 | .838 | .076 | .860 | .865 | .047 | .805 | .836 | .057 | .930 | .907 | .033 |
| AFNet | 38.4 | 37.10 | .935 | .913 | .042 | .863 | .849 | .070 | .863 | .867 | .045 | .797 | .826 | .057 | .925 | .905 | .036 |
| GateNet | 108.3 | 100.02 | .945 | .919 | .040 | .869 | .857 | .067 | .888 | .885 | .040 | .818 | .838 | .055 | .933 | .915 | .033 |
| LGSN | 102.18 | 68.69 | .946 | .923 | .036 | .872 | .855 | .068 | .898 | .893 | .033 | .827 | .846 | .049 | .937 | .921 | .029 |
| Ours_{T1} | 2.11 | 12.75 | .943 | .923 | .036 | .861 | .856 | .065 | .873 | .881 | .039 | .814 | .840 | .056 | .931 | .916 | .033 |

Table 1. The quantitative comparison with others state of the art models in terms of F-measure, S-measure and MAE

where W_j signifies the attentive features for each respective sub-branch. We multiply attentional features with sub-features and concatenate the outputs, shuffling them to yield L_d . The dual outputs are concatenated for the final feature vector.

$$M_d = \text{Shuffle}(\text{Concat}[W_j * H_j])$$

$$L_d = \text{Concat}(M_{d1}, M_{d2})$$

Leveraging multi-scale representations across distinct sub-features empowers this aggregation strategy to construct a hierarchical feature, enhance contextual information, address scale discrepancies, and optimize computational efficiency.

3.4 Shallow to Deep Fusion Mechanism

We use a streamlined fusion strategy, extracting global contextual features with Global Average Pooling, applying element-wise multiplication, concatenating, and using depthwise separable convolution for refinement. Shallow features are merged for boundary-aware details, followed by additional DW layers for the final output.

4. Experiment Results

4.1 Performance Comparison

Datasets: ECSSD, HKUIS, DUTOMRON, DUTS-TE, and PASCAL-S. Metrics: F-measure, MAE, and S-measure. Computational efficiency is measured using parameters, FLOPS, and FPS.

Notably, as shown in Table 1, our architecture outperforms BASNet in F-measure, S-measure, E-measure, and MAE while reducing FLOPS by 98% and parameters by 77%. Compared to GateNet, our model excels in S-measure on the ECSSD dataset, with similar reductions in

FLOPS and parameters. In Table 2, our FPS rate is double that of EGNet and BASNet and 75 times faster than RFCN.

Table 2. FPS comparison

| Models | RFCN | NLDF | DSS | PicaNet |
|--------|---------|-------|--------|---------|
| FPS | 0.4 | 18.5 | 7.0 | 5.6 |
| Models | C2S | EGNet | BASNet | AFNet |
| FPS | 25 | 12.7 | 12.5 | 22 |
| Models | GateNet | DGRL | Ours | |
| FPS | 18 | 3.6 | 30 | |

5. Conclusion

We introduce an novel SOD architecture with pyramid pooling transformers and a lightweight decoder. It utilizes a Wavelet Transform-Centric Aggregation Mechanism for boundary information and a Parallel Partial Semantic Aggregation Mechanism for multi-scale semantics, achieving SOD performance akin to large frameworks while reducing FLOPS by 98% and parameters by 77%.

References

- [1] Wang, W., Lai, Q., Fu, H., Shen, J., Ling, H. and Yang, R., 2021. Salient object detection in the deep learning era: An in-depth survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(6), pp.3239-3259.
- [2] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, et al., "An image is worth 16x16 words: Transformers for image recognition at scale," *arXiv preprint arXiv:2010.11929*, 2020.
- [3] Wu, Y.H., Liu, Y., Zhan, X. and Cheng, M.M., 2022. P2T: Pyramid pooling transformer for scene understanding. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.

A Study on Problems of the Analysis of AI Medical Images

Sun-Kuk Noh¹

¹ SW Convergence Education Institute, CHOSUN University, Gwangju, Republic of Korea
[e-mail:nsk7078@chosun.ac.kr]

Abstract

Artificial intelligence (AI) is critical to data-driven decision-making to increase resource utilization, operational performance, and service quality in various industry sectors, especially healthcare. The application of AI is essential for collecting and analyzing data such as personal genetic information, medical information, and lifestyle information. Deep learning (DL) is a new AI technology applied to the medical imaging field to analyze and diagnose medical images of organs. In this paper, we looked into problems with medical image data in medical image analysis using AI, and analyzed the configuration of a medical information system for transmitting medical image data using AI.

Keywords: artificial intelligence (AI), deep learning (DL), medical image analysis, picture archiving and communication system (PACS)

1. Introduction

The fourth industrial revolution (4IR) of the 21st century is underscored by the integration of artificial intelligence (AI) and computer-based information communication technology (ICT) that improves productivity in most industries. In particular, AI processes big data beyond human capabilities and facilitates accurate, efficient, and fast decision-making. Therefore, AI is emerging as a solution to productivity improvement and problems in various fields [1-3].

AI is critical for data-driven decision-making in various industries, especially healthcare. Massive amounts of data, rapidly advancing computations, and high-performance models validated against natural images are delivering results that exceed the diagnostic and interpretive capabilities of doctors. Using AI to increase resource utilization, operational performance, and quality of service in healthcare operations can significantly improve care outcomes and increase patient satisfaction while reducing costs. A study by Bennett et al. reported that using AI algorithms to diagnose diseases improved diagnostic performance by 41.9% and reduced healthcare costs by 58.5% [4].

In the medical imaging field, machine learning (ML) and deep learning (DL) are applied to analyze medical images of organs, and this field constitutes an active area of research. DL algorithms and model structures have been introduced to DL-based medical image analysis, and studies on lesion detection, quantification, and classification are ongoing [5]. Indeed, healthcare is shifting towards personalized approaches, such as proactive diagnosis and prevention of individual diseases. Therefore, the collection and analysis of large-scale big data such as genetic, medical, and living information by applying AI are essential [6]. In Korea, research on AI medical software termed Doctor Answer has been ongoing since 2018 with the participation of governments, hospitals, and companies to identify a medical approach suitable for the Korean population. The development of Doctor Answer 2.0 is currently in progress [7]. In this paper, we looked into problems with medical image data in medical image analysis using AI, and analyzed the configuration of a medical information system for transmitting medical image data using AI.

2. Related Work

2.1 AI

AI technology is classified into ML and DL, as shown in Fig. 1.

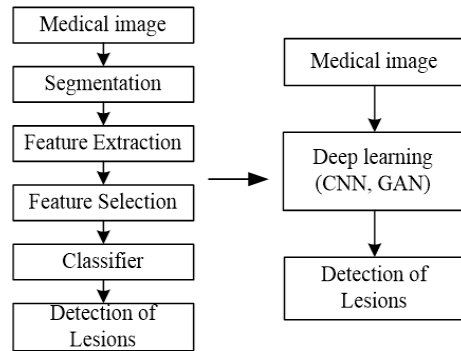


Fig. 1. Machine learning (ML) and deep learning (DL)

2.2 Medical field and AI

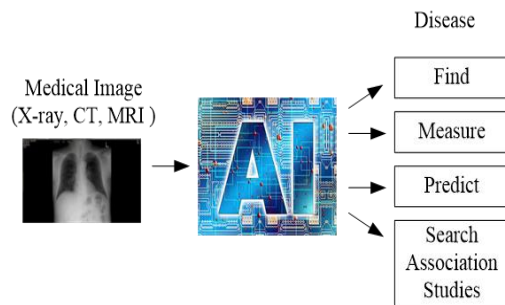


Fig. 2. AI-based medical field

The analysis and diagnosis of diseases using AI are presented in Fig. 2. According to a STAT News report in 2019, Google AI has outperformed six radiologists in detecting lung cancer, identifying 5% more cancers and reducing false positive rates by 11% [8]. According to a paper published in Nature Medicine in 2019, 6716 cases of lung cancer detected using AI-based computed tomography (CT) imaging exhibited an accuracy of 94%. The accuracy of magnetic resonance imaging (MRI) for brain tumors, which are difficult to judge, was >85%. The accuracy of Google AI-based diagnosis of breast and lung cancers was 99% and 95%, respectively; and the accuracy for the diagnosis of metastatic cancer was significantly higher. Moreover, a paper published in Nature in 2020 reported that Google's DeepMind algorithm may be more accurate than real-life doctors for identifying breast cancers [9].

3. Medical Imaging System

3.1 Medical imaging equipment

Medical imaging methods for image acquisition are listed in Table 1.

Table 1. Medical imaging equipment

| Source | Method | Use and characteristics |
|-------------------|--|--|
| Visible ray | Endoscopy | Used for intestinal endoscopy, with a light source and CCD sensor in front of the endoscope |
| | Microscopy | While examining tissue, microscope eyepieces and objective lenses are used to magnify small objects |
| Radiation | X-ray | Light source is emitted from an X-tube and passes through the body; image is created by the transmitted x-ray intensity difference of tissue at each site |
| | CT | Captures x-ray images from various angles around cross-section of the human body, integrates two or more 2D x-ray images into one 3D image, and generates 3D images within a few seconds |
| | Positron emission tomography (PET) | A radioactive drug (radioactive glucose) emitting positrons is injected into the body and is transformed into 3D images after 360° analysis of the human body |
| Magnetic field | MRI | Based on a magnetic field, hydrogen atoms in the body generate 3D images based on the difference in the speed of transmission through muscles and fat |
| | Ultrasonic wave | Used for heart, stomach, lung, and fetal tests |
| Digital pathology | Virtual microscopy WSI (whole-slide imaging) | Sub-field of pathology that focuses on data management based on information generated from digitized specimen slides. Using computer-based technology, digital pathology utilizes virtual microscopy or WSI. Can be used for feature detection of mitotic figures, epithelial cells, or tissue-specific structures such as lung cancer nodules, glomeruli, or vessels; or estimation of molecular biomarkers such as mutated genes, tumor mutational burden, or transcriptional changes [16–18]. |

3.2 Picture Archiving and Communication System (PACS)

The Picture Archiving and Communication System (PACS) is a large-scale system in radiology that coordinates image information throughout the hospital and is used in connection with most departments. To build a PACS environment, technologies such as Image Display and Processing, Data Communication and Networking, Database system, Information Management, User Interface, and Data Storage/Archive must be integrated. PACS stores images obtained by advanced imaging devices such as general X-ray, CT, MRI, DSA, ultrasound, and nuclear medicine in a computer as digital images and transmits them to each terminal for simultaneous image search anywhere in the hospital.

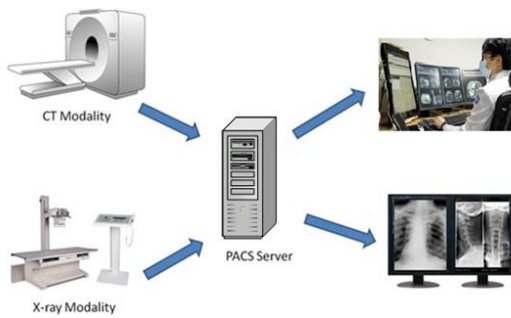


Fig. 4. PACS

4. Problems of Medical Imaging Data

Medical imaging datasets include a vast amount of data used to train DL to classify specific diseases. It is common to use data for pre-learning and analyzing for performance verification of AI Model. Medical imaging datasets exhibit the characteristics of volume, diversity, velocity, reliability, and value, which are the five major factors in big data. Obtaining big data from medical images with annotations is challenging because it requires professional manpower and is time-consuming to read and comment on medical images. In addition, patient consent is required for the use of big data with patient medical image. The limitations of medical imaging data is listed in Table 2.

Table 2. Limitations of medical imaging data

| Clause | Content |
|------------------------|--|
| Quality | Difficulty in obtaining large amounts of labeled data |
| Pixel size | There are many 3D images, and image size is quite large |
| Small object size | Object size is relatively small |
| Compensating | The same tissue image of the same disease requires correction for age |
| Additional information | Use other information such as sex, smoking, and drinking for classification analysis |

5. Conclusion

In this paper, we looked into problems with medical image data in medical image analysis using AI. Sharing datasets is crucial in DL-based medical image analysis. A dataset constitutes a large medical imaging data labeled by a trusted expert group. However, obtaining labeled medical imaging data for all diseases is difficult. In addition, label processing by reading medical images requires professional manpower, is time-consuming, and may result in differing opinions among medical staff. Additionally, ethical concerns arise since datasets contain patients'

personal information. In this regard, anonymization and limited access to datasets must be considered.

ACKNOWLEDGMENTS

This study was supported by research fund from Chosun University (2022)

References

- [1] I. Arel, D. C. Rose, and T. P. Karnowski, "Research frontier: deep machine learning a new frontier in artificial intelligence nceresearch," IEEE Computational Intelligence Magazine, vol. 5, no. 4, 2010.
- [2] R. Kang Hyun, "Composition of visual feature vector pattern for deep learning in image forensics," IEEE Access, vol. 8, pp. 188970–188980, 2020.
- [3] S-K Noh, "Recycled Clothing Classification System Using Intelligent IoT and Deep Learning with AlexNet," Computational Intelligence and Neuroscience Volume 2021, Article ID 5544784, 8 pages <https://doi.org/10.1155/2021/5544784>
- [4] AI Framework Predicts Better Patient Health Care and Reduces Cost, Casey Bennett, By INDIANA UNIVERSITY FEBRUARY 12, 2013, <https://scitechdaily.com/ai-framework-predicts-better-patient-health-care-and-reduces-cost/>
- [5] Rav D, Wong C, Deligianni F, Berthelot M, Andreu Perez J, Lo B, et al. Deep learning for health informatics. IEEE J Biomed Health Inform. 2017 Jan;21(1):4-21.
- [6] Geert Litjens et. al., "A Survey on Deep Learning in Medical Image Analysis." arXiv:1702.05747, 2017. 6. 4
- [7] Haleem A et al., Current status and applications of Artificial Intelligence (AI) in medical field: An overview, Current Medicine Research and Practice, <https://doi.org/10.1016/j.cmrp.2019.11.005>
- [8] Google AI improves lung cancer diagnosis accuracy, May 21, 2019, <https://www.massdevice.com/report-google-ai-improves-lung-cancer-diagnosis-accuracy/>
- [9] Scott Mayer McKinney et al., International evaluation of an AI system for breast cancer screening, Nature, pp 89–94, Vol 577, 2 January 2020.

Image encryption based on the maximum-periodic chaotic sequence generator

Sung-Won Kang¹⁾ and Kyung-Hyune Rhee²⁾

¹⁾ Department of Information Security, Pukyong National University
Busan, South Korea
jsm2371@hanmail.net

²⁾ Division of Computer and AI Engineering, Pukyong National University
Busan, South Korea
yisecure@gmail.com

Abstract

Due to the development of internet social network service technology and AI, various issues such as image leakage, unauthorized use, and information exposure continue to occur online. Researchers are developing various security technologies to prevent these issues. Image encryption technology is the technique that can be applied as part of security technology among image processing technologies. Recently, image encryption based on a chaotic sequence generator is actively being studied to efficiently process in a large image database. The chaotic sequence generator has the characteristic of quickly generating sequences due to the high randomness of the simultaneously generated sequences, furthermore it has high resistance to statistical attacks. However, even if it is guaranteed that the periodicity of the sequence generated in the continuous function space is infinite in the theory, there is a risk that the generated sequence may have a short period due to round-off problems when operating through an application. In this paper, we propose a new chaotic sequence generator that always generates a maximum-length period sequences. And we construct image encryption based on the generator and analyze the results.

Keywords: Chaotic sequence generator, Image encryption, Maximum-length sequence

1. Introduction

Recently, image encryption technology using a chaos sequence generator is being actively researched to efficiently process large amounts of images uploaded and downloaded to the Internet [1]. The chaos sequence generator can quickly generate a sequence, and because the sequence created at the same time has the characteristic of having high randomness, it has very strong resistance to statistical attacks [2-3]. Taking advantage of these characteristics, image encryption technologies using chaos sequence generators are being actively researched. However, when the periodicity of the sequence generated by the chaos sequence generator is

actually operated through an application, there is a problem that the generated sequence may have a short period due to the round-off problem, and discussion is needed to resolve this. In this paper, we propose a new generator that generates a chaotic sequence that always has the maximum length cycle, create image encryption based on this generator, and analyze the results.

2. Construction of the proposed system

2.1 The maximum-periodic chaotic sequence generator

If the characteristic polynomial of the discrete function that generates the sequence is a

primitive polynomial, it is always possible to generate a sequence with the maximum period. For example, if you want to generate a sequence that falls within the range of numbers with a size of 256 bits, it is possible to always generate a sequence with the maximum period using a generation function that has a 256th degree primitive polynomial as a characteristic polynomial. In this case, the period is $2^{256}-1$. Discrete functions suitable for the maximum period sequence generator include companion matrices and 90/150 cellular automata.

Fig. 1 is a flow chart showing the sequence generator. It is possible to create a chaotic sequence with a large period at any time by simply combining the sequence generated using the chaos function and the sequence generated using the maximum period sequence generator with xor. The period of the sequence created by the linear combination of two sequence is the greatest common multiple of the period of the two sequences. And because the properties of chaotic functions are maintained through linear combination with linear functions, the sequence created by combining two sequences still has chaotic properties.

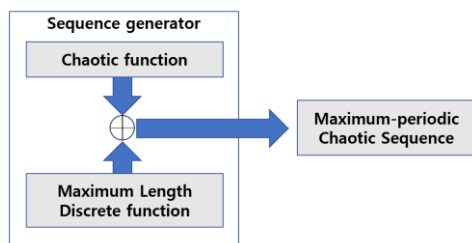


Fig. 1. The proposed sequence generator that produces a maximum periodic chaotic sequence

2.2 The process of the image encryption

The proposed image encryption based on the chaotic sequence generator is the process of generating a key image using the generated sequence and XORing it to the original image. The decryption method is to obtain the original image by XORing the key image with the encrypted image. **Fig. 2** is a result of the image encryption process.

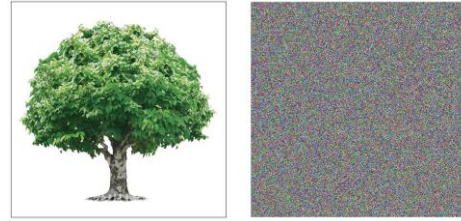


Fig. 2. A result of the image encryption

3. Conclusions

A variety of research on image encryption is being conducted to prevent image-related security issues that frequently occur in online services such as social networking services. And recent image encryption technology is designed based on chaos functions, but there is a risk that chaos function-based technology will have a small cycle due to the round-off problem. To solve this problem, this paper proposed a new sequence generator that generates a sequence with a period of at least $2^{256}-1$, and showed that the results of image encryption.

Acknowledgement

This research was supported by the Republic of Korea's MSIT (Ministry of Science and ICT), under the 2022 technology commercialization capability enhancement project (2022-BS-RD-0034) supervised by the INNOPOLIS.

References

- [1] Y. Pourasad, R. Ranjbarzadeh, A. Mardani, "A New Algorithm for Digital Image Encryption Based on Chaos Theory," *Entropy*, Vol 23, No 3, pp. 341-355, 2021.
- [2] W. Alexan, M. Elkandoz, M. Mashaly, E. Azab, A. Aboshousha, "Color Image Encryption Through Chaos and KAA Map," *IEEE Access*, Vol 11, pp. 11541-11554, 2023.
- [3] S. Zhou, X. Wang, Y. Zhang, "Novel image encryption scheme based on chaotic signals with finite-precision error," *Information Sciences*, Vol 621, pp. 782-798, 2023.

Privacy-preserving credential smart contracts using Zokrates

Geunyoung Kim, Hongseok Seo and Jaecheol Ryou

Department of Computer Science, Chungnam National University
Daejeon, South Korea

[e-mail: gykim@cnu.ac.kr, kei3824@naver.com, jcryou@home.cnu.ac.kr]

*Corresponding author: Jaecheol Ryou

Abstract

The need for secure user authentication in blockchain-based applications has been growing with the increased adoption of Decentralized Identity (DID) credentials in blockchain. In this study, we propose a smart contract system that securely validates DID credentials to overcome the limitations of traditional centralized authentication systems. This system ensures the safe identification of users within blockchain-based applications by authenticating their identities in a trusted manner within the blockchain. As the demand for user authentication in blockchain rises, this paper emphasizes the significance of a blockchain-based identity verification system that guarantees both privacy and security. Leveraging the Zero-Knowledge Proof method and utilizing the Zokrates tool, this innovative approach aims to provide solutions for the digital identity verification process, thereby expanding the scope of blockchain technology applications.

Keywords: Blockchain, Privacy-preserving, Credential, Zero-knowledge proof

1. Introduction

The emergence of Bitcoin[1] in 2009 ushered in the era of blockchain technology, followed by Ethereum[2] in 2014, which expanded the scope of smart contracts. Among them, NFT(Non Fungible Token) have come to play a particularly important role, and the concept of NFTs was first introduced in early 2017 through CryptoPunks, a variant of ERC-20[3], and Crypto Kitties, which utilise ERC-721[4], became so popular that they paralysed the Ethereum network. Meanwhile, virtual asset exchanges are implementing Know Your Customer(KYC) procedures for anti-money laundering(AML) and countering the financing of terrorism(CFT) at home and abroad. Since KYC is a mandatory procedure under the Act on the Reporting and Use of Specified

Financial Transaction Information, the Financial Action Task Force (FATF) has recently recommended that it is important to verify the identity of users not only in virtual asset exchanges but also in Web3 fields such as DeFi (Decentralized Finance) and NFTs due to their decentralized nature[5]. However, there is a trend. In the blockchain ecosystem, there is a disadvantage that it is difficult to verify identity due to the anonymous nature of wallet addresses. That's why we've recently been using DID to authenticate users, but we need to use off-chain to verify the DID. Therefore, even blockchain has the disadvantage of requiring a centralised authority. In this paper, we propose a method for on-chain verification of DIDs, which can guarantee user privacy and can be extended in various ways by authenticating users on Web3.

This work was supported by Institute for Information & communications Technology Planning & Evaluation(IITP) grant funded by the Korea government (MSIT)(No.2022-0-01200, Convergence security core talent training business(Chungnam National University))

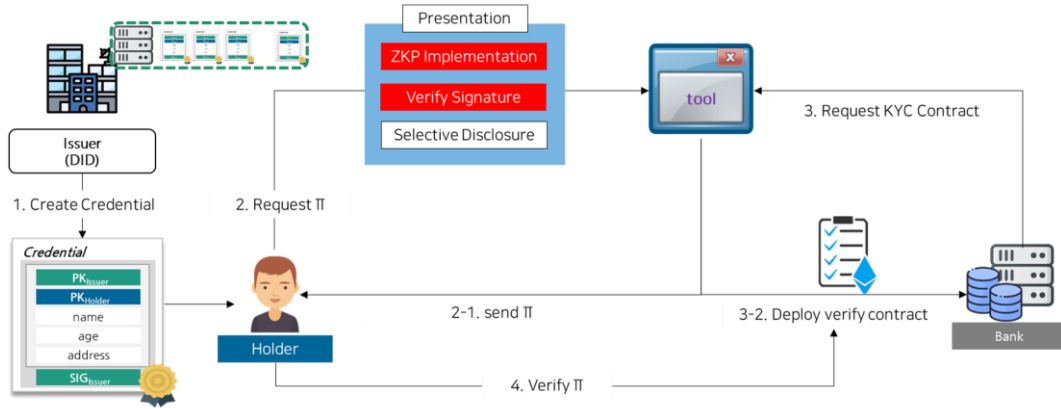


Fig. 1. Privacy-preserving model

2. Privacy-preserving Credential Model

Using Zokrates, the on-chain Privacy-preserving Model is depicted in Fig. 1.

1. User receives a credential generated by the DID Issuer
2. User generates Proof(π) based on Credential using zokrates tools.
3. Create a smart contract capable of verifying the Proof(π).
4. Verify the Proof(π) created through the smart contract.
5. Simultaneously with the verification of Proof(π), the user is verified by using the service or issuing additional NFTs.

2.1. Generating Proof(π)

Generate a proof by selecting only the information desired from the user's presentation. Initially, employ a range proof to verify the Credential Number, age to confirm that the age is over 19 for specific adult verification, and the signature of the Issuer who generated the Credential in Fig. 2.

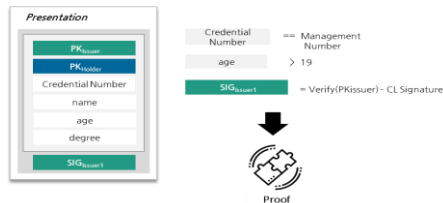


Fig. 2. Privacy-preserving model

Zokrates will be used to generate proofs that can guarantee these contents. In this case, we use CL Signature[6] to utilise Selective Disclosure. CL Signature uses a discrete algebra problem to mask the user's value. By using this, the value that the user wants to hide is raised to an index, and it can be verified like Fig. 3.

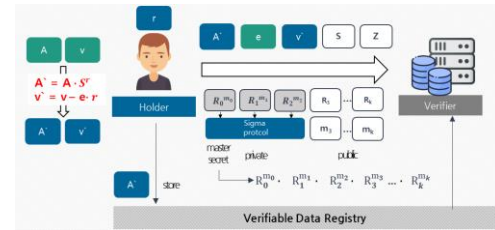


Fig. 3. Privacy-preserving model

2.2. Generating verification contract

To verify the signature of the Issuer that generated the Credential, Zokrates plugin will be used to generate a proof ensuring this. In Algorithm 1, the verifyProof function within the verify contract procedure authenticates the user's presentation proof. This function is invoked by supplying the proof generated, along with the issuer and user's DID present in the presentation, and the credential number as inputs. After the completion of this function, the user's identity has been authenticated, and various operations can be performed within the smart contract.

Algorithm 1. verifyProof

Input : calldata proof, calldata input, calldata issuer, calldata user, uint numOfCred

```

If (verify(proof, input) == true)
    .. // service start
    ▷ verify presenta-tion proof using Zokrates
    ▷ if proof & input is ture, start the web3

```

For instance, after the verification of proof as outlined in Algorithm 2, one can validate the authenticated identity in Web3 by invoking a function that generates an NFT.

Algorithm 2. createNFT()

Input : string tokenurl, address user

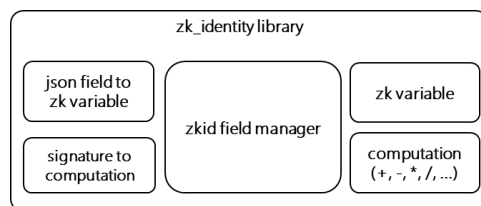
```

tokenIds.increment();
_mint(user , newItemId);
▷ Issues to the user, including the NFT number
_setTokenURI(newItemId , tokenURI) ;
▷ Setting up NFT and Token URI

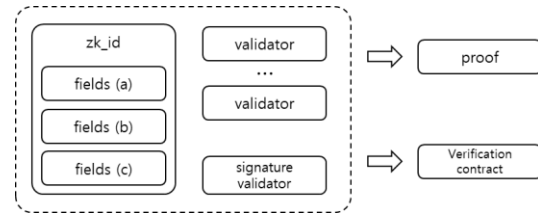
```

3.3. Usage of the ZK-Identity Tool

To generate proofs and contracts, we need a tool that takes credentials as input and generates proofs and contracts. This tool uses the Zokrates library, and the component by component is shown in Fig. 4.

**Fig. 4.** zk_identity library

It is responsible for parsing JSON to create zkvariables and managing the created files. In addition, there are components that not only create signatures as proofs, but also compute using various range proofs, etc. The tool that plays the above role is shown in Fig. 8. It receives the credential and goes through the parsing process, delivers the proof to the user who executed the programme, and delivers the contract to the verifier. The resulting appearance is depicted in Fig. 5.

**Fig. 5.** zk_identity process**4. Conclusion**

The advantage of our proposed tool is that it can generate zero-knowledge proofs from DID credentials, which Zokrates cannot handle. The tool also has a zero-knowledge proof library tailored to ZK-SNARK, which parses the user's credentials to generate proofs and contracts. However, similar to other research, this tool does not represent a fully decentralized system. In this schema, user identity authentication relies on utilizing DID Credentials, but it requires the assistance of a trusted Issuer when creating DID Credentials. Secondly, a drawback is that the current smart contract operates only on the Ethereum Virtual Machine (EVM), limiting its support to the Ethereum blockchain and not allowing for compatibility with various blockchain platforms. Therefore, we aim to explore the web3 ecosystem where this tool can be applied and set the development of DApps as a future goal.

References

- [1] NAKAMOTO, Satoshi; BITCOIN, A. A peer-to-peer electronic cash system. Bitcoin.—URL: <https://bitcoin.org/bitcoin.pdf>, 2008, 4.
- [2] BUTERIN, Vitalik, et al. Ethereum white paper. GitHub repository, 2013, 1: 22-23.
- [3] ERC-20, <https://github.com/ethereum/EIPs/blob/master/EIPS/eip-20.md>
- [4] ERC-721, <https://github.com/ethereum/EIPs/blob/master/EIPS/eip-721.md>
- [5] FATF, Virtual Assets and Virtual Asset service providers, 2021
- [6] Camenisch, J.L., Lysyanskaya, A.: Signature schemes and anonymous credentials from bilinear maps. In: Franklin, M. (ed.) CRYPTO 2004. LNCS, vol. 3152, pp. 56–72. Springer, Heidelberg (2004)

Digital Vulnerabilities in QR-Based Photo Access: An Analysis and Countermeasures for Self-Photo Studio Systems

Hong Seok Seo and Jina Hong and Jaecheol Ryu

Chungnam National University

Deajeon ,South Korea

[e-mail: kei3824@naver.com, jina05@naver.com, jcryou00@gmail.com]

*Corresponding author: Jae-Cheol Ryou

Abstract

The recent evolution of the self-photo studio industry has been marked by the advent of digital photo access services via QR codes. This study critically addresses multiple security vulnerabilities inherent to such services. Notably, many self-photo studio systems fail to implement web-based URL encryption. This oversight, coupled with insufficient session management and inadequate protection against path traversal, allows unauthorized users to access not only others' photos but also other files stored on the server, simply by manipulating the URL. These web vulnerabilities significantly elevate the risk of personal data breaches, underscoring the urgent need for web security enhancements. In this paper, we provide an in-depth analysis of the technical characteristics of these web vulnerabilities and their potential implications. Additionally, we explore and propose security countermeasures to ensure the safekeeping of user data within self-photo studio services.

Keywords: self-photo studio, URL encryption, Web vulnerabilities, Web security enhancements

1. Introduction

The advent of the digital era has brought revolutionary changes to the photography industry, with self-photo studios at the forefront of this transformation. Emerging as a new phenomenon in urban culture, self-photo studios have established themselves as a hub for social interaction and self-expression among the MZ[1] generation[2]. These studios redefine the concept of traditional photo studios by offering modern services that allow customers to take their own photos and access digital copies via QR codes.

The rapid growth of this industry has also had a clear impact on the startup market, with sales in

2022 increasing by 271% compared to the previous year[3]. However, beneath this growth and popularity lies the serious issue of security vulnerabilities in web-based services. The expanding market and increasing security threats amplify concerns over the protection of users' personal data.

This study focuses on analyzing the web security vulnerabilities of self-photo studio services. Many self-photo studio systems overlook basic security measures[4], including web-based URL encryption, which, combined with insufficient session management and inadequate protection against path traversal, opens the door for

| Self-studio | Static QR Code Link | Insufficient QR Code Authentication | Predictable URL Patterns | Inadequate Security (CSP) | Content Policy |
|-------------|---------------------|-------------------------------------|--------------------------|---------------------------|----------------|
| Studio A | ✓ | ✓ | ✓ | ✓ | |
| Studio B | ✓ | ✓ | X | X | |
| Studio C | ✓ | ✓ | X | X | |
| Studio D | ✓ | ✓ | ✓ | X | |
| Studio E | ✓ | ✓ | X | X | |
| Studio F | ✓ | ✓ | ✓ | X | |
| Studio G | ✓ | ✓ | X | ✓ | |

Table 1. Self-Studio Photo Service Security Vulnerabilities

unauthorized users to manipulate URLs and access not only other people's photos but also other files stored on the server³. These web vulnerabilities significantly increase the risk of

personal data breaches, underscoring the urgent need for enhancements in web security. Therefore, this paper provides an in-depth analysis of the technical characteristics of these web vulnerabilities and their potential implications, and explores and proposes security countermeasures to ensure the safekeeping of user data within self-photo studio services.

2. Background

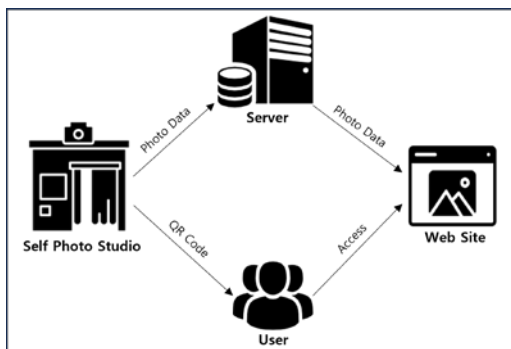


Fig. 1. The Process of Photo Shooting and Downloading in a Self-Photo Studio

2.1 Self-Service Photo Studio: Operational Workflow

The basic operation mechanism of a self-service photo studio is as illustrated in [Fig. 1](#). It is based on a system where users independently conduct photo shoots in the studio and immediately acquire digital results. This process consists of

the following steps:

Photo Shooting: Users take photos using the photo booth installed inside the self-service photo studio.

Photo Upload: The photos taken are automatically uploaded to the photo studio's server immediately after the shooting ends. This server typically uses cloud-based storage to manage and store photo data.

QR Code Generation: Once the photo upload is complete, the system generates a QR code that can be used to access the photos. This QR code is either printed directly at the photo studio or displayed on the photo booth's screen.

Photo Download: Users scan the QR code with a smartphone or other digital device camera and access the web page where the photos are stored via a web browser. This web page provides an interface for users to view, select, and download their photos.

3. QR Code Security Flaws in Photo Services

In [Table 1](#), an exhaustive analysis of web security vulnerabilities within self-photography studio services is presented, identifying four critical areas of concern: Static QR Code Linking, Insufficient QR Code Authentication, Predictable URL Patterns, and the Lack of Implementation of Content Security Policy (CSP). This comprehensive analysis, conducted using the Burp Suite, a leading web security testing tool, spanned seven web services, from Company A to Company G. The study revealed that the static linking of QR codes inherently

permits once-generated QR codes to remain indefinitely valid and reusable, thus posing a significant security risk. This is particularly concerning when QR codes are shared via social media, enabling unauthorized access to photographs.

Furthermore, the research highlighted the absence of additional authentication processes for photo access via QR codes in self-photography studio services. This security oversight allows anyone to access photos during the designated 2-3 day storage period, thereby posing a substantial threat to the protection of personal information. In terms of URL predictability, Companies A and E exhibited predictable patterns in URLs generated through QR codes. Company A's failure to encrypt URL parameters facilitated attackers in easily guessing links to other photo files. In contrast, Company E's application of Base64 encoding was found to be easily circumventable using basic decoding tools, thus exposing a security vulnerability. This aspect of security vulnerability aligns with the concerns raised by Samuel and Erlingsson regarding the risks of data confusion and the importance of proper data parsing in preventing security breaches[5].

Lastly, the study identified that Companies A and G had not implemented the Content Security Policy (CSP) header. CSP serves as a vital web security standard that restricts the sources of content executable on a web page, effectively preventing cross-site scripting (XSS) attacks and data injection attacks. The absence of CSP implementation heightens susceptibility to XSS attacks and undermines the overall security integrity of the web application.

4. Conclusions

This study significantly enhances the understanding of web security vulnerabilities in self-photo studio services, a new phenomenon in the digital era. The findings reveal critical vulnerabilities such as static QR code linking, lack of authentication procedures, URL predictability, and the non-implementation of Content Security Policy (CSP), clearly illustrating the serious threats these issues pose to user privacy and data security.

Beyond merely identifying problems, this research offers concrete solutions and improvement strategies to strengthen the security of self-photo studio services and protect user data. It is expected to aid service providers in recognizing and addressing these security vulnerabilities. Additionally, it contributes to raising awareness among users about safer service usage practices to protect their personal information. Therefore, this study is anticipated to play a significant role in achieving two crucial objectives: enhancing the security of self-photo studio services and safeguarding user privacy.

References

- [1] dic.hankyung, "Generation MZ", dic.hankyung.com, <https://dic.hankyung.com/economy/view/?seq=14683>, (accessed October 5, 2023)
- [2] Na, Geon-woong. "Unmanned Operation 'Attractive'... Market Saturation 'Concern'." Mael Business Newspaper Economy, <https://www.mk.co.kr/economy/view.php?sc=50000001&year=2023&no=173060> (Accessed November 11, 2023.)
- [3] Park Sungdae. KB Kookmin Card reports "Last year's growth industries: unmanned photo studios and craft beer pubs." Straight News. Retrieved from <https://www.straightnews.co.kr/news/articleView.html?idxno=224373> (Accessed January 26, 2023)
- [4] Hong Seo-yoon; Kim Seung-in. A Study on the Content Expandability of Generation Z Self-Photo Studios. 2023.
- [5] SAMUEL, Mike; ERLINGSSON, Úlfar. Let's Parse to Prevent Pwnage. In: 5th USENIX Workshop on Large-Scale Exploits and Emergent Threats (LEET 12). 2012.

Comparative Analysis of Blockchain-Based Voting Systems Using Machine Learning Techniques

**Uzma Jafar, Mohd Juzaidin Ab Aziz, Zarina Shukur, Hafiz Adnan Hussain,
and Zulkefli Mansor**

Faculty of Information Science & Technology, Universiti Kebangsaan Malaysia, Selangor, Malaysia
[e-mail: uzmajafar@gmail.com, juzaidin@ukm.edu.my, zarinashukur@ukm.edu.my,
hafizadnan87@gmail.com, kefflee@ukm.edu.my]

*Corresponding author: Uzma Jafar

Abstract

Integrating Blockchain technology into electronic voting systems promises to enhance security, transparency, and efficiency in electoral processes. However, the performance and reliability of these systems vary significantly, which requires a comprehensive evaluation. This research conducts a comparative analysis of various Blockchain-based voting systems using machine learning techniques to assess their performance, security, and user-friendliness. Several Blockchain-based voting systems were meticulously selected, and machine learning algorithms facilitated the assessment of system performance metrics, security vulnerability detection, and user experience evaluation. The study's findings reveal significant system efficiency, scalability, and robustness differences, with distinct correlations between Blockchain architecture and overall system performance. This paper provides empirical insights into the capabilities and limitations of current Blockchain-based voting systems, highlighting the critical role of machine learning in enhancing system analysis. The results offer valuable guidance for developing more secure, scalable, and user-friendly voting systems, paving the way for their broader adoption in democratic processes. Furthermore, the study underscores the need for ongoing research and innovation in this rapidly evolving field, particularly in addressing the challenges of system scalability and security.

Keywords: Blockchain, Electronic Voting Systems, Machine Learning, Voting Security, System Scalability, Digital Democracy

1. Introduction

Blockchain technology has gained much attention in the digital age for its potential to revolutionize various sectors, especially electronic voting systems. Blockchain technology is known for its decentralization, immutability, and transparency. It was initially used to facilitate Bitcoin transactions but has

now expanded its influence to areas where trust and security are critical fields [1]. While not a new concept, electronic voting faces challenges such as electoral fraud and voter disenfranchisement, where Blockchain-based voting systems (BBVS) promise improvements through a secure, transparent, and anonymous mechanism [2]. However, the effectiveness of BBVS depends not only on Blockchain but also on its design and integration within the electoral

ecosystem. This study aims to use machine learning (ML) to conduct a comparative analysis of BBVS, examining aspects like efficiency, security, and user satisfaction and addressing the increasing interest in their use for official elections [3]. The research aims to enhance understanding of BBVS, offer insights into their practical applications, and inform stakeholders on optimizing these systems [4]. The study delves into how different Blockchain architectures affect voting system performance, considering technical elements such as consensus mechanisms and node distribution [5].

2. Related Work

Integrating blockchain technology into electronic voting systems has garnered widespread attention due to its potential for enhancing security and transparency. Initial research has highlighted the blockchain's ability to create a decentralized, immutable ledger, making it ideal for recording votes that are tamper-proof. This aspect is critical in the fight against electoral fraud and in upholding the integrity of the democratic process, as noted by researchers like [6]. Further exploration into this field has led to examining various blockchain implementations, including public, private, and consortium models, each offering distinct advantages and challenges for voting systems as outlined by [7]. However, these innovations have their complications. The security and privacy of BBVS remain prominent concerns, with studies revealing both the cryptographic strengths of blockchain and potential vulnerabilities, such as susceptibility to 51% attacks [8]. The balance between voter anonymity and electoral transparency and auditability presents a complex dilemma, as investigated by [9]. The practicality of BBVS in large-scale elections has been a topic of debate, especially when considering scalability and efficiency issues. Technical challenges, particularly in transaction throughput and system latency, pose significant obstacles to their widespread adoption, as noted by [10]. Innovations like layer-two protocols and alternative consensus mechanisms have been proposed to mitigate these challenges, with varying levels of success [10]. Machine learning (ML) has emerged as a valuable tool for improving BBVS in recent years. Techniques

such as anomaly detection and predictive modeling are used to analyze blockchain data, identify patterns, and foresee system behavior under different conditions, an approach that has been instrumental in detecting fraudulent voting activities [11]. Despite extensive research on various facets of BBVS, there needs to be more comprehensive comparative studies that employ ML to evaluate these systems across performance, security, and usability. This research aims to fill this gap, providing new perspectives on the efficacy of BBVS and showcasing the role of ML as a fundamental tool for their analysis and enhancement.

3. Methodology

Our methodology uses machine learning techniques to analyze Blockchain-based voting systems (BBVS) comprehensively.

3.1 Security Evaluation Framework for Blockchain-Based Voting Systems

To ensure the robust security of BBVS, our study introduces a comprehensive security evaluation framework. This framework addresses 51% of attacks, node security vulnerabilities, and consensus mechanism weaknesses. It incorporates specific assessment criteria, including network decentralization, mining power distribution, and cybersecurity standards compliance. We also implement quantitative evaluation metrics, such as the frequency of security breaches and network resilience, supported by rigorous testing procedures like penetration testing and smart contract auditing. This security evaluation system not only assesses the current security status of BBVS but also provides recommendations for enhancements and ensures adaptability to evolving threats. By regularly updating our evaluation criteria and metrics, we maintain the relevance and effectiveness of our security analysis in the dynamic Blockchain environment.

3.2 System Selection

The BBVS were meticulously chosen based on various criteria to ensure diversity and representation of different blockchain-based voting approaches. The selection process considered factors such as the architecture type, extent of use, and geographical distribution of

these systems. By doing so, we were able to identify and include BBVS that represent a variety of different contexts, from government elections to organizational voting. It ensures that our analysis and research of these systems are comprehensive and truly representative of the blockchain-based voting landscape.

Table 1. Categorization of Selected Blockchain-Based Voting Systems

| BBVS | BC Type | Scale of Use | Geographic Distribution |
|-----------------------|------------|-----------------------|-------------------------|
| Estonia's e-Residency | Public | National | Estonia |
| Voatz | Hybrid | State/Local | USA, Various Locations |
| Agora | Consortium | International/Local | Multiple Countries |
| Follow My Vote | Public | Experimental/Local | USA, Various Locations |
| Polys | Public | Corporate/Small Scale | International |

3.3 Data Collection

Data was collected using a multi-faceted approach. Performance metrics were sourced from system logs and operational reports. Security incident reports were gathered from official documentation and third-party audits as shown in [Table 2](#). User feedback was solicited through surveys and interviews to gauge user-friendliness and satisfaction as shown in [Table 3](#). This qualitative data complemented the quantitative data.

Table 2. Performance Comparison for the selected BBVS

| BBVS | Transaction Speed (tps) | Block Time | System Latency (sec) | Security Incidents (last year) |
|-----------------------|-------------------------|------------|----------------------|--------------------------------|
| Estonia's e-Residency | 1500 | 2.0 | 0.5 | 2 |
| Voatz | 700 | 3.0 | 1.0 | 5 |
| Agora | 1000 | 2.5 | 0.8 | 3 |
| Follow My Vote | 600 | 4.0 | 1.2 | 4 |
| Polys | 800 | 3.5 | 0.9 | 2 |

Table 3. User Feedback and Expert Opinions for the Selected BBVS

| BBVS | User Feedback | Expert Opinions |
|-----------------------|--|--|
| Estonia's e-Residency | High satisfaction; reliable and secure. | Pioneering national-scale model. |
| Voatz | Mixed reviews; concerns over usability. | Innovative hybrid approach; security concerns. |
| Agora | Positive; transparent and international. | Consortium model raises questions on decentralization. |
| Follow My Vote | Interest in experimental stages; usability issues. | Potential future model; needs real-world testing. |
| Polys | Favorable for corporate use; easy to use. | Flexible application; scalability concerns. |

3.4 Machine Learning Approaches

The machine learning analysis framework in our study on BBVS began with comprehensive data preprocessing, utilizing normalization and transformation to standardize diverse datasets. We employed a range of machine learning algorithms due to their relevance in assessing BBVS. Classification methods like decision trees and random forests were pivotal for categorizing systems based on performance and security, classifying voting patterns, and detecting anomalies. It enhanced the reliability and security of BBVS. Support Vector Machines (SVM) and neural networks played a crucial role in analyzing complex patterns and predicting voting behaviors, contributing significantly to evaluating security and performance. Clustering algorithms like K-Means and anomaly detection using Isolation Forest and One-Class SVM provided insights into user behaviors and identified security outliers, aiding in the assessment of user-friendliness. Principal Component Analysis (PCA) was also utilized to reduce data complexity, facilitating the identification of key variables impacting BBVS performance. The analysis employed Python's sci-kit-learn and R's caret package, aiming for an exhaustive comparative evaluation and providing robust, multi-faceted insights. Each method brought a unique strength to our analysis,

collectively ensuring a comprehensive evaluation of BBVS in terms of performance, security, and user-friendliness.

4. Implementation

The study involved a systematic analysis of various BBVS using ML. To select the BBVS, the researchers conducted literature reviews and industry reports and finalized a diverse selection based on the blockchain architecture, deployment scale, and geography. Data collection entailed sourcing performance metrics from system logs or reports, gathering security data from documentation and audits, and obtaining user feedback via surveys and interviews. The researchers used Python for data processing and ML model implementation, R for statistical analysis, SQL databases, and Excel for data management. ML models were developed and trained to classify, cluster, and detect anomalies in the data. The results were interpreted to identify scalability, security, and user satisfaction patterns. This phase was crucial for translating methodology into practical steps, enabling comprehensive data collection and analysis, and providing a nuanced understanding of BBVS through ML.

5. Results

Machine learning analysis of BBVS revealed significant performance variations. Private blockchain architectures generally outperformed public ones in transaction speed. Scalability assessments showed differing capacities to handle peak voting volumes. Security evaluations highlighted vulnerabilities, particularly in consensus mechanisms, and underscored the importance of regular security updates. User feedback favored systems with simpler interfaces, directly impacting voter satisfaction, while accessibility varied across systems. Comparative analysis linked blockchain architecture to performance and security, highlighting its critical role in BBVS design. These results emphasize the necessity for continuous monitoring and adaptation in BBVS deployment to ensure effectiveness and security, focusing on the architecture's impact.

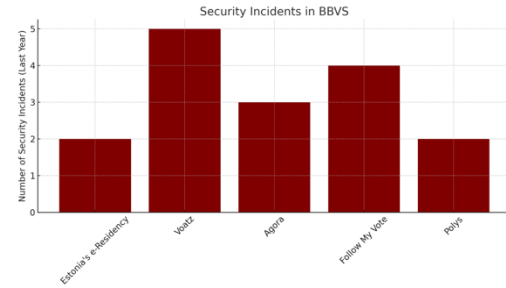


Fig. 1. Security Incidents in BBVS

The above **Fig. 1** presents the number of security incidents reported last year for each BBVS. It visually compares the security incidents, highlighting which systems experienced more incidents. Voatz, for example, shows a higher number of incidents than others.

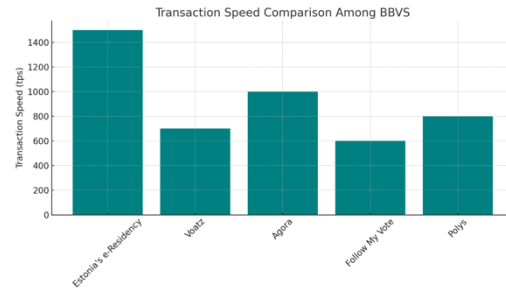


Fig. 2. Transaction speed comparison among BBVS

Fig. 2 illustrates the transaction speed (in transactions per second) for each BBVS. It clearly shows the differences in transaction speeds across the systems, with Estonia's e-Residency leading in speed.

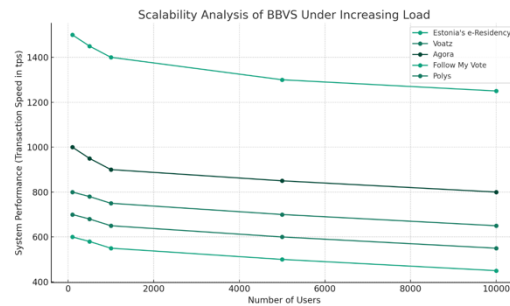


Fig. 3. Scalability analysis of BBVS under increasing load

In **Fig. 3**, scalability analysis of various BBVS is shown under increasing user load. The graph

shows how the performance of Estonia's e-Residency, Voatz, Agora, Follow My Vote, and Polys varies with the increase in user numbers, demonstrating their scalability to handle larger voting populations.

6. Discussion

The analysis of BBVS using ML techniques has yielded several critical insights, highlighting the nuances and complexities of integrating Blockchain technology into electoral processes.

6.1 Performance Variability Across BBVS

Our study revealed significant performance differences among various BBVS. Notably, systems with private Blockchain architectures generally offered higher transaction speeds than those with public Blockchains. This underscores an essential trade-off in BBVS design: the efficiency of private Blockchains versus the transparency and decentralization of public ones. Private Blockchains may be more suitable for large-scale electoral applications where efficiency is crucial. However, this choice impacts the democratic characteristics of voting systems, such as openness and voter trust.

6.2 Security Concerns in BBVS

The security analysis highlighted vulnerabilities in certain BBVS, especially those with weaker consensus mechanisms. The recurring security incidents in some systems emphasize the critical need for ongoing vigilance and regular updates to security protocols. This aspect is particularly crucial in voting systems, where integrity and trust are paramount. Our findings suggest that maintaining a secure voting environment requires a balance between strong security measures and the transparency necessary in electoral processes.

6.3 User Experience and Accessibility

User feedback pointed towards a preference for simpler, more intuitive interfaces in BBVS, impacting overall voter satisfaction. This highlights an often-overlooked aspect in the technical design of voting systems - user-friendliness. The disparity in accessibility and inclusivity across different systems calls for

more attention to ensuring that voting systems are readily usable by all eligible voters. This is not just a technical requirement but a fundamental aspect of ensuring democratic participation.

6.4 Impact of Blockchain Architecture

The comparative analysis underlined the significant role of blockchain architecture in determining the performance and security of BBVS. Public, private, or consortium architectural choices bring specific advantages and challenges. This finding indicates that the selection of blockchain architecture should be a deliberate decision based on the unique requirements of each electoral context.

7. Conclusion

Our investigation into BBVS via ML techniques has culminated in crucial insights, underscoring the profound impact of blockchain architecture on BBVS performance and security. This research highlights that while blockchain technology bolsters the integrity and transparency of the voting process, specific architectural choices are pivotal in determining a system's efficiency and robustness. Key findings reveal that system strengths and vulnerabilities vary widely across BBVS, emphasizing the need for ongoing development and refinement. Security, paramount in electronic voting, is not just a function of initial design but also requires continuous monitoring and updates. User experience, crucial for voter engagement and trust, largely hinges on interface design and accessibility. This study contributes significantly to the discourse on blockchain's role in democratizing voting processes, providing empirical data that can inform future system designs and policy decisions. The use of ML in analyzing complex data has been instrumental in revealing patterns and correlations that enhance our understanding of these systems. However, the study has limitations. The rapid evolution of blockchain and ML technologies means these findings need revisiting.

Additionally, the scope of the available data was a constraint, pointing to the need for more comprehensive data gathering and reporting in future BBVS implementations. Future research should include longitudinal studies to track the

long-term performance and viability of BBVS. Exploring the integration of advanced AI and ML algorithms could yield more profound insights into predictive analytics and anomaly detection, further securing and streamlining BBVS. Expanding the research to encompass a more comprehensive array of BBVS across different geopolitical landscapes would offer a global view of their effectiveness and scalability.

8. References

- [1] M. Conti, E. S. Kumar, C. Lal, and S. Ruj, "A survey on security and privacy issues of bitcoin," *IEEE communications surveys & tutorials*, vol. 20, no. 4, pp. 3416-3452, 2018.
- [2] F. Chaabane, J. Ktari, T. Frikha, and H. Hamam, "Low power blockchained e-vote platform for university environment," *Future Internet*, vol. 14, no. 9, p. 269, 2022.
- [3] P. Khanpara, S. Patel, and S. Valiveti, "Blockchain-based E-Voting Technology: Opportunities and Challenges," in *2022 7th International Conference on Communication and Electronics Systems (ICCES)*, 2022: IEEE, pp. 855-861.
- [4] Y. Liu, F. R. Yu, X. Li, H. Ji, and V. C. Leung, "Blockchain and machine learning for communications and networking systems," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 2, pp. 1392-1431, 2020.
- [5] S. Tanwar, N. Gupta, P. Kumar, and Y.-C. Hu, "Implementation of blockchain-based e-voting system," *Multimedia Tools and Applications*, pp. 1-32, 2023.
- [6] U. Jafar, M. J. Ab Aziz, Z. Shukur, and H. A. Hussain, "A Systematic Literature Review and Meta-Analysis on Scalable Blockchain-Based Electronic Voting Systems," *Sensors*, vol. 22, no. 19, p. 7585, 2022.
- [7] U. Jafar and M. J. A. Aziz, "A state of the art survey and research directions on blockchain based electronic voting system," in *Advances in Cyber Security: Second International Conference, ACeS 2020, Penang, Malaysia, December 8-9, 2020, Revised Selected Papers 2*, 2021: Springer, pp. 248-266.
- [8] U. Jafar, M. J. A. Aziz, and Z. Shukur, "Blockchain for electronic voting system—review and open research challenges," *Sensors*, vol. 21, no. 17, p. 5874, 2021.
- [9] U. Jafar, M. J. Ab Aziz, Z. Shukur, and H. A. Hussain, "A Cost-efficient and Scalable Framework for E-Voting System based on Ethereum Blockchain," in *2022 International Conference on Cyber Resilience (ICCR)*, 2022: IEEE, pp. 1-6.
- [10] H. A. Hussain, Z. Mansor, Z. Shukur, and U. Jafar, "Ether-IoT: A Realtime Lightweight and Scalable Blockchain-Enabled Cache Algorithm for IoT Access Control," *Computers, Materials & Continua*, vol. 75, no. 2, 2023.
- [11] J. El-Gburi, G. Srivastava, and S. Mohan, "Secure voting system for elections," *International Journal of Computer Aided Engineering and Technology*, vol. 16, no. 4, pp. 497-511, 2022.

Artifact Analysis of Android Device connected to Hardware Wallet

Jina Hong, Hongseok Seo and Jaecheol Ryu

Department of Computer Engineering, Chungnam University
Daejeon, South Korea

[e-mail: jina05@naver.com, kei3824@naver.com, jcryou00@gmail.com]

*Corresponding author: Jaecheol Ryu

Abstract

In cryptocurrency transactions, transactions are recorded in blockchain distributed ledger publicly but more and more cases are being used for crimes by exploiting the characteristic of anonymity that the subject of the transaction participants cannot be identified. By using hardware wallets with an Android device, an database file in the application data folder is created. Analysis of this artifact provides a way to obtain cryptocurrency-related information in case of obtaining an Android device from a criminal. In this regard, we have implemented a tool called CryptoInfoGetter that can be used in Windows10 environment to analyze and obtain cryptocurrency-related information.

Keywords: Forensics, Android, Hardware Wallet, Blockchain, Cryptocurrency

1. Introduction

Cryptocurrencies built on blockchain technology enable transactions through digital signatures, and the transaction details are recorded on a distributed ledger, visible to everyone. However, the participants in the transactions(sender and receiver) remain anonymous. This anonymity has been exploited for activities such as concealing or laundering criminal funds, as well as facilitating transactions related to drugs and child pornography. In domestic centralized exchanges, personal verification is mandatory, and they can assist with investigations when necessary, making tracking relatively straightforward. However, with hardware wallets, they do not require personal verification, which creates challenges for tracking. Additionally, hardware wallets are the storage of critical information, such as private keys and public keys, directly by the user, rather than third parties, making it likely for criminals to opt for cryptocurrency wallets.

To obtain cryptocurrency-related information from devices used to connect hardware wallets, it is necessary to analyze artifacts generated by each hardware wallet's application. In this research, we analyzed the artifacts generated on Android devices when hardware wallets, D'cent Biometric Wallet and Ledger Nano S, were connected and used. We used the results of this analysis to develop CryptoInfoGetter, a tool for obtaining cryptocurrency-related information.

In Section 2, we describe the environment and results of acquiring and analyzing artifacts generated after using an android device with hardware wallets. Section 3 outlines the methods of implementing CryptoInfoGetter using these artifacts and usage. In Section 4, we present the conclusion and outline future research tasks.

2. Aritifact Analysis

2.1 Analysis Environment

In this paper, we use the rooted Samsung Galaxy

S10 5G with Android 12 and two hardware wallets, namely the D'cent Biometric hardware wallet and Ledger Nano S. Subsequently, we utilized the Android-specific applications, D'cent Wallet for D'cent and Ledger Live for Ledger, to perform interactions. We conducted an analysis on the generated database files within the /data/data/<package_name> path, identifying meaningful data. Transactions were created using Bitcoin Testnet and Ethereum Testnet (ETH-GOERLI). The extraction and analysis of the database files were carried out in a Windows 10 environment using the Android Debug Bridge (ADB) tool.

Table 1. Full Specifications of the Devices Used in the Study

| | | |
|---------|----------------|----------------------------------|
| Android | Device Name | Galaxy S10 5G |
| | Version | Android 12 |
| D'cent | Device Name | D'cent Biometric hardware wallet |
| | Kernel Version | 2.25.2.83c3 |
| | KSM Version | 1.0.0.1139 |
| Ledger | Device Name | Ledger Nano S |
| | MCU Version | 2.1.0 |
| | SE Version | 1.12 |
| PC | OS | Windows 10 Pro |

Table 2. Tools and Applications used

| Software Name | Version | Usage |
|----------------------|---------|--|
| D'cent Wallet | 5.24.1 | Android Application for D'cent Hardware Wallet |
| Ledger Live | 3.20.1 | Android Application for Ledger Hardware Wallet |
| Android Debug Bridge | 33.0.3 | Android File Acquisition |

2.2 D'cent

The package name for the D'cent Wallet is com.kr.iotrust.dcent.wallet, and it stores data in the default.realm file within the files folder. Obtainable information from the default.realm file includes cryptocurrency wallet details, hardware wallet and pending transactions.

Among the cryptocurrency wallet details, the wallet label can be utilized to infer the purpose of use, while the wallet address can be employed to retrieve transaction information from the blockchain network. Regarding hardware wallet information, deducing the purpose of use is

possible by utilizing the hardware wallet label, and identifying the cold wallet owned by the user can be valuable for suspect investigations. Pending transaction details, accessible only from the Android device and mempool, serve as evidence of transactions generated on the specific Android device. This enables the confirmation of the time of transaction creation, facilitating the construction of a timeline.

2.3 Ledger

The package name for Ledger Live is com.ledger.live, and it stores data in the AsyncStorage file within the databases folder. It is a Key-Value database file and value is a json format. Meaningful information was obtained primarily from the Storage Table. The data includes details on each cryptocurrency wallet, hardware wallet, transactions, pending transactions, and the initial execution date of the application.

Specifically, information about each wallet is meticulously stored under the accounts.active.js: 2:<blockchain_network_name>:wallet_address key. The wallet label can be used to infer the purpose of use. Hardware wallet information can be valuable for suspect investigations by identifying cold wallets owned by the user. In contrast to D'cent, Ledger stores all transaction information in the database file, enabling the construction of a timeline using this data. Pending transaction details confirm transactions generated on the specific Android device, providing the actual time of transaction creation for timeline construction.

3. Implementation and Utilization of the tool

3.1 Implementation of the tool

Using the artifact analysis results in Section 2, we implemented a tool named CryptoInfoGetter to acquire cryptocurrency-related data. This tool was developed using C++ in Visual Studio 2019 and is compatible with the Windows operating system. To read the realm file containing D'cent's application data, we utilized the open-source Realm Core. Additionally, for reading the AsyncStorage file containing Ledger's application data, we employed the open-source SQLite3. To use CryptoInfoGetter, execute it in the

command prompt (cmd). If intending to retrieve information from D'cent, input the option -dcent; for Ledger information, input -ledger. Additionally, provide the path where the files from the Android device are copied.

```

Device Information
-----
Device Id : 4790050382116351580272230061930744470100000011591837000000000000
Device Name : D'CENT-ID-52098
Device Label : My D'CENT
Firmware Version : 2.25.2.83c3
KSM Version : 1.0.0.1139
Device Type : ble_dongle

Account Information
-----
Label : Jina
Address Path : m/44'/0'/0'
Coin Type : BITCOIN
Device Id : 4790050382116351580272230061930744470100000011591837000000000000
Address Balance
1EpPeKIKjKDBMqkd813E4yEund100TPmwU 0
1MXPBsSeZecKuJz08hc45vP3c1DyK1k8 0
1A0Qs1hdxab7z0hQ85ULfjje2Zuruxg 0
1MoB1bw7Rg8W7ASn6B8Xju5Jq81YusJD 0
18kqSatnJTKpGdTeVYvs2DsUYNNF1Hq3 0
13SmJsaF18V85ggy1sL36Qx785GazvPrNG 0
17b1TdMjJkHmPB3kLEyZgsqr16c42tALD 0
1Mf1Pj3pEck0twucbVSc7g9mTq7RHC3PTJ 0
1HoT0aermlPzrGjK84cQ7dK7G2nm6zNp 0
1Dexx0Nyth1gBNH6ALr31LmLVlWAra3QXG 0
1Gw0T0qgUJEzNRPv1oLDV42QW19g1Lyc 0
1JGvH85G4K66P21KQ71Mb6OmXV1L1UK9R 0
1CFMajv37D4CoZRRvfm1TCuyh2KPxH2LbV 0
1HcpQd2Rh1VP5cS4YPrv1MfZWz1dHFrSw 0
18PQ3qg1QceM6SRMzygT1vn568LZAsLxr4 0

```

Fig. 1. Execution Result of CryptoInfoGetter with Dcent option

```

Transaction4
Transaction Hash : 0x2c3e2bc90ceb6454917d1c703b1c402b8a1c75cc7013306ad917aac030e1fa4
Transaction Type : IN
Transaction Date : 2023-04-27T06:25:00.001Z
Senders : 0x0b980f5a39e5e79f45550e777b94ab18c99a
Recipients : 0x8cE701b2014e03e314d06823409ab089c888b85
Transaction Value : 10000000000000000
Transaction Fee : 10502000000000000

Transaction5
Transaction Hash : 0x08e372599af048cd4eac7f1217dac4281cf06b178cf77423865191ef91b8c52
Transaction Type : IN
Transaction Date : 2023-04-26T13:19:36.001Z
Senders : 0xA5700352420Efe7Fd1992d7810A0829af55b38a
Recipients : 0x8cE701b2014e03e314d06823409ab089c888b85
Transaction Value : 10000000000000000
Transaction Fee : 7770000000000000

Pending Transaction1
Transaction Hash : 0xcae478b1a8c4d352b019d3eadd3c3a177181f2aad0c2db8e36fb413894a5a8d
Transaction Type : OUT
Transaction Date : 2023-05-19T08:04:37.783Z
Senders : 0x8cE701b2014e03e314d06823409ab089c888b85
Recipients : 0xA5700352420Efe7Fd1992d7810A0829af55b38a

```

Fig. 2. Execution Result of CryptoInfoGetter with Ledger option

3.2 Utilization of the tool

Forensic investigators can execute CryptoInfoGetter in a Windows operating system environment, after copying all files from the Android device for analysis.

The acquired cryptocurrency data from CryptoInfoGetter can be cross-verified by querying the blockchain network for validity. This information can then be utilized to construct a crime timeline or serve as evidence. Particularly noteworthy is the ability to infer criminal intent using data not recorded on the blockchain network, such as wallet labels, pending transactions, and hardware wallet details. Investigations can also be conducted based on

the user's usage of hardware wallets, including information on the types and quantities used.

This tool thus proves valuable in forensic analysis, offering a means to gather, verify, and utilize cryptocurrency-related data in the context of criminal investigations.

4. Conclusions

In this study, we analyze database files from Android devices used with hardware wallets, the D'cent Biometric Hardware Wallet and Ledger Nano S. We described the information that can be obtained and the methods of utilizing these artifacts. Additionally, for Android devices utilizing D'cent and Ledger wallets in a Windows 10 environment, we implemented CryptoInfoGetter to extract details such as wallet addresses, balances, hardware wallet information, and transactions.

In the future, based on this research, we plan to investigate whether artifacts exist that provide data related to the deletion of D'cent wallet and Ledger Live applications. Furthermore, we aim to analyze data generated in memory during various scenarios while using both applications.

References

- [1] MinTaek Lim, JeoungYoon Kang, JunSung Park, MoonGyu Lee, HyeonDeok Jeong, JungTaek Seo, "A Study on the Extraction of Cryptocurrency Transaction Information Based on Memory Analysis", Journal of Digital Forensics, 16(1), pp. 99-117, 2022.
- [2] JungUn Ahn, SangJin Lee, "The analysis of bitcoin wallet artifact", Journal of Digital Forensics, 16(2), pp. 74-85, 2022.
- [3] Haight, T., Breiting, F., Baggili, I., "If I Had a Million Cryptos: Cryptowallet Application Analysis and A Trojan Proof-of-Concept", LNICST 259, pp. 45-65, 2019.

HPC Cluster-based Customized Container Image Manager and Builder

Gukhua Lee*

Supercomputing Infrastructure Center, Korea Institute of Science and Technology Information
Daejeon, South Korea
[e-mail: ghlee@kisti.re.kr]

*Corresponding author: Gukhua Lee

Abstract

This paper presents the implementation of a custom container image manager and builder in a High-Performance Computing(HPC) cluster-based environment. This development offers users the convenience of creating and saving their own personalized Artificial Intelligence(AI) service platforms within a container-based HPC service environment. To achieve this, we have created an image manager and builder. The image manager is capable of processing image requests from multiple users concurrently, distributing tasks to image builders running on the compute nodes. On the other hand, the image builder can receive queued tasks from the image manager, generate customized container images based on running instances, and store them in the private container registry. We have verified the functionality of this system by applying it to HPC cluster-based systems, including the supercomputer Nurion and GPU system Neuron. Additionally, we have established an interoperability architecture and mechanism to integrate seamlessly with container-based supercomputing environments.

Keywords: High-Performance Computing, Customized container image, Image manager, Interoperability, Supercomputing environment

1. Introduction

As the utilization of AI technology continues to grow in computational science, there is a rising demand for AI service platforms[1,2]. Concurrently, there is an increasing need for HPC hardware capable of supporting these service platforms. Specifically, within the HPC domain, where machine learning or deep learning is applied to models performing extensive calculations with substantial data sets, the provision of customized development environment services has transitioned from being optional to imperative.

Various organizations, including JSC (Jülich Supercomputing Center), CSCS (Swiss National Supercomputing Center), and NERSC (National Energy Research Scientific Computing Center),

provide Jupyter services on supercomputers, and Google offers the Colab service[3] to its users. The foundational infrastructure configurations for these services are container-based[4]. However, these configurations pose the inconvenience of necessitating users to install and configure essential packages and frameworks each time they access the service, executing it with an initialized container image[5].

In this paper, we address the aforementioned challenge by introducing an architecture for an HPC-based customized container image manager and builder. We propose an integration mechanism designed for HPC cluster-based systems and deliver services through a web-based portal named MyKSC (<https://my.ksc.re.kr>). This approach has been applied to the KISTI-5 supercomputer, known as

Nurion, and the GPU system Neuron.

2. Architecture

2.1 Image Manager

As shown in Fig. 1, the image manager enhances processing efficiency by incorporating a task queue. This allows simultaneous image-saving requests from multiple users to be handled through distributed and parallel processing in a distributed environment on the HPC cluster-based system. The image manager comprises a RESTful API (Application Programming Interface) implemented with Flask, responsible for processing requests from a web portal. Another component sends tasks to the image builder on the compute nodes through the Redis message broker. Additionally, the image manager features a function to record the image processing status in the database.

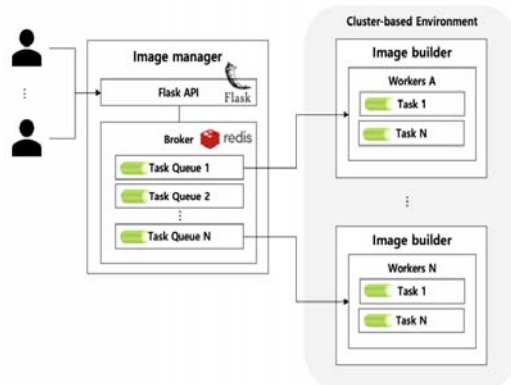


Fig. 1. Image manager and builder architecture

2.2 Image Builder

The image builder executes user requests received from the image manager through the Celery framework. The task involves committing a running container into a container image, transferring the created image to the container's private registry named MyKSC Hub (<https://myhub.ksc.re.kr>), and recording the image status in the database. Installed on each compute node, the image builder is configured to handle multiple tasks concurrently and can be dynamically expanded based on the resource utilization of the compute node.

3. Implementation

3.1 Integration Components

Within the AI services, every application runs as a Kubernetes-based service pod. We have created a specially designed image manager and builder for RStudio and Jupyter pods. The workflow's six components are used to construct the integration architecture using Kubernetes, as shown in Fig. 2.

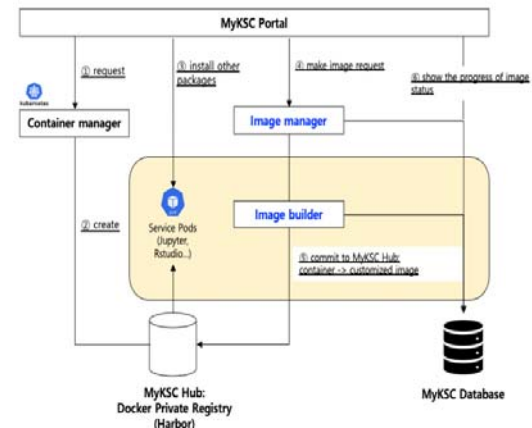


Fig. 2. Integration workflow

① Through the container manager (Kubernetes), users submit requests to create service applications (Jupyter and RStudio).

② Pre-configured YAML scripts are used by the container manager to deploy service application pods.

③ Using Jupyter or RStudio, users install AI model packages or libraries required for development straight into the service pods (containers).

④ Following installation, users use the MyKSC portal to submit a request to the image manager to save the environment that is currently running. The request will be sent in multiple formats via this portal in JSON data format. Using the parameters that are passed, the image manager sends requests to the image builder of a particular node.

⑤ After receiving commands from the image manager, the image builder uses the NERDCTL tool to commit the running container to the customized image, which is then uploaded to the MyKSC Hub.

⑥ Through the MyKSC portal, users can view and manage their customized images.

3.2 Functional Components

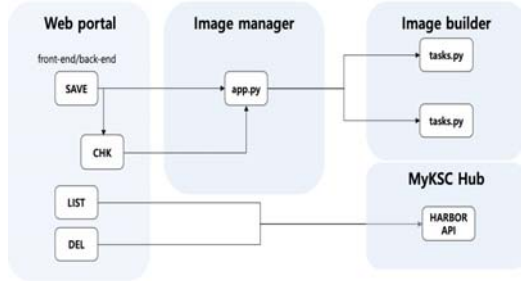


Fig. 3. Functional components

The web portal's back end has been integrated with four functions (shown in Fig. 3). The SAVE function makes it easier to send user data and the required parameter values in JSON data format to the image manager. The state of image processing is continuously observed by the CHK function. Depending on the resource type, images are divided into GPU and CPU categories. App.py contains the code for the SAVE and CHK functions, which make use of Flask, Celery, and Redis. Tasks.py defines the tasks for every builder in app.py. The Harbor API is used by the LIST and DEL functions to list users' images and make it possible to delete each user's image.

4. Conclusions

The custom container image manager and builder for HPC clusters is presented in this paper, enabling users to construct and store customized AI service platforms in a container-based environment. While the image builder creates personalized container images based on queued tasks and stores them in a private registry, the image manager effectively manages concurrent user requests by allocating tasks to image builders on compute nodes. The system's effectiveness and smooth integration potential with container-based supercomputing environments are highlighted by its successful validation on HPC cluster systems, such as the Nurion supercomputer and Neuron GPU system.

In the future, we hope to launch the service in a supercomputing environment and refine features in response to user input and operational experience.

References

- [1] Diogo R. Ferreira, and JET Contributors, "Using HPC Infrastructures for Deep Learning Applications in Fusion Research," *Plasma Physics and Controlled Fusion*, Vol 63, No. 8, 2021.
- [2] Estela Suarez, Norbert Eicker, Thomas Moschny, Simon Pickartz, Carsten Clauss, Valentin Plugaru, Andreas Herten, Kristel Michielsen, Thomas Lippert, "Modular Supercomputing Architecture," *White Paper of a Success Story of European R&D*, 2022.
- [3] I. Ali, A. Khan and M. Waleed, "A Google Colab Based Online Platform for Rapid Estimation of Real Blur in Single-Image Blind Deblurring," *2020 12th International Conference on Electronics, Computers and Artificial Intelligence (ECAI)*, Bucharest, Romania, pp. 1-6, 2020.
- [4] A. J. Younge, K. Pedretti, R. E. Grant and R. Brightwell, "A Tale of Two Systems: using Containers to Deploy HPC Applications on Supercomputers and Clouds," *2017 IEEE International Conference on Cloud Computing Technology and Science (CloudCom)*, Hong Kong, China, pp. 74-81, 2017.
- [5] J. Hursey, "Design Considerations for Building and Running Containerized MPI Applications," *2020 2nd International Workshop on Containers and New Orchestration Paradigms for Isolated Environments in HPC (CANOPIE-HPC)*, Atlanta, GA, USA, pp. 35-44, 2020.

Data-Driven Service Design Education Framework for New Product Development

Seohee Kim[†], Hyemin Park[†], Yu Cheng, Inseop Lee and Sanghun Sul*

Service Design Interdisciplinary Major, Sungkyunkwan University
Seoul, South Korea

[e-mail: kju6330@skku.edu, hyemin525@g.skku.edu,
chengyu91@skku.edu, lisszm@naver.com, sanghunsul@skku.edu]

*Corresponding author: Sanghun Sul [†]These authors contributed equally to this work

Abstract

In the big data era, companies emphasize innovation to improve their competitiveness. The new product development (NPD) process has a noteworthy effect on corporate innovation, with the stages of idea development and business analysis making significant contributions to enhancing innovation performance. Data-driven service design (DDSD) can positively impact the above stages through the collection and analysis of customer requirements, market opportunities, buying processes, and customer preferences. Despite the significant role of DDSD, only a few university courses concentrate on data in service design (SD), complicating the development of the data-related competencies demanded by modern enterprises. Consequently, data visualization, analysis, and obtaining insight from data in SD are persistent challenges in DDSD education. This study introduces a DDSD educational framework for enhancing innovation performance in the early stages of NPD. The framework is based on the double diamond process and comprises four phases: (1) Preparation, (2) Guidance, (3) Practice, and (4) Evaluation. The framework trains students with SD and data-driven approaches, including Business Intelligence. Its effectiveness was demonstrated by training 40 students with a home appliance brand. This led to enhancements in the students' ability to think critically based on SD, conduct data analysis, and generate insights.

Keywords: Data-driven service design education, Educational framework, New product development, Innovation performance, Service design, Business intelligence

1. Introduction

In today's high-tech economy, innovation is critical for a company's competitiveness, with new product development (NPD) playing a crucial role. NPD is instrumental in helping companies meet customer demands and gain a competitive edge [1]. Early NPD stages, idea development and business analysis, significantly impact innovation performance [1]. Data-driven service design (DDSD) positively influences these stages by collecting and analyzing customer needs, market potential, purchase processes, and preferences. While service design

(SD) education has grown, there is a global lack of emphasis on DDSD education.

This paper introduces a DDSD educational framework applicable in the initial stages of NPD. The framework's efficacy was validated through a study involving 40 students.

2. Related Works

2.1 NPD and Innovation Performance

NPD involves idea development, business analysis, product design, testing, and commercialization [1]. Meeting customer

expectations throughout all stages is crucial for the company's long-term survival [2], which is also core in SD. Analyzing NPD reveals that the initial stages, idea development and business analysis, impact innovation performance [1]. Previous studies emphasize big data's role in NPD [3], and data enhances innovation performance by understanding customer needs and optimizing processes [4]. These studies stress the significance of applying DDS in early NPD stages to boost innovation.

2.2 SD and DDS Education

SD creates socio-economic value through an interdisciplinary approach. SD education aims to cultivate design thinking, encompassing technical and non-technical skills. Current education curricula highlight the need for data processing skills [5]. Despite progress, challenges persist in DDS education, particularly in data visualization, analysis, and insight extraction, hindering optimal data utilization in the design process.

3. Framework

This paper introduces a DDS educational framework for early NPD stages. The framework consists of four phases: (1) Preparation, (2) Guidance, (3) Practice, and (4) Evaluation based on the double diamond process (Fig. 1).

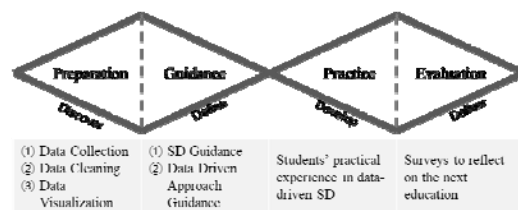


Fig. 1. DDS educational framework

The Preparation consists of three steps. First, educators select and gather data. Second, they clean the data using established rules. Lastly, they utilized business intelligence (BI) to visualize the data and facilitate better comprehension by students from diverse fields.

The Guidance consists of two steps. First, educators introduce the concept of SD and its

methodology. Second, they instruct data-driven approach and how to utilize the prepared BI.

The Practice involves team collaboration on topics, making solution deliverables. With proper training during the Guidance, students can effectively apply SD process and data.

During the Evaluation, students reply to a survey evaluating the education. It will be incorporated following education.

4. Experiment

An DDS training was conducted with a South Korean home appliance brand. 40 multidisciplinary undergraduate students formed eight teams. The brand's design theme was NPD in the washing and drying domain.

4.1 Preparation

Instagram posts were chosen as a data source. We selected 23 search keywords based on brand requests and gathered recent posts. We removed ad posts and those unrelated to home appliances. We eliminated meaningless hashtags and merged similar ones. After cleaning, the dataset had 1,775 posts with 12,363 hashtags. BI visualizes data using the number of hashtags, search keyword frequency, hashtag frequency, hashtag network, process mining, and page URL.

4.2 Guidance

We instructed our students on the importance of focusing on customer needs and introduced the double diamond process. We equipped them with three SD methodologies: persona, lotus blossom, and scenario. Additionally, we explained the prepared BI with feature examples.

4.3 Practice

We examined students' deliverables. All teams developed personas. Seven teams utilized the lotus blossom, with three in discover phase and four in develop phase. All teams wrote scenarios describing situation after the solution.

Teams relied on data, particularly during the discover and define phases. In the discover phase, they utilized various methods, such as analyzing hashtag network of prepared BI or checking market potential through reports. In the define

phase, they examined multiple posts and reports to define the needs and pain points. During the develop phase, they analyzed relevant products. Finally, in the deliver phase, they created a solution through prototyping, using reference materials such as papers and laundry sizes.

4.4 Evaluation

Data literacy, data use, and BI UEQ (User Experience Questionnaire) were examined through a survey. Each was measured on a 5-point Likert scale (1 = Strongly Disagree; 5 = Strongly Agree). Data literacy ($\alpha = .96$, $M = 3.84$, $SD = 0.75$) was measured by 20 items, which comprises data understanding, data analysis, data evaluation, data tool usage, data expression, and communication with data [6]. Data use ($\alpha = .90$, $M = 4.19$, $SD = 0.77$) referred to the use of data in Practice, which comprises four items: discover, define, develop, and deliver. BI UEQ ($\alpha = .94$, $M = 3.73$, $SD = 0.85$) was measured by a modified 12 items, which comprises attractiveness, efficiency, perspicuity, dependability, stimulation, and novelty [7]. The attractiveness is the parent concept among the others.

We conducted RM ANOVA with post-hoc paired t-test pairwise analysis on data literacy, data use, and BI UEQ for each dimension. In data literacy, results from the six-dimension comparison showed a significant difference in at least one dimension ($F = 4.84$, $p < .01$). Data tool usage was significantly less than data understanding, data evaluation, data expression, and communication with data (all $p < .03$), and data analysis was significantly less than data expression and communication with data (all $p < .04$). Other pairs were all non-significant. In data use, results from the four-dimension comparison showed no significant difference ($F = .70$, $p = .56$). In BI UEQ, results from the six-dimension comparison showed a significant difference in at least one dimension ($F = 3.10$, $p = .01$). Attractiveness and efficiency were significantly less than perspicuity, dependability, and stimulation (all $p < .05$), and novelty was significantly less than stimulation ($p < .01$). Other pairs were all non-significant. It seems low-efficiency resulted in low-attractiveness.

5. Conclusions

This paper presents a DDS educational framework designed to improve innovation performance in NPD, specifically regarding idea generation and business analysis. The framework consists of four phases: (1) Preparation, (2) Guidance, (3) Practice, and (4) Evaluation. It underwent validation with 40 undergraduate students. The students employed the SD methodologies and fully utilized the data in each double diamond phase. The students' low level of data tool usage and data analysis among data literacy demonstrates the appropriateness of providing BI. However, enhancing the BI efficiency is necessary for the next education.

References

- [1] N. Tzokas et al., "Navigating the New Product Development Process," *Industrial Marketing Management*, Vol 33, No. 7, pp. 619-26, 2004.
- [2] I. O. Ugboro, K. Obeng, "Top Management Leadership, Employee Empowerment, Job Satisfaction, and Customer Satisfaction in Tqm Organizations: An Empirical Study," *Journal of quality management*, Vol 5, No. 2, pp. 247-72, 2000.
- [3] A. McAfee et al., "Big Data: The Management Revolution," *Harvard business review*, Vol 90, No. 10, pp. 60-68, 2012.
- [4] G. George et al., "Big Data and Management," *Academy of management Journal*, Vol 57, No. 2, pp. 321-326, 2014.
- [5] B. Ferster, *Teaching machines: Learning from the intersection of education and technology*, JHU Press, 2014.
- [6] Y. Chung and J. You, "Validation of Data Literacy Scale for University Students and Analysis of Freshmen Data Literacy – The Case of A University," *Korean Journal of General Education*, Vol 16, No.5, pp. 245-260, 2022.
- [7] B. Laugwitz et al., "Construction and evaluation of a user experience questionnaire," in *Proc. of HCI and Usability for Education and Work: 4th Symposium of the Workgroup Human-Computer Interaction and Usability Engineering of the Austrian Computer Society*, 2008.

Derivation of Home Appliances Features with Stretchable Displays based on UX

Seung Eun Chung, Youjin Seo, and Han Young Ryoo*

Department of Content Convergence, Ewha Womans University,
Seoul, South Korea

[e-mail: ilovechse@ewha.ac.kr, youjinseo1@gmail.com]

*Corresponding author: Han Young Ryoo

Abstract

This research is to derive the key features on the products that satisfy the user's needs in the field of home appliances based on the commercialization characteristics of the stretchable display. Recently, products that can be decorated and reflect the individuals' preferences are being developed in the home appliances field. When the stretchable displays are applied, they will transform into freeform and be able to apply to various home appliances, furniture and objects; thus, UX can be provided innovatively. In this research, user expected experiences on the home interior have been investigated and analyzed. Scenarios have been derived with these results, that provide the user expected experience through the stretchable display technology features. Based on the scenarios, the home appliance features that fulfill the user's expectations, utilizing the high-resolution image display features as well as the complex three-dimensional deformation features of the stretchable display, have been defined and examples are provided.

Keywords: Home Appliances, Stretchable Display, Display of Things, User Experience

1. Introduction

Along with the fourth industrial revolution, the popularization of the ICT technologies brought advancement of the residence, and the development of the technologies that allow personalization caused the higher needs of the home appliances that can reflect the user's tastes as part of the interior design[1]. The recent appliances allow the users to modify the colors of the exterior panels to match with the surrounding interior designs. There is a growing recognition of the home appliances as one of the interior design factors as well as increasing demands of the products that can be customized. In this situation, stretchable displays are being in the

spotlight as the next display technologies for its elasticity and stretchability, because they can be transformed into freeform displays[2]. In 2022, LG Display succeeded to materialize the 12-inch stretchable display with 20% stretchability, featuring full-color RGB and a high resolution of 100ppi which is similar level of regular monitors[3]. The stretchable displays that are under development at the moment have full-color high resolution which allows to reproduce the vivid image, and due to the stretchability, they can cover the atypical surfaces. Therefore, they can be applied to various 3D-shaped everyday objects, and they are receiving attention as DOT, Display of Things. To commercialize the display that has the innovative form factor like this, researches on

This work was supported by the Technology Innovation Program (or Industrial Strategic Technology Development Program) (No.20010744, Design and development of stretchable display products for electronic and smart devices) funded By the Ministry of Trade, Industry & Energy(MOTIE, Korea).

the user interaction and interface techniques are being actively conducted[4][5], [6][7].



Fig. 1. Image of the Stretchable Display Prototype developed by LG Display in 2022

If stretchable displays are applied to home appliances, users stand to benefit from the ability to showcase vivid images and alter the display's shape. Consequently, this innovation can effectively meet users' diverse needs by not only being applicable to existing home appliances but also extending to various furniture and other objects within the home.

In this research, the applicability of the stretchable displays to atypical surfaces will be investigated. Also, innovative and effective products that can not only be applied to the home appliances but also to the surface of the objects that can be found in home space. In order to examine the direction of the commercialization of the stretchable display, the technology development status will be checked and the user expected experiences that need to be taken into consideration in the home appliances field will be derived. The objectives of this research is to develop user scenarios with the stretchable displays applied to the home appliances, and to propose innovative products.

2. User Expected Experience on Home Interior Products

2.1 User Survey Targets and Methods

In this section, the consumption and usage behaviors related to expressing preferences and decorating home were examined and the expected experiences were investigated. One on one interview was conducted where participants were asked on what objects, including the existing home appliances and furniture, they change and how in order to get a different look of home, followed by the in-depth questions as to what the reasons and needs are. The participants for the in-depth interviews were selected if their annual spending for home interior is more than

KRW5,000,000 and the number of purchases are more than four times a year. They were consisted of female and male in between 20s and 30s from single households to four-person households.

2.2 User Expected Experience Analysis

The analysis was executed through open coding by converting the recorded interviews into scripts. First, we have extracted the significant phrases that are related to the current behaviors and needs, and organized them into 53 types by grouping the similar phrases in meaning and removing the duplicates. Secondly, the linguistic data has been organized through the Affinity diagram that is used to have similar concepts by mutual affinity. A group of nine graduate students was asked to remove the phrases, out of the 53 types, that are deemed not to fall under the attributes of the users' value or expected experience as well as to combine the similar concepts with significant contents. Lastly, we asked the participants to group the ones that are thought to be related in meaning into higher categories. The users' expected experiences with home interior products have been derived into four factors, encompassing 17 detailed attributes. These factors include Environmental Aesthetics(5 attributes), Customization of Emotion(3 attributes), Presentation of Preferences(4 attributes), and Functionality of Space and Life Support(5 attributes).

Table 1. User Expected Experience factor derived from user research

| Factors | Definition |
|---|---|
| Environmental Aesthetics | Expect to have the design, materials, and colors of the product to harmonize with the surrounding environment |
| Customization of Emotion | Expect to create the desired atmosphere and experience a change in mood through product modifications |
| Presentation of Preferences | Expect the product to align with personal preferences and allow for differentiation from others |
| Functionality of Space and Life support | Expect functional utility that supports the characteristics of living spaces and usage scenarios |

3. Innovative Product Derivation with Stretchable Display Based on UX Scenario

3.1 Guideline for UX Scenario

The scenario-based design methodology in this research is used by the design team in order to

apprehend the user's needs and derive the appropriate concepts prior to the development of the products or services. In the UX scenario, how the user will use the products or services is anticipated in the detailed context, and based on this, further designs will be built up for the specific functions, information, and interactions [8]. The companies use the scenario techniques to develop future preceding products, and also derive the ideas for future products by combining the future variable factors that will affect the future innovative products and the user context expectations [9].

In this research, technical attributes of stretchable displays, such as the ability to display high-resolution images and 3D deformation, have been set as variables for application in near-future products. Scenario development guidelines have been organized to provide users with expected experiences derived beforehand. The organized detailed information is in the **Table 2** below.

Table 2. Guidelines to Offer User Expected Experience Based on Technical Features

| Features | Description |
|---------------------------------|---|
| Display High Resolution Image | <ul style="list-style-type: none"> - Colors and patterns in interior design enhance aesthetics and environmental harmony. Easily modifiable designs allow for emotional refreshment and atmosphere customization. Digitalized design data enables previewing, testing, and sharing to showcase the user's taste. - Visual information, conveyed through texts, graphics, or videos, enables smart service functions that enhance space functionality and support the user's needs. Additional smart functions consider the original product functions, arrangement, and usage behavior. |
| Display Physical 3D deformation | <ul style="list-style-type: none"> - Applicable to existing home appliances and furniture, the application doesn't require full product coverage but can be partial. Utilizing actuators, it enables dynamic shape changes based on situations, showcasing environmental aesthetics, emotional customization, and user preferences. - Physical form variations convey information, support functions, and enhance smart services for space functionality and life support. |

3.2 UX Scenario on Interior Home Appliances with Stretchable Display

Based on the above guidelines, scenario ideas were collected and expanded, and UX scenarios have been developed.

3.2.1 Idea Cards

In the context of using the home interior and home appliances, an ideation has been conducted to fulfill the expected experiences considering the form factor attributes of the stretchable display. Since the stretchable displays are

applicable to the physical objects due to its flexibility, furniture and interior materials were also included in the ideation process. To collect the ideas, below categories are established, and the participants of graduate students and undergraduate students were requested to write out idea cards.

Different from the way that the existing products could provide, ideas were not only needed to fulfill the user expected experience from the home interior, but also to expand the functions of the product itself.

3.2.2 UX Scenario

The ideas that are not appropriate for the proposed guidelines are removed, and the similar ones are combined or grouped together to streamline the main ideas. The applicable products and the usage situation are defined, and the gathered ideas are further developed within the context. The scenarios per product are consisted of five to six scenes. The detailed context of the home interior appliance usage and the flow are elaborated, and the needed features and the interactions are materialized. The finalized scenarios are presented with an explanation and representative images. The functions, interface, and interactions of the products that offer the experiences that user is expecting are described in each scene. The derived main scenarios are as follows.

Scenario #1 for Smart Interior Home Appliances

In this scenario, the user is able to prepare the meals using the built-in home appliances and to clean up after a home party with friends. The stretchable display has been applied to the control area of small appliances. The key features and prototype design are as in the table below.

Table 3. Key Features of Smart Interior Home Appliances

| | |
|---|--|
| Features to Display High Resolution Image | <ul style="list-style-type: none"> - Unified variations suitable for the surrounding interior color and patterns - Display of the online recipe or video within the display area of the built-in appliances - Reactive interface that can link with the device nearby and expand the display. |
| Features to Display Physical 3D deformation | <ul style="list-style-type: none"> - Seamless interior without any button when not in use of the home appliances - Calling the 3D function buttons by a simple touch - Deformation of the physical handle according to the situation |

Scenario #2 for Smart Interior Sofa

In this scenario, the user is able to change the sofa design that matches with the seasons as well as to control the home appliances, such as TV or air conditioning, while taking a rest on the sofa. The stretchable display has been partially applied to the armrest of the sofa. The key features and prototype design are as in the table below.

Table 4. Key Features of a Smart Interior Sofa

| | |
|---|--|
| Features to Display High Resolution Image | <ul style="list-style-type: none"> - Different options of the sofa armrest in shape, color, and pattern - Selection of the smart home devices, such as TV, Air conditioning, Air purifier, etc. and add more devices to link |
| Features to Display Physical 3D deformation | <ul style="list-style-type: none"> - Offering a multi-controller interface at the position of the hand on the armrest - Deformation of the engraved control buttons for softkey - Accurate button operations through the tactile senses of a hand - Changes of the shape, number, and functions of the controller buttons according to the linked home IoT devices |

6. Conclusions

Although the stretchable display techniques have been revealed, it is unclear how the techniques are required to apply on which products in order to be accepted in the market. Thus, this research proposed the concepts and features of home appliances that fulfill the consumer's potential needs through the form factor attributes of the stretchable display. Moreover, the home appliance field, expanded to the digital transformation of objects through the applicability of the display on the surface, has been suggested. In the future, a research and analysis on the influencing factors of the user's purchase, intentions and decisions to use will be conducted to verify the product validity through a technical acceptance survey on the products with the stretchable displays proposed in this research.

References

- [1] C. Hyun, and C. Park, Stretchable display apparatus and display control method and apparatus of the stretchable display apparatus (U.S. 9,804,699 B2). U.S. Patent and Trademark Office. 2017.10.31.
- [2] C. W. Kim, G. W. Kim, and G. G. Kang, "Practical measures of three-dimensional stretchability on stretchable displays," *Journal of Information Display*, pp. 1-9, 2023.
- [3] H. Jung, C.I. Park, M.B. Gee et al., "High-resolution active-matrix micro-LED stretchable displays," *Journal of the Society for Information Display*, Vol 31, No. 5, pp. 201-210, 2023.
- [4] C. Hyun, and C. Park, Stretchable display apparatus and display control method and apparatus of the stretchable display apparatus (U.S. 9,804,699 B2). U.S. Patent and Trademark Office. 2017.10.31.
- [5] J. Chi, and S. Eim, Stretchable display device and operating method thereof (U.S. 9,928,571 B2). U.S. Patent and Trademark Office. 2018.3.31.
- [6] J. Park, and J. Kim, Three-dimensional tactile display apparatus using stretchable light-emitting material and manufacturing method of thereof (KR 10-2088807). The Korean Intellectual Property Office. 2020.3.9.
- [7] H.Y. Ryoo, and S.E. Chung, Solid content output device and method using stretchable display (KR 10-2302498). The Korean Intellectual Property Office. 2021.9.9.
- [8] A. Cooper, R. Reimann, and D. Cronin, "About face3: the essentials of Interaction design," *NJ, USA: John Wiley & Sons*, 2014.
- [9] G. Kim, and W. Kim, "Future context scenarios for user-centered advanced design," *Archives of Design*, Vol 33, No. 1, pp. 93-110, 2020.

Analyzing the effects of alcohol-related image types on drinking efficacy and intentions on Instagram: An Eye-Tracking Approach

Min hee Lee¹, Kyung Han You¹, Du Mengqi¹, Yifan Wang¹, and Jeong Kyu Lee²

¹Department of Media & Communication Studies, Jeonbuk National University
Seoul, South Korea

[e-mail: minhee@jbnu.ac.kr, kyunghanyou@gmail.com, dumengqi08@gmail.com,
yf970616@gmail.com]

² Department of Public Health, University of Oklahoma, The United States
[e-mail : lee.jeongkyu@gmail.com]

*Corresponding author : Kyung Han You

Abstract

Influencer marketing's growing importance on social media, particularly its impact tied to follower count, is notable. The rise of alcohol-related content and ads on social platforms is linked to user drinking habits. Despite this, insufficient research exists on alcohol marketing on social media, with existing studies relying on surveys. This study used experimental methods and eye-tracking devices to measure visual attention on social media content. The aim was to understand the relationship between attention, alcohol efficacy, and purchase intention. Results revealed diverse attention levels to influencer alcohol ads, influencing alcohol expectations and purchase intentions.

Keywords: Influencer marketing, Six Wheel, Alcohol advertisements, Eye-Tracking, Alcohol intentions, Alcohol expectancies

1. Introduction

Influencer marketing is gaining importance on social media, particularly for connecting with younger demographics through authentic and captivating information. The visual nature of social media platforms, with images and videos, makes influencer posts a significant marketing strategy. Influencers, as sources of informal information, wield substantial social and cultural influence, classified by research into categories based on follower count. Social media marketing, combining commercial ads with user-generated content, enhances efficacy. Increased exposure to alcohol advertising correlates with more frequent alcohol use, prompting alcohol manufacturers to shift

resources to online platforms. While studies have emphasized the link between alcohol consumption trends and social media advertising, there is a scarcity of research investigating the visual perception's influence. This study aims to evaluate the effectiveness of influencer marketing, particularly in alcohol ads, using eye-tracking technology to analyze attention-grabbing elements.

2. Methods and Research Questions

2.1 Methods

Involving 62 participants (average age 25.5 years, 34 females, and 28 males), the study utilized the Tobii Pro Nano eye tracker and a 24-inch monitor. Participants, categorized into celebrities,

influencers, and the general public, viewed Instagram posts for 15 seconds, focusing on elements like follower count and model faces. Eye tracking measures, assessing total durations and fixations, were employed to analyze commercial attention. A follow-up survey explored participants' exposure to alcohol advertising on social media and drinking patterns, covering type, frequency, quantity, daily intake, monthly consumption, and binge drinks.



Fig. 1. Setting AOI

2.4 Research questions

Research Question 1: Which specific elements of alcohol advertisements on social media capture the subjects' attention?

Research Question 2: Does attention have an impact on drinking behavior or perceived efficacy?

Research Question 3: Does the influence of demographic variables (gender, country, kind of drinking) on drinking behavior or perceived efficacy vary? (Assess the presence of interaction effects)

3. Results

The study employed eye-tracking and post-survey queries to investigate elements in alcohol-related social media ads influencing individuals' intentions and expectations about alcohol. Tests compared posts featuring celebrities, influencers, and social groups. In eye movements, the Kruskal-Wallis test revealed significant differences in attention among celebrities, influencers, and the ge-

neral population regarding account information, inspicture content, ad messages, and faces. Celebrities focused more on account information, while the general population engaged more with inspicture content and faces. Influencers showed greater fixation on ad messages. The discussion on drinking intentions and alcohol expectancies highlighted significant differences among celebrity models, especially in 'be a better lover' items. Celebrity models exhibited a higher prevalence of alcohol consumption, notably being urged by a friend to drink ($p=0.043$). In differences by demographics, gender variations showed higher occurrences of feeling dizzy ($p=0.010$) and being clumsy ($p=0.001$) in women. No significant indicators were found in eye tracking. Nationality and drinking types revealed no notable differences in eye tracking or post-survey items. Interaction effects between gender and alcohol expectancies showed reduced sensations of dizziness in women with increased alcohol consumption, while men experienced heightened dizziness. A positive correlation between alcohol consumption and the desire to improve as a lover was observed in women, while men's 'be a better lover' decreased with increased alcohol consumption. No significant interaction effects were found for gender and alcohol type on other measures..

4. CONCLUSION

The study analyzed alcohol-related Instagram posts categorized by celebrities, influencers, and social accounts, exploring attention levels, Alcohol Expectancies, and drinking intentions. Findings revealed higher interest in account information of celebrities and influencers with substantial followers. Influencers paid more attention to ads but had no impact on perceived effectiveness or drinking habits. Celebrities influenced drinking effectiveness, notably in the item "I would be a better lover." Photos of gatherings consuming alcohol drew focus on facial features, but impact was unspecified. A correlation was found between gender, alcohol type, and effects like dizziness and perceived sexual prowess, with women showing reduced dizziness and enhanced perceived prowess, while men experienced the opposite.

References

- [1] L. Hudders, S. De Jans, and M. De Veirman, "The commercialization of social media stars: a literature review and conceptual framework on the strategic use of social media influencers," *International Journal of Advertising*, vol. 40, no. 3, pp. 327-375, 2021.
- [2] Seongjae Lee, Sungwoo Cho, & Soosun Cho, "A Study on the semantic information analysis and classification for SNS image," *한국정보처리학회 학술대회논문집*, vol. 19, no. 2, pp. 507-509, 2012.
- [3] A. P. Schouten, L. Janssen, and M. Verspaget, "Celebrity vs. Influencer endorsements in advertising: the role of identification, credibility, and Product-Endorser fit," *Leveraged Marketing Communications*, pp. 208-231: Routledge, 2021.
- [4] P. K. Salimian, R. Chunara, and E. R. Weitzman, "Averting the perfect storm: addressing youth substance use risk from social media use," *Pediatric annals*, vol. 43, no. 10, pp. e242-e247, 2014.
- [5] P. Anderson, A. De Bruijn, K. Angus *et al.*, "Impact of alcohol advertising and media exposure on adolescent alcohol use: a systematic review of longitudinal studies," *Alcohol & Alcoholism*, vol. 44, no. 3, pp. 229-243, 2009.
- [6] E. P. Cabrera-Nguyen, P. Cavazos-Rehg, M. Krauss *et al.*, "Young adults' exposure to alcohol-and marijuana-related content on Twitter," *Journal of studies on alcohol and drugs*, vol. 77, no. 2, pp. 349-353, 2016.
- [7] A. E. Barry, A. A. Padon, S. D. Whiteman *et al.*, "Alcohol advertising on social media: Examining the content of popular alcohol brands on Instagram," *Substance use & misuse*, vol. 53, no. 14, pp. 2413-2420, 2018.

Lightweight Deep Learning-based Product Object Classification Scheme

Dongchil Kim, Kyeongun Seo and Chang Mo Yang

Information Media Research Center, Korea Electronics Technology Institute (KETI)
Seoul, South Korea

[e-mail: dckim@keti.re.kr, ke.seo@keti.re.kr, cmyang@keti.re.kr]

*Corresponding author: Dongchil Kim

Abstract

This paper proposes a lightweight deep learning-based product object classification scheme to operate on the edge server. In order to develop a lightweight deep learning model, we optimized the hyper-parameters of the Vision Transformer (ViT) model. Also, the proposed model was trained, verified, and tested by selecting 6 classes from the ImageNet Large Scale Visual Recognition Challenge 2012 (ILSVRC2012). Through implementation results, the performance of the proposed lightweight deep learning model was confirmed in terms of lightweight efficiency, object classification speed and accuracy.

Keywords: Lightweight deep learning, Product object classification, Vision transformer, Mobile device, Edge server

1. Introduction

Recently, with the acceleration of changes in the non-face-to-face economic environment, consumption patterns are swiftly transitioning from traditional offline purchases to online transactions [1]. Reflecting these environmental changes, many applications are being developed that automatically classify product objects using Artificial Intelligence (AI) technology on mobile devices. However, as the types of mobile devices diversify, it is difficult to develop a product object classification scheme optimized for mobile device characteristics. Accordingly, there is a need to develop an edge server-based product object classification scheme that can classify product objects without being affected by mobile device characteristics. To classify objects regardless of the characteristics of the mobile device, a product object classification structure is

defined as shown in Fig. 1. The mobile device captures and verifies the video, transmitting it to the edge server. It receives images from the mobile device and analyzes the objects. Subsequently, the classified object information is transmitted to the mobile device.

In this paper, we propose a lightweight deep learning-based object classification scheme to classify products on the edge server in an environment where product objects are detected by integrating a mobile device and edge server as shown in Fig. 1. In order to develop a lightweight deep learning model, we optimized the hyper-parameters of the ViT model [2]. Also, the proposed model was trained, verified, and tested using ILSVRC2012 dataset [3].

The remainder of this paper is organized as follows: Section 2 describes product object classification scheme. In Section 3, we present

the implementation results. Finally, Section 4 concludes our scheme.



Fig. 1. Structure for product object classification based on the integration mobile device and edge server

2. Product Object Classification Scheme

2.1 Dataset configuration

To create a deep learning model for product object detection, the ILSVRC2012 dataset was used as shown in **Fig. 2**. Among these datasets, 6 classes were used, and out of a total of 1500 data for each class, 70% was used for training, 15% for verification, and 15% for testing as shown in **Table 1**.



Fig. 2. Dataset examples of ILSVRC2012

Table 1. Configuration of datasets

| Class | Training | Validation | Test |
|-------------|----------|------------|----------|
| Acorn | 975(70%) | 195(15%) | 195(15%) |
| Banana | 975(70%) | 195(15%) | 195(15%) |
| Lemon | 975(70%) | 195(15%) | 195(15%) |
| Orange | 975(70%) | 195(15%) | 195(15%) |
| Pineapple | 975(70%) | 195(15%) | 195(15%) |
| pomegranate | 975(70%) | 195(15%) | 195(15%) |
| Total | 5850 | 1170 | 1170 |

2.2 Optimized hyper-parameters

To classify products on the edge server, we modified the ViT-base model to create a lightweight deep learning model. ViT model splits an image into fixed-size patches, linearly embed each of them, add position embeddings, and feed the resulting sequence of vectors to a standard transformer encoder. In order to perform classification, it adds an extra learnable classification token to the sequence [2]. As shown in **Table 2**, we optimized the hyper-parameters of the ViT-base model. The input patch size was set to 16×16 in the existing ViT-base model, the hidden size was set to 120, the MLP size was set to 512, the number of Heads was set to 12, the number of Layers was set to 5, and the Attention dropout rate was set to 0.1. We conducted an implemental test and set hyper-parameters so that object classification can be performed in an edge server environment with low hardware specifications.

Table 2. Optimized hyper-parameters

| Hyper-Parameters | Values |
|------------------------|----------------|
| Patches | 16×16 |
| Hidden size | 120 |
| MLP size | 512 |
| Heads | 12 |
| Layers | 5 |
| Attention dropout rate | 0.1 |

3. Implementation Results

To evaluate the proposed lightweight product object classification model, we used the programming language of Python 3.9. **Fig. 3** shows the implementation results of the product object classification model. Three methods were used to evaluate the performance of the proposed model. First, Frame Per Second (FPS) for object classification were evaluated. As shown in **Fig. 3**, the FPS of the proposed model was 120.43, confirming that object classification is possible at high speed. In addition, the top-1 accuracy was confirmed to be 71.45% through the ILSVRC2012 test dataset, which is also related to NetScore [4], a lightweight efficiency measurement metric. Also, FPS and Top-1 Accuracy showed similar performance compared to [5]. The NetScore metric for assessing the performance of a deep neural network for practical usage can be defined as Equation (1). $a(N)$ is the accuracy of the network, $p(N)$ is the

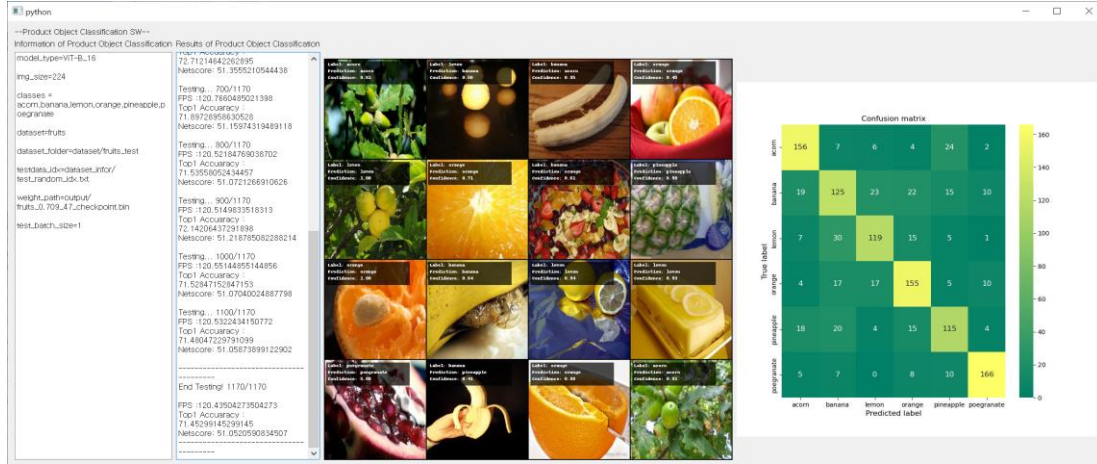


Fig. 3. Test results of the product object classification model

number of parameters in the network, $m(N)$ is the number of multiply-accumulate (MA) operations performed during network inference, and α , β , γ are coefficients that control the influence of accuracy, architectural complexity, and computational complexity of the network on Ω . The NetScore of the proposed model was 51.05%, slightly lower than the 56.17% in [6]. However, the overall model performance can be operated on a lightweight edge server.

$$\Omega(N) = 20 \log \left(\frac{a(N)^\alpha}{p(N)^\beta m(N)^\gamma} \right) \quad (1)$$

4. Conclusions

In this paper, we propose a lightweight deep learning-based object classification scheme. To operate on the edge server, we optimized the hyper-parameters of the ViT model. Additionally, we trained, validated, and tested the model using the ILSVRC2012 dataset. Through implementation results, the implementation results confirm the performance of the proposed model in terms of lightweight efficiency, object classification speed and accuracy. In the future, we plan to further improve lightweight efficiency through model optimization.

References

[1] D. Kim, K. Seo, and C. Yang, "Method for Product Object Classification Based on Mobile Devices and Edge Server Integration," in *Proc. of Institute of*

Embedded Engineering of Korea Fall Conference, 2023.

- [2] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit, and N. Houlsby, "An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale," in *Proc. of International Conference on Learning Representations (ICLR)*, 2021.
- [3] O. Russakovsky*, J. Deng*, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei (* = equal contribution), "ImageNet Large Scale Visual Recognition Challenge," *International Journal of Computer Vision (IJCV)*, pp.1-43, 2015.
- [4] A. Wong, "NetScore: Towards Universal Metrics for Large-scale Performance Analysis of Deep Neural Networks for Practical On-Device Edge Usage," in *Proc. of International Conference on Image Analysis and Recognition (ICIAR)*, 2019.
- [5] S. Mehta and M. Rastegari, "Mobilevit: light-weight, general-purpose, and mobile-friendly vision transformer," in *Proc. of International Conference on Learning Representations (ICLR)*, 2022.
- [6] Y. Xiong, H. Liu, S. Gupta, B. Akin, G. Bender, Y. Wang, P. J. Kindermans, M. Tan, V. Singh, and B. Chen, "MobileDets: Searching for Object Detection Architectures for Mobile Accelerators", in *Proc. of Computer Vision and Pattern Recognition (CVPR)*, 2021.

Smart Video Surveillance System using 5G Mobile Communication and Edge Server

Chang Mo Yang, Dongchil Kim, and Kyeongseun Seo

Smart Media Research Center, Korea Electronics Technology Institute
Seoul, South Korea

[e-mail: cmyang@keti.re.kr, dckim@keti.re.kr, ke.seo@keti.re.kr]

*Corresponding author: Chang Mo Yang

Abstract

In this paper, we propose a new smart video surveillance system that uses video devices equipped with 5G mobile communication technology, edge servers using video analysis technology, and VMS(Video Management System). The 5G video devices such as fixed 5G CCTV and mobile 5G camera for vehicles transmit videos and device location information to the edge server. The edge server is responsible for analyzing received videos in real time. Additionally, the edge server plays the role of delivering the received videos, device location information, and video analysis information to VMS. VMS stores the transmitted videos, device location information, and video analysis information in the system, or performs retrieval and display of the stored videos. The experimental results show that the proposed smart video surveillance system is suitable for video surveillance services using 5G video devices.

Keywords: Smart video surveillance system, 5G video device, edge server, VMS, 5G mobile communication

1. Introduction

Recently, interest in smart video surveillance systems using 5G mobile communication technology and artificial intelligence technology for video analysis is increasing [1,2]. The existing smart video surveillance system has a structure in which videos captured by CCTV are transmitted to VMS through a wired network, and then VMS performs analysis, storage, retrieval, and display of the videos [3,4]. On the other hand, the smart video surveillance system using 5G video devices uses 5G mobile communication technology and video devices equipped with portable power sources such as batteries or solar power. Additionally, an edge server is used to relay videos between 5G video

devices and VMS.

In this paper, we propose a smart video surveillance system consisting of 5G video devices, edge servers, and VMS. In the proposed system, 5G video devices such as fixed 5G CCTV and mobile 5G camera for vehicles are used. 5G video devices are responsible for transmitting captured videos and device location information to the edge server. The edge server analyzes the videos received in real time and then delivers the videos, device location information, and video analysis information to VMS. VMS automatically classifies and stores transmitted videos, device location information, and video analysis information, and performs retrieval for stored videos. Additionally, VMS performs a GIS-based display for the videos.

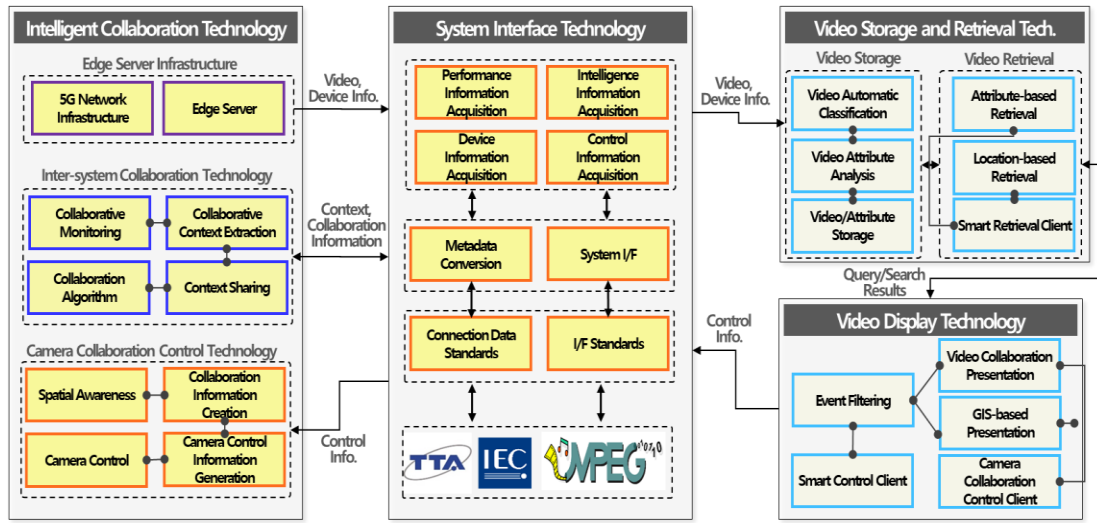


Fig. 1. Technical structure of the proposed smart video surveillance system

2. Proposed Video Surveillance System

2.1 Technical Structure of the System

Fig. 1 shows the technical structure of the proposed smart video surveillance system. As shown in Fig. 1, the proposed system consists of intelligent collaboration technology, system interface technology, video storage and retrieval technology, and video display technology.

Intelligent collaboration technology is a technology for performing collaboration between subsystems of the proposed system, and is composed of edge server infrastructure, inter-system collaboration technology, and camera collaboration control technology. Edge server infrastructure consists of 5G network infrastructure and edge server that can transmit and receive data through 5G networks. Inter-system collaboration technology is defined as a technology for stable load management on edge servers through collaboration of edge servers depending on the resource status. The inter-system collaboration technology consists of collaboration monitoring module, collaboration context extraction module, collaboration context sharing module, and collaboration algorithm. Camera collaboration control technology is defined as a technology for controlling multiple cameras based on spatial coordinates. Camera

collaboration control technology consists of a spatial information recognition module, camera collaboration information generation module, camera control information generation module, and camera control module.

System interface technology is the interface technology between 5G video devices, edge servers, and VMS that make up the video surveillance system. ONVIF (Open Network Video Interface Forum) standard is used as the 5G video device interface, and Korea TTA (Telecommunications Technology Association) standard is applied to the interface between video surveillance systems.

Video storage and retrieval technology is a technology for automatically classifying, storing and retrieving videos transmitted from edge servers. It consists of video automatic classification module, video attribute analysis module, video and attribute information storage module, attribute-based retrieval module, location-based retrieval module, and smart retrieval client.

Video display technology is defined as a technology for displaying videos which are transmitted from edge server or stored in video storage on VMS. It consists of event filtering model, smart control client, video collaboration display module, GIS-based display module, and camera collaboration control client.



(a) Fixed 5G CCTV



(b) Mobile 5G camera for vehicles

Fig. 2. 5G video devices

2.2 5G Video Device

Fig. 2 shows 5G video devices such as fixed 5G CCTV and mobile 5G camera for vehicles. Since 5G does not provide fixed IP services to date, 5G video devices use an edge module to receive video from a camera module and then transmit the received video to the edge server using 5G mobile communication. At the same time, the small edge module receives device location information using the GPS module mounted on the 5G video devices and then transmits it to the edge server.

2.3 Edge Server

The edge server performs real-time analysis of videos received from 5G video devices, while transmitting videos, device location information, and video analysis information to VMS. The edge server consists of an edge cluster with one master server and multiple worker servers. Master server performs software package distribution, settings, control, and functions for worker servers. For this purpose, the master server uses Docker and Docker Repository to store or download containers, and uses Kubernetes to control the worker server or perform operation settings. Worker server uses Docker to download containers from master server and uses Kubernetes for communication and control with master server.

2.3 VMS(Video Management System)

VMS automatically classifies and stores videos, device location information, and video analysis information transmitted from the edge

server, and performs retrieval of the stored videos. In addition, it plays videos transmitted in real time from the edge server or stored videos, or performs GIS-based display using the device location information and the transmitted videos.

3. Implementation Results

Fig. 3 and **Fig. 4** show the implementation results of the proposed smart video surveillance system. For the proposed system, eight 5G video devices, including one fixed 5G CCTV and seven mobile 5G cameras, three edge servers, and one VMS were used. For video analysis on the edge server, an artificial intelligence model was used to analyze the location of objects.

Fig. 3 shows the results of receiving videos from the edge server. The first video in **Fig. 3** shows video captured through one fixed 5G CCTV, and the remains show videos captured through seven vehicle 5G cameras.

Fig. 4 shows the implementation result of receiving video, device location information, and video analysis information from the edge server in the VMS, and then displaying the location information of objects on the GIS map. The dotted part of the GIS map shows the location of objects, and the function is implemented to display thumbnail images of objects when objects appear in the video.

4. Conclusions

In this paper, we proposed a smart video



Fig. 3. Implementation result of playback for videos received from edge server



Fig. 4. Implementation result of GIS-based object location display

surveillance system consisting of 5G video devices, edge servers, and VMS. The proposed system uses 5G mobile communication to transmit videos captured by 5G video devices and the device location information to the edge server. The edge server retransmits the received videos and device location information to VMS, while performing video analysis based on artificial intelligence and transmitting the analysis results to VMS. The implementation results show that the proposed system is suitable for smart video surveillance services using 5G video devices.

References

- [1] N. Dilshad, J. Hwang, J. Song, and N. Sung, "Applications and Challenges in Video Surveillance via Drone: A Brief Survey", Int'l Conf. on Info. and Comm. Tech.
- [2] H. Kim, Y. Cha, T. Kim, and P. Kim, "A Study on the Security Threats and Privacy Policy of Intelligent Video Surveillance System Considering 5G Network Architecture", Int'l Conf. on Electronics, Info., and Comm., pp. 1-4, April 2020.
- [3] J. Park, and S. Yi, "Development of Video Database and a Video Annotation Tool for Evaluation of Smart CCTV System," Journal of Korea Institute of Electronic Comm. Science, vol. 9, no. 7, pp. 739-745, July 2014.
- [4] L. Li, W. Huang, I. Gu, R. Luo and Q. Tian, "An Efficient Sequential Approach to Tracking Multiple Objects Through Crowds for Real-Time Intelligent CCTV Systems", IEEE Trans. on Systems, Man, and Cybernetics, Part B (Cybernetics), vol. 38, no. 5, pp. 1254 - 1269, Oct. 2008.

Convergence, pp. 728-732, Oct. 2020.

Autonomous driving service effectiveness verification technology in perspective of laws and regulations

Kyungwon Kim and Jongbin Park

Information Media Research Center, Korea Electronics Technology Institute
Seoul, South Korea

[e-mail: kwkim@keti.re.kr, jpark@keti.re.kr]

*Corresponding author: Kyungwon Kim

Abstract

At a time when autonomous driving is getting closer to reality, innovative technologies related to autonomous driving are being researched. Current autonomous driving researches are mainly focused on the technological development to enhance the performance of autonomous vehicles and research to evaluate the performance of autonomous vehicles. However, there is a shortage of research that assesses whether autonomous vehicles comply with regulations and their impact on the road environment when operating on actual roads. This paper proposes technology for evaluating whether autonomous vehicles actively comply with current laws and regulations. This approach enables the verification of various driving maneuvers of autonomous vehicles during the autonomous driving service, ensuring their safe execution within the bounds of regulatory frameworks.

Keywords: autonomous driving, laws and regulations, simulation scenarios, accident risk rate, road environment adaption rate, Traffic law compliance rate

1. Introduction

The current focus of autonomous driving technology development primarily revolves around enhancing the performance of autonomous vehicles in terms of recognition, judgment, and control[1]. However, Considering the situation where autonomous vehicles and regular vehicles coexist on real roads, it becomes imperative to assess the ability of autonomous vehicles to operate in compliance with existing traffic laws and whether other vehicles consider the potential risk of accidents posed by autonomous counterparts. Insufficient attention has been given to research evaluating autonomous vehicles from the perspective of

their impact on overall traffic dynamics and safety.

In this paper, we implement scenarios applying regulations and laws related to autonomous driving in virtual simulations. Through this, we aim to proactively recreate potential autonomous driving hazards that may occur in reality, validate the feasibility of regulations and laws related to autonomous driving through the analysis of autonomous driving simulation results, and propose a platform technology for autonomous vehicle developers to verify performance metrics that should be considered from a regulatory perspective.

This work was supported by Institute for Information & communications Technology Planning & Evaluation(IITP) grant funded by the Korea government(MSIT) (No.2021-0-01352, Development of technology for validating the autonomous driving services in perspective of laws and regulations)

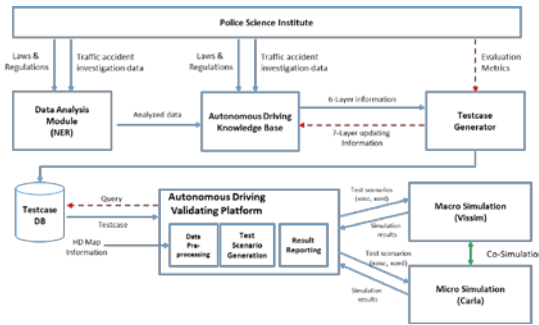


Fig. 1. Configuration of autonomous driving regulatory compliance validating platform

2. Autonomous Driving Regulatory Compliance Validation Platform

The proposed regulatory compliance validation platform for autonomous driving services in this paper comprises the construction of a knowledge base through the analysis of autonomous driving regulations and traffic accident investigation data. It is built upon this foundation for generating testcases, generating test scenarios for autonomous driving simulations, and includes Macro/Micro simulation and result reporting modules.

2.1 Knowledge base construction and testcase generation

To validate autonomous driving services from a regulatory perspective, we establish a knowledge base based on current autonomous driving related-laws and analysis of traffic accident data. This knowledge base is structured following the 6-layer format proposed by the Pegasus project[2] for test case composition, with additional definitions and considerations for regulatory elements.

The previously collected traffic accident data obtained through the Police Science Institute(PSI), the regulatory authority overseeing traffic regulations, is transformed into the 6-layer format using Named Entity Recognition (NER) algorithms and is then incorporated into the knowledge base. Subsequently, diverse test cases are generated based on evaluation metrics proposed by PSI and the knowledge base. These generated test cases are stored in a database format, and regulatory information (7-Layer) for each testcase is re-added to the knowledge base. This knowledge base, structured as a separate service, offers a

platform for general users to search for autonomous driving-related laws/regulations, and associated evaluation metrics.

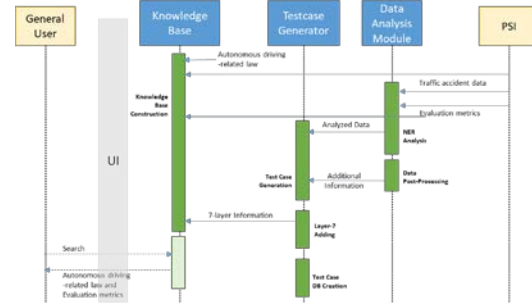


Fig. 2. Process of knowledge base construction and testcase generation

2.2 Simulation scenario generation

To enable autonomous driving simulations, the creation of scenarios executable within the simulator is essential. Initially, user input is gathered to obtain information about the target autonomous driving service for validation. The validation target segment is then subdivided, and detailed information for each segment is extracted from the high-definition map database. Using this information, regulatory evaluation factors are incorporated, and an existing test case database is queried to extract potential test case candidates for verification in each target segment. Subsequently, based on these test cases, test scenario sets for each target segment are automatically generated in a format suitable for autonomous driving simulations.

The simulation scenarios are generated in accordance with the international standard ASAM[3] format, which is a recognized format for autonomous driving simulation. Therefore, the generated scenarios can be used across various simulators.

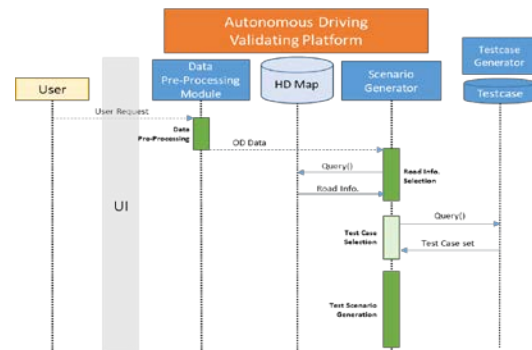


Fig. 3. Process of simulation scenario generation

2.3 Simulation and result reporting

To evaluate the effectiveness of autonomous driving services from a regulatory standpoint, two simulators are applied. A Macro simulator is used for validating autonomous driving services from a traffic flow perspective, while a Micro simulator, reflecting the performance of autonomous vehicles, is employed to assess autonomous driving services from a vehicle-centric viewpoint.

The Macro simulator generates specified test elements (geometry, traffic control scenarios, individual vehicle movements, etc.) within the simulation through a scenario-based driving environment generation module. It performs simulations from a traffic flow perspective by creating elements such as networks, interfaces, and trajectories. The Micro simulator validates autonomous driving performance from the perspective of autonomous vehicles based on the generated scenarios. The results of each simulation are provided as outputs, evaluating metrics such as accident risk and regulatory compliance of the target autonomous driving software, using an automated analysis indicator calculation module.

The simulation results derived from the two simulators are analyzed based on regulatory evaluation metrics such as accident risk, compliance rate with regulations, and adaptability to road environments. The analysis outcomes are presented in both statistical and graphical report formats.

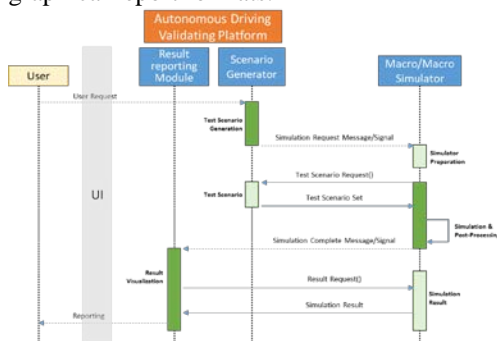


Fig. 4. Process of simulation and result visualization/reporting

3. Implementation Results

Fig. 5 illustrates the prototype implementation results of the simulation result reporting interface

for the proposed platform in this paper. The interface incorporates a user-friendly presentation of simulation information, including a summary, simulation videos, snapshots of key event scenes, and numerical values for evaluation metrics. Furthermore, it offers a time-series visualization of the movements and events of objects that compose the simulation scenario to be validated in the specified segment. Detailed information such as speed, position, wheel angle, etc., for each object is also provided.



Fig. 5. Simulation result reporting

4. Conclusions

In this paper, we introduce a simulation-based effectiveness verification technology to assess the proactive compliance of autonomous driving services utilizing autonomous vehicles with current laws and regulations. Through this, it is anticipated that potential shortcomings in autonomous driving algorithms occurring in real-world settings can be identified, thus contributing to technical improvements and enhancing the overall stability of existing technologies.

References

- [1] E. Thorn, et al. "A framework for automated driving system testable cases and scenarios," *United States. Department of Transportation. National Highway Traffic Safety Administration*, 2018.
- [2] Pegasus Project, <https://www.pegasusprojekt.de/en/>, 2021.09.24.
- [3] ASAM(Association for Standardization of Automation and Measuring Systems), <https://www.asam.net/>, 2021.09.24.

A Study for 3D Cross-Modal Retrieval using Self-supervised Learning

Yeon-Seung Choo, Hyun-Sik Kim, and Yong-Suk Park

Content Convergence Research Center (CCRC), Korea Electronics Technology Institute (KETI)
Seoul, South Korea

[e-mail: piksal@keti.re.kr, hskim@keti.re.kr, yspark@keti.re.kr]

*Corresponding author: Yeon-Seung Choo

Abstract

3D cross-modal retrieval is a task which retrieval 3D objects across the modalities, such as images, meshes, and points. Conventional supervised learning-based retrieval methods are usually searching their classes based on label prediction which is mostly used in supervised learning. However, most 3d objects acquired in-the-wild, are assigned no ground truth labels that can be used for significant class-based feature clustering. To address this issue, we do practical experiments using contrastive learning without distinguishable information like labels, to compare with supervised method. As a result, we can result a possibility to retrieve 3D cross-modal retrieval using self-supervised learning.

Keywords: Cross-modal, Object Retrieval, Supervised Learning, Self-Supervised Learning, Contrastive Learning,

1. Introduction

From the development of deep learning, 3D information based applications are commonly used. For a instance, using lader sensors, the user can obtain 3d object informations constructed with meshes and point cloud. However, to retrieve 3d objects in the real world, it is essential to extract cross-modal object features regardless of label criteria.

Actually, self-supervised learning based cross-modal retrieval has been studied mainly in the area of text-to-image, or vice versa [1-2]. Even though there are similar researches are existed, however, 3D object retrieval over 3 modalities not have been studied enoughly. To address this issue, we do experiments to compare with supervised learning and self-supervised learning method using contrastive learning. SimCLR,

which trains to extract distinguishable features from positive and other negatives in the batch with unlabeled data, is normally used as one of the baseline in self-supervised learning [3].

2. Proposed Methods

2.1 Traditional Method

To expand SimCLR for 3D Cross-Modal Retrieval (3DCMR) task, we simply replace multi-modal samples to positive samples. As in detail, SimCLR extracts image features using ResNet and projection head [4]. Applying embedding features from projection head, SimCLR can extract and compare semantic features using a pair of augmented images from same source image. Equation (1) presents their contrastive loss, called InfoNCE loss. In Equation (1), \mathbf{z} and τ denote embedding features

and scale parameter. And \mathbb{I} is a conditional indicate function, and **sim** means cosine similarity, respectively.

$$\ell_{i,j} = \log \frac{\exp(\text{sim}(z_i, z_j)/\tau)}{\sum_{k=1}^{2N} \mathbb{I}_{[k \neq i]} \exp(\text{sim}(z_i, z_k)/\tau)} \quad (1)$$

As a result, the positive pair z_i and z_j attract each other and the others are contrary in the batch.

2.2 cross-modal SimCLR

Likewise, cross-modal SimCLR attracts multi-modal features. Equation (2) presents cross-modal SimCLR loss.

$$\ell_{i,j} = -\frac{1}{|M|} \sum_{m=1}^M \log \left(\frac{\exp(\text{sim}(z_i, z_m)/\tau)}{\sum_{k=1}^{2N} \mathbb{I}_{[k \neq i]} \exp(\text{sim}(z_i, z_k)/\tau)} \right) \quad (2)$$

In Equation (2), the features z_m from multi-modal data M attract each other, such as positive samples of SimCLR.

3. Experimental Results

We compare between mostly employed self-supervised 3DCMR baseline, Cross-Modal Center Loss, and our cross-modal SimCLR [5].

3.1 Dataset

Following the baseline, we do experiments on ModelNet40 dataset, acquired from <https://github.com/LongLong-Jing/Cross-Modal-Center-Loss> [6].

3.2 Experimental Results

Table 1 shows the experimental results. As a result, the baseline outputs better performance while it uses label prediction. However, cross-modal SimCLR also can extract features to some extent, while using only projected embedding features.

Table 1. Experimental Result

| Retrieval Task (batch size: 48) | | | Baseline [5] | cross-modal SimCLR |
|------------------------------------|---|-------|-----------------|-----------------------|
| Image | → | Image | 85.64 | 60.34 |
| Image | → | Mesh | 86.94 | 59.39 |
| Image | → | Point | 85.59 | 58.83 |
| Mesh | → | Image | 88.91 | 59.33 |
| Mesh | → | Mesh | 86.50 | 59.35 |

| | | | | |
|----------|---|-------|-------|-------|
| Mesh | → | Point | 86.67 | 55.31 |
| Point | → | Image | 85.44 | 58.91 |
| Point | → | Mesh | 84.67 | 55.57 |
| Point | → | Point | 86.62 | 58.71 |
| Mean mAP | | | 86.33 | 58.42 |

4. Conclusions

In this paper, we do experiments to compare supervised learning and self-supervised learning based methods in 3DCMR task. For the test, we expand traditional SimCLR to cross-modal SimCLR replacing augmented positive samples to multi-modal samples. The experimental result verifies applying labeled data outputs better performance as we expected, but also cross-modal feature extraction using self-supervised learning can extract modal-consistent features across modalities to some extent.

References

- [1] L. Zhen, P. Hu, X. Wang, and D. Peng, “Deep Supervised Cross-modal Retrieval,” in *Proc. of IEEE Computer Vision and Pattern Recognition (CVPR)*, 2019.
- [2] P. Hu, X. Peng, H. Zhu, L. Zhen, and J. Lin, “Learning Cross-Modal Retrieval with Noisy Labels,” in *Proc. of IEEE Computer Vision and Pattern Recognition (CVPR)*, 2021.
- [3] T. Chen, S. Kornblith, M. Norouzi, and G. Hinton, “A Simple Framework for Contrastive Learning of Visual Representations,” in *Proc. of International Conference on Machine Learning (ICML)*, 2020.
- [4] K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recognition,” in *Proc. of IEEE Computer Vision and Pattern Recognition (CVPR)*, 2016.
- [5] L. Jing, E. Vahdani, J. Tan, and Y. Tian, “Cross-Modal Center Loss for 3D Cross-Modal Retrieval,” in *Proc. of IEEE Computer Vision and Pattern Recognition (CVPR)*, 2021.
- [6] Z. Wu, S. Song, A. Khosla, F. Yu, L. Zhang, X. Tang, and J. Xiao, “3D ShapeNets: A Deep Representation for Volumetric Shapes,” in *Proc. of IEEE Computer Vision and Pattern Recognition (CVPR)*, 2015.

Design and Implementation of 3D Visualization Service Architecture for 2D CAD Data

Seulki Song, Minju Cho and Jungwook Wee

Korea Electronics Technology Institute

Seoul, South Korea

[e-mail: seulkieemma@keti.re.kr, cmj1228@keti.re.kr, jwwwee@keti.re.kr]

*Corresponding author: Jungwook Wee

Abstract

In this paper, a 3D visualization service architecture is designed and implemented to visualize 3D objects modeled from 2D CAD data in multi-device environments. The service architecture consists of a 2D CAD file parser, a 3D generator, and a hybrid transmission system for the visualization of 3D objects. The 2D CAD parser extracts section images and shape length information necessary for 3D generation. The 3D generator automatically creates 3D models based on the calculated 2D shape information. The hybrid transmission system transfers these 3D models to various devices via individual connections. Moreover, we have established a database (DB) to facilitate the integration of each component. The implemented service architecture has the advantage of visualizing 3D objects generated from 2D CAD data on various devices in the industrial field.

Keywords: 2D CAD, 3D Visualization, Service Architecture

1. Introduction

Virtual space service technologies such as XR, digital twin, and metaverse are evolving at a rapid pace. In particular, it is expected that it will be used in the manufacturing industry or industrial education field by producing 3D data and sharing it in the field [1]. Additionally, there is a growing expectation for technologies that can effectively utilize the 2D CAD held by industries in field operations. Recently, many studies have been conducted on technologies such as automatic conversion of 2D CAD into 3D models, real-time transmission of 3D models, and XR visualization. However, there is still insufficient of research on the service architecture for easily converting 2D CAD into

3D models and effectively utilize it in the industrial field.

In this paper, we designed and implemented a service architecture that parses 2D CAD files and automatically converts them into 3D models and can be visualized in a multi-device environment. We first analyzed related technologies, designed a service architecture to link these technologies, and then implemented and verified it.

2. Related Research

2.1 Automatic 3D Modeling

A method of automatically generating 2D CAD into a 3D model was proposed in [2], utilizing CAD images for modeling process. First, it

performs the function of extracting the shape, position information, and numerical data characters located next to the shape within one section of the CAD image. The extracted shapes and data are sorted into an array list. The sorting method initially categorizes the shapes into groups based on their positions and then sorts them according to the largest shape within each group. Using the information from the list, it calculates details such as the rotation, 3D position, and the presence of embossing for individual shapes, and generates the 3D model based on these calculations. In the proposed service architecture, a component was designed to parse 2D CAD files and create images of sections to generate 3D models based on 2D CAD files. Furthermore, the numerical data extracted from the image was validated through a comparison with the numerical data extracted from the parsing data.

2.2 Hybrid Transmission System

A hybrid transmission system based on the web for sending the 3D model to various devices is proposed [3]. It is possible to share 3D models in real-time regardless of the quality of 3D models, hardware performance of devices, and network environment. In particular, the method of streaming 3D models as 2D projection video on the server reduces the initial delay. The server creates data channels for object control on the user device and manages them with a peer-to-peer connection to the device. The proposed service architecture is designed and implemented to allow not only 3D object control on the device but also seamless integration with the 3D generation component.

3. 3D Visualization Service Architecture

The proposed service architecture initially parses 2D CAD data. It then automatically generates 3D models based on the images and parsed information extracted from the 2D CAD files. The generated 3D models are visualized on various multi-devices through a web-based hybrid transmission system. The service architecture is designed as shown in Fig. 1. The 2D CAD file parser loads and parses 2D CAD files. A 2D CAD file may contain one or more

parts with one or more surfaces. A single 3D model can be created from one or more CAD files. Thus, by setting a 'target name' of the 3D model before loading CAD files, the CAD files are organized and stored under the same 'target name' for the generation process. When one of the loaded files is selected, the CAD is parsed, and it is drawn in the drawing area. If a user drags a single section in the drawing area, the length, angle, position, and the image of the section are generated. When the user specifies a location name (e.g., Top, Bottom, Left-side) of the dragged area, the generated image and data, along with the section name, are uploaded to the database under the 'target name'. The 3D Model Generator fetches stored images from the database, analyzes the images, and extracts shapes and numerical data. The extracted data is compared and verified against the numerical data stored in the database. Once validation is complete, the entire calculation is performed, and the 3D model is generated. The generated data and the 3D storage address are uploaded to the database, while the 3D model is uploaded to the server. The server and devices are managed through the peer-to-peer connection, and a monitoring system has been implemented. The monitoring system not only oversees device connections but also allows for the monitoring of device control information and modification details.

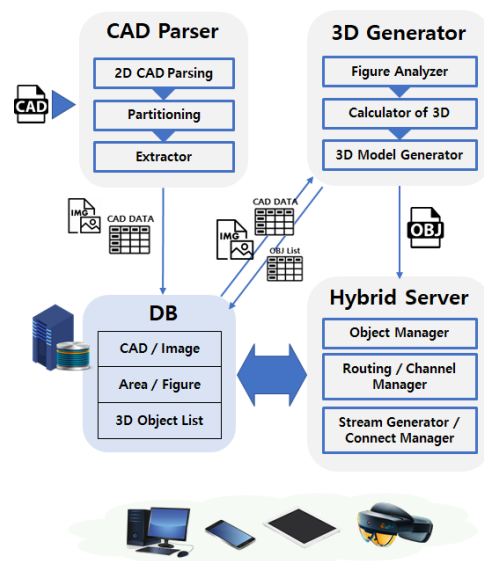


Fig. 1. 3D Visualization Service Architecture

4. Experimental Results

The 2D Parser is implemented in Javascript and Three.js, and it operates within a browser. The 3D Generator is based on Unity and functions as a Windows program. The hybrid transmission system is implemented using nodeJs, Javascript, Electron, three.js, and a STUN server. User devices have been implemented to operate within various device resolutions and function within web browsers. HoloLens, an AR device, is implemented and operates based on Unity.

Fig. 2 shows an example screen of the 2D Parser. It has loaded CAD files of the fire standpipe system. The left list denotes the loaded CAD list. The central area displays one of the CAD from the list and allows dragging. The right area indicates a list of section images of the CAD that have been dragged. The image list is stored in the database along with the location names. **Fig. 3** shows the monitoring system of the hybrid transmission system. This system enables the validation of connections to the server by devices like PCs, phones, and tablets acting as clients. Additionally, it confirms that HoloLens2 has

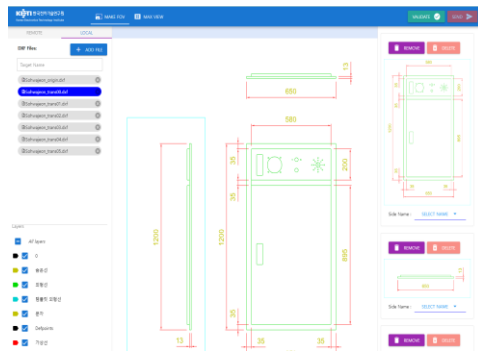


Fig. 2. 2D CAD Parser

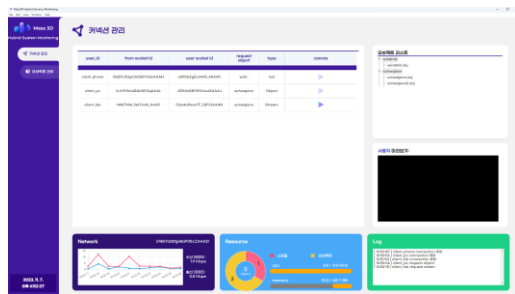


Fig. 3. Monitoring System



Fig. 4. Hololens 2 Visualization

successfully downloaded 3D objects from the server. **Fig. 4.** displays the visualization of the matched 3D model on the HoloLens 2 screen, corresponding to the target model.

4. Conclusions

In this paper, a 3D visualization service architecture was designed and implemented to visualize 3D objects modeled from 2D CAD data in multi-device environments. The implemented service architecture has the advantage of visualizing 3D objects generated from 2D CAD data on various devices in the industrial field.

References

- [1] S. Doolani, C. Wessels, V. Kanal, C. Sevastopoulos, A. Jaiswal, H. Nambiappan, F. Makedon, "A Review of Extended Reality (XR) Technologies for Manufacturing Training," *MDPI Technologies*, Vol 8, Issue 4, 2020.
- [2] M. Seo, Y. Lee, H. Choi, H. Lim, Y. Choi and J. Jeon, "Design of Automatic 3D Object Method Based on 2D CAD Image Analysis," *Korea Society Broadcasting Media Magazine*, Vol 27, No. 3, pp. 48-59, 2022.
- [3] J. Wee, M. Cho and Y. Lee, "Implementation of WebRTC-based Hybrid Transmission System for 3D Visualization," *Fourteenth International Conference on Ubiquitous and Future Networks (ICUFN)*, pp. 724-726, 2023.

Deep Learning-Based Video Editing for Privacy Protection

Yong-Suk Park, Yeon-Seung Choo, and Hyun-Sik Kim

Contents Convergence Research Center, Korea Electronics Technology Institute
Seoul, South Korea

[e-mail: yspark@keti.re.kr, piksal@keti.re.kr, hskim@keti.re.kr]

*Corresponding author: Yong-Suk Park

Abstract

The popularity of video-sharing social media platforms is invading the privacy of individuals by capturing images of people without their consent. Manually selecting information that needs to be de-identified during the media editing process is laborious and time-consuming. Incorporating artificial intelligence techniques in media editing can simplify the task and increase the process efficiency. In this paper, we propose the use of deep learning-based de-identification and object-removal techniques in video editing to prevent unwanted or unnecessary exposure of information.

Keywords: De-identification, privacy protection, video editing

1. Introduction

The ability to de-identify or edit out a particular object in media editing tools is essential. This functionality provides the basis for privacy protection. When filming in public places, people in the background are often captured as well. The laws from many countries require obtaining agreement or consent from people to be included in the media if the filmed content is to be used commercially. In most cases, it is impossible or inefficient to obtain individual consent for filming from them. Therefore, de-identification is performed during the editing process to prevent exposure of information and protect the identity of these people. Security is another critical issue. The media content may capture information, such as street addresses, license plate numbers, name tags, etc., that can challenge data privacy.

Currently, there are diverse media editing tools in the market providing de-identification

functionality. However, the process of finding the de-identification target is not automated, and it is still up to the editor to specify what needs to be de-identified. Checking the media content for de-identification targets frame by frame is a repetitive and time-consuming task, requiring a lot of effort and attention. As a result, some targets are omitted and editing errors become apparent.

Using artificial intelligence-based techniques in media editing can greatly improve productivity [1]. In this paper, we propose a method that combines several techniques based on deep learning for the de-identification of targets in video editing tasks.

2. Proposed Method

The process pipeline of the proposed de-identification method is shown in Fig. 1. “Target Registration” registers the object to de-identify. The target object may be a person or a thing. People are identified using face

recognition. If many random people in the background need to be de-identified, we register the target to identify, instead. During the search process, any face mismatch will be marked as targets to de-identify. De-identification targets may also include things, such as license plates or specific texts. “Object Detection” detects the registered target classes in each video frame. “Object Tracking” checks whether or not the objects detected in the current frame have been detected in the previous frame. If the objects from the current frame are different from those of the previous frame, we assume that a scene change has occurred. “Image Segmentation” extracts the outline segmentation mask of the target objects. The mask defines the area of the frame to apply the de-identification effect. The de-identification effect, such as blur and pixelation, is applied in the “Effect Application” stage. Depending on the need, the target may be removed entirely from the scene using inpainting.

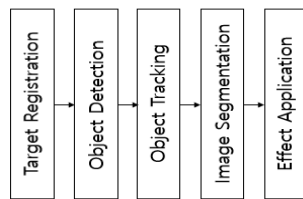


Fig. 1. The proposed de-identification pipeline for video editing.

3. Implementation and Results

The proposed de-identification method is implemented using the following components. To recognize and identify people, FaceNet face recognition and Multi-task Cascaded Convolutional Networks (MTCNN) face detection models are used. FaceNet and MTCNN are good at identifying even at varying face angles and rotations. To detect text areas within the frames, the Efficient and Accurate Scene Text Detector (EAST) model is used. EAST can detect unclear or tilted text, as well as partially occluded text. To extract the actual text from the detected text areas, Tesseract Optical Character Recognition (OCR) is used. For object detection, segmentation, and tracking, YOLOv8 is used [2]. Independent algorithms may be used for each of the tasks, but YOLOv8 provides results for the 3 tasks in a single process. For inpainting,

Large-Mask Inpainting (LaMa) was used [3]. LaMa performs fast inpainting and supports arbitrary image sizes. The method was implemented in Windows PC using NVIDIA GeForce RTX 3090 GPU. Some de-identification results in video editing are shown in Fig. 2.

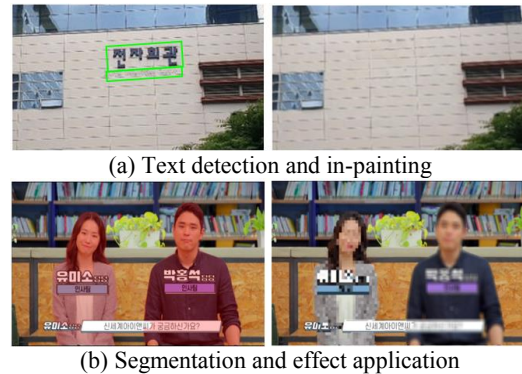


Fig. 2. Example of de-identification results.

4. Conclusions

De-identifying sensitive personal information is essential in video editing to protect privacy and avoid legal problems. In this paper, we incorporate AI-based techniques, such as face detection and recognition, text detection and recognition, object detection and tracking, image segmentation, and inpainting, to enable de-identification in video editing and production.

References

- [1] T. H. Soe, “AI Video Editing Tools. What Editors Want and How Far is AI from Delivering?” *arXiv:2109.07809 [cs.HC]*, pp. 1-7, 2021.
- [2] Ultralytics, “Ultralytics YOLOv8,” *Ultralytics Github*. <https://github.com/ultralytics/ultralytics> (accessed Nov. 11, 2023).
- [3] R. Suvorov et al., “Resolution-robust Large Mask Inpainting with Fourier Convolutions,” in *Proc. of 2022 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, pp. 3172-3182, 2022.

Interaction Recognition in Video for Edge Devices

Yong-Suk Park, Hyungsung Kim, and Hyun-Sik Kim

Contents Convergence Research Center, Korea Electronics Technology Institute
Seoul, South Korea

[e-mail: yspark@keti.re.kr, hyungsung@keti.re.kr, hskim@keti.re.kr]

*Corresponding author: Yong-Suk Park

Abstract

Fast and accurate recognition of human interaction is critical in many service applications, such as violence and crime prevention. Accurate recognition from video processing accompanies heavy computation costs. This paper proposes a lightweight method to detect human activity and interaction in video for resource-constrained edge devices. The proposed method uses distance estimation and interaction detection based on pose estimation. The distance between people is estimated in every frame, and interaction detection is executed only when they become close in proximity. A distance grid is used to estimate the location of people, which enables fast distance estimation.

Keywords: Activity recognition, interaction recognition, edge devices

1. Introduction

Action recognition in videos is a field in computer vision, which allows analysis of actions or activities of people appearing in the scene. Action recognition has applications in various areas, such as security and monitoring, public safety, sports, games, etc. The process of action recognition requires high computational resources, and it generally consists of data collection and pre-processing, feature extraction, time series predictive model, and data post-processing [1]. Such requirements make it difficult for action recognition to be embedded into edge computing devices, such as smart CCTV. Interaction recognition, which needs to analyze the actions of multiple people, is even more challenging [2]. If interaction recognition can be easily enabled in edge devices, real-time criminal activity detection becomes feasible, providing solutions to many social problems that

we are facing today. In this paper, a method for providing interaction detection using edge devices is proposed. The proposed method uses distance estimation and interaction detection based on pose estimation.

2. Proposed Method

The basic process of the proposed method is shown in Fig. 1.

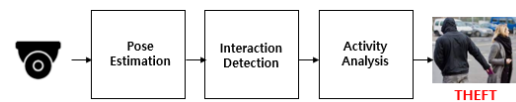


Fig. 1. Interaction Recognition Process.

Input frames are received from the camera. The “Pose Estimation” block uses object detection and pose estimation models to get body keypoints and body bounding box areas. The

“Interaction Detection” block estimates the distance between the subjects and verifies the possibility of interaction occurrence. If the subjects are at a distance capable of interaction, the “Activity Analysis” block uses a pre-trained custom pose estimation model to detect pre-defined interaction activities.

3. Implementation and Results

To estimate the location and distance between the subjects, a distance grid is generated through a calibration process. Keypoints detected from pose estimation can be used to create a bounding box for a person. The calibrator uses his height information and the computed bounding box height information to determine the correspondence between image pixels and actual distance metrics. Using this correspondence, a distance grid can be generated for the current scene.

When people appear on the scene, the keypoint on their feet are compared with the grid to estimate their locations, and simple grid counting is used to estimate their distance. This is illustrated in Fig. 2. If there are 7 grids between the two people, and if each grid represents 50 cm, then the distance is approximately 3.5 m.

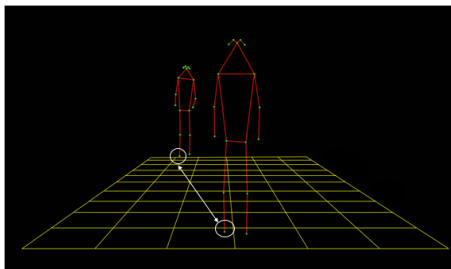


Fig. 2. Distance estimation using the distance grid.

For pose estimation and activity analysis, TensorFlow’s MoveNet was used [3]. MoveNet is a bottom-up estimation model that can localize 17 body keypoints. MoveNet uses MobileNet as the backbone, making it suitable for edge devices. The multi-pose version of MoveNet also allows simultaneous detection of up to six people.

For custom model training, human interactions, such as making conversation, kicking, hitting with a hand, high-five, and hand-shaking, were initially considered. 800 videos and images were

collected for training through web crawling. The current accuracy of the trained model for interaction identification is approximately 65% due to the low number of training samples. We plan to improve and enhance the custom model performance in the future.

The model was trained and tested on a PC and then converted to TensorFlow Lite format to deploy on Android devices. Example results for interaction recognition are shown in Fig. 3.



Fig. 3. Examples of interaction recognition results.

4. Conclusions

In this paper, a method is presented to recognize the interaction activity of people in edge computing environments. The multi-pose version of MoveNet was used for the pose estimation. A distance grid was used to estimate the proximity between people to determine if interaction needs to be analyzed.

References

- [1] C. Wang and J. Yan, “A Comprehensive Survey of RGB-Based and Skeleton-Based Human Action Recognition,” *IEEE Access*, Vol. 11, pp. 53880-53898, 2023.
- [2] S. Zhang and V. Callaghan, “Real-time Human Posture Recognition Using an Adaptive Hybrid Classifier,” *International Journal of Machine Learning and Cybernetics*, Vol. 12, pp. 489-499, 2021.
- [3] TensorFlow, “Next-Generation Pose Detection with MoveNet and TensorFlow.js,” *TensorFlow Blog*. <https://blog.tensorflow.org/2021/05/next-generation-pose-detection-with-movenet-and-tensorflowjs.html> (accessed Nov. 11, 2023).

Metaverse Asset Creation Tool for Copyright Protection

Yong-Suk Park, Yeon-Seung Choo, and HyoungSung Kim

Contents Convergence Research Center, Korea Electronics Technology Institute
Seoul, South Korea

[e-mail: yspark@keti.re.kr, piksal@keti.re.kr, hyoungsung@keti.re.kr]

*Corresponding author: Yong-Suk Park

Abstract

Virtual spaces and metaverse are gaining once again popularity with the emergence of space computing. Just as in any other digital space, virtual assets created in the metaverse are vulnerable to copyright infringement. In this paper, we present a 3D model or asset creation tool that incorporates deep learning-based similarity comparison to prevent copyright infringement. The tool provides information on how similar the created asset is compared with other existing assets, letting the creator know about the likelihood of dispute before the created asset is published.

Keywords: 3D asset, copyright, metaverse, content protection

1. Introduction

The COVID-19 pandemic has accelerated the adoption of remote work and online socializing, prompting people to seek more immersive and engaging ways to connect and interact virtually. Virtual worlds in games have long explored the idea of interconnected spaces, contributing to the emergence of the metaverse. The metaverse is defined as a spatial computing platform that provides digital experiences as an alternative to or a replica of the real world [1].

There is a growing acceptance and integration of digital experiences into daily life. People are becoming more comfortable in living parts of their lives in digital environments. This has led to the emergence of the creator economy. In the creator economy provided by the metaverse, individuals generate income by creating and monetizing digital content or goods. Creators in

the metaverse can leverage various platforms to showcase their content, often referred to as assets, which may include art, music, games, etc. Many creators generate profit by selling digital assets, such as accessories for avatars and decorations for virtual spaces.

With the influx of people into the creator economy, plagiarism and copyright infringement in the metaverse have become a major concern. People started replicating avatars and virtual objects, copying scripts and code, displaying original digital art, and replaying music without permission or giving proper attribution.

Such acts of plagiarism can diminish the value of original creations and undermine the efforts of creators. To build a sustainable creator economy, copyright protection must be upheld. However, identifying and addressing plagiarism in virtual spaces is complex due to the ease of replicating digital content. In this paper, we present a

metaverse asset creation tool that can check for copyright infringement from the initial stages of asset creation based on similarity comparison.

2. Methodology

Assets in the metaverse exist in different forms or modalities, such as images, text descriptions, or 3D models. 3D models can also be in many formats, such as meshes or point clouds. For asset similarity comparison, we use 3D cross-modal retrieval [2]. 3D cross-modal retrieval is a process used in the field of computer vision and information retrieval, specifically for retrieving information across different modalities in the context of 3D content. It involves searching and retrieving information from a dataset that contains various types of data in different modalities. The “cross-modal” aspect refers to searching across these different types of data, aiming to find relevant information regardless of the format it's stored in. We trained a model using different asset datasets for 3D cross modal retrieval based on center loss. Fig. 1 illustrates the process and shows sample retrieval results.

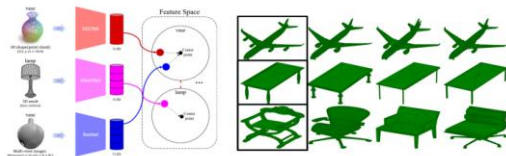


Fig. 1. Asset similarity comparison using 3D cross-modal retrieval.

3. Implementation and Results

For the asset creation tool, we use Blender [3]. Blender is a powerful open-source 3D creation suite offering a comprehensive set of tools for 3D modeling, animation, rigging, simulation, rendering, etc. It supports various file formats used in the metaverse, allowing seamless integration with other software and platforms.

Blender add-ons are additional tools, scripts, or functionalities created by developers to extend Blender’s capabilities. 3D cross-modal retrieval model is made accessible into the tool using the add-on interface.

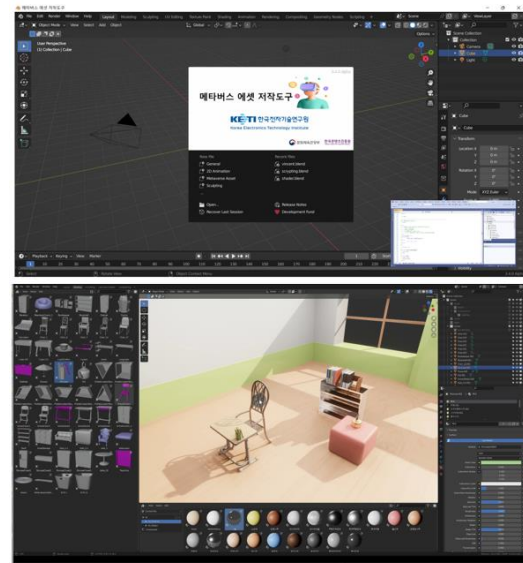


Fig. 2. Asset similarity search result in the metaverse asset creation tool.

An example of asset similarity search result is shown in Fig. 2.

4. Conclusions

In this paper, a metaverse asset creation tool is introduced that can be used to prevent copyright infringement. The tool provides a similarity comparison functionality that checks the similarity of the asset being created against pre-existing assets in the database, letting the user know if the asset design may be disputable.

References

- [1] P. Que, Y. Zeng, and F. Gao, “The Current Situation and Prospect of the Development of Metaverse Technology,” *2022 4th International Conference on Applied Machine Learning (ICAML)*, pp. 1-5, 2022.
- [2] L. Jing, E. Vahdani, J. Tan, and Y. Tian, “Cross-Modal Center Loss for 3D Cross-Modal Retrieval,” *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 3141-3150, 2021.
- [3] V. M. Venâncio, “Blender 3D Asset Creation for the Metaverse,” Packt Publishing, 2023.

Emotion Recognition Enhancement with Gesture Interpretation

Yong-Suk Park, HyoungSung Kim, and Yeon-Seung Choo

Contents Convergence Research Center, Korea Electronics Technology Institute
Seoul, South Korea

[e-mail: yspark@keti.re.kr, hyoungsung@keti.re.kr, piksal@keti.re.kr]

*Corresponding author: Yong-Suk Park

Abstract

Facial expression recognition is one of the most widely used methods to classify human emotions. However, relying only on facial expressions for human emotion recognition often gives inaccurate results. People can manipulate their facial expressions to hide their true feelings. In addition, emotions from facial expressions can be interpreted differently depending on the scene context. Gesture analysis can provide additional information about the scene context. It can also help determine the true underlying emotions since body gestures often happen inadvertently, without the subject being aware of them. In this paper, a method to reclassify emotions recognized from facial expression recognition is presented. Gesture recognition is used to reclassify emotions and improve the accuracy of human emotion recognition.

Keywords: Emotion recognition, facial expression recognition, gesture interpretation

1. Introduction

Human emotional state recognition is being actively researched by integrating different technologies, such as video processing, artificial intelligence, big data analysis, etc. If computer systems can interpret human emotions accurately, new service models will be enabled to provide personal care and deal with diverse social situations.

Humans express themselves mostly by using verbal communication, facial expressions, and body gestures. In interpersonal communication, facial expressions provide rich emotional information that reflects the person's state of mind. Human facial expressions can be generalized, but the complexity and variability of facial features due to racial differences, along

with the influence of sociocultural environments, make accurate emotional state determination difficult. Therefore, emotions from facial expressions can be interpreted differently depending on the individual or environmental context as illustrated in [Fig. 1](#). In addition, people can also manipulate their facial expressions to suppress or hide their true feelings intentionally. On the other hand, body gestures are often made inadvertently, without being aware of them. Body gestures can help in better understanding and decoding people's moods and emotions.

In this paper, we leverage the accuracy of human emotion recognition by using gesture interpretation in addition to facial expression recognition. Gesture analysis helps in determining the context of the scene, which enables more accurate emotion classification.



Fig. 1. Facial expression recognition with gesture interpretation enables more accurate emotion classification.

2. Emotions Represented in Gestures

Annotated videos from various datasets were used to classify different gestures that are representative of specific emotions as shown in Fig. 2 [1, 2].

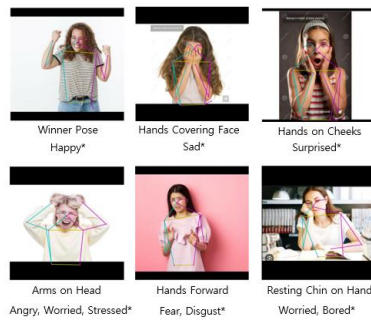


Fig. 2. Examples of gestures and their corresponding emotions (*).

The MoveNet pose estimation model was used to train a custom model that can distinguish among the different emotion gestures [3].

3. Emotion Reclassification

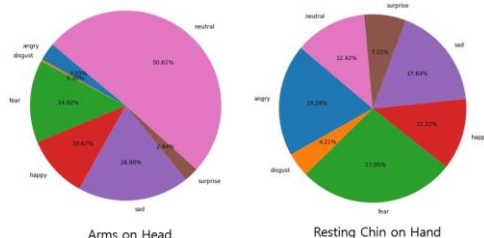


Fig. 3. Analysis of facial expressions associated with gestures.

The frequency of facial expressions that appear for each emotion gesture was analyzed using a

facial expression recognition model trained with MobileNetV2. Fig. 3 shows the distribution of facial expressions for some of the gestures. In the proposed method, the emotion is initially independently determined using facial expression recognition and gesture recognition models. If the emotion resulting from facial expression recognition is part of the facial emotions associated with the gesture, then the emotion represented by the gesture becomes the final emotional state. If the emotion resulting from facial expression recognition is not part of the facial emotions associated with the gesture, then the facial emotion becomes the final emotional state. An example is shown in Fig. 4.

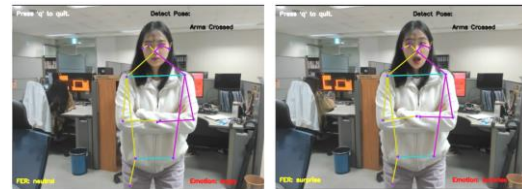


Fig. 4. Example of emotion reclassification.

4. Conclusions

Performing emotion recognition using only facial expressions has limitations. In this paper, gesture interpretation is performed together with facial expression recognition to classify emotion states more accurately.

References

- [1] F. Noroozi, C. A. Corneanu, D. Kamińska, T. Sapiński, S. Escalera, and G. Anbarjafari, "Survey on Emotional Body Gesture Recognition," *IEEE Transactions on Affective Computing*, Vol. 12 No. 2, pp. 505-523, 2018.
- [2] J. Lee, S. Kim, S. Kim, J. Park, and K. Sohn, "Context-Aware Emotion Recognition Networks," in *Proc. of 2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 10142-10151, 2019.
- [3] TensorFlow, "Next-Generation Pose Detection with MoveNet and TensorFlow.js," *TensorFlow Blog*. <https://blog.tensorflow.org/2021/05/next-generation-pose-detection-with-movenet-and-tensorflowjs.html> (accessed Nov. 11, 2023).

Design and Implementation of a Communication Repeater Operating on UAV Frequencies for Unmanned Vehicles

Sung-Chan Choi, Sungwook Jung, Jong-Hong Park, Ilyeop Ahn

Autonomous IoT Research Center, Korea Electronics Technology Institute

Seoul, South Korea

[e-mail: csc@keti.re.kr]

*Corresponding author: Sung-Chan Choi

Abstract

This study presents a drone-mounted communication repeater system using a 5GHz band, enhancing disaster response by bridging communication gaps when infrastructure is down. Leveraging the Qualcomm Atheros chipset for compliance with the IEEE 802.11n standard and Korean frequency use guidelines, the device supports extended 25 km coverage. Its array of high-gain antennas is tested for 15 km range efficacy in transmitting control and video data. Future validations aim to expand its utility across varied terrains and operations, improving the reliability of emergency communications for ground and aerial unmanned vehicles in diverse rescue scenarios.

Keywords: Communication Repeater, UAV Frequency, Unmanned Vehicles

1. Introduction

This study focuses on the innovative use of aerial communication networks to enhance the capabilities of unmanned aerial vehicles (UAVs) and robots in disaster relief efforts. In scenarios where human intervention is unfeasible, such as during natural or human-made disasters, these advanced networks play a crucial role. Central to this research is the development of a communication repeater system that can be attached to drones, forming a vital component of a mission-oriented dynamic network. [1]

The primary objective of this research is to establish a reliable communication network, tailored for disaster scenarios, that enables UAVs to conduct real-time surveillance and robots to assist in on-site recovery operations. The study explores the use of specific frequency ranges, namely 5.030-5.090 GHz and 5.091-5.150 GHz,

which are sanctioned for drone communications in Korea. These frequencies are dedicated to both control and mission-related operations for UAVs. This paper provides an exploration of the design process, development stages, and field validation of a communication repeater system, specifically engineered for these designated frequency bands.

The remainder of this study is organized as follows: Section 2 discusses system design and development, Section 3 covers field trials, and Section 4 concludes the paper.

2. System Design

2.1 Baseband and RF System

The communication repeater's modem utilizes the Qualcomm Atheros 9322 chipset, which adheres to the IEEE 802.11n standard. It operates within the Industrial, Scientific, and Medical

This research was supported by Unmanned Vehicles Core Technology Research and Development Program through the National Research Foundation of Korea(NRF) and Unmanned Vehicle Advanced Research Center(UVARC) funded by the Ministry of Science and ICT, the Republic of Korea (NRF-2020M3C1C1A02086427)

(ISM) radio bands, specifically from 5.030 to 5.150 GHz. This spectrum is allocated for unmanned vehicle use at the lower end of the 5 GHz range and is compatible with the chipset. [2]

Furthermore, to enhance the reach of the repeater's signal, each antenna path of the repeater is engineered to accommodate an amplifier with an output of 5W, theoretically enabling coverage of distances up to 25 kilometers. The design incorporates synchronized timing between the modem's transmission and reception cycles and the amplifier's RF filter. This synchronization reduces latency attributed to discrepancies in the clocks of the modem and amplifier throughout the communication process.

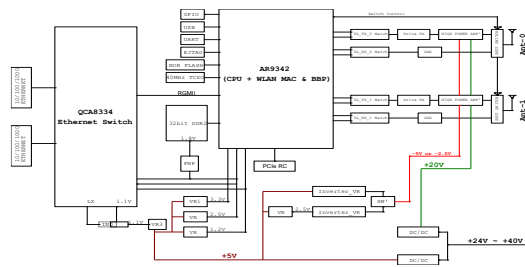


Fig. 1. Block Diagram of Communication Repeater Baseband and RF Systems

2.2 Antenna Configuration

The design of the communication repeater includes a main unit and a set of antennas that are configured for drone attachment. This setup arranges the antennas in a quad-array pattern, facilitating layered communication from aerial to terrestrial points. The array features eight directional antennas, each with a gain of 13 dBi, distributed with two antennas facing each of the four cardinal directions.



Fig. 2. Integration of Drone Systems with Communication Repeater Antennas

3. Experiments

In our practical evaluation, we confirmed the repeater's capacity to facilitate communication over distances up to 15 kilometers. This was achieved through a test scenario where drones were deployed in three different directions from the repeater's central location, and CCTV footage was streamed. The drones were managed remotely, with both their control data and video feeds transmitted via the repeater. These transmissions were then received and verified at a ground control station in communication with the repeater.

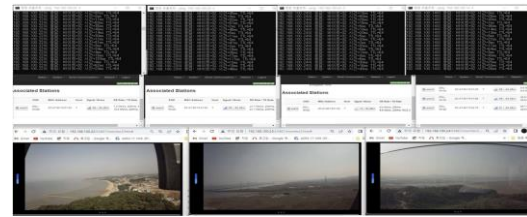


Fig. 3. Field Test Screen

4. Conclusions

This study presents the design of a drone-compatible communication repeater designed to provide temporary communication solutions in disaster-stricken areas. We have successfully demonstrated its use in extending communication ranges in mission-specific zones. Future work will involve comprehensive testing of the repeater across various terrains, including terrestrial, maritime, and aerial regions, and executing real-world operational trials in environments with multiple unmanned vehicles.

References

- [1] D. Liu, Q. Jiang, H. Zhu, and B. Huang, "Distributing UAVs as wireless repeaters in disaster relief via group role assignment," *International Journal of Cooperative Information Systems*, Vol. 29, No. 01n02, 2040002, 2020.
- [2] M. Rethfeldt, T. Brockmann, R. Eckhardt, B. Beichler, L. Steffen, C. Haubelt, and D. Timmermann, "Extending the FLExible Network Tester (Flent) for IEEE 802.11 s WLAN Mesh Networks," in *Proceedings of the 2022 IEEE International Symposium on Measurements & Networking (M&N)*, pp. 1-6, IEEE, 2022.

Designing a Pipeline to Leverage NFTs for Access Control in Web3.0

Hyungsung Kim, Yong-Suk Park and *Hyun-Sik Kim

Korea Electronics Technology Institute
Seoul, South Korea

[e-mail: hyungsung@keti.re.kr, yspark@keti.re.kr, hskim@keti.re.kr]

*Corresponding author: Hyun-Sik Kim

Abstract

Non-Fungible Tokens (NFTs), based on the ERC-721 standard, have primarily been employed to represent digital collectibles. This has impacted artists seeking to monetize their digital art, as well as cryptocurrency holders looking to profit through trading. However, this narrow focus on digital collectibles underestimates the broader utility of NFTs. At their core, NFTs serve as digital representations of rights, making them versatile tools in the emerging Web3.0 landscape. Unlike its predecessors, Web 1.0 and Web 2.0, where service providers retained data ownership, Web3.0 is decentralized and grants data ownership to users. This shift necessitates new methods for managing ownership rights, and NFTs have emerged as a promising solution. This paper proposes a pipeline that leverages NFTs for access control within Web3.0. In this framework, NFTs are used to authenticate user identities and to grant services based on the rights represented by the NFTs.

Keywords: Blockchain, Ethereum, Non-Fungible Token

1. Introduction

Non-Fungible Tokens (NFTs) [1], based on the Ethereum [2] standard token interface ERC-721, have gained popularity as digital collectibles. However, the true essence of NFTs lies in representing digital ownership, making them significant in the emerging Web 3.0 ecosystem. Web 3.0 represents a new paradigm of web services where users have ownership rights over their data. In Web 2.0, service providers typically hold the rights to user-generated data. In contrast, Web 3.0 empowers users with ownership rights over their own data. These ownerships can be represented as NFTs.

A further distinction of Web 3.0 is its fundamental approach to user data; it typically

does not store personal information. Instead, it retrieves data from a blockchain network using a user's wallet address. Although the adoption of Web 3.0 services is currently in its initial stage, primarily within decentralized finance, its potential applications are vast. As the ecosystem evolves, there will be an increasing need for service providers to verify user identities and grant access rights. This paper proposes a design for a verification pipeline that utilizes NFTs for user recognition and access control in Web 3.0 services. By leveraging the inherent properties of NFTs, we can establish a secure and decentralized mechanism for user authentication and authorization across various platforms.

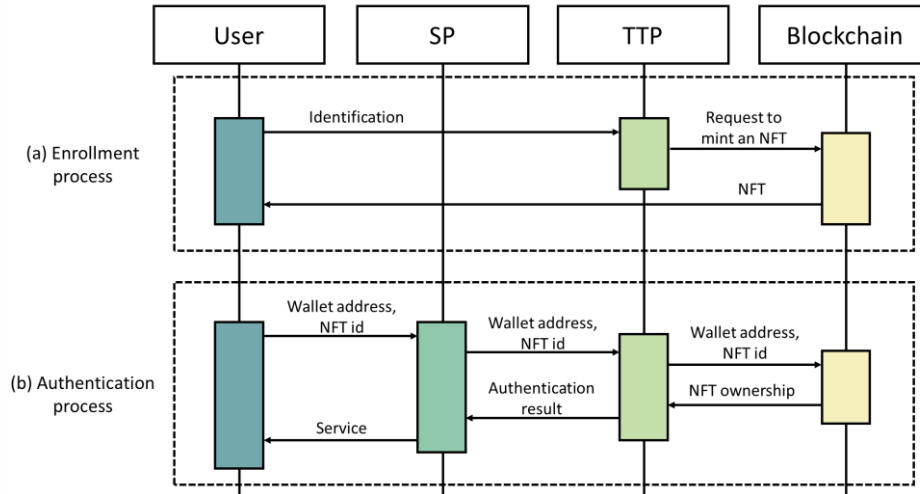


Fig. 1. Overview of the Pipeline Design. In process (a), the user submits identification information to the TTP and receives an NFT. process (b) depicts the SP authenticating the user with the NFT minted in (a).

2. Pipeline Design

In the design, we utilize a Trusted Third Party (TTP), operating under the assumption that the Web 3.0 Service Provider (SP) trusts the credentials issued by the TTP. **Fig. 1** presents an overview of the pipeline design, which is divided into two main processes: the Enrollment process and the Authentication process. During the Enrollment process, a user submits identification information to the TTP, which then requests the minting of an NFT on the Blockchain. This step signifies that the TTP retains the necessary identification information, and the minted NFT serves as a badge confirming that the user has been authenticated by TTP.

The Authentication process is a critical step where the SP verifies the user's identity. In this phase, the user transmits their wallet address, which holds the NFT, along with the NFT Id acquired during the Enrollment process. Upon receiving this information, the SP consults the TTP to confirm whether the wallet address indeed possesses the NFT. The TTP then queries the Blockchain to verify the ownership of the wallet address. If the Blockchain confirms ownership, it responds with a 'true' status to the TTP. Subsequently, if the TTP validates the user's identity based on the information provided during the Enrollment phase, it communicates a

'true' or 'false' result to the SP. The SP, upon receiving a 'true' result, proceeds to grant the user access to the service.

3. Conclusions

One of the notable aspects of this system is the Blockchain's role and the fact that the TTP only provides a response in the form of 'true' or 'false'. This mechanism bears a resemblance to zero-knowledge proofs, where the verifier (in this case, the SP) does not learn any additional information about the prover beyond what is strictly necessary; simply receiving a 'true' or 'false' without access to the user's identification details. However, it is not a zero-knowledge proof in the strictest sense, as the SP does learn the NFT Id that the user possesses. Despite this, the feature of the SP only receiving a binary response aid in safeguarding against brute force attacks by a potentially malicious SP.

References

- [1] W. Entriken, D. Shirley, J. Evans and N. Sachs, EIP-721: Non-Fungible Token Standard, Jan. 2018, [online] Available: <https://eips.ethereum.org/EIPS/eip-721>.
- [2] V. Buterin, "A next-generation smart contract and decentralized application platform," [Online]. Available: <https://ethereum.org/en/whitepaper/>.

Efficient Hierarchical Platform Structure Designs for Smart Aids-to-Navigation in Marine Environments

Jaeho Lee, Sungyoon Cho, Wongi Jeon, Byoungchul Song, Yangsub Kim, Donghyun Kim, Junsu Kim and Kiwon Kwon

Korea Electronics Technology Institute (KETI)
Seongnam-si, South Korea

[e-mail: {smitelims15, sycho, jeonwg, songbc, likant, ditel89, jesus0084, kwonkw} @keti.re.kr]

*Corresponding author: Kiwon Kwon

Abstract

This paper proposes a hierarchical smart Aids-to-Navigation (AtoN) structure consisting of main platform with independent sub-platforms such as interface, power and communication platform to efficiently manage and operate various marine sensors, lights, and power modules mounted within AtoN. Based on the proposed structure, it is expected that the ease of expanding and installing not only existing equipment and sensors, but also new sensors according to new marine navigation service requirements, and the stability of the entire system can be dramatically improved through independent operation of each sub-platform. Moreover, efficient data exchange protocols supporting various types of sensor message and interface connectors between platforms robust in maritime environments are suggested for stable multi-platform operations. Last, we propose a selective multi-mode communication platform for transmitting various types of data such as numbers, voice, and video, where it is expected to establish a seamless information exchange system between smart AtoN and land information center.

Keywords: Hierarchical Platform, Interface, Structure, Multi-mode Communication, Smart Aids-to-Navigation (AtoN)

1. Introduction

The AtoN systems have played a crucial role in ensuring safe and efficient sea route, struggling to adapt to the dynamic and unpredictable maritime environment. Recently, it has given rise to ongoing efforts in integrating cutting-edge technologies such as artificial intelligence, big data analytics, and advanced communication protocols into the development of next-generation AtoN solutions.

Through field investigation and analysis of current operational challenges and future requirements for AtoN systems, we specify the technological development items, as follows.

- (Upgrade Facility) It is required to support stable power supply system with light weight battery system for AtoN and smart facility operation through efficient multi-functional system such as various sensor data collection and AI-based fault diagnosis to recognize damage and (light-out) accident of AtoN.

- (Stabilize Information Transmission System) It is required to establishing a stable marine communication infrastructure providing the optimal information transmission system for the harsh marine environment and building the exclusive communication network for AtoN

Based on above development issues on AtoN in marine environments, we suggest the hierarchical platform based AtoN structure,

supporting efficient management and preventing any risk of AtoN replacement. Moreover, the proposed design scheme targets to various maritime data service collection for expanding new marine services.

2. Proposed Smart AtoN Structure

In this section, a smart AtoN structure consisting of hierarchical multiple platforms, which strengthens easy installation and management of AtoN as well as new marine data exchange is proposed as follows.

2.1 Platform based Smart AtoN Design

The structure of integrated platform for smart AtoN, which divides independent roles of sensor, communication and power into each sub-platform and then provides hierarchical connection between sub-platforms, are suggested, as described in Fig. 1. The proposed scheme consists of basic operation, interface, power, communication and structure platform, where its operation is verified by demo system represented in Fig. 2.

- **Basic Operation Platform:** Design module-type lightweight multi-structure OS on Docker Container framework for effective control and independent operation of multiple platforms supporting basic monitoring & control of status AtoN and providing AI-based failure diagnosis [1].
- **Interface Platform:** Design flexible physical interface hub and expansible data processing blocks for existing sensor and new sensor to be installed.
- **Power Platform:** Suggest power management & control module based on maximum power point tracking (MPPT) and power saving module based on li-ion battery which provides 50% weight reduction of currently deployed battery module with same capability.
- **Communication Platform:** Suggest supporting selective 3 different communications System (AIS, NB-IoT, LTE) for various data transmission such as number, voice and video, allowing up to 100km communication coverage within the coast [2].
- **Structure Platform:** Suggest the design in consideration of AtoN shape structure, position control and visibility.



Fig. 1. Hierarchical Platform Structure of Smart AtoN



Fig. 2. Implementation of Demo System of Smart AtoN

3. Conclusions

In this paper, we proposed platform based smart AtoN design that provides a stable communication system for AtoN marine information, smart facility operation through AI-based fault diagnosis and high-efficient light weight power supply technology. From proposed design, it is expected to secure field worker's safety through platform-based real time digital management and allow new types of marine data collection, which effectively support MASS and various smart marine services to become a leader of marine technology and global standards.

References

- [1] Ashley Davis, "Bootstrapping Microservices with Docker, Kubernetes, and Terraform: A project-based guide", Manning, 2021.
- [2] S. -W. Jo and W. -S. Shim, "LTE-Maritime: High-Speed Maritime Wireless Communication Based on LTE Technology," in IEEE Access, vol. 7, pp. 53172-53181, 2019.

Designing a Blender Add-On for Managing Creative Commons Zero Licensed 3D Assets

HyounSung Kim, Hyun-Sik Kim and *Yong-Suk Park

Korea Electronics Technology Institute

Seoul, South Korea

[e-mail: hyoungsung@keti.re.kr, hskim@keti.re.kr, yspark@keti.re.kr]

*Corresponding author: Yong-Suk Park

Abstract

Creative Commons Zero (CC0) is a unique form of Creative Commons License (CCL) that relinquishes all copyright and related rights, effectively donating the work to the public domain. While CC0 was not widely adopted in traditional web services like Web 1.0 and Web 2.0, it has gained prominence in the decentralized Web 3.0 environment for sharing rights within public communities. Blender, a popular free open-source 3D computer graphics software, allows for the integration of third-party add-ons due to its open-source nature. This paper presents a design for an add-on that facilitates the management of CC0-licensed 3D assets in Blender. Specifically, we propose a system that leverages an API key from non-CC0 3D asset marketplaces to identify users and utilizes this key in conjunction with the InterPlanetary File System (IPFS) to upload 3D assets.

Keywords: 3D asset, Blender, Creative Commons License, InterPlanetary File System

1. Introduction

The InterPlanetary File System (IPFS) [1] is a peer-to-peer file-sharing protocol that aims to enable the storing and sharing of files in a distributed file system similar to BitTorrent [2]. IPFS consists of a network of nodes, with each node storing a collection of hashed files. Contrary to the statement that each node stores complete files, IPFS nodes actually store only the chunks of files they are interested in or have agreed to pin, which means to keep persistently. Unlike BitTorrent, where files are typically downloaded from multiple nodes each holding different parts of the file, IPFS retrieves files based on their unique hash. If a single node in IPFS holds the complete file corresponding to a requested hash, it can serve the entire file, which

enhances the system's robustness against node failures.

While BitTorrent struggles to serve files if a node with a unique part of the file goes offline, IPFS can retrieve the content from any node that has it, making it more resilient and suitable for persistent services. However, because IPFS is designed to make files widely available across the network, it is not inherently private. This openness, though, can be advantageous for distributing files under public licenses such as Creative Commons Zero (CC0), where the intention is to share rights freely within public communities. In this paper, we suggest a design

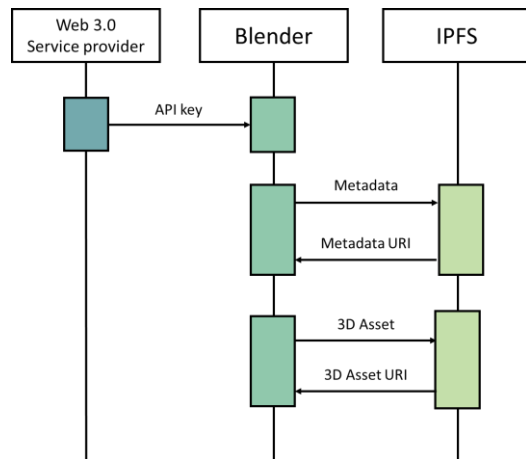


Fig. 1. Overview of the Blender Add-on Design. The Web 3.0 Service Provider issues an API key for creator authentication. The Blender add-on utilizes this key to interact with IPFS.

of Blender add-on to handle CC0 licensed 3D assets using IPFS.

2. Design

Blender is an open-source tool for creating 3D assets. Its open-source nature allows users to develop add-ons to enhance its functionality. Unlike proprietary software, Blender does not have an in-built authorization mechanism like a login system, necessitating additional services for user recognition. It is important to note that even when assets are released under CC0 and dedicated to the public domain, the original creator is not necessarily anonymous.

Fig. 1 presents an overview of how a Blender add-on can be utilized in conjunction with a Web 3.0 Service Provider (SP) and the IPFS. The SP issues an API key, which acts as a representation of the user, and the user sends this key along with the metadata to IPFS. IPFS then provides a Metadata Uniform Resource Identifier (URI), enabling access to the metadata through the IPFS gateway and its URI. Upon receiving the metadata URI, Blender can then upload the 3D asset to IPFS and, in a similar fashion, obtain a URI for the 3D asset itself.

The proposed design is poised to significantly impact the Web 3.0 service ecosystem. As we transition into Web 3.0, users are empowered not

only with the ability to consume content but also with the ownership rights over their data. Our design acknowledges and facilitates this paradigm shift by enabling the secure encoding and transfer of rights for user-generated content within Blender, with the potential integration of Non-Fungible Tokens (NFTs) [3]. Furthermore, the synergy between IPFS and Blender's flexibility ensures that creators can maintain control over their work, streamlining both access and distribution. This method enhances the security of 3D assets and opens up innovative avenues for creators to exploit their creations. It promotes community engagement with digital assets in a transparent and fair manner, laying the groundwork for a dynamic and interactive digital ecosystem.

3. Conclusions

In this paper, we have introduced a design for a Blender add-on that leverages the IPFS to facilitate the distribution of 3D assets under the CC0 license. By integrating an API key issued by an authenticator, such as a Web 3.0 Service Provider, our design enables creators to retain management over their assets, despite the open-license nature of CC0. This system is designed to be seamlessly compatible with NFTs, which are increasingly used to represent ownership rights on Web 3.0 platforms. The implications of this design extend beyond simple asset management, offering a new model for asset control and rights representation in the evolving digital landscape.

References

- [1] Juan Benet, "IPFS - Content Addressed, Versioned, P2P File System (DRAFT 3)", [Online] Available: <https://raw.githubusercontent.com/ipfs/papers/master/ipfs-cap2pfs/ipfsp2p-file-system.pdf>.
- [2] BitTorrent Foundation, "BitTorrent (BTT) White Paper", [Online]. Available: [https://www.bittorrent.com/btt/btt-docs/BitTorrent_\(BTT\)_White_Paper_v0.8_7_Feb_2019.pdf](https://www.bittorrent.com/btt/btt-docs/BitTorrent_(BTT)_White_Paper_v0.8_7_Feb_2019.pdf).
- [3] W. Entriken, D. Shirley, J. Evans and N. Sachs, EIP-721: Non-Fungible Token Standard, Jan. 2018, [online] Available: <https://eips.ethereum.org/EIPS/eip-721>

A Novel Design of Shipment Decision Framework in a Land-based Fish-farm Management with Big-data Based Aquaculture Productivity, Price, and Costs Modeling

Jaeho Lee¹, Juhyoung Sung¹, Jongho Paik², Kiwon Kwon¹, Yangsub Kim¹, Kyungwon Park¹,
and Sungyoon Cho¹

¹Korea Electronics Technology Institute (KETI)
Seongnam-si, Republic of Korea

[e-mail: {smitelims15, jh.sung, kwonkw, likant, kwpark, sycho}@keti.re.kr]

²Department of Software Convergence, Seoul Woman's University
Seoul, Republic of Korea

[e-mail: paikjh@swu.ac.kr]

*Corresponding author: Sungyoon Cho

Abstract

In this paper, we propose decision-making solutions for the optimal timing of fish species shipment in land-based fish farm management and the corresponding software framework implementation. While existing decision-making method for the timing of shipment mainly relies on the fish farm owner's experiential judgment or the current production level, our proposed algorithm integrates fish growth prediction model with external environmental estimation models, such as market price and operation cost, to determine an appropriate shipment timing from the perspective of net revenue. Particularly, proposed prediction models were designed by considering their own characteristics, where Gaussian Process Regression (GPR) method reflecting different growth rate of each fish individual, LSTM-based prediction engine with time series statistic and the statistical forecast model applying seasonal effects are exploited to the corresponding fish growth, market price and operation cost prediction model, respectively. Moreover, we adopt each model as an independent engine module and then structure the modules to implement in web-based software framework, which allows easy modification of system operation and extensions to different fish species and land-based fish farm environments.

Keywords: Aquaculture, Decision Making, GPR, LSTM, Profit Prediction, Statistical Analysis

1. Introduction

Aquaculture is increasingly recognized as a pivotal solution to address the growing global demand for seafood [1]. Nonetheless, devising optimal operation strategies for aquaculture proves challenging due to unforeseen internal and external events. Current efforts to enhance

productivity and profitability involve the integration of smart fish-farm technologies, leveraging AI and IoT applications. These technologies include AI-driven feeding for fish price prediction or Recirculating Aquaculture Systems (RAS) promoting sustainable aquaculture with growth prediction [2]. However, addressing the intricacies of optimal shipping timing remains elusive, necessitating the

incorporation of sophisticated growth, price, and cost models, as underscored in [3]. This paper proposes a decision-making algorithm for fish shipment timing, amalgamating AI-based fish growth prediction, cost and price estimation models, and a statistical database. The models function as independent engine modules within a web-based software framework. The focus is on halibuts, the most abundantly produced fish species in Korea, with potential applicability to a broader range of aquatic species.

2. Shipment Decision Framework and Decision-making System Model

2.1 Shipment Decision Framework

The proposed scheme, represented in Fig. 1, is to provide the analytical results of an optimal timing for fish species shipments in terms of net revenue. The corresponding decision-making model is designed as a combination of fish growth, market price and operation cost prediction to calculate net revenue, where each model exploits Gaussian Process Regression (GPR), Long Short-Term Memory (LSTM), and a statistical forecasting, respectively.

With these models and algorithm, the framework is implemented as shown in Fig. 2. In Fig. 2, a) presents the overall status of fish farms and information about such as fish prices and operation costs. Meanwhile, b), c), d), and e) in Fig. 2 show prediction modules of growth, prices, net profits by weight class, and trends of net profits suggesting the optimal timing of shipment respectively.

2.2 Proposed Growth, Price, and Cost Models

Due to the nature of land-based aquaculture farms, where well-grown fish and slow-grown fish are classified and raised separately, fish growth prediction models require both predicted average weight and its distribution at a certain future time. To make a decision of sorting and shipment, GPR is adopted for growth prediction model since it provides prediction value of average weight and deviation [4]. The GPR based growth prediction model is trained by the measured growth data from Jeju [5], where the simulation results are represented in Fig. 3. As shown in Fig. 3, the expected weights are

represented at a certain future time when the measured weight is given at the current time. Moreover, its standard deviation and 95% prediction interval provided as 8.02 g and ± 15.735 g, respectively.

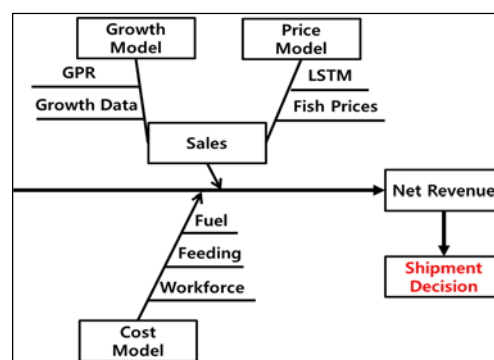


Fig.1. Fishbone diagram of a shipment decision algorithm



Fig. 2. Web framework for shipment decision

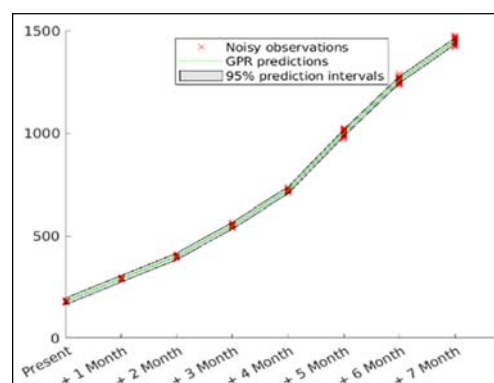


Fig. 3. GPR prediction for halibut growth

For market price prediction, many-to-one LSTM is selected as it is widely used for time-series prediction. The corresponding LSTM models are designed to predict different market prices for multiple representative weights such as 500 g, 1 kg, and 1.5 kg, respectively, due to the characteristic that the market price per unit weight varies depending on the weight of halibut. Local market prices of halibuts from Korea Maritime Institute are utilized for LSTM modeling and its accuracy of prediction is compared by changing the input sequence length to 3, 6, and 12 which correspond to the previous 3, 6, and 12 months respectively, as presented in **Table 1**.

In aquaculture, given the influence of seasonal external factors and other variables, monitoring intervals that are excessively short may fail to capture these dynamic processes adequately. On the other hand, overly long intervals may result in the averaging of diverse factors, leading to a reduction in accuracy. Furthermore, depending on growth rates of halibuts, sorting and shipment of halibuts proceeds within 3, 6 or 12 months with seed stocking typically occurring twice a year on an annual basis. Thus, among sequence lengths, the lengths of 3, 6, 12 for datasets are selected for market price prediction. And the length of 6 months suggests the lowest MSE as presented in **Table 1**.

Table 1. Design parameters of LSTM prediction model and performance

| | | | |
|--|----------|-------------|-----------|
| LSTM type | | Many-to-one | |
| Number of LSTM layers | | 2 | |
| Number of LSTM hidden units | | 10 | |
| Split ratio of training, validation, and test set | | 60:20:20 | |
| Number of training epochs | | 1000 | |
| Learning rate | | 0.005 | |
| Training batch size | | 64 | |
| Cost function | | L2 Loss | |
| Optimizer | | Adam | |
| Sequence Length | 3 | 6 | 12 |
| MSE | 24.549 | 24.091 | 26.939 |

Lastly, based on the growth model and market price model designed earlier, we propose a revenue and net profit model with a cost model as follows:

V_m represents the total revenue of halibuts m months later, formulated as:

$$V_m = \sum_i \int_{k_{i-1}}^{k_i} cdf_m(X) dX \times N_m \times PM_{i_m} \quad (1)$$

where cdf_m signifies the cumulative distribution function derived from GPR predictions, k_i denotes the border value of an interval within weight class i , X represents the average weight of halibuts, PM_{i_m} is the price corresponding to weight class i , and N_m is the total number of halibuts in a targeted fish farm.

GP_m presents the net profit of halibuts m months, formulated as follows:

$$GP_m = V_m - C_m \quad (2)$$

where C_m means operational costs, including workforce, feeding, and fuel costs, as meticulously outlined in [6].

3. Conclusions

In conclusion, our paper proposes an integrated framework using GPR, LSTM, and statistical forecasting to optimize fish species shipment timing in land-based fish farms. This solution addresses information imbalances among fish farmers, ensuring efficient operations, minimizing resource waste, and maximizing productivity and profitability. The framework's adaptability to various fish species and modular structure further enhance its utility.

References

- [1] R. Subasinghe, D. Soto, J. Jia, "Global aquaculture and its role in sustainable development," *Reviews in Aquaculture*, Vol 1, No 1, pp. 2-9, 2009.
- [2] V. Lugert et al., "A review on fish growth calculation: Multiple functions in fish production and their specific application," *Reviews in Aquaculture*, Vol 8, No 1, pp. 30-42, 2014.
- [3] Maribeth P. Alapan et al., "Factors affecting the market price of fish in the northern part of Surigao del Sur, Philippines," *Journal of Environment and Ecology*, Vol 7, No. 2, pp. 34-, 2016.
- [4] Schuzl, E. et al., "A tutorial on Gaussian process regression: Modelling, exploring, and exploiting functions," *Journal of Mathematical Psychology*, Vol 85, pp. 1-16, 2018.
- [5] Younsoo, L., "Productivity comparison to the management of olive flounder, *Paralichthys olivaceus* aqua-farms in Jeju and Wando," *Jeju National University*, Dissertation, pp. 1-23, 2019.
- [6] Yeongsu, O. et al., "A Study on the Price Structure of Flatfish in Korea," *Korea Maritime Institute*, pp. 1-140, 2006.

High Speed Template Matching System for Wafer Alignment based on FPGA

Jae Hyuk So, Dong Hyun Lee, Tae Min Hwang, Min Joon Kim and Hye Won Kang

Data Convergence Platform Research Center, Korea Electronics Technology Institute
Seongnam, South Korea

[e-mail: sojh0124@keti.re.kr, donghyun2@keti.re.kr,
taemin.hwang@keti.re.kr, mjoon@keti.re.kr, hwkang.1998@keti.re.kr]

*Corresponding author: Jae Hyuk So

Abstract

This paper presents the results of designing wafer alignment system using a FPGA for application to vision inspection equipment based on template matching algorithm software. As semiconductor processes become finer, a high-speed template matching system for vision inspection equipment is required to process high-resolution images repeatedly and quickly. Current software-based vision inspection systems have limitations in processing high-resolution images at high speed. To improve these problems, we propose a FPGA-integrated automatic wafer alignment system has low-latency, high-speed image template matching method optimized for real wafer alignment mark image. we change the conventional OpenCV-based software with an algorithm optimized on a hardware architecture and implement to the FPGA, then presents verification results.

Keywords: Wafer alignment system, Normalized Cross Correlation, Template matching, Vision inspection, FPGA-integrated

1. Introduction

In semiconductor manufacturing, wafer alignment system is an important subsystem for wafer probing, it is a high-precision equipment that can detect wafer alignment key using optical image processing and correct wafer position automatically [1,2]. Recently, as semiconductor processes become finer, the demands of higher and faster automatic vision inspection method have increased significantly as the quality control of manufacturing process had become increasingly important. Using the vision method, the aligner could reduce the process steps and time required for wafer inspection, as well as unexpected problems caused by the operators during manufacturing.

The accuracy of the template matching algorithm for locating the marker and the data processing speed of the software play an important role. However, this induces a significant increment in computational cost because several similarity calculations considering the targeted angle steps with rotated template version must be done at each position of the detected region.

In this paper, we propose a dedicated hardware architecture has been designed to implement the hardware-processed template matching algorithm applying it to an FPGA. To simplify the calculation procedure, an optimization strategy based on target image size has been applied in the full exhaustive search template matching procedure to check the primary search area and increase by stages to the entire search

area. We build wafer alignment system equipment and present the results of implementation through tests using FPGA board.

2. Hardware Architecture of Template Matching

This chapter describes the procedure of converting detailed template matching algorithms used based on OpenCV into hardware architecture. Fig. 1 shows the process structure of template matching algorithm.

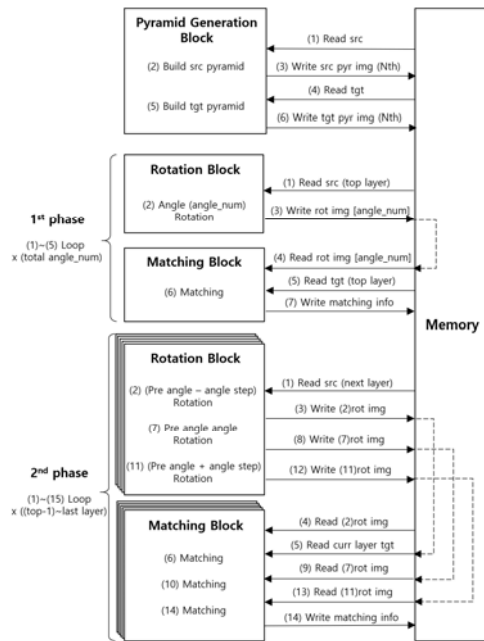


Fig. 1. Process structure of template matching algorithm

We apply a Gaussian pyramid to captured image of wafer in real-time and stored align mark images and compare the features. It performs a first phase called a top layer search, by rotating the input image through various angles to extract approximate navigation points at low resolution. The selected angle in the smallest resolution image determines the number of searches, followed by affine transformation, template matching, and normalization. Second phase, it performs a pyramid search, which is a precise comparison of the pyramid images based on the results of the adjacent angle of the first phase. Therefore, angle generation is an important process that determines the amount of computation.

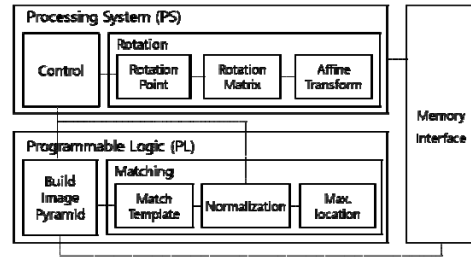


Fig. 2. PS/PL block of FPGA

In order to improve the overall processing speed, the build image pyramid and the matching block were processed by FPGA's PL, and rotation block was processed by the PS, as shown in Fig. 2. The build image pyramid and the matching block can process each pixel of the Image sequentially, allowing high-speed memory access and parallel processing of consecutive pixels, which can increase speed when processed by PL. The rotation block is advantageous to process on PS due to its high number of modifications and poor continuity of memory access coordinate values.

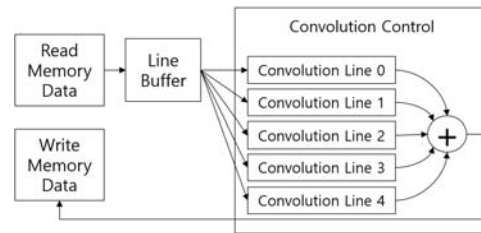


Fig. 3. Block of convolution control

Reading image data from memory is stored in line buffer. There are 5 convolution line blocks, so multiplication of 5 lines can be processed simultaneously. Even within the convolution line, multiplication of 5 pixels is processed at once and added. So, one pixel is output for each clock as shown in Fig. 3. Using the pyramid image generation block above, create pyramid images in several layers as shown in Fig. 4.

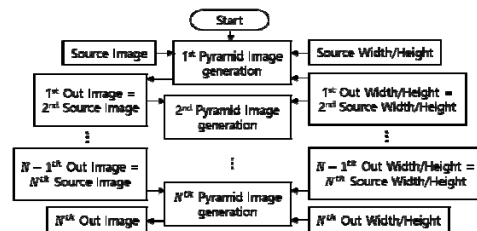


Fig. 4. Pyramid image layer generation

Fig. 5 shows process of matching block that is optimized for reducing computation. First, read the target image from memory and stored in the buffer. Next, only the size of the target image's height required for matching, not the entire source image, is read from memory and stored in the buffer. Matching algorithm is applied by multiplying and adding each pixel by the entire target image of the buffer and the target image size of the source image.

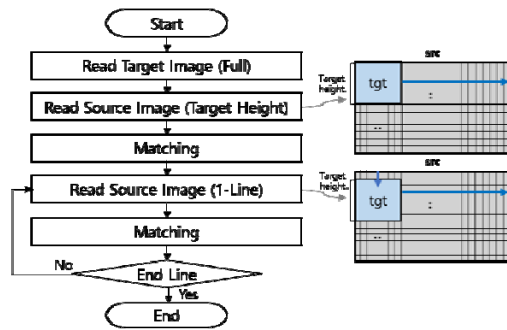


Fig. 5. Process of template matching

3. Hardware Implementation

We developed system implementation apparatus which consists of vision module, 3-axis motorized stage and controller, as shown in Fig. 6. The NCC template matching system is implemented and evaluated on Xilinx ZCU104 development board.

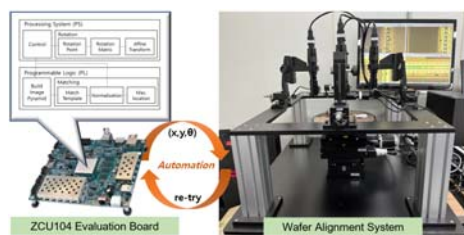


Fig. 6. Configuration of wafer alignment system

In the alignment system of Fig. 6, the alignment result values between the conventional reference software and the FPGA designed in this paper were summarized for five items. **Table 1** shows the result of commercial SW and **Table 2** shows the result of FPGA that can be seen that the processing time is faster. In addition, as a result of optimizing the angle step for the removal of the function and implementation of the hardware, the angle deviation was decreased.

Table 1. Alignment result of commercial SW

| Actual Angle (°) | X | Y | Measured Angle (°) | Angle Deviation (°) | Process Time(s) |
|------------------|---------|---------|--------------------|---------------------|-----------------|
| -0.2 | 717.843 | 538.492 | -0.12 | 0.07 | 0.507 |
| -0.3 | 715.856 | 536.346 | -0.35 | 0.05 | 0.550 |
| -0.4 | 713.991 | 535.304 | -0.35 | 0.04 | 0.610 |
| -0.5 | 713.104 | 534.033 | -0.45 | 0.04 | 1.370 |
| -0.6 | 711.041 | 533.288 | -0.58 | 0.01 | 0.619 |
| -0.7 | 710.036 | 532.254 | -0.58 | 0.11 | 0.662 |
| -0.8 | 707.032 | 530.048 | -0.91 | 0.11 | 0.823 |
| -0.9 | 706.048 | 529.048 | -0.91 | 0.01 | 0.534 |
| -1.0 | 704.996 | 528.332 | -0.90 | 0.09 | 0.767 |

Table 2. Alignment result of FPGA

| Actual Angle (°) | X | Y | Measured Angle (°) | Angle Deviation (°) | Process Time(s) |
|------------------|---------|---------|--------------------|---------------------|-----------------|
| -0.2 | 721.000 | 540.250 | -0.24 | 0.04 | 0.092 |
| -0.3 | 719.000 | 537.750 | -0.36 | 0.06 | 0.079 |
| -0.4 | 717.000 | 536.750 | -0.36 | 0.04 | 0.081 |
| -0.5 | 716.500 | 535.750 | -0.48 | 0.02 | 0.084 |
| -0.6 | 714.000 | 535.250 | -0.60 | 0 | 0.076 |
| -0.7 | 712.500 | 533.750 | -0.72 | 0.02 | 0.087 |
| -0.8 | 710.000 | 532.250 | -0.84 | 0.04 | 0.076 |
| -0.9 | 708.500 | 530.500 | -0.96 | 0.06 | 0.086 |
| -1.0 | 708.500 | 530.250 | -1.02 | 0.02 | 0.090 |

4. Conclusions

In this paper, we design a hardware architecture applying in FPGA based on the template matching software used in the wafer alignment system and present the results of experiments on actual equipment. In order to accurately and quickly process high-resolution images, it has limitation as software-based process. Therefore, high-speed template matching based on FPGA is required, and experimental results have confirmed its fast operation with minimal angular deviation.

References

- [1] J. KIM, "New Wafer Alignment Process Using Multiple Vision Method for Industrial Manufacturing," *Electronics*, Vol 7, No. 39, pp. 39, 2018.
- [2] M. Cong, X. Kong, Y. Du, J. Liu, "Wafer Pre-Aligner System Based on Vision Information Processing," *Information Technology Journal*, Vol 6, No.8, pp. 1245-51, 2007.

AI model deployment management capabilities for oneM2M based AIoT system

SeungMyeong Jeong and Chungjae Choe

Autonomous IoT Research Center, Korea Electronics Technology Institute
Seongnam, South Korea

[e-mail: sm.jeong@keti.re.kr, cjchoe1@keti.re.kr]

*Corresponding author: SeungMyeong Jeong

Abstract

IoT middlewares are now supporting AI features and being used in AIoT systems. As a core AI feature, AI model management functionality has been designed and implemented for model registration and deployment to a field device in AIoT systems. Compare to the previous works, in this work several aspects have been enhanced. Firstly, multi-model deployment is supported for one device. Secondly, REST resources are defined to manage model deployments which includes model running status. Also, implementations provided improved AI model management is a viable solution for AIoT systems.

Keywords: AI, AIoT, oneM2M, Radar Sensor, Radar Eye

1. Introduction

IoT (Internet of Things) has been converging with other technologies such as AI and digital twin. oneM2M as the IoT middleware global de facto standard is also evolving as AIoT (AI + IoT) platform in the upcoming Release 5. Many use cases were collected and new requirements were derived [1].

One of the key use cases that AIoT platform can offer is dataset management. To build AI-enabled services, datasets which are merge of multiple IoT sensor measurements are needed to train AI/ML models and inference inputs from the sensors [2]. An AIoT platform also provide an AI/ML model management feature. For AI-based services, IoT data is used for AI models performance measurements, and after some enhancements, new models are deployed again [3]. As the extension to oneM2M standard platform, AI model repository and deployment management capability for AIoT devices has been implemented [4]. However, it considered a

single AI model deployment to a device as the first prototype.

In this paper, we propose enhanced AI model deployment mechanism which can support multiple models for a device with one or more IoT sensor input(s).

2. AI Model Deployment Interface

2.1 Requirements

In the previous work, an AI model deployment command was simply represented with the resource identifier of the AI model resource under the AIoT device representing resource. In this work, multiple AI models can be deployed to a remote AIoT device which may have one or more IoT sensors attached. A deployment information can indicate the input data for inference which is a data sharing resource (e.g., <container> resource) and inference result storage as well as AI model resource identifier.

2.2 Resource design and procedures

The two types of resource types are defined for the model deployment: each deployment information and its collection. The resource type for the collection of deployments is *modelDeployments* (mdp) which has meta information for number of deployed, running, stopped models individually. The *deployment* (dpm) is the resource type for a deployment.

Table. 1. Attributes of an *<deployment>* resource

| attribute | description |
|----------------------|--|
| modelID (moid) | ID of the AI model resource |
| inputResource (inr) | ID of the inference input resource |
| outputResource (our) | ID of the inference result storage resource |
| modelStatus (mds) | AI model status (deployed, running, stopped) |

Table 1 summarizes the attributes of the *deployment* resource type. Like the previous work, AI model is contained in a separate resource and a deployment resource refers that model resource with *modelID*. Input resource can indicate several resource types. When *<container>* and *<timeSeries>* resources are indicated, their latest *<contentInstance>* and *<timeSeriesInstance>* resources are served as input to the model. This can be a raw data instances from an IoT sensor but also can be a dataset merging several sensor reading resources [2]. When a new *<deployment>* resource is created, the *modelStatus* is set as deployed by default. After that as a management command, user can send an update request for the resource which sets the status as “run” or “stop” then the status gets updated as “running” or “stopped”, respectively.

3. Implementation and Deployment

The new resource APIs implementation enabled interworking between the AIoT platform and the service portal. The AIoT device equipped one radar sensor and receives new deployments with the existing subscription/notification feature.

Fig. 1 shows how the new resource APIs can be used for AIoT devices. A preprocessed radar sensing data instance in “sensor1” resource is the same input data for three different model deployments.

This example describes the benefits of multiple model deployment support capability. With the same sensor measurement, material detection and counting model can separately run. Also, two counting models with different training parameters or datasets can run together and their inference results stored in different target containers. Then the performance can be compared later with the results stored in the AIoT platform.

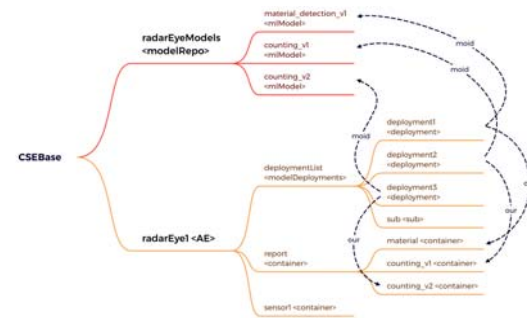


Fig. 1. Example deployment resources usage

4. Conclusions

In this paper, we present more flexible and practical AI model deployment management method than previous work. This was implemented as the additional oneM2M APIs and adopted into the AIoT platform. With this platform, more than one model can be deployed to an AIoT device which may have more than one sensor attached. Also, the inference input data can link to any data sharing resources including a dataset of multiple sensor measurement mashup as well as individual sensor reading resource.

References

- [1] oneM2M, “AI enablement to oneM2M (TR-0068) v0.10.0,” available at: <http://www.oneM2M.org>
- [2] S. Jeong, Y. Lee, D. H. Park, “Machine Learning dataset creation common service for real-time data platform,” *International conference on Internet 2021*.
- [3] C. Benzaïd, T. Taleb, J. Song, “AI-based autonomic and scalable security management architecture for secure network slicing in Beyond 5G Networks,” *IEEE Network*, Vol 36, No. 6, pp.165-174,

2022.

- [4] S. Jeong, D. Heo, C. Choe, “AI-enabled AGV service management with AIoT platform,” in *Proc. of Asia Pacific International Conference on Information Science and Technology (APICIST)*, 2023.

Sinusoidal Embedding for WiFi CSI-based Human Action Recognition

Taehyeon Kim[†] and Seho Park

Contents Convergence Research Center,
Korea Electronics Technology Institute,
Seoul, South Korea

[e-mail: taehyeon.kim@keti.re.kr, sehohpark@keti.re.kr]

*Corresponding author: Dr. Taehyeon Kim

Abstract

The ubiquity of WiFi networks presents an unobtrusive opportunity for human activity recognition, a significant facet in the progression of intelligent systems. This paper introduces an innovative technique that integrates Sinusoidal embedding with Multilayer Perceptron (MLP) neural networks to interpret Channel State Information (CSI) for the purpose of detecting human actions. Sinusoidal embedding transforms low-dimensional CSI subcarrier values into a high-dimensional feature space, capturing intricate signal variations, which in turn facilitates a more detailed representation for neural network processing. We evaluate the proposed method on the NTU-HAR dataset, consisting of six distinct human activities captured via WiFi CSI. The results exhibit a significant improvement, where the sinusoidal embedding-augmented MLP model achieved an accuracy of 98.7%, outperforming the conventional vanilla MLP model's accuracy of 96.9%. This enhancement underscores the potency of leveraging high-dimensional feature spaces to improve pattern recognition within CSI data. The demonstrated efficiency of our proposed model suggests its potential for seamless real-world implementation in intelligent environments for accurate and non-invasive action recognition, paving the way for future advancements in real-time and dynamic applications.

Keywords: WiFi sensing, channel state information, human action recognition, deep learning, neural network

1. Introduction

As the digital era advances, wireless communication systems are experiencing an ever-increasing demand for data traffic. In response, Multiple Input Multiple Output (MIMO) technology [1], coupled with Orthogonal Frequency Division Multiplexing (OFDM) [2], has evolved, delivering exceptionally high spectral efficiency. A prime benefit of this technology is the provision of Channel State Information (CSI), which intricately captures the signal path quality

between transmitting and receiving antennas. The comprehensive details within CSI assist in mapping signal propagation through various environments, leading to the development of sophisticated WiFi-based radar technologies that detect human movement using advanced signal processing and machine learning techniques [3].

Among the suite of tools for interpreting signals in wireless systems, the Multilayer Perceptron (MLP) is distinguished as a foundational neural network architecture. The MLP comprises several fully connected layers, where each neuron's output is processed through an

activation function. The journey begins with the input layer, which readies data for the hidden layers that follow, thereby incrementally refining the signal. This process culminates in the output layer, which categorizes the inputs into recognizable outputs.

This paper presents an innovative approach that utilizes Sinusoidal embedding techniques to convert CSI subcarrier values into high-dimensional features for MLP processing, particularly aimed at CSI-based action recognition. This method ingeniously overcomes the challenges posed by the lower-dimensional structure of conventional WiFi CSI subcarrier information, significantly enhancing the system's ability to identify human actions.

2. Proposed Method

The foundation of our method is the application of sinusoidal embedding to augment Channel State Information (CSI) derived from WiFi systems for human action recognition. This section elaborates on our proposed method, describing the embedding transformation process and the subsequent classification using a Multilayer Perceptron (MLP).

2.1 Sinusoidal Embedding of CSI Data

Sinusoidal embedding is a technique that projects low-dimensional data into a higher-dimensional space using sinusoidal functions. When it comes to WiFi CSI, each subcarrier's value is traditionally presented in a lower-dimensional space, potentially failing to capture the complex variability necessary for accurate action recognition. To rectify this, we apply a sinusoidal function to each subcarrier value, transforming it into a high-dimensional feature space. This embedding involves mapping the real and imaginary components of the CSI onto a sinusoidal wave, thereby encoding positional information in a way that preserves the data's periodicity and phase information.

The transformation is defined mathematically by:

$$F(x) = [\sin(x), \cos(x), \dots, \sin(kx), \cos(kx)],$$

where x is the CSI subcarrier value and k determines the dimensionality of the embedding. The choice of k is pivotal as it dictates the

granularity and the resolution of the embedding, balancing between computational efficiency and the richness of the feature representation.

2.2 WiFi Human Action Sensing

After transforming the CSI data via sinusoidal embedding, the high-dimensional features are fed into an MLP for classification. The MLP architecture consists of multiple layers: an input layer that receives the embedded features, several hidden layers that extract and abstract the complexities of human actions, and an output layer that associates the learned features with specific action categories.

The MLP utilizes a supervised learning approach to identify patterns within the sinusoidal-embedded CSI data that correlate with distinct actions. During training, the network adjusts its weights through backpropagation, reducing a loss function that measures the discrepancy between the predicted action categories and the actual labels.

2.3 Implementation and Training

In our implementation, the MLP includes three hidden layers with ReLU activation functions to introduce non-linearity, which enables the model to capture more complex functions. The output layer employs a softmax activation function, offering a probabilistic interpretation of the action predictions. The network is trained using a categorical cross-entropy loss function and optimized with an Adam optimizer. The hyperparameters are carefully tuned to optimize the balance between learning speed and convergence stability.

Through this proposed methodology, we demonstrate that sinusoidal embedding significantly enriches the feature space of WiFi CSI data, thereby allowing the MLP to detect subtle patterns indicative of human actions with increased accuracy. The subsequent sections will present experimental results that support the effectiveness of this method.

3. Experimental Result

We evaluated the sinusoidal embedding-based Multilayer Perceptron (MLP) model for human action recognition using WiFi Channel State

Information (CSI) on the NTU-HAR dataset, which consists of WiFi CSI measurements for six different human activities.

3.1 Dataset Description

The NTU-HAR dataset includes WiFi CSI data collected from three antennas, with each CSI measurement consisting of 114 subcarriers sampled over 500 time intervals. The dataset is categorized into six classes that represent different actions: boxing, circling, cleaning, falling, running, and walking. This classification allows for effective training and evaluation of action recognition models. For our experiments, the dataset was divided into 936 training samples and 264 testing samples, ensuring ample data for the model to learn from and a separate set for assessing performance.

3.2 Model Performance and Comparison

We compared two models: the proposed WiFi CSI-based human action recognition model incorporating sinusoidal embedding, and a conventional vanilla MLP model. The proposed model was designed to transform the 3 x 114 x 500 CSI data into a high-dimensional feature space prior to classification by the MLP, whereas the vanilla MLP processed the CSI data without any embedding.

The experimental results are summarized as follows:

- ✓ The proposed sinusoidal embedding-based MLP model achieved an impressive action recognition accuracy of 98.7% on the test set.
- ✓ The vanilla MLP model, without sinusoidal embedding, demonstrated a lower accuracy of 96.9%.

These results demonstrate a significant advantage of the proposed method over the conventional approach, underscoring the effectiveness of sinusoidal embedding in enhancing the representation of human actions from WiFi CSI data.

3.3 Experimental Result Analysis

The improvement of 1.8 percentage points in accuracy with the proposed model highlights the potential of sinusoidal embedding to enrich the

feature representation of WiFi CSI data. The generated high-dimensional space from the embedding allows for a more detailed and complex representation of the data, enabling the MLP to differentiate subtle distinctions between various actions. The precision improvement is especially important in applications where accuracy is critical, such as fall detection systems for the elderly.

Additionally, the high level of accuracy achieved by the proposed model suggests its suitability for real-world deployment in scenarios where non-intrusive action recognition is beneficial. The model's robustness in accurately differentiating actions promises advancements in the development of intelligent environments that can respond to human activities.

4. Conclusions

In conclusion, this work has verified the enhanced capability of a sinusoidal embedding-augmented MLP over a vanilla MLP in recognizing human actions from WiFi CSI data, improving accuracy from 96.9% to 98.7%. This underscores the value of high-dimensional feature spaces in capturing complex patterns within CSI signals. The practical outcomes of these findings offer promising applications in intelligent environments where precise and non-intrusive action recognition is essential. Future endeavors will focus on optimizing this approach for real-time applications and expanding its applicability to more complex and dynamic settings.

References

- [1] Lu, Lu, et al. "An overview of massive MIMO: Benefits and challenges." *IEEE journal of selected topics in signal processing* 8.5 (2014): 742-758.
- [2] Armstrong, Jean. "OFDM for optical communications." *Journal of lightwave technology* 27.3 (2009): 189-204.
- [3] Yang, Jianfei, et al. "SenseFi: A library and benchmark on deep-learning-empowered WiFi human sensing." *Patterns* 4.3 (2023).

Melody Extraction for Non-Vocal Music based on Tone-Octave Network

Dalwon Jang, Daeun Chae, Jaewon Lee, and JongSeol Lee

Communications and media R&D Division

Korea Electronics Technology Institute

[e-mail: dalwon@keti.re.kr, daeun@keti.re.kr, jwlee0121@keti.re.kr, leejs@keti.re.kr]

*Corresponding author: Dalwon Jang

Abstract

This paper presents an experimental results of melody extraction for non-vocal music. Many melody extraction algorithms are interested in singing melody, but our concern is non-vocal music. To extract melody of non-vocal music, experiments using various training sets and an existing algorithm are performed. In our experiments, it is verified that changing training set with the algorithm, which is developed for singing melody, lead to performance enhancement for non-vocal music.

Keywords: Melody extraction, non-vocal music

1. Introduction

Automatic melody extraction extracts frequency-based representation of the dominant melody (i.e., frequency values of the frames) from music clips, generally polyphonic music recordings [1,2]. The challenges and approaches of early researches can be found in [2]. The early researches commonly are based on spectral representation and peak selection [2]. The recent researches are based on training-test structure [1,3,4]. Since there are many melodies in polyphonic music, for successful melody extraction both finding dominant melody and extracting frequency value from the melody are essential. The definition of dominant melody is not clear. It can be differed depending on the music. However, recent researches focus singing melody extraction [3,4], thus their performance for non-vocal music will not be acceptable. With the state-of-the-art algorithm which aims to extract singing melody, we tested melody

extraction of non-vocal music clips. We used Tone-Octave Network (TONet) algorithm [3], and it is thought that its property of extracting singing melody comes from training set, not the structure of the algorithm. Thus, we tested various training sets including non-vocal music with the algorithm, and the results are presented in this paper.

2. TONet algorithm

TONet algorithm was proposed in 2022, and it improves both tone and octave perceptions introducing three new ideas [3]. First, based on the combined frequency and periodicity representation (CFP), Tone-CFP (TCFP) is introduced. With TCFP, convolutional layer can capture harmonic relationships in its receptive field. Second, new scheme that deals with tone and octave separately is proposed. Third, a variety of existing backbones can be applied to TONet, and the TONet allows the use of three

different backbones: MCDNN [8], MSNet [9], and FTANet [4]. The results in [3] show TONet improve the singing voice extraction performance. However, it can not find the results related to non-vocal melody extraction.

3. Experimental results

We used various combinations of existing datasets for training, and non-vocal music clips in ADC2004 and MIREX05 datasets are used for test. Various training sets used in our experiments are shown in Table 1. Among existing datasets, MIR-1K is much bigger than others and it does not have non-vocal music [5]. There are vocal and non-vocal music in Medley DB [6]. ORCHSet[7] and TinySOL have only non-vocal music data, but their sizes are relatively very small. Training set T1 is same as training set used in [3], which means T1 is suitable for vocal melody extraction.

Table 1. Training dataset (V means Vocal, and NV means Non-vocal.)

| | | |
|----|------|--------------------------------|
| T1 | V | MIR-1K + Medley(V) |
| T2 | NV | Medley(NV) |
| T3 | V+NV | MIR-1K + Medley(all) |
| T4 | NV | ORCHSet+TinySOL |
| T5 | NV | ORCHSet+ Medley(NV) |
| T6 | V+NV | ORCHSet+ Medley(NV)+ MIR-1K |

In our experiments FTANet was used as a backbone model of TONet. In [3], results for three backbone models are presented, and the performance using FTANet is better than others. In general, various metrics including raw pitch accuracy (RPA), raw chroma accuracy (RCA), voicing recall (VR) and voicing false alarm (VFA) have been used in melody extraction. But, in this paper only overall accuracy (OA) is used. In the literature, OA is often considered more important than other metrics [2,3]. The experimental results are shown in Table 2. The case using T1, which is well performed for vocal melody extraction [3], shows bad result. Since data size is too small T4 leads to the worst result. Training with only non-vocal data leads to the better results than training with T1. It is varified that the property of training data, either vocal of non-vocal, has an effect on the performance. However training data size also

plays a prominent part. We cannot obtain sufficient non-vocal data, thus T6 which combines MIR-1K and T5 gets the best result.

Table 2. Overall accuracies

| Training set | OA |
|--------------|---------|
| T1 | 62.112 |
| T2 | 66.8845 |
| T3 | 70.6898 |
| T4 | 52.9815 |
| T5 | 71.2585 |
| T6 | 73.1451 |

4. Conclusions

In this paper, we present the experimental results of non-vocal melody extraction. Based on TONet algorithm, which aims to extract singing melody, the performance of extracting non-vocal melody is studied. The original version of TONet, which is trained using vocal dataset leads to unsatisfactory result. But the performance gets better with training of non-vocal music dataset. Merging the more datasets and verifying the results after training with the dataset are our future works.

References

- [1] J. Lee, D. Jang, and K. Yoon, "Automatic melody extraction algorithm using a convolutional neural network," KSII Trans. Internet Inf. Syst., vol. 11, no. 12, pp. 6038–6053, 2017.
- [2] J. Salamon, E. Gomez, D. P. W. Ellis, and G. Richard, "Melody extraction from polyphonic music signals: Approaches, applications, and challenges," IEEE Signal Processing magazine, 2014
- [3] K. Chen, S. Yu, C.-i. Wang, W. Li, T. Berg-Kirkpatrick, and S. Dubnov, "Tonet: Tone-octave network for singing melody extraction from polyphonic music," Proc. ICASSP, 2022
- [4] S. Yu, X. Sun, Y. Yu, and W. Li, "Frequency-temporal attention network for singing melody extraction," in ICASSP 2021.
- [5] C.-L. Hsu and J.-S. Jang, "On the improvement of singing voice separation for monaural recordings using the MIR-1K

- dataset,” *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 18, no. 2, pp. 310–319, Feb. 2010.
- [6] R. M Bittner, J. Salamon, M. Tierney, M. Mauch, C. Cannam, and J. P. Bello, “Medleydb: A multitrack dataset for annotation intensive mir research,” in *ISMIR 2014*.
 - [7] J. J Bosch, R. Marxer, and E. Gómez. “Evaluation and combination of pitch estimation methods for melody extraction in symphonic classical music,” *Journal of New Music Research*, 45(2):101–117, 2016.
 - [8] S. Kum, C. Oh, and J. Nam, “Melody extraction on vocal segments using multicolumn deep neural networks,” in *ISMIR 2016*.
 - [9] T.-H. Hsieh, L. Su, and Y.-H. Yang, “A streamlined encoder/decoder architecture for melody extraction,” in *ICASSP 2019*.

Study on Video Surveillance System Data Interface for Interoperability with Heterogeneous Physical Security System

Sungjoo Park and Dongchil Kim
Information Media Research Center, KETI
Seoul, South Korea
[e-mail: bpark@keti.re.kr, dckim@keti.re.kr]
*Corresponding author: Sungjoo Park

Abstract

This paper proposes a video surveillance system data interface for interoperability of data, video data and event analysis results, with a heterogeneous physical security system that can't acquire the video-related data. In this paper, the heterogeneous physical security system means two systems: biometrics-based physical access control system and security sensor system. In detail, it describes the range, data, and requirements for interoperability between video surveillance systems and heterogeneous physical security systems, and defines the architecture, processes, and messages. The interoperability process is defined by key steps and functions, and the detailed specifications of messages are defined in data formats for request and response data for data interoperability.

Keywords: Video surveillance system, interoperability, data interface, physical security system

1. Introduction

Existing video surveillance systems have a single system structure that provides security services by analyzing objects and events using CCTV images. Although a deep learning has improved its performance significantly, false alarm and non-detected errors still occur. In order to overcome these performance limitations and provide more diverse and intelligent security services, a linked system structure that utilizes externally data and control information based on connections with heterogeneous sensors, devices, and platforms is attracting attention. In order to implement such an interconnected security system, it is necessary to transmit the data held by individual systems to the outside in a standardized way, and at the same time receive necessary information from the outside and apply it to new services [1].

This paper proposes a data interface standard for interoperability of video and video analysis data and defines data interface specifications for collection, analysis, control, and authority between video surveillance systems and heterogeneous physical security systems. In detail, it describes the scope of interoperability, interoperability information, and requirements between video surveillance systems and heterogeneous physical security systems, and defines interoperability structure, procedures, and messages. The interoperability procedure is defined by major steps and functions, and the interoperability message defines the detailed specifications of messages sent and received for data interoperability.

2. Interoperability of Video Surveillance System

2.1 Interoperability Information

Interoperability information in the video surveillance system is divided into raw data, analysis data, and control information, as shown in Fig. 1. The collected data from video camera devices is images, and information such as camera ID and video access URL is interconnected. Analysis data provides video-based event analysis information for abnormal situations such as arson and loitering, and information such as camera ID, event occurrence time, event name, and event UUID is interconnected. Control and management information refers to control/management information of a controllable video camera device. Control information includes detailed control function information such as PTZ control, and management information includes information such as camera ID and camera installation location.

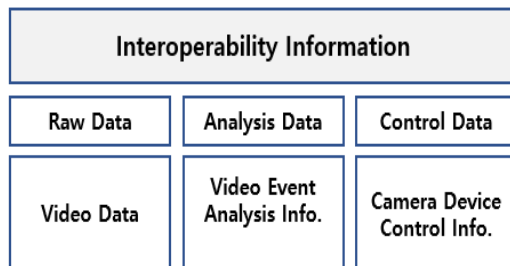


Fig. 1. Interoperability Information

2.2 Interoperability Structure

The interoperability structure for data interoperability between video surveillance systems and heterogeneous physical security systems is structured as shown in Fig. 2 based on interoperability interfaces and communication protocols. Defined procedures and message formats are used for interoperability based on the HTTP protocol, and device interconnect interfaces and product protocols supported by the camera are used for interoperability between video surveillance systems and video camera devices.

2.3 Interoperability Procedure

The detailed procedure for data interoperability

between video surveillance system and heterogeneous physical security systems consists of four steps, as shown in Fig. 3. The interoperability process begins after exchanging registration and interoperability information between systems that support interoperability. Step 1 is to authenticate the user for interoperability, Step 2 is to exchange system information, Step 3 is to exchange video data, and Step 4 is to exchange video analysis information.

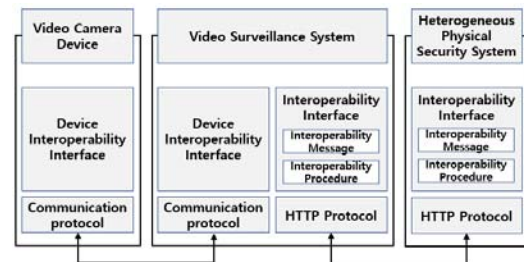


Fig. 2. Interoperability Structure

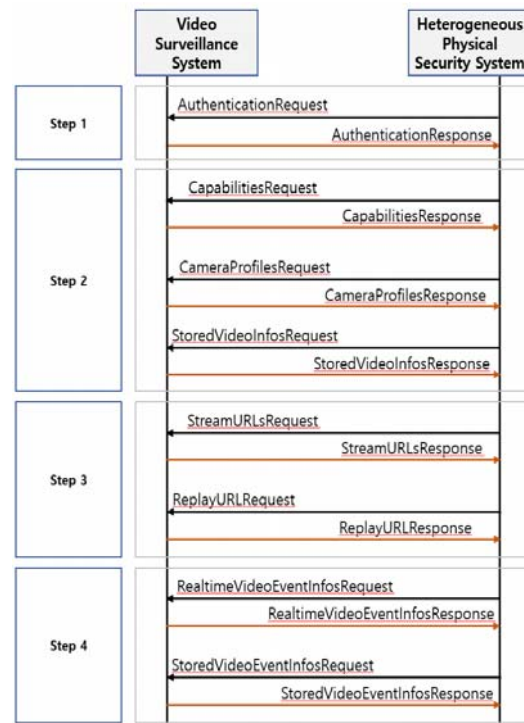


Fig. 3. Interoperability Procedure

2.4 Interoperability Message Format

Messages used to interoperability video surveillance systems are defined as shown in

Table 1. The processing results of detailed messages in this paper are defined using the response codes defined in **Table 2**.

Table 1. HYPER-PARAMETERS VALUES EXPLORED

| Functions | Message Format |
|--------------------------------|---------------------------------|
| User authentication | AuthenticationRequest |
| | AuthenticationResponse |
| Transfer support features | CapabilitiesRequest |
| | CapabilitiesResponse |
| Camera list | CameraProfilesRequest |
| | CameraProfilesResponse |
| Stored video list | StoredVideoInfosRequest |
| | StoredVideoInfosResponse |
| Stored video transmission | StreamURLsRequest |
| | StreamURLsResponse |
| Real-time video transmission | ReplayURLRequest |
| | ReplayURLResponse |
| Real-time video event analysis | RealtimeVideoEventInfosRequest |
| | RealtimeVideoEventInfosResponse |
| Stored video event analysis | StoredVideoEventInfosRequest |
| | StoredVideoEventInfosResponse |

Table 2. COMMON RESPONSE CODES

| Code | Description |
|------|----------------------|
| 200 | Success |
| 201 | No information |
| 400 | Wrong request |
| 401 | Authentication error |
| 402 | Token expiration |
| 403 | No permission |
| 404 | No device |
| 405 | Not supported |
| 501 | Server error |
| 502 | DB error |
| 999 | Other errors |

3. Conclusions

In this paper, we proposed a data interface standard for interoperability of video and video analysis data and defined data interface specifications for collection, analysis, control between video surveillance systems and heterogeneous physical security systems. And we described the scope of interoperability, interoperability information and interoperability structure and messages formats.

Based on the interoperability standards defined in this paper, it is expected that an efficient data interoperability will be possible. It is expected that this will enable the implementation of more diverse and intelligent security services.

Acknowledgment

This work was supported by the ICT R&D program of MSIT/IITP. [2019-0-00203, Development of 5G-based Predictive Visual Security Technology for Preemptive Threat Response]

References

- [1] "Interface for Interoperability between Video Management System and External Video Surveillance System," *TTA, TTAK.OT-10.0406*, 2016.

A Study on Reactive Power Characteristics and Compensation According to Voltage Regulation in Low Voltage Premises System

Seunghyeon Park, Kiwoong Kwon, Sanghun Kim, and Yongho Kim

Energy IT Convergence Research Center, Korea Electronics Technology Institute
Seongnam, South Korea

[e-mail: {park.seunghyeon, kiwoong.kwon, ksh7150, and yhkim}@keti.re.kr]

Abstract

In this paper, we talk about reactive power characteristics and compensation according to voltage regulation in low voltage premises systems. Through 10% voltage regulation simulation, the change in reactive power required by the load is examined and the need for reactive power supply is confirmed. In addition, methods of reactive power supply are examined, and in the process, what form of reactive power supply is currently being performed in the premises system, and how the experimental environment is configured and compared. Finally, we briefly introduce the design of the reactive power compensator to be tested directly in the future work.

Keywords: Reactive Power, Low Voltage, Conservation Voltage Reduction, Static Var Generator, Inverter

1. Introduction

In recent years, there has been an increasing demand for a conservation voltage reduction technology for controlling load equipment to be operated at an optimal efficiency voltage for the purpose of improving energy efficiency in the industry. At this time, it is also important to maintain the operating characteristics of the load while taking advantage of power saving through voltage regulation. Therefore, in this paper, We would like to introduce research on reactive power characteristics according to voltage regulation in low voltage premises systems(industry, builing) and talk about reactive power compensation, which is one method of voltage regulation.

2. Reactive Power Characteristics According to Voltage Regulation

The Conservation Voltage Reduction (CVR) is a principle in American National Standards Institute (ANSI) Standard C84.1 that sets the voltage range of auxiliary terminals of transformers so that an acceptable voltage band can operate without harming household appliances [1]. For the simulation of normal operation, the induction motor most widely used in industrial sites was selected as the representative system for the industrial premises. The load was simulated by dividing it into two cases, a constant impedance load and a constant torque load, and the configuration is shown in Figure 1 below.

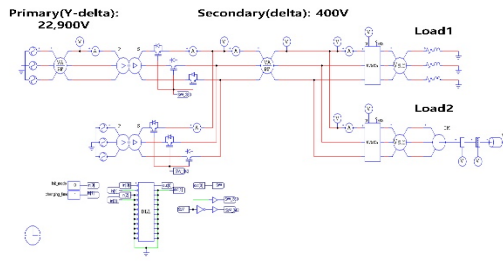


Fig. 1. PSIM Simulation schematic

As a result of the simulation, voltage regulation is achieved and the benefit of reducing the active power under constant impedance load can be seen in **Fig. 2**. At this time, when 22.9kV is stepped down to 400V and controlled at 360V through 10% voltage regulation, it is a condition of the premises system where an induction motor with an output of 200kW is driven.

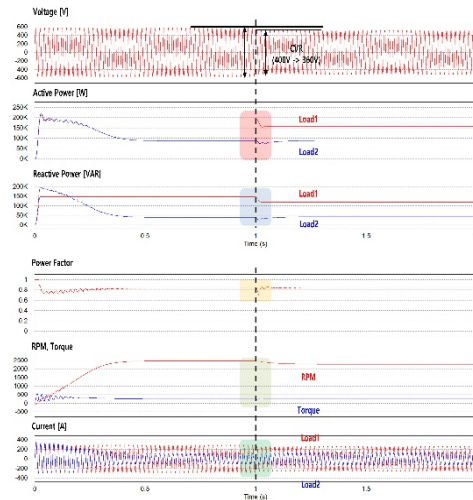


Fig. 2. PSIM Simulation result graph

At this time, it is necessary to supply increased reactive power from the reactive power compensator to maintain a constant torque. In the case of an induction motor with constant torque operation, reactive power increases, which is a characteristic that appears due to an increase in reactive current to maintain a constant torque. It can also be seen that the power factor of the entire premises system may be deteriorated due to the increased reactive power, and thus reactive power is required to be supplied at the same time at the transformer stage where voltage regulation is performed.

3. Currently Operating Reactive Power Compensation Method

3.1 Capacitor Bank

A capacitor bank applied to the building's premises system is first installed at the substation equipment stage to supply reactive power, and a capacitor bank additionally installed in the equipment with a large load at the distribution board to compensate for the reactive power last.

3.2 Limitation of Capacitor Bank

As a result of checking the energy diagnosis report of the building, which is the premises system, it was managed with an average power factor of 99% and 98% in 2019 and 2020, respectively. However, the power quality between the distribution board and devices in the distribution board of the air conditioning room where the air conditioning facilities on the 4th floor are in operation. We can see that the power factor drops to about 0.7 after the power factor improvement point.

4. Proposed Reactive Power Compensation Method

4.1 Compensation Plan

Recently, inverter type SVG (Static Var Generator) devices are increasing. It is located close to the load and has the advantages of fast response to required reactive power, high efficiency, continuous linear compensation, power factor improvement, as well as distortion and harmonics improvement [3]. The SVG reference model was installed close to the air handling unit AHU (Air Handling Unit) to observe changes in reactive power. In Figure 3 below, the power factor was 0.715 during AHU operation, and it can be seen that the power factor is improved to about 0.971 when the target value of SVG is set to 0.95.



Fig. 3. SVG reference model operation status

4.2 Determination of Range

Among domestic low-voltage induction motors, the maximum output of motors in the economical range was set at 200kW. In addition, the reactive power adjustment range and capacity were based on the case where the efficiency and power factor values before and after improvement were substituted.

$$Q = P(\tan \theta_1 - \tan \theta_2)$$

(At this time, since P corresponds to the input of the motor, not the motor output, the efficiency must be reflected.)

$$Q = 200 \times 1.1 \times (0.90 - 0.80) = 22 \text{ [kVAR]}$$

If a reactive power compensator with a capacity of 50 kVAR is determined, reactive power equivalent to two 200 kW motors can be supplied.

4.3 Specifications of Compensator

Compared to the 2-level inverter, the 3-level inverter has a smoother output waveform, lower switching loss and higher level, which has advantages in waveform, voltage, and loss due to the operating characteristics of the elements, but also has the disadvantage of increasing the number of elements [2]. An Active Neutral Point Clamped (ANPC) inverter has a structure in which the neutral point diode of an NPC (Neutral Point Clamped) inverter is changed to a switching element such as an IGBT, and can actively determine the current path.

Therefore, in this paper, the ANPC structure, which can reduce the advantages of multi-level and harmonics and heat, is used as an inverter topology. Assuming that the rated voltage on the DC side is 800 V in determining the IGBT specifications, design the specifications with an IGBT module rated at 650V and 150A with a safety margin.

The harmonic reduction design through the LCL filter to meet the voltage distortion standard is shown in below. Each element represents the inductance of the inverter side, the inductance of the grid side, the capacitance of the filter, and the equivalent series resistance, and is calculated based on the maximum allowable range in calculating the capacity of the elements.

$$L_1 = 43.86 \times 10^{-6} \text{ [H]}$$

$$C = 23.21 \times 10^{-6} \text{ [F]}$$

$$L_2 = 2.08 \times 10^{-6} \text{ [H]}$$

$$F_{res} = 23.44 \times 10^3 \text{ [Hz]}$$

It can be seen that the resonant frequency F_{res} falls between 10 times the grid frequency and 1/2 times the switching frequency.

5. Conclusion

In this paper, we talked about reactive power characteristics and compensation according to voltage regulation in low voltage premises systems. The required reactive power increased according to voltage regulation, and among various methods, it was confirmed that the SVG device capable of supplying reactive power close to the load linearly compensated the power factor up to the terminal device the best.

In the future work, with the selected topology and filter, we intend to fabricate a voltage inverter type reactive power compensator using power electronic devices such as IGBT (Insulated Gate Bipolar Transistor). In addition, with the manufactured reactive power compensator, various control methods and scenarios are planned to expand into research on cooperative control of reactive power for effective CVR in low voltage premises systems.

References

- [1] ANSI, "ANSI Standard C84.1-1995 Electric Power Systems and Equipment Voltage Ratings (60 Hz)," 1995.
- [2] R. Teichmann and S. Bernet, "A comparison of three-level converters versus two-level converters for low-voltage drives, traction, and utility applications," in *IEEE Transactions on Industry Applications*, vol. 41, no. 3, pp. 855-865, May-June.
- [3] M. Wu, A. Tang and Z. Cheng, "Design and Implementation of a New Type of Static Var Generator," *2021 IEEE 5th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*, Chongqing, China, pp. 668-672, 2021.

Enhancing Blockchain Platform Security through the Analytics of User-Centric Psychometric Paradigm

Jihye Park, Juno Lee and Hangbae Chang*

Department of Security Convergence, Chung-Ang University

Department of Industrial Security, Chung-Ang University*

Seoul, South Korea

[e-mail: pjh06023@cau.ac.kr, iamjuno95@cau.ac.kr, hbchang@cau.ac.kr*]

*Corresponding author: Hangbae Chang

Abstract

Blockchain, a key technology in the Fourth Industrial Revolution(4IR), is advancing by providing users with high reliability through decentralized ledger management and consensus. While its application is expanding in various fields, including finance, logistics, healthcare, and public sectors, it brings about new security risks. Despite the adoption of blockchain as a means to protect user privacy, security risks related to user behavior and psychological states persist.

In this context, this study focuses on the psychological characteristic of risk perception among users. Through the analysis of the psychometric paradigm, practical strategies and methods to enhance security from the usability perspective of blockchain platforms are proposed. Additionally, by categorizing users into experts, laymen, and students, diverse security enhancement measures are compared and suggested. The aim is to understand the psychological states and behaviors of various users, laying the foundation to design security mechanisms applicable to blockchain platforms.

Ultimately, this research contributes academically by combining psychological measurement analysis with security research, presenting a multidisciplinary perspective. It also aims to practically offer user-centric security enhancement strategies in blockchain platforms, contributing to the development of user-friendly security policies in future corporate and organizational contexts.

Keywords: Security, Blockchain, Psychometric Paradigm, User-Centric Approach, User Experience

1. Introduction

Blockchain, a core technology of the Fourth Industrial Revolution, plays a crucial role in ensuring the integrity of data. Particularly in the traditional financial sector, it serves as a means to address security issues by providing users with a high level of reliability through decentralized ledger management and consensus mechanisms

[1]. As the utilization of blockchain expands into various domains, the innovation of blockchain technology concurrently introduces new security challenges [2]. While blockchain is employed as a tool to protect user privacy, security risks stemming from user behavior and psychological states persist.

Recognizing the importance of a proactive approach to enhance user-centric platform

“This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the ITRC(Information Technology Research Center) support program(IITP-2023-2018-0-01799) supervised by the IITP(Institute for Information & communications Technology Planning & Evaluation)”

security, this research focuses on the perception of risk among users and acknowledges the need for practical strategies. In doing so, we aim to present practical approaches and methods to enhance the usability of blockchain platforms through an analysis of psychological measurement paradigms. Additionally, by categorizing users into experts, the laymen populace, and students, we compare and analyze users, proposing diverse security enhancement measures. Through these efforts, we aspire to contribute to the establishment of user-friendly security policies and strategies in future blockchain platforms.

2. Background of Research

2.1 Security Risks in Blockchain Platforms

Public and private blockchain systems, exemplified by platforms like Ethereum, Hyperledger Fabric, and Corda, are considered to enhance security. However, associated security issues have revealed inherent limitations [2]. Notably, it is crucial to consider that the primary target of security threats in blockchain platforms is the 'user.'

For instance, public blockchains like Bitcoin (BTC) have the drawback of storing confidential information in ledgers accessible to anyone, highlighting unsuitability for safeguarding sensitive data. Additionally, certain blockchains present potential risks related to the protection of personal information.

2.2 Risk Perception from the Perspective of the Psychometric Paradigm

Risk is understood as the combination of the likelihood and magnitude of undesirable outcomes. In this context, risk is subjectively defined by individuals influenced by psychological, social, and institutional factors, based on their subjective perceptions of the severity of potential concrete events, the probability of occurrence, personal vulnerability, and other related factors.

Additionally, the psychological concept of Perceived Risk involves users of blockchain platforms defining their attitude towards the intensity and likelihood of risky events they may experience [3,4,5].

3. Scope and Methods

3.1 Scope and Methods

This study aims to explore the psychological aspects of security risks arising from blockchain platforms and, for this purpose, utilizes a scale derived from previous research to measure users' risk perceptions. The study focuses on experts, the laymen public, and students as participants, assuming them to be users of blockchain platforms. Students include those from relevant majors, while the laymen public encompasses both non-experts and non-student individuals [2,7,10,11].

The research methodology involves a literature review based on prior studies, followed by a Google survey to conduct quantitative research using concepts derived from the literature analysis. Utilizing STATA ver 17.0, the study performs descriptive statistics analysis, including basic statistics, correlation, and factor analysis. Finally, hypothesis testing through comparative analysis (t-test) will be conducted to analyze differences among the three groups.

4. Proposed Research Design through Literature Review

4.1 Identification of Security Risk Factors in Blockchain Platforms

This study adopted eight sub-risk factors derived from the top four factors (Transaction Validation and Consensus, Privacy and Access Control, Blockchain Software Security, Blockchain Service Security) based on previous research on security risks in blockchain platforms [1,2,6].

4.2 Risk Perception Scale from a Psychometric Perspective

The Psychometric Paradigm is recognized as the most influential model in the field of risk analysis. Consequently, the scale for risk perception in the psychological dimension has been partially adopted in the context of this study, structured using a Likert 5-point scale. This incorporates aspects from 10 psychological traits (Spontaneity, Immediacy, Personal Knowledge, Scientific Knowledge, Controllability, Familiarity, Fear, Severity, Intensity of

Perceived Risk, Degree of Social benefit) referenced in prior research [2,7,8,9].

4.3 Research Questions and Hypotheses

The Research Questions(RQ) for the Hypotheses(H) to be investigated in this study are as follows: <Table 1.>.

Table 1. The Research Questions for Hypotheses

| Category | RQs for Hs |
|----------|---|
| RQ 1. | In terms of the perception of security risks in blockchain, are there any differences in psychometric measurements among societal groups (experts, laymen, students)? |
| RQ 2. | In the perception of security risks in blockchain, are there any differences in the perceived intensity of risk and understanding of social benefits among societal groups (experts, laymen, students)? |

5. Conclusions

5.1 The Expected Results

The comparative analysis of perceived risks on blockchain platforms among research subjects suggests differences in risk perception levels between experts with direct experience in blockchain technology and individuals, including students, who may lack a comprehensive understanding of severity.

5.2 The Expected Implications

The anticipated effects of this study are as follows:

Firstly, by addressing 'Security Risks in Blockchain Platforms' as an academic topic, we expect to provide academic insights by extracting critical security risk factors from a usability perspective.

Secondly, through the comparative analysis among user groups(experts, laymen, and students), we aim to find commonalities representing security risks in blockchain technology among these three groups. By balancing practicality and academia, we anticipate offering 'applicable' practical implications and presenting pragmatic solutions.

References

- [1] Dae Sung Lee, "Next Generation Blockchain Trends," The Journal of Korea Institute of Information and Communication, Vol 27, No. 10, pp. 1308-1311, 2023.
- [2] Hee Youl Kim, "Analysis of Security Threats and Countermeasures on Blockchain Platforms," The Journal of Korean Institute of Information Technology, Vol 16, No. 5, pp. 103-112, 2018.
- [3] Slovic, P. "The Perception of risk," Earthcan, London, 2000.
- [4] Korea Occupational Safety & Health Agency, "Safety Culture Guide 2, (Chapter 3. Creating a safety culture through risk perception management)," 2022.
- [5] Darker, C., and Whittaker, A. C., "Risk perception," In Encyclopedia of behavioral medicine, pp.1-3, 2018.
- [6] Jung Sook Kim, "Service Status and Problem Analysis Based on BlockChain," The Society of Convergence Knowledge, Vol. 6, No. 1, pp.135-140, 2018.
- [7] Young Ai Lee, Na Keung LEE, "Psychological Dimensions of Risk Perception," Journal of Cognitive Science, Vol. 16, No. 3, pp.199-211, 2005.
- [8] Siegrist, M., Keller, C., & Kiers, H. A. "A new look at the psychometric paradigm of perception of hazards," Risk Analysis: An International Journal, Vol. 25, No. 1, pp.211-222, 2005.
- [9] Fischhoff, B., Slovic, P., Lichtenstein, S., Read, S., & Combs, B. "How safe is safe enough? A psychometric study of attitudes towards technological risks and benefits," Policy sciences, Vol. 9, No. 0, pp.127-152, 1978.
- [10] Ik Jae Chung, So Won Ahn, "Analysis of Risk Perception and Cognitive Map of Risk: Application of Facilitated Modelling," Korean Journal of Psychology: General, Vol. 35, No. 3, pp.481-503, 2016.
- [11] Slovic, P., Fischhoff, B., & Lichtenstein, S., "Behavioral decision theory perspectives on risk and safety," Acta psychologica. Vol. 56, No. 1-3, pp.183-203, 1984.

Experimental Verification of Temporal Activity Sequencing Anomaly Detection Algorithm in Business Process

Do-Hyeon Lee¹, Jin-Hyeong Lee¹, Eunbi Cho², Tsagaantsooj Batzaya², Thanh-hai Nguyen³, and Kwanghoon Pio Kim⁴

¹Department of Computer Engineering, Kyonggi University

²Department of Public Safety Bigdata, Kyonggi University

³Department of Computer Science, Kyonggi University

⁴Division of AI Computer Engineering, Kyonggi University

Seoul, South Korea

[e-mail: {ehgus8, jaims2289, eunbee0508, tsbaachka95, nguyenthanhhai, kwang}@kyonggi.ac.kr]

*Corresponding author: Kwanghoon Pio Kim

Abstract

This paper addresses the increasingly important domain of anomaly detection within business process management through the lens of process mining. We use an LCS (Longest Common Subsequence) algorithm-based approach for measuring the similarity between business process log data and subsequently calculating an Anomaly Score. Our methodology leverages a normalized LCS to determine deviations in process execution logs effectively. By applying statistical methods, specifically the Prediction Interval, we classify the log data into normal and anomalous categories. We demonstrate the effectiveness of our approach by comparing the Information Control Network structures before and after the classification process. The results show a significant reduction in the complexity of the structures, indicating that our anomaly detection method can successfully refine business process data for better clarity and understanding. Our findings suggest that the LCS-based anomaly detection framework is a powerful tool for enhancing the quality of process mining and reengineering efforts, leading to more streamlined and efficient business processes.

Keywords: Business Process, Process Mining, Anomaly Detection

1. Introduction

In recent times, enterprises and organizations have been achieving informatization of their organizations using Business Processes. This process-centric informatization technology consists of Business Process Management and Work Flow technology. A Business Process signifies the workflow of enterprises and organizations, which is recorded as log format data. With the informatization and enterprise information management being based on Business Processes, i.e., Work Flow, there has

arisen a need to address issues of productivity and efficiency improvement through Business Process Reengineering and automation. Therefore, enterprises and organizations achieve productivity gains by reengineering and automating processes, based on the analysis of process log data that includes information about the events, timing, and performers of the work process. To structurally analyze Business Processes, it is possible to create visualized models of processes through process mining from the process event log data generated by enterprises and organizations. Inefficiencies in

the existing process model can be identified through analysis based on the visualized model obtained through existing process mining techniques. However, the existing process execution history data contains anomalies, which complicates the process visualization model in the mining process. Therefore, this study aims to solve the problem of complex structures that make analysis difficult in the visualized model by refining the log data based on a Business Process Anomaly Detection algorithm, and performing process mining on the refined data. This approach will be verified through the study.

2. Related Work

2.1 Business Process Anomaly Detection

In this study, the method from a paper by author Ashish Sureka was used for anomaly detection in business processes.[1] The algorithm used in this paper is the Longest Common Subsequence (LCS) algorithm. By utilizing this algorithm, the similarity between business process log data can be measured, and from this similarity, an Anomaly Score can be calculated, which serves as a criterion to classify data as anomalous.

2.2 Process Modeling

Through process mining from process event log data, an Information Control Network of the process can be created by various algorithms, e.g., rho algorithm.[2] An Information Control Network is a theoretical mathematical model for process modeling.[3] There are various static and dynamic analysis techniques centered around the Information Control Network model. Alongside the algorithms used in this study for generating Information Control Networks from process event log data, it's also crucial to recognize foundational methods like the Alpha algorithm. These methods form the basis for developing Petri Net models in process mining.[4]

3. Data Preprocessing

3.1 Dataset

The datasets used in this study are commonly used in process analysis and process mining, namely Control summary and PrepaidTravelCost.

These datasets are in the XES (eXtensible Event Stream) format. In such XES format files, there are traces, which are records of a single process, and within each trace, there are multiple events. There can be thousands to tens of thousands of such traces. Since a trace contains multiple events in chronological order, it represents a single sequence.

3.2 Data Preprocessing

Here, data preprocessing refers to using the LCS algorithm to calculate the similarity between traces, and using this similarity to compute the anomaly score. The similarity is not simply LCS but the normalized LCS (nLCS) as proposed in the paper by Ashish Sureka.[1]

$$nLCS(S_p, S_q) = \frac{nLCS(S_p, S_q)}{\sqrt{|S_p||S_q|}}$$

(Equation 1)

The specific preprocessing method is as follows: use the above formula to calculate the similarity between each trace, then sort them in descending order. Afterward, take the similarity of the arbitrarily chosen constant kth closest trace. This kth similarity is used to calculate the anomaly score of that trace. The anomaly score is defined as the inverse of the similarity. Since applying this algorithm to all traces to calculate similarity is time-consuming, this study saves the results of the similarity calculation as a matrix in a file.

4. Data Classification

In this study, an algorithm to classify data into normal and anomalous based on the anomaly score, which serves as evidence for identifying anomalous data, was employed. This classification uses Prediction Intervals commonly applied in statistics for outlier detection, specifically the one-sided 95% and 80% intervals. The following is the formula for the upper confidence limit.

$$x_0 < \bar{x} + t_{\alpha} * s * \sqrt{1 + 1/n}$$

(Equation 2)

Here, \bar{x} represents the mean of the anomaly scores, s is the sample standard deviation of the anomaly scores, and n denotes the number of data (traces, sequences). It is assumed that the distribution of anomaly scores follows a normal distribution. After calculating the upper confidence limit, data with anomaly scores exceeding this limit are classified as anomalous.

Table 1. parameters for classification

| Parameter Dataset | K | Mean | Standard Deviation | Upper Confidence Limit |
|-----------------------|------|-------|-----------------------|------------------------------|
| Control summary | 2500 | 1.018 | 0.06885 | 1.1308 |
| Prepaid TravelCost | 20 | 1.05 | 0.1334 | 1.1621 |

In this context, the 95% interval is used for the Control Summary dataset, while the 80% interval is applied to the PrepaidTravelCost dataset.

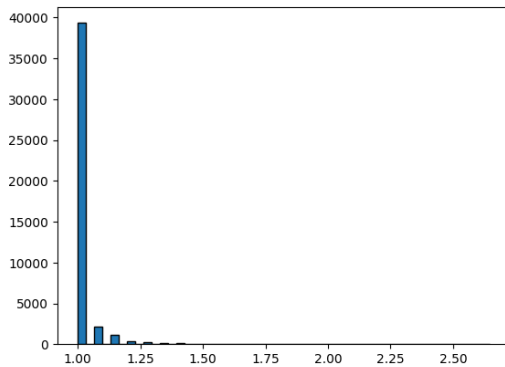


Fig. 1. Histogram of Anomaly scores of Control summary

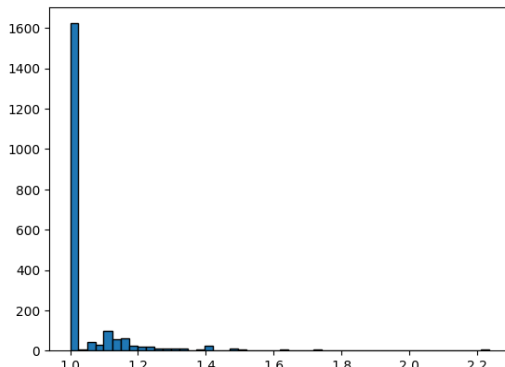


Fig. 2. Histogram of Anomaly scores of PrepaidTravelCost

Table 2. result of classification

| Dataset | The number of normal traces | The number of abnormal traces | The number of total traces |
|--------------------|-----------------------------|-------------------------------|----------------------------|
| Control summary | 41545 | 2263 | 43808 |
| Prepaid TravelCost | 1875 | 224 | 2099 |

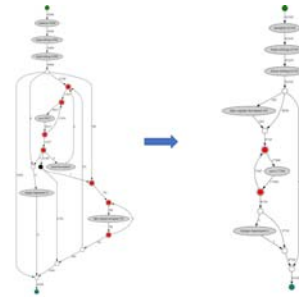


Fig. 3. ICN model of Control summary before and after classification

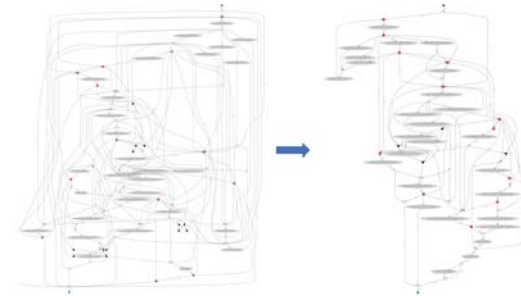


Fig. 4. ICN model of PrepaidTravelCost before and after classification

Upon examining **Figs 3** and **4**, it becomes evident that the ICN model, which was complex and challenging to interpret before data refinement, has been simplified after the refinement process, making it easier to analyze.

5. Conclusions

This study implemented the LCS algorithm for anomaly detection in business processes, determining similarity and anomaly scores to classify data as normal or anomalous using Prediction Intervals. The comparison of the ICN structures before and after data classification demonstrated that complex structures were effectively simplified, proving the efficacy of the LCS-based anomaly detection. This process significantly enhances the analysis of business

processes by refining data, ensuring more accurate and efficient interpretations. These findings not only validate the approach taken in this research but also lay the groundwork for future studies in business process analysis.

References

- [1] Ashish Sureka, "Kernel Based Sequential Data Anomaly Detection in Business Process Event Logs," *arXiv*, 2015.
- [2] K. -S. Kim, D. -L. Pham and K. P. Kim, "p-Algorithm: A SICN-Oriented Process Mining Framework," *IEEE Access*, vol. 9, pp. 139852-139875, 2021.
- [3] Kwanghoon Kim, Clarence A Ellis, "ICN-based Workflow Model and its Advances," *IGI Global*, pp. 142-171, 2001.
- [4] DE MEDEIROS, AK Alves, et al., "Process mining: Extending the alpha-algorithm to mine short loops," *Technische Universiteit Eindhoven*, 2004.

A Study on the Utilization of Blockchain for Ensuring Corporate Sustainability Based on the Protection of Personal Information

Dabin Lee, Yeji Choi, Jaewook Byun and Hangbae Chang*

Department of Security Convergence, Chung-Ang University
Seoul, South Korea

e-mail: {ilwoldabin97, yji9783, bjw0328}@cau.ac.kr

Department of Industrial Security, Chung-Ang University
Seoul, South Korea

e-mail: *hbchang@cau.ac.kr**

*Corresponding author: Hangbae Chang

Abstract

The ability to access extensive information has been a key determinant in wealth disparity throughout human history. At the dawn of the Fourth Industrial Revolution, also known as the information economy era, the economic value of personal information has surged. Consequently, there has been an increase in crimes aimed at misappropriating personal information for financial gain, with corporate customer databases becoming a prime target. Breaches of personal information, or even the mere suggestion of such incidents, have repeatedly led to corporate devaluation, highlighting security breaches as a significant business risk. Given the current corporate practices of collecting, storing, and utilizing customer data, the inherent risk of information leaks remains, and security measures without structural reform continue to exhibit similar vulnerabilities. This study analyzes cases from literature, both domestic and international, where such leaks have led to corporate devaluation, drawing implications and analyzing the causes behind these security incidents. This paper proposes the adoption of blockchain technology as a protective measure for personal information and as an enhancer of corporate sustainability, outlining its potential benefits. This research is significant for its proactive measures in protecting personal information within the context of the Fourth Industrial Revolution and offers strategic guidance for improving corporate sustainability.

Keywords: Personal Information, Blockchain, Corporate Sustainability

1. Introduction

With the onset of the information economy era following the Fourth Industrial Revolution, various data began to be considered “the new oil of the 21st century,” with personal information gaining significant economic value. As a result, the theft and criminal use of personal information have increased. The number of reported personal information leaks escalated

from about 1,000 cases in 2019 to over 7,000 in 2021, a more than sevenfold increase, and the trend of personal information leak consultations has been rising annually[1]. Corporate customer information is also stolen and leaked in many cases. The problem is that incidents of personal information leaks, or even the suggestion of such, can affect a company's sustainability, leading to drops in share prices or customer losses.

The current structure, where customer information is stored in a centralized manner and

This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the ITRC(Information Technology Research Center) support program(IITP-2023-2018-0-01799) supervised by the IITP(Institute for Information & communications Technology Planning & Evaluation)

it's challenging for customers to track their information usage, inherently poses security vulnerabilities. If no change is made to this structure with inherent vulnerabilities, similar types of risks will persist.

Hence, this study analyzes cases from both domestic and international literature where personal information leaks led to corporate devaluation. It explores the use of blockchain technology's decentralization feature to protect personal information and proposes it as a method to enhance corporate sustainability.

2. Terminology

2.1 Sustainability

The rapidly changing society, marked by events like the Fourth Industrial Revolution, the outbreak of COVID-19, and the climate crisis, has brought fast-paced changes to the corporate environment, making sustainability a key topic for businesses.

While the concept of sustainability and sustainable development has recently been linked with environmental issues and used in the context of environmental sustainability, this study uses the term in reference to the continued existence of businesses in the future.

2.2 Blockchain

Blockchain is a data integrity and tamper-proof technology based on distributed computing, which stores data in a chain format. It is designed to prevent arbitrary manipulation while allowing public access.

One of the characteristics of blockchain is its decentralization, differing from traditional centralized structures, enabling P2P transactions without intermediaries. This allows network members to share data in real-time through the ledger. The transparency is also assured as each network node provides an identical record of transaction histories[2].

Blockchain technology plays a role in returning personal information stored in each company back to the individuals and provides an environment for users to manage their own personal information.

3. Analysis of the Relationship Between Personal Information and Corporate Sustainability

Incidents related to the protection of personal information have led to decreases in both corporate trust and reputation, inflicting economic damages on companies.

In 2018, Facebook was revealed to have illegally collected and utilized personal information, leading to an estimated loss of around 40 billion dollars in just two days[3]. Earlier this year, the South Korean telecommunications company LG Uplus announced a leak of information of about 180,000 customers, which led to a roughly 1% drop in its stock price[4]. These cases demonstrate how incidents related to personal information protection can lead to a decline in corporate value, thereby threatening corporate sustainability. Especially, cases where leaks occurred through centralized corporate data servers suggest that the current structure of centrally owning customer information is vulnerable to such incidents and could lead to major accidents when breaches occur.

In this trend, movements emphasizing personal information protection are emerging.

Nasdaq includes 'data privacy' in its ESG disclosure area, marking compliance with data privacy policies and GDPR regulations[5]. The 'K-ESG Guidelines' released by the Ministry of Trade, Industry, and Energy of South Korea include the establishment of information protection systems and remedies for personal information infringement under the social(S) diagnostic items[6].

In this situation, Blockchain, by decentralizing and thus securing personal information stored in central data servers, can contribute to protecting personal information and thereby help companies secure sustainability.

4. Results

This study presents potential benefits and implementation methods of blockchain technology as a means to protect personal information and enhance corporate sustainability, based on case studies from domestic and international literature.

The Netherlands is currently using DigiD, a mobile digital identity card application based on

blockchain's Decentralized Identifier (DID) technology. Unlike physical IDs, which may be limited in selectively providing information, DigiID allows users to only provide information relevant to the situation, minimizing personal information exposure[7].

In South Korea, a telecommunications company Medibloc used blockchain's decentralization feature to ensure medical data sovereignty for patients. It stores patients' medical records and health information on the blockchain, allowing patients to manage their medical records and share them with professionals when necessary[8].

Thus, blockchain's decentralization allows individuals to directly manage their information and share it without intermediaries when necessary. This could solve the problems related to personal information that arise when collecting information in a centralized manner.

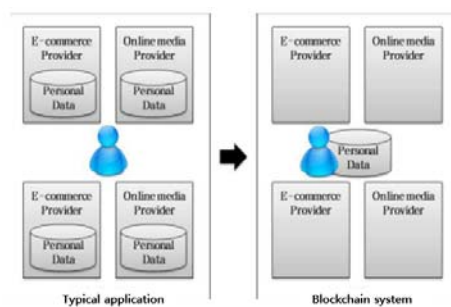


Fig. 1. Comparison between typical application and blockchain system[9]

5. Conclusions

This study aims to propose the adoption of blockchain technology as a means of protecting personal information and enhancing corporate sustainability in the rapidly changing environment, where the significance of corporate sustainability is increasingly. We explored the importance of personal information protection in the context of corporate sustainability and suggested the direction for securing sustainability through blockchain technology.

However, blockchain, characterized by transparency through shared transaction ledgers among network participants, has points of conflict with personal information protection laws that primarily focus on protection. This study acknowledges its limitation in not addressing these conflicting issues. Therefore,

future research will aim to find technical and managerial solutions to address these challenges.

References

- [1] Y. Seo, "Personal Information Infringement Report Increased 7 Times in 1 Year," *Information and Communication Newspaper*, Oct. 6, 2022, <https://www.koit.co.kr/news/articleView.html?idxno=104009>.
- [2] Y. Choi, J. Byun, J. Moon, and H. Jang, "Exploratory Study on Blockchain Introduction for ESG Management in the Distribution Industry," *Knowledge Management Research*, Vol 24, No.3, pp. 217-237, 2023.
- [3] D. Seok, "Facebook Shares Fall, Affected by Expanded Investigation into Data Leak Scandal," *Digital Today*, July 4, 2018, <https://www.digitaltoday.co.kr/news/articleView.html?idxno=201044>.
- [4] M. Yang, "LG U+ Leaks Personal Information of 180,000 Customers...Down in early trading," *e-newspaper*, Jan. 11, 2023, <https://www.etnews.com/20230111000010>.
- [5] Nasdaq, "ESG Reporting Guide 2.0", 2019.
- [6] MOTIE, "K-ESG Guide line v1.0", 2021.
- [7] K. Kim, "[Blockchain Field] 'Making Overseas Travel Without Passports' Dutch Blockchain ID Experiment," *Yonhap News*, June 27, 2018, <https://www.yna.co.kr/view/AKR20180626167400002>.
- [8] J. Jang, "When I move to the hospital, do I take my medical records with me? Simple solution with personal health information platform", *DBR*, Issue 1, June 2018, https://dbr.donga.com/article/view/1203/article_no/8660.
- [9] W. Jeon, "a study of personal information management and transaction technology using blockchain", *informatization research*, Vol 16, No.1, pp. 87-96, 2019.

The effects of users' perceived platform indispensability on their actual usage: A comparative study of Korea and Finland

Moonkyoung Jang¹ and Seongcheol Kim^{2,*}

¹Business School, Gachon University, Seongnam-Si, South Korea

²School of Media and Communication, Korea University, South Korea

[e-mail: mkjang@gachon.ac.kr, hiddentrees@korea.ac.kr]

*Corresponding author: Seongcheol Kim

Abstract

Digital platforms are not an option anymore but rather an essential tool and the core of the digital ecosystem. Accordingly, it is necessary to pay attention to how users evaluate the indispensability of the dominant platform. In addition, we may need to examine the effects of users' perceived indispensability of digital platform on their real usage of the platform. It seems that there is no study on digital platform indispensability in a specific country. We conducted an online survey in both countries. Two identical online survey questionnaires were used to collect data. The results showed that Finnish users perceive their search engine more indispensable in their lives than Koreans do. Finnish users use the search engine more frequently during the day even though it is Google's platform. In addition, users' perceived indispensability of digital platform has a direct impact on actual usage of digital platform in both countries.

Keywords: Digital platform, Indispensability, Actual usage, Korea, Finland

1. Introduction

Digital platforms are not an option anymore but rather an essential tool and the core of the digital ecosystem. Accordingly, it is necessary to pay attention to how users evaluate the indispensability of the dominant platform. In addition, we may need to examine the effects of users' perceived indispensability of digital platform on their real usage of the platform. It seems that there is no study on digital platform indispensability in a specific country. This study will conduct a comparative study between Korea and Finland.

2. Research questions and methodology

2.1 Research questions

The term "platform" itself indicates essential infrastructures for electronic commerce, online communication, and digital social relationships [1]. The value of platforms cannot be ignored; the role of platforms is expanding to connecting and facilitating any kind of transaction within society. With their integrated nature, starting from fintech to e-commerce, digital platforms remove boundaries in the digital economy and have become "the operating system of our lives" [2]. Accordingly, it is necessary to pay attention to how users perceive the indispensability of the

digital platform in their daily lives [3]. In addition, it would be meaningful to examine how users' perceived indispensability of digital platform would impact on their actual usage of the platform. In the meantime, Korea becomes a platform battlefield where local digital platforms compete with so-called global big tech platforms. In Korea, Naver still ranks first in the search platform, but its share is decreasing. Google's share in search platform is growing rapidly. Europe, on the other hand, lacking its own digital platform, suffers from a so-called "platform gap" and is trying to cope with the current situation through internal cooperation and active regulations [4]. In Finland, Google's share in search platform is dominant. The levels of the presence of local digital platforms, in particular search engine platform, are different in Korea and Finland, so this study focuses on a comparison between two countries.

2.2 Methodology

For this study, Korea and Finland were chosen as our research subjective. These two countries are one of the leading digital powers in Asia and Europe, respectively, and countries which survive global digital ecosystem led by the U.S. and China. In addition, these countries have the leading information technologies (IT) manufacturers such as Samsung Electronics and Nokia. Those countries also have global tech-startups. For example, there global mobile apps such as lineage M in Korea and Angry birds and Crash of Clan in Finland. IT infrastructure of two countries, such as internet usage and smartphone penetration, is rated as the world's top. However, explained before, the levels of the presence of local digital platforms are different in Korea and Finland. But while Korea has its home-grown local platforms such as search engine, mobile messenger, social media and so on, Finland lacks its own digital platforms including search engine. They normally use Google for Internet search in Finland. Thus, it is suitable for the international comparative study because its different situation with similar level of ICT infra.

We conducted an online survey in both countries. Two identical online survey questionnaires were used to collect data. In Korea, the data were collected in November 2022, in cooperation with a professional research agency Micromill Embrain. In Finland, the data

were collected between May and July 2023. The Finnish respondents were recruited through multiple channels, such as university notice boards, student mailing lists, and the authors' social media networks.

We obtained 153 useable responses from Finland and 214 from Korea. The Korean sample contained 142 (66%) females and 72 (34%) males. The Finnish sample consisted of 80 (52%) females, 68 (44%) males and 5 who identified as "others".

Prior to conducting the main analysis, Composite Reliability (CR) and Average Variation Extracted (AVE) were measured for all respondents and Korean/Finnish respondents to verify the convergent and discriminant validity of measurement for variables. The composite reliability level of each variable is 0.7 or more, and the average variance extracted value is 0.5 or more, indicating that the reliability of the measurement is satisfied.

3. Results

To respond to our research question 1, we conducted an independent two-sample t-test to analyze the differences in the perceived indispensability of digital platform and actual usage. According to Table 1, the perceived indispensability of digital platform was significantly higher in Finnish responses (mean = 3.873) compared to Korean responses (mean = 3.617). This result implies that Finnish users perceive their search engine more indispensable in their lives than Koreans do. In terms of actual usage, the daily use frequency of digital platform are significantly higher in Finnish responses (mean of Daily use frequency = 2.510) compared to Korean responses (mean of Daily use frequency = 1.911). Finnish users also use the search engine more frequently during the day even though it is Google's platform.

Table 1. Country differences in digital platform

| | Finland M (S.D.) | Korea M (S.D.) | Mean difference |
|------------------------|---------------------|-------------------|--------------------|
| Indispensability | 3.873 (0.533) | 3.617 (0.767) | 0.256*** |
| Daily use frequency | 2.510 (1.182) | 1.911 (0.892) | 0.599*** |
| Daily use duration | 2.308 (0.074) | 1.766 (0.048) | 0.541*** |

Note: Using t-test; * $p < .05$, ** $p < .01$, *** $p < .001$

To respond to our research question 2, we used the SEM analysis for each sample. The SEM results showed that users' perceived indispensability of digital platform has a direct impact on actual usage of digital platform in both countries (See Fig. 1). Regarding the impacts of digital platform characteristics on users' perceived indispensability of digital platform, the SEM results showed that security risk and usefulness have a direct impact on Finnish users' perceived indispensability of digital platform. However, for Korean users, only usefulness has a direct impact on users' perceived indispensability of digital platform. This result indicates that there is no significant negative impact of perceived security risk on the indispensability of digital platform in Korea.

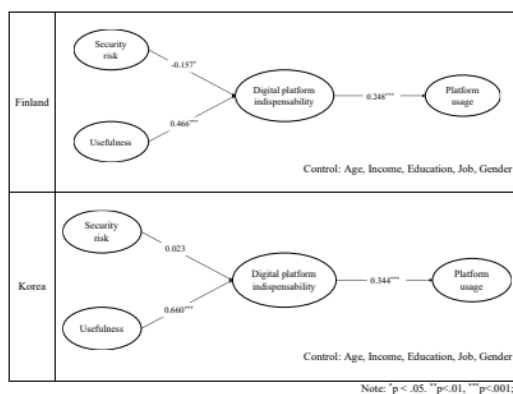


Fig. 1. SEM results of two countries

To respond to our research question 3, we ran a multigroup analysis (MGA) to find differences in path coefficients between the Korean and Finnish respondents. The intention was to examine whether the impact of DL, IL, or the four constructs of UTAUT on intention to use digital technology for learning differed between Korean and Finnish respondents.

Conclusions

In this paper, we examined how users' perceived indispensability of digital platform would impact on their actual usage of the platform in Korea as well as Finland.

The results showed that Finnish users perceive their search engine more indispensable in their lives than Koreans do. Finnish users use the search engine more frequently during the day even though it is Google's platform. In addition,

users' perceived indispensability of digital platform has a direct impact on actual usage of digital platform in both countries. However, when the path relationships between users' perceived indispensability and actual usage were assessed, no significant differences between the Korean and Finnish respondents were confirmed. The results showed that both Korean users and Finnish users are different in their perceived indispensability of digital platform and actual usage, but they are similar in the path relationships between their perceived indispensability and actual usage.

This study is not without limitations. Owing to the context of this research (Korea and Finland), further research is needed in other countries to examine our research questions. Moreover, we did not examine demographic differences between the Korean and Finnish people.

References

- [1] A. Quarta, "Narratives of the digital economy: How platforms are challenging consumer law and hierarchical organization," *Global Jurist*, Vol 20, No 2, pp. 1–21, 2020.
- [2] S. Vaidhyanathan, *Antisocial media: How Facebook disconnects us and undermines democracy*. Oxford University Press, 2022
- [3] D. L. Hoffman, T. P. Novak, & A. Venkatesh, A. (2004). "Has the Internet become indispensable?" *Communications of the ACM*, Vol 47, No 7, pp. 37-42, <https://doi.org/10.1145/1005817.1005818>, 2004.
- [4] S. Hermes, E. K. Clemons, M. Schrieck, S. Pfab, M. Mitre, M. Böhm, M. Wiese, H. Krcmar, "Breeding Grounds of Digital Platforms: Exploring the Sources of American Platform Domination, China's Platform Self-Sufficiency, and Europe's Platform Gap. In and others," in *Proc. of the European Conference on Information Systems*, pp. 1–18, 2020.

A Structural Anomaly Detection Approach in Process Mining

Minjae Park¹ and Kwanghoon Pio Kim^{2*}

¹ Department of Applications Software, Division of Computer and Information
DAELIM UNIVERSITY COLLEGE
29, Imgok-ro, Dongan-gu, Anyang-si, Gyeonggi-do, 13916, Republic of KOREA
e-Mail: mjpark@daelim.ac.kr

² Data and Process Engineering Research Lab.
Division of AI Computer Science and Engineering
KYONGGI UNIVERSITY
154-42 Kwangkyosan-ro, Youngtong-gu, Suwon-si, Gyeonggi-do, 16227, Republic of Korea
e-Mail: kwang@kgu.ac.kr

*Corresponding Author: Kwanghoon Pio Kim

Abstract

In the research field of the business process life-cycle management through mining and reengineering, it is important for the mining systems to discover the structurally safe and complete business process models. However, these systems are faced with various types of anomalies that can derive inaccurate and unstructured models from business process enactment event log datasets. A peculiar one out of such anomalies, which is the core challenging issue of this paper, is the structural anomalies that affect to the structural quality of the discovered business process models. This paper, therefore, proposes a structural anomaly detection approach that can detect and prevent those structural anomalies in performing the business process mining activities. Finally, we practically prove the conceptual excellency and functional feasibility of the proposed approach through a couple of experimental verifications with the real business process enactment log datasets available in the 4TU Center for Research Data.

Keywords: workflow and business process enactment event sequence; temporal workcase; activity-workitem event sequence; performer event sequence; structural anomaly detection; structural anomaly prevention; process logging; process mining; IEEE-XES event stream standardized format

1. Introduction

In recent years, many data and process scientists have been interested in the field of process mining research and system developments [1][2], discovering and rediscovering business process (which is called as process from now on.) models and their process-related knowledge from process enactment event-log datasets. Generally, the core competitiveness of the process-aware enterprises and organizations ought to depend on the power of business process life-cycle management scheme based upon the automated process discovery and knowledge mining

functionality. At this stage, however, we need to take attention to the qualities of the process discovery approaches, the level of which especially depends on how to deal with the various types of operational anomalies [3], such as modeling anomaly, operating anomaly, logging anomaly, and discovering anomaly, that can produce their corresponding inaccurate event logs – *Noises* [3][5]. In this paper, we focus on a special type of operational anomalies named as structural anomalies that violate the structural rules [4][6][7] within the control-flow gateway building-blocks of a discovered process model. It is very important to detect those structural

anomaly cases and their associated process event log traces and properly treat them as a structural anomaly prevention scheme, because a small number of those associated process event log traces may easily and negatively affect the overall levels of qualities of the automated process discovery algorithms and systems [8][9] in terms of the structural formations of control-flow gateway building-blocks composing of selective XOR, parallel AND, and iterative LOOP gateways.

3. Structural Anomaly Detection

In the paper, the emphasis is placed on keeping the structural quality of the discovered process model by using the typical SICN-oriented process mining approach, ρ -Algorithm [9]. The algorithm theoretically and experimentally proved that all the combinational building blocks of four process routing patterns can be discovered from a process enactment event log dataset. At the same time, it explored for that the event log traces embed a certain quantity of hidden noises that may seriously affect the correctness as well as the syntactical and structural quality of the underlying process mining algorithm and system, and it also built an

anomaly-treatment algorithm that can detect and remove those hidden noises caused by the temporal activity-sequencing anomalies. In this paper, we try to explore for another type of hidden noises caused by the so-called structural anomalies and propose their detection approach. This section won't describe the details of the proposed structural anomaly detection approach due to the page limitation.

In order to verify the feasibility of the proposed approach, the authors carried out an experiment on a typical process execution event log dataset named as the Customer Summary process dataset, which is containing the total 43,808 event log traces. After fulfilling the experiment on the dataset, at last we successfully found out that all the event log traces are classified into the total 59 event log trace-clusters and each trace-cluster is containing a certain number of traces having the identical activity-event sequences. Note that the top five trace-clusters, in terms of the holding number of event log traces, contain 31806, 5004, 1328, 2525, and 778 traces, respectively. We also obtained the aggregate information on the dataset, such as the number of event log traces (*i.e.*, 43,808 traces), the number of activity-events (*i.e.*, 161,296 events), the number of trace-clusters (*i.e.*, 59 trace-clusters), and others.

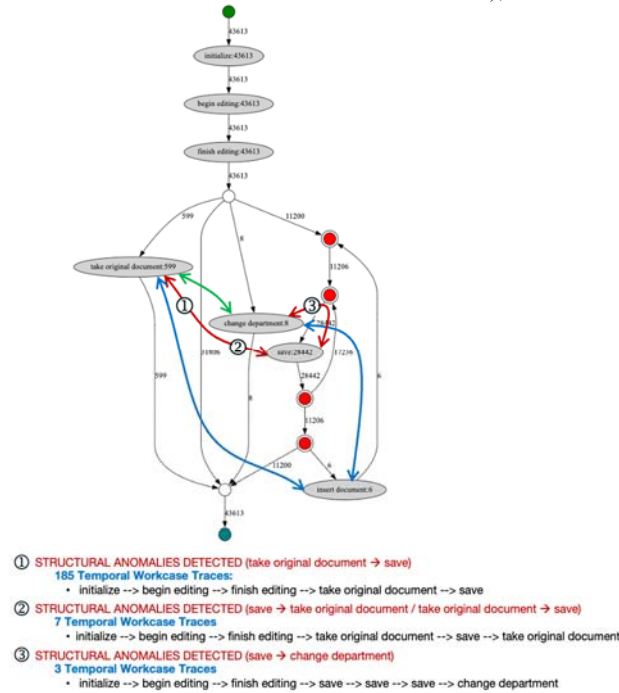


Fig. 1. Detection Results of Process Event Log Traces Caused the Structural Anomalies

Fig. 1 illustrates the discovered process model from the dataset from 48,613 traces out of the total 48,808 event log traces. We can easily recognize that the structural formation of the discovered process model is perfectly satisfied with the structural requirement of matched-pairing and proper-nesting properties, and so it is sound and safe information control net process model. As you can see, only the three trace-clusters (*i.e.*, 185-trace, 7-trace, and 3-trace) and their 195 event log traces were detected as the structural anomalies breaking the structural formation of the on-going discovered process model. Consequently, we confirmed and proved that the operational feasibility of the structural anomaly detection approach be a tangible, operable, and reasonable solution for discovering sound and safe process model from a process event log dataset through this verificational experiment.

4. Conclusion

This paper addresses a novel approach of process mining and preprocessing technique focusing on structural anomaly that possibly breaks the structural formation of the discovered process model. Therefore, these structural anomalies must be detected and removed for discovering a sound and safe information control net process model, and their detection algorithm should be taken into consideration in developing and implementing the process mining algorithms and systems. This paper excogitated the algorithmic approach and its experimental verification of detecting and removing the event log traces causing these structural anomalies and proved its effectiveness in improving the qualitative degree of the process mining algorithm and system by applying the typical SICN-oriented process mining algorithm (ρ -Algorithm) and system embedding the proposed approach to an ideal dataset of process enactment event logs as well.

Acknowledgements

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT, Ministry of Science and ICT), Republic of Korea (Grant No. NRF-2022R1A2C2093002).

References

- [1] P. K. G, K. R. R. Babu, S. U, R. S. Thomas, Saigopika, S. Walter, S. Thomas, Price forecasting on a large scale data set using time series and neural network models, *KSII Transactions on Internet and Information Systems*, 16 (12), 3923–3942, 2022. doi:10.3837/tiis.2022.12.008.
- [2] X. Wang, Z. Cao, R. Zhan, M. Bai, Q. Ma, G. Li, Density-based outlier detection in multi-dimensional datasets, *KSII Transactions on Internet and Information Systems*, 16 (12), 2022. doi:10.3837/tiis.2022.12.002.
- [3] Y. Lee, I.-S. Chang, S. Oh, Y. Nam, Y. Chae, G. Choi, G. Park, Hierarchical flow-based anomaly detection model for motor gearbox defect detection, *KSII Transactions on Internet and Information Systems*, 17 (6) 1516–1529, 2023. doi:10.3837/tiis.2023.06.001.
- [4] M. N. Lacheheb, N. Hameurlain, R. Maamri, Resources consumption analysis of business process services in cloud computing using petri net, *Journal of King Saud University - Computer and Information Sciences*, 32 (2020) 408–418. doi: https://doi.org/10.1016/j.jksuci.2019.08.005.
- [5] K. Kim, A model-driven workflow fragmentation framework for collaborative workflow architectures and systems, *Journal of Network and Computer Applications*, 35 (1), 97–110, 2012. doi:10.1016/j.jnca.2011.03.029.
- [6] H. Ahn, K.P. Kim, Formal approach to workflow application fragmentations over cloud deployment models, *Computers, Materials & Continua*, 67 (3), 3071–3088, 2021.
- [7] W. V. der Aalst, T. Weijters, L. Maruster, Workflow mining: discovering process models from event logs, *IEEE Transactions on Knowledge & Data Engineering*, 16 (9), 1128–1142, 2004.
- [8] K.-S. Kim, D.-L. Pham, K.P. Kim, ρ -algorithm: A sicn-oriented process mining framework, *IEEE Access*, 9, 139852–139875, 2021. doi:10.1109/ACCESS.2021.3119011.
- [9] C. A. Ellis, A. J. Rembert, K.-H. Kim, J. Wainer, Beyond workflow mining, in: *International Conference on Business Process Management*, Springer, pp. 49–64, 2006.

A Case Study on Security Courses for Information Security Personnel

Eunsun Choi¹, Jooyeon Han², and Namje Park³

Creative Education Center, Jeju National University, Jeju, South Korea¹,
Department of Convergence Information Security, Graduate School, Jeju National University, Jeju, South Korea²,
Department of Computer Education, Teachers College, Jeju National University, Jeju, South Korea³
[e-mail: choi910624@jejunu.ac.kr, jooyeonhan@branksome.asia, namjepark@jejunu.ac.kr]
*Corresponding author: Namje Park

Abstract

Information security has become an essential element of safety and stability in modern society. In particular, information security is of paramount importance for the military and law enforcement. Therefore, in this paper, we studied the case of developing and operating security courses for security personnel in military and public institutions. The contents of the training included the work of the security officer, the theory of security inspection, recent hacking cases, and security methods. In addition, the practice was conducted to enhance the understanding of participants, and it was organized as time to explore ways to use metaverse technology for security. Participants in the course said they got an opportunity to communicate with experts and that the practice was helpful, and opinions were collected that they hoped the training time for security accident cases would be expanded.

Keywords: Information security course, personnel training, security academy, metaverse, industrial technology protection

1. Introduction

Cyber threats are becoming increasingly intelligent and organized worldwide due to the daily use of telecommuting and changes in the work environment after the COVID-19 pandemic [1]. Therefore, it is important to respond preemptively and systematically. Private companies with large organizations meet demand by hiring skilled personnel through internal training and in-house job conversion systems. In contrast, the public sector is having difficulty hiring information protection personnel due to relatively weak wage competitiveness, which results in a lack of guaranteed expertise among the information protection personnel in their offices [2].

In addition, cyber threats are diversifying according to changing technologies, so it is important for information protection personnel in the public sector, including the military and police, to continuously research and establish strategies to identify and respond to recent trends in cyber threats. In this paper, we analyzed the case of designing and operating information protection courses for military and police information security personnel at J University.

2. Security Courses for Information Security Personnel

2.1 Security Academy for Security Personnel in Military

This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2022S1A5C2A04092269), and this work was also supported by an Electronics and Telecommunications Research Institute (ETRI) grant funded by the Korean government. [23ZD1160, Regional Industry ICT Convergence Technology Advancement and Support Project in Daegu-Gyeongbuk (Mobility)].

The security academy was designed for military security personnel and consisted of a three-day course in September 2022. The educational topics consisted of topics that could be helpful in the practice of security personnel in the military, such as predictive security technology for preemptive risk response, understanding the defense security paradigm, cyber operation technology, and major security issues. In addition, the instructors were led by expert team teaching such as a professor of convergence information security at J University, a researcher at the Korea Institute for Defense Analyses, a researcher at the National Security Research Institute, and representatives of cybersecurity IT companies. The designed training contents are shown in [Table 1](#).

Table 1. Training program for military security personnel

| Day | Time | Contents |
|---------|-------------|---|
| 1st day | 10:00~10:50 | Predictive security technologies for proactive risk mitigation: Blockchain, Quantum Cryptography, Artificial Intelligence |
| | 11:00~11:50 | Transition of national defense security paradigm and understanding K-Risk Managemt Framework (K-RMF) |
| | 14:00~16:50 | Document security |
| 2nd day | 09:00~09:50 | Review of security measures and security assessments |
| | 10:00~10:50 | Roles and responsibilities of security personnel |
| | | Core principles of future technologies |
| | 11:00~11:50 | Legal improvements for the establishment of a national classified information protection system |
| 3rd day | 13:00~16:50 | PC security checks (theory and practice) |
| | 09:00~09:50 | Information and communication security |
| | 10:00~10:50 | Cyber warfare technology and key security issues as observed in the Ukraine conflict |

| | |
|-------------|---|
| 11:00~11:50 | Recent cases of cyberattacks in defense and responses |
| 13:00~13:50 | Comprehensive practical exercises |
| 14:00~15:50 | Comprehensive assessment |
| 15:00~16:50 | Lecture evaluation and survey |

2.2 Metaverse Security Course for Military Officers

Metaverse security training was conducted for military commanders, staff, sheriff's officers and other officers. It was designed as a four - instalment program in October 2022. The educational topic consisted of the subject of exploring the applicability of the defense metaverse in line with the changing times. The instructors included professors in the department of convergence information security at J University, Army educators, and researchers at the Defense Security Institute. In addition, panel discussions and expert exchange time were added after the lecture to provide opportunities for mutual communication to learners who participated in the course. Details of metaverse security education are shown in [Table 2](#).

Table 2. Training program for military officers

| Time | Contents |
|-------------|--|
| 14:20~15:10 | Metaverse security framework |
| 15:20~15:50 | Applicability and future of defense metaverse |
| 15:50~16:20 | Viewing Metaverse through the lens of spatial changes |
| 16:20~16:50 | Security measures for the application of Defense Metaverse |
| 17:00~17:30 | Panel discussion for defense metaverse security |
| 17:30~18:00 | Meeting with experts |

2.3 Industrial Technology Protection Academy for Public Sector Employees

The Industrial Technology Protection Academy consisted of specialized lectures on

cybersecurity at public institutions in Jeju. The system was developed to cater to public institution employees, law enforcement personnel, as well as executives and staff in various companies. This academy course was developed as a four-lesson course to teach how to respond to information infringement occurring in various fields and to elicit knowledge about it. The education course consisted of topics such as industrial technology leakage prevention education, types of unfair competition, and intellectual property law. Furthermore, we have expanded the platform for meetings between experts and practitioners by incorporating dedicated consultation sessions with experts. The detailed course content is as outlined in [Table 3](#).

Table 3. Training program for government officials

| Time | Contents |
|-------------|---|
| 14:00~14:50 | Industrial technology leak prevention education |
| 15:10~16:00 | Types and cases of unfair competition |
| 16:00~16:50 | Intellectual property law and legal structure business |
| 16:50~18:00 | Industrial technology leak and legal structure consultation |

3. Conclusions

Managing national secrets and maintaining security is considered one of the essential factors that influence a country's competitiveness. This is why it is imperative to provide education to enhance information protection and security capabilities for the military and police, who handle sensitive information [3]. The cyberspace is often referred to as the fifth battleground without borders. The training of advanced cyber experts requires highly specialized training and can be time-consuming [4]. Therefore, in-depth research is needed to develop strategies for nurturing cyber warriors effectively. In this paper, we analyzed a research case where information security courses were conducted for military and public sector information security professionals at J University. There were three different cases of education: a 3-day Security Academy, a 1-day Metaverse Security Course, and the Industrial Technology Protection Academy. All educational programs had a high

level of expertise, including professors from the Department of Convergence Information Security at J University, researchers from government-affiliated research institutes, and cybersecurity Information Technology company representatives, serving as instructors. Furthermore, the inclusion of time for interactions with experts and panel discussions provided opportunities for mutual communication.

Participants in the education program found the practical exercises and the education on secrets and security-related topics used in real-world scenarios to be beneficial. However, some expressed a desire for increased practical training time and additional education on security incident cases to improve the program. Additionally, they emphasized the need for diversified promotion to attract more public sector employees to participate in the education.

In the future, it is necessary to collect and analyze the opinions of education participants and consider redesigning security education programs. Furthermore, there is a need for effectiveness analysis research to determine if the education contributed to enhancing participants' security awareness and expertise after the education programs were conducted.

References

- [1] D. -W. Kim, S. -W. Chai and J. -C. Ryou. "A Study on Domestic Information Security Education System," *Journal of the Korea Institute of Information Security and Cryptology*, Vol. 23, No. 3, pp. 545-559, 2013.
- [2] S. -H. Lee, H. -J. Jun, and T. -S. Kim, "Evaluation of Public Information Security Training Programs : A Case Study," *Journal of Information Technology Services*, Vol. 19, No. 1, pp. 173-185, 2020.
- [3] D. S. Yun. Learning model development of utilizing a tutor to study cyber-terrorism," *Convergence security journal*, Vol. 15, No. 6, pp. 49-57, 2015.
- [4] S. Kim. "[Addressing Cyber Threats and Security Knowledge] Recommendations for the efficient development of military cyber experts and the enhancement of cyber education and training systems," *Military Journal*, Vol. 2021, No. 4/5, pp. 62-65, 2021.

Analysis of Hallyu research trends using text mining

Miyoung Kim and Sunju Park

Suwan Middle School

Dept. of Computer Science Education, Gwangju National University of Education

Gwangju, South Korea

[e-mail: chemi20@naver.com, sjpark@gnue.ac.kr]

*Corresponding author : Sunju Park

Abstract

This paper analyzed changes and research trends from Hallyu 1.0 to 4.0. KCI listed and candidate papers were collected and quantitative research methods such as word cloud, topic modeling, and network analysis were used. As a result of the analysis, we need to focus on developing Hallyu content through research that deeply understands consumer behavior and consumption trends, and on developing new strategies and platforms that respond to the digital age. This study provides important indicators for the continued growth and development of the Hallyu and suggests directions that can further solidify the Korean Wave's position as a global cultural content.

Keywords: Hallyu, research trends, quantitative analysis, topic modeling

1. Introduction

This paper seeks to explore the growth process of the Hallyu by analyzing related research papers by period. Through this, we aim to deeply understand how the Hallyu developed into a global cultural phenomenon. In addition, based on the results obtained through research, we aim to provide insight into future trends and development possibilities in the field of Hallyu research.

The goal of this paper is to provide a deeper and broader understanding of the Hallyu by presenting a new perspective on Hallyu research and exploring important elements that previous studies have missed. To this end, we would like to examine trends in Hallyu research using text mining techniques targeting academic papers related to the Hallyu and suggest implications for the direction of future development.

2. Text Mining

Recently, Korean wave research using big data is increasing[6]. Among them, a lot of research is being conducted using text mining. Text mining is a technique to find valuable and meaningful information from text data[3]. In this paper, text mining refers to the analysis of texts such as titles, keywords, and abstracts of Hallyu-related papers in KCI-listed and candidate papers from 1997 to May 2023 in RISS.

3. Results

To select papers, we used the database of the academic information service (<http://www.riss.kr>, RISS) provided by the Korea Education and Research Information Service(KERIS). We searched for papers on the topic of 'Korean Wave' as of May 31, 2023, and as a result, a total of 1,437 papers were collected from 1997 to May 2023. Among these, after excluding 117 papers, including papers without

abstracts (if there were English abstracts, they were converted to Korean abstracts) and papers that did not fit the topic, a total of 1,320 papers were selected for final analysis. The number of papers on Hallyu by year is as shown in (Fig. 1).

| year | No. of papers | Percentage (%) | year | No. of papers | Percentage (%) |
|------|---------------|----------------|-------|---------------|----------------|
| 1997 | 1 | 0.25 | 2012 | 80 | 19.66 |
| 1999 | 1 | 0.25 | 2013 | 113 | 27.76 |
| 2000 | 4 | 0.98 | 2014 | 121 | 29.73 |
| 2001 | 4 | 0.98 | 2015 | 81 | 19.90 |
| 2002 | 19 | 4.67 | 2016 | 112 | 27.52 |
| 2003 | 8 | 1.97 | 2017 | 73 | 17.94 |
| 2004 | 17 | 4.18 | 2018 | 79 | 19.41 |
| 2005 | 67 | 16.46 | 2019 | 63 | 15.48 |
| 2006 | 56 | 13.76 | 2020 | 58 | 14.25 |
| 2007 | 80 | 19.66 | 2021 | 50 | 12.29 |
| 2008 | 34 | 8.35 | 2022 | 59 | 14.50 |
| 2009 | 40 | 9.83 | 2023 | 24 | 5.90 |
| 2010 | 27 | 6.63 | total | 1,320 | 100 |
| 2011 | 49 | 12.04 | | | |

shown in (Fig. 2). Through this, we will analyze the trends of Hallyu research in different periods and environments.

| category | period | No. of research papers | Percentage (%) |
|------------|-------------|------------------------|----------------|
| Hallyu 1.0 | 1997 ~ 2004 | 54 | 4 |
| Hallyu 2.0 | 2005 ~ 2010 | 304 | 23 |
| Hallyu 3.0 | 2011 ~ 2021 | 879 | 67 |
| Hallyu 4.0 | 2022 ~ | 83 | 6 |



Fig. 2. Number of papers on Hallyu by periods.

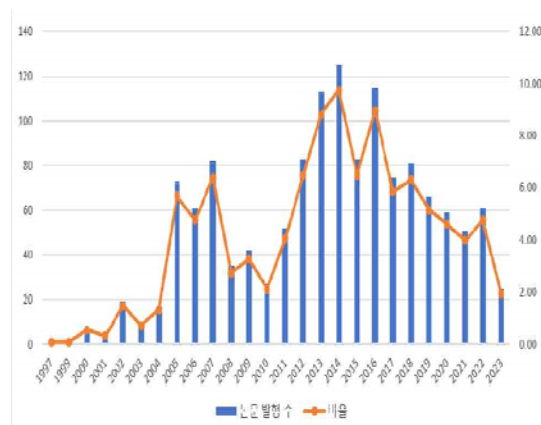


Fig. 1. Number of papers on Hallyu by year

Based on the graph in (Fig. 1), the period from 1997 to 2004, when the number of papers first appeared, was defined as Hallyu 1.0, and the period from 2005 to 2010, when the number of papers increased and then rapidly decreased, was defined as Hallyu 2.0. After that, the number of papers increased again, and the period from 2011 to 2021 was divided into Hallyu 3.0, and from 2022 to the present was set as Hallyu 4.0. The number of research papers by Hallyu period is as

4. Conclusions

This paper focuses on the changes from Hallyu 1.0 to Hallyu 4.0 and examines the research trends of these changes. The results of word cloud, topic modeling, and network analysis using Jupyter Notebook and Python on selected papers are as follows.

The main research topic of Hallyu 1.0 is the cause and characteristics of Hallyu phenomenon. Hallyu 2.0 confirms the continuity of Hallyu and discusses the revitalization and expansion of Hallyu industry. During this period, the Hallyu grew into a global phenomenon, and the importance of the Korean language was emphasized. Hallyu 3.0 research shows changes that focus on consumers rather than producers of Hallyu. The research direction of Hallyu 4.0 will be the effective use of digital media, deep understanding of the public, and response to changing contents.

We must focus on developing Hallyu content through research that deeply understands consumer behavior and consumption trends, and on developing new strategies and platforms that respond to the digital age. By doing this, it is

expected that the Hallyu will be able to further solidify its position as a global cultural content.

References

- [1] BBC news, Why 26 Korean words have been added to Oxford English Dictionary [5.OCT.2021], <https://www.bbc.com/news/newsbeat-58749976/2023.02.21>
- [2] Hwang, S. I., Park, J. B., “‘Korean Wave’ News Analysis Using News Big Data,” *Journal of the Korea Entertainment Industry Association*, Vol 14, No. 5, pp. 1-14, 2020.
- [3] Jo, J. H., & Kang, H. J., “A study on the activation of cultural product exports through the spread of Hallyu culture,” *Journal of Customs Economics and Management*, Vol 14, No. 1, pp. 271-296, 2013.
- [4] Kim, H. K., Global Hallyu Issue Magazine, Vol. 35, pp. 9-12, 2020.
- [5] Kim, S., & Lee, W.S., “Network text analysis of medical tourism in newspapers using text mining: The South Korea case,” *Tourism Management Perspectives*, Article 31: pp. 332-339, 2019.
- [6] Lee, K. Y., & Park, S. H., “Korea Foundation for International Culture Exchange,” 2023 Overseas Hallyu Survey, 2023.
- [7] Lee, K. Y., & Park, S. H., “Topic modeling analysis for exploring new fields of Hallyu”, *Journal of the Korean Society for Entertainment Industry*, Vol 17, No. 3, pp. 13, 2023.
- [8] Lee, J. M., Lee, S. Y., Jung, N. H., & Ku, C. M., “A study on the national trends and perception differences of Hallyu based on Hofstede's cultural dimensions theory,” *Journal of Hotel and Tourism Research*, Vol 18, No. 2, pp. 21-40, 2016.
- [9] Lee, J. W., & Kim, Y. K., “A study on the relationship between the Korean Wave and national brand attitude: Focused on the moderating effect of cultural dimensions,” *Journal of Advertising and Public Relations Research*, Vol 11, No. 2, pp. 123-148, 2018.
- [10] Son, S. H., “A meta-analysis of academic papers: 10 years of Hallyu observed through the press and society,” *Media and Society*, Vol 17, No. 4, pp. 122-153, 2009.
- [11] Yun, N. R., Song, J. E., “Qualitative meta-analysis of Korean wave research - Focusing on content analysis of research participants,” *The Journal of Culture Contents*, Vol 10, No. 7-38, 2017.

Comparative Analysis of Chinese Personal Information Protection Textbooks

Eunsun Choi¹, Pureum Kim², Mijin Kim³, Yujin Jung³, Jinsu Kim², and Namje Park⁴

Creative Education Center, Jeju National University, Jeju, South Korea¹,
Department of Convergence Information Security, Graduate School, Jeju National University, Jeju, South Korea²,
Institute of Convergence Science, Technology, and Society, Jeju National University Jeju, South Korea³,
Department of Computer Education, Teachers College, Jeju National University, Jeju, South Korea⁴
[e-mail: choi910624@jejunu.ac.kr, kpr0016@stu.jejunu.ac.kr, mjkim@jejunu.ac.kr,
yujinjung@jejunu.ac.kr, kimjinsu@jejunu.ac.kr, namjepark@jejunu.ac.kr]

*Corresponding author: Namje Park

Abstract

As the development of Chinese digital technology is rapidly accelerating, information security is emerging as an important topic in China's digital environment. The Chinese government is making efforts to increase awareness by strengthening laws and regulations related to information security and conducting information security education. The purpose of this paper is to provide implications for research on the development of information security textbooks by analyzing three textbooks to raise awareness of personal information protection for teenagers and their parents in China. According to the analysis of the textbook contents based on the learning goal of "Information Society Responsibility" presented by the Chinese national curriculum, it was found that all textbooks included privacy-related topics; however, none of them covered artificial intelligence ethics. In particular, it was analyzed that most of the 'Youth Information Security Practice' textbooks were biased toward data security threats.

Keywords: Personal information security, textbook research, information society responsibility, data security threats, self-management skill

1. Introduction

China, which boasts a population of over 1.4 billion, is suffering from hacking and personal information leakage incidents. Accordingly, the "Decision of the Standing Committee of the National People's Congress on Strengthening Network Information Protection 2012" was announced in December 2012 [1]. Since then, personal information has been protected through laws such as the "Law of the People's Republic of China on the Protection of Consumer Rights and Interests" and the "Network Security Law of

the People's Republic of China". However, until recently, there was no single law for the protection of personal information, but efforts are being made in earnest to protect personal information and prevent infringement by enforcing "Personal Information Protection Law of the People's Republic of China" in November 2021 [2].

As people spend more time online in the future, various problems such as Internet addiction, cyber violence and invasion of privacy are occurring. Therefore, it is important to educate children from a young age to be alert and vigilant about various things that happen on the

This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2022S1A5C2A04092269), and this work was also supported by an Electronics and Telecommunications Research Institute (ETRI) grant funded by the Korean government. [23ZD1160, Regional Industry ICT Convergence Technology Advancement and Support Project in Daegu-Gyeongbuk (Mobility)].

Internet. Personal information protection education includes information on preventing information leaks as well as intellectual property rights, cyberbullying, information protection laws, and cyber etiquette [3].

This paper aims to provide implications by comparing and analyzing three information security textbooks designed for children, adolescents, and parents in China. Therefore, we analyzed the internal system of the textbook and made suggestions based on the analysis results.

2. Methodology

This paper was written with the purpose of comparing and analyzing Chinese information security textbooks to find the characteristics of each textbook and provide implications. To this end, the study was conducted in three stages. Initially, we collected Chinese information security textbooks. Subsequently, we established textbook evaluation standards. Finally, we conducted a comparative analysis of the internal structures of the chosen textbooks.

In this paper, textbooks mainly targeting elementary and middle school students were used for analysis, excluding professional books published as university textbooks. The textbooks used in the analysis are ‘Internet Safety and Informatization’, ‘Becoming Good Friends with the Internet’, and ‘Practicing Youth Information Security’. ‘Internet Safety and Informatization’ was written under the leadership of a national organization, and ‘Becoming Good Friends with the Internet’ was published by a publisher affiliated with the National Air Force Medical University. This textbook was created as a health growth series for elementary and middle school students, and is designed so that parents can also read it together. ‘Youth Information Security Practice’ is a textbook used for elective subjects in elementary and middle schools in Beijing.

To analyze textbooks, we used the analysis frame of reference used by Kim (2011) in her comparative study of computer textbooks [4]. This analysis standard consists of an external system that includes the textbook's appearance, publishing system, content selection, and editing, and an internal system that includes unit composition and learning content for each unit. Among these, we conducted an internal system analysis to compare the contents of the textbooks in detail.

3. Comparative Analysis

In order to analyze the internal system of the textbook, the structure of the units and the learning content for each unit were analyzed.

First, when examining the composition criteria of the unit, the unit title was composed of the contents of ‘Internet Safety and Informatization’ and ‘Becoming Good Friends with the Internet’, but ‘Practice of Youth Information Security’ is applied to each unit with a consistent composition system. Although not all units are the same, it is analyzed to follow the format of mission outline - basic knowledge - mission practice - read.

Second, the analysis of learning contents for each unit is as follows. ‘Internet Safety and Informatization’ is characterized by explaining how informatization is progressing in daily life, etiquette to be observed online, and precautions to be taken when using the Internet in a storytelling format using easy-to-understand examples. ‘Becoming Good Friends with the Internet’ specifies the pros and cons of the Internet and lists in detail the risk factors that students may encounter while using the Internet. In particular, it provides various ways to guide students who are addicted by citing real or imaginary cases of addicted students. Unlike ‘Internet Safety and Informatization’ and ‘Becoming Good Friends with the Internet,’ ‘Youth Information Security Practice’ is a textbook that mainly focuses on cyber security technology practice. Therefore, it is designed to explain various attacks and related terms such as password cracking, spoofing, sniffing, malware, and web attack techniques, and to help students learn cyber defense techniques. **Table 1** represents the learning content for each unit of the three textbooks.

Table 1. Internal system of Chinese information security textbooks

| Textbook | Element |
|---|--|
| Intenet Safety and Informatization | Informatization and social life |
| | Internet ettiquete |
| | Use the Internet safely |
| Becoming Good Friends with the Internet | Know the dangers of the Internet and stay away from Internet traps |
| | How can I save you, a child addicted to the internet? |

| | |
|--|---|
| Youth Information Security Practice | Correctly instruct children on “green” Internet access |
| | Build an “Internet firewall” through psychological counseling |
| | Network scanning and sniffing |
| | Cryptography decryption techniques |
| | Spoofing |
| | Vulnerability exploitation techniques |
| | Denial of service attack techniques |
| | Malware techniques |
| | Web attack techniques |
| | Information protection techniques |

In addition, we analyzed the learning content of the three textbooks based on the learning goal of ‘Information Society Responsibility’ presented in China’s 2022 revised curriculum. It can be seen that ‘Internet Safety and Informatization’ has a relatively more even distribution of educational contents than ‘Becoming Good Friends with the Internet’ and ‘Youth Information Security Practice’, and among them, Internet etiquette has the highest frequency (4 times/17.39%). ‘Becoming Good Friends with the Internet’ was described with the greatest emphasis on self-management skills (33 times/40.74%), which comes from suggesting a variety of ways to escape Internet addiction. It was analyzed that these two textbooks did not account for a large amount of explanations related to data security threats or security technologies compared to ‘Youth Information Security Practice’. In ‘Youth Information Security Practice’, content related to data security threats was recorded with the highest frequency, accounting for more than half of all textbooks (23 times/53.49%), and cyber security technology was also introduced the most among the three books (5 times/11.63%). All three textbooks did not cover artificial intelligence ethics and algorithm-related content.

4. Conclusions

Due to digital transformation and the development of information and communication technologies throughout the international community, new technologies such as the

Internet, cloud, mobile, Internet of Things, and blockchain are constantly being developed. China, one of the most populous countries in the world, is one of the countries with a very fast pace of informatization. Therefore, the protection and security of sensitive information such as personal information and financial information is emerging as a very important issue.

This paper compared and analyzed the internal structure of information security textbooks developed for elementary and middle school students in China. As a result of the analysis, the textbooks selected as the textbook for school elective subjects were developed for practicing cyber attack and security techniques. And the remaining two textbooks targeting many children and parents were developed to teach Internet etiquette and self-management skills. The content was mainly about how to use it safely. Through the analysis results, this paper can make the following suggestions. Unless the textbook has been professionally developed for practice, it should be designed to include content from a variety of areas. Furthermore, it is necessary to organize the content according to the standards presented in each country and validate it through expert validation.

References

- [1] Seo, E. K. “A Study on the Legislation of Personal Information Protection in China - Focusing on the Cyber security law,” *Journal of Sinology and China Studies*, Vol. 72, No. 0, pp. 131-155, 2017,
- [2] Korea Internet & Security Agency. “Status of China's personal information protection law administrative system and major violations,” Korea Internet & Security Agency, 2021.
- [3] Park, S. S., and Han, S. K. “KASP-Integrated Learning Model for the Information Security Education in Elementary Education,” *Journal of The Korean Association of information Education*, Vol. 14, No. 2, pp. 157-164, 2010.
- [4] Kim. Y. M. “Comparative Study on the Computer Textbooks between India and Korean High school – for the Academic high school,” Master Thesis. Konkuk University, Seoul, Korea, 2011.

Perception Study on Digital Literacy Teachers Should Possess Through Text Mining Analysis

Eunsun Choi¹, Hyunjeong Lim², Keunbae Choi³, Heuipi Kim⁴, Ji-Yun Kim⁵, Hosoo Lee³,
Yeon-Soo Joo², and Namje Park⁵

Creative Education Center, Jeju National University, Jeju, South Korea¹,
Institute of Convergence Science, Technology, and Society, Jeju National University Jeju, South Korea²,
Department of Mathematics Education, Teachers College, Jeju National University, Jeju, South Korea³,
Department of Practical Arts Education, Teachers College, Jeju National University, Jeju, South Korea⁴,
Department of Computer Education, Teachers College, Jeju National University, Jeju, South Korea⁵
[e-mail: choi910624@jejunu.ac.kr, celloart11@naver.com, kbchoe@jejunu.ac.kr, khp@jejunu.ac.kr,
jiyunkim@jejunu.ac.kr, hosoo@jejunu.ac.kr, pcc762cnd@jejunu.ac.kr, namjepark@jejunu.ac.kr]

*Corresponding author: Namje Park

Abstract

The aim of this study was to explore perceptions of the digital literacy that teachers in the era of intelligence information technology should possess. We conducted interviews with educational experts, future technology specialists, and elementary and middle school teachers to investigate the digital competencies and teaching, learning, and assessment skills that educators teaching the digital native generation need to develop. The interview data were preprocessed and analyzed using text mining techniques. We examined the top 10 words based on TF-IDF weights. Keyword frequency analysis revealed that understanding of “technology” (Frequency = 7, TF-IDF = 9.09) was perceived as the most important. Centrality analysis results indicate that “students” (Degree Centrality=0.57, Closeness Centrality=0.66, Betweenness Centrality=0.26) exhibit the highest degree centrality, closeness centrality, and betweenness centrality. This suggests that teachers' competencies are more geared towards fulfilling a vital role in assisting students in acquiring literacy skills rather than solely focusing on enhancing their own competencies.

Keywords: Digital literacy, teacher competency, text mining, TF-IDF, centrality

1. Introduction

In light of the evolving educational landscape, the proficiency of educators in harnessing digital media for instructional purposes is gaining significance [1]. Furthermore, there is an increasing emphasis on their capacity to rework, generate, and apply information in conjunction with pedagogical expertise. However, as most of

the research related to digital literacy targets students and the general public [2]. Therefore, there is a paucity of research concerning educators' digital literacy. In this paper, we conducted focus group interviews with education experts, technology experts, and elementary and secondary teachers. The aim of this study was to analyze the digital literacy skills that teachers should possess in the era of intelligent information technology through text mining.

This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2022S1A5C2A04092269), and this work was also supported by an Electronics and Telecommunications Research Institute (ETRI) grant funded by the Korean government. [23ZD1160, Regional Industry ICT Convergence Technology Advancement and Support Project in Daegu-Gyeongbuk (Mobility)].

2. Methods

This study aimed to examine experts' perceptions of teachers' digital literacy using text mining techniques. Text mining involves extracting valuable words from an unstructured collection of texts, computing their frequencies, and deriving meaning [3]. To achieve this, we conducted a Focus Group interview (FGI) involving two professors from the Department of Education, three IT professionals, one researcher, four elementary school teachers, and one middle school teacher. We collected their opinions through a Google form. For data analysis, we employed the text analysis tool Textom. To prepare the collected data for analysis, we conducted preprocessing tasks such as tokenization, stopword removal, and synonymization. We also selected key terms by calculating the frequency of refined data through text mining. Furthermore, we assessed the Term Frequency-Inverse Document Frequency (TF-IDF) values to determine word importance within the document [4]. Additionally, a centrality analysis was performed to identify influential nodes.

3. Results

In order to analyze the perception of digital literacy that teachers should have, we asked 11 subjects to freely inform them of their opinions on the digital competencies and digital-based teaching, learning, and evaluation capabilities that teachers teaching the digital native generation should cultivate. In this paper, their opinions were collected and keyword frequency analysis and centrality analysis were conducted.

3.1 Keyword Frequency Analysis

71 words were found in the opinions of the collected experts, and the top 24 words were analyzed based on the frequency of appearance and importance of keywords. the top 10 words with the highest frequency of occurrence were as follows: "technology" (Frequency = 7, TF-IDF = 9.09), "students" (Frequency = 5, TF-IDF = 7.41), "literacy" (Frequency = 4, TF-IDF = 6.82), "ability" (Frequency = 4, TF-IDF = 6.82), "knowledge" (Frequency = 4, TF-IDF = 6.82), "information" (Frequency = 3, TF-IDF = 4.80), "platforms" (Frequency = 3, TF-IDF = 7.19),

"skill" (Frequency = 3, TF-IDF = 5.98), "capability" (Frequency = 3, TF-IDF = 5.98), and "aspects" (Frequency = 2, TF-IDF = 4.80).

Among these, "platforms" exhibited a higher TF-IDF value compared to the word appearance frequency. It was identified as a keyword carrying a key message. In addition, although the word "devices" is not very frequent, the TF-IDF value is relatively high (Frequency = 2, TF-IDF = 6.18). Therefore, the word is also considered an important word. Table 1 shows the results of the keyword frequency analysis of the top 10 words.

Table 1 Results of keywords analysis of the top 10 words

| Rank | Word | Frequency | Word | TF-IDF |
|------|-------------|-----------|------------|--------|
| 1 | technology | 7 | technology | 9.09 |
| 2 | students | 5 | students | 7.41 |
| 3 | literacy | 4 | platforms | 7.19 |
| 4 | ability | 4 | knowledge | 6.82 |
| 5 | knowledge | 4 | ability | 6.82 |
| 6 | information | 3 | literacy | 6.82 |
| 7 | platforms | 3 | devices | 6.18 |
| 8 | skill | 3 | ideas | 6.18 |
| 9 | capability | 3 | capability | 5.98 |
| 10 | aspects | 2 | skill | 5.98 |

3.2 Centrality Analysis

To quantify the relative importance of specific nodes in the refined words, we conducted the degree centrality, closeness centrality, and betweenness analysis. The degree centrality represents the connection relationship of words, and it can be understood that the higher the value, the higher the influence on other words. "Students" (Degree Centrality = 0.57), "literacy" (Degree Centrality = 0.48), "technology," "knowledge," and "teachers" (Degree Centrality = 0.38) were ranked in the following order.

Closeness centrality is an indicator of how short the path one node has to reach another node, and the higher the proximity centrality, the more central it is to the network. Closeness centrality was found to be highest for "student," "literacy" (Closeness Centrality = 0.66), "teachers" (Closeness Centrality = 0.62), "information" (Closeness Centrality = 0.60), and "technology" (Closeness Centrality = 0.58) in that order.

Betweenness centrality is a measure of the extent to which a particular node acts as an

intermediary between other nodes in the network, reflecting its role as a mediator. Other nodes heavily depend on nodes with high betweenness centrality. High betweenness centrality was observed in the following order: "students" (Betweenness Centrality = 0.26), "literacy" (Betweenness Centrality = 0.15), "knowledge" (Betweenness Centrality = 0.12), "technology," and "platforms" (Betweenness Centrality = 0.10). The centrality analysis results for the top 10 words are as shown in **Table 2**.

Table. 2 Results of centrality analysis of the top 5 words

| Rank | Word | DC | Word | CC | Word | BC |
|------|--------------|------|--------------|------|-------------|------|
| 1 | students | 0.57 | students | 0.66 | students | 0.26 |
| 2 | literacy | 0.48 | literacy | 0.66 | literacy | 0.15 |
| 3 | knowledge | 0.38 | teachers | 0.62 | knowledge | 0.12 |
| 4 | technology | 0.38 | information | 0.60 | technology | 0.10 |
| 5 | teachers | 0.38 | technology | 0.58 | platforms | 0.10 |
| 6 | information | 0.33 | knowledge | 0.57 | teachers | 0.08 |
| 7 | platforms | 0.29 | platforms | 0.55 | information | 0.07 |
| 8 | intelligence | 0.29 | intelligence | 0.54 | ability | 0.07 |
| 9 | ability | 0.24 | capability | 0.54 | usage | 0.04 |
| 10 | usage | 0.24 | ideas | 0.54 | tools | 0.03 |

*D.C. = Degree Centrality, C.C. = Closeness Centrality, B.C. = Betweenness Centrality

4. Conclusions

Digital literacy, a term first introduced by Gilster (1997), was initially defined as critical thinking, information comprehension, and utilization in addition to the ability to use computers [5]. Since then, the definition has diversified with the development of digital devices and technologies, but in recent years, the ability to share data in the web space, communicate and reconstruct data has been emphasized along with the basic utilization ability of digital devices.

Teachers' digital literacy should be added with educational significance to generally defined digital literacy. For teachers who teach the digital native generation, digital literacy is becoming an essential competency to be cultivated.

In this study, we investigated and analyzed various experts' perception of digital literacy that teachers should have.

As a result of keyword frequency analysis, it was found that "technology" and "students" had a high frequency of appearance in documents. This means that not only the ability to use digital media but also the understanding of various technologies is an important part of digital literacy, and that students must be at the center of cultivating teachers' competencies. Additionally, "devices" may not have a high overall frequency of occurrence, but it appeared frequently in specific documents. It means that some experts have seen that teachers' ability to select and handle digital devices needed in class is regarded as a teacher's digital literacy.

After conducting centrality analysis, it was determined that "student" exhibited the highest values in degree centrality, closeness centrality, and betweenness centrality. It is in line with the keyword frequency analysis results, suggesting that teachers' digital literacy is meaningful only when students' digital competencies are improved. It should be noted that the centrality value of "technology" was generally high. In the end, it means that teachers' approach to technology is a key element of digital literacy, as in the results of keyword frequency analysis.

References

- [1] D. Roh, "Elementary school teachers' reflection on the class in the Covid-19 situation – Focusing on digital literacy," *The Korean Journal of Literacy Research*, Vol. 13, No. 1, pp. 171-202, 2022.
- [2] J. Kim. "A Study on the digital literacy influencing on teaching quality," Doctoral Thesis, Chungnam National University, 2021.
- [3] K. E. Voss, E. R. Spangenberg, and B. Grohmann, B. "Measuring the hedonic and utilitarian dimensions of consumer attitude," *Journal of Marketing Research*, Vol. 40, No. 3, pp. 310-320, 2003.
- [4] L. C. Freeman, D. Roeder, and R. R. Mulholland, "Centrality in social networks: II. Experimental result," *Social Networks*, Vol. 2, No. 2, pp. 119-141, 1979.
- [5] P. Gilster, "Digital literacy," Wiley Computer Pub., New York, 1997.

Analysis of the Effectiveness of SW·AI Education Camp for Elementary School Students

Jaekwoun Shim, Wooyeol Kim and Youngkwon Bae

Dept. of Computer Education, Daegu National University of Education
Daegu, South Korea

[e-mail: jkshim@dnue.ac.kr, jhon@dnue.ac.kr, bae@dnue.ac.kr]

*Corresponding author: Youngkwon Bae

Abstract

Contrary to its image as an IT powerhouse, Korea is being criticized for being negligent in information education. In addition, software and artificial intelligence education started relatively late compared to other countries, so the improvement of the education system and environment is delayed. In Korea, the Digital Sprout Camp has been carried out nationwide since December 2022 as part of securing educational time in schools, strengthening teachers' capabilities, and preparing educational environment to provide information education in earnest. The Digital Sprout Camp, which was operated during the 2023 winter vacation period, ran 11,549 programs in 90 locations including elementary, middle, and high schools, universities, companies, and public institutions across the country, and a total of 197,262 people participated. This provided educational opportunities at a national level to improve computational thinking and digital capabilities. In this study, we analyze the perceptions of teachers and students who experienced the Digital Sprout Camp and draw implications for software and artificial intelligence education policies.

Keywords: Software Education, Artificial Intelligence Education, Elementary School Student Awareness, Teacher Awareness, Camp

1. Introduction

In the advanced information society, the ability to utilize artificial intelligence technology is a very important competency. Therefore, there is a need to teach students who will live in the future society the ability to use information technology. The purpose of K-12 education is focused on cultivating character and creativity. Additionally, it is also important to develop the capabilities necessary for future society. Therefore, in K-12, there is a need to provide education that utilizes technologies related to the 4th Industrial Revolution, such as artificial intelligence and software.

Many countries are already providing students with education using computers, artificial intelligence, and software. In particular, the importance of software education based on computational thinking is emphasized. In order to prepare for the future society, efforts are being made to revise the curriculum or improve the educational environment by using digital more actively. In the United States, Association for Computing Machinery(ACM) and Computer Science Teachers Association(CSTA) published K-12 Computer Science Standards [1]. Computer Science Standards attempted to teach principles and concepts related to computer science along with digital citizenship using educational programming tools such as Scratch. In the UK, the emphasis was on students learning

how to create their own programs rather than using programs implemented by others. In order to achieve this goal, starting in 2014, the existing ‘ICT’ subject was changed to ‘computing’ subject, and it was included in the required basic curriculum for K-12 so that all students can complete computer programming education [2]. China developed an artificial intelligence textbook in K-12, tested it in pilot schools, and is now expanding and operating it nationwide [3]. Japan announced the ‘New Learning Guidelines’ in 2017 and made programming and coding education mandatory starting in 2020. The educational goal was not to train programming and coding experts, but to improve students’ computational thinking [4].

Each country is making various efforts to improve computational thinking, but students and parents perceive that change in education is slow. In Korea, the importance of artificial intelligence education and the need for software education are fully understood, but public education does not provide education or the quality of education is unsatisfactory, leading to dependence on private education [5]. This response from parents is due to insufficient training time, teacher competency, and preparation of the educational environment to systematically implement software education. The Ministry of Education has made various efforts to make up for the shortcomings in software education and artificial intelligence education, such as making coding education mandatory in K-12, increasing the number of education hours, and providing teacher training [6].

Among various policies, Digital Sprout Camp has been carried out nationwide in Korea since December 2022 as part of securing information education time in schools, strengthening teachers’ capabilities, and preparing the educational environment [7]. The Digital Sprout Camp, which was operated during the winter vacation period in 2023, ran 11,549 programs in 90 locations including elementary, middle, and high schools, universities, companies, and public institutions across the country, and a total of 197,262 people participated [8]. This provided educational opportunities at a national level to improve computational thinking and digital capabilities.

The Digital Sprout Camp, which was held nationwide in Korea, was a meaningful

educational venue where high-quality software and artificial intelligence education was provided to K-12 students. This study aims to analyze teacher and student satisfaction with the Digital Sprout Camp held in the winter of 2023 and derive policy implications.

2. Method

2.1 Experiment Participant and Process

The subjects of the study were 1,305 first to sixth grade elementary school students who participated in the camp. We recruited 216 first and second graders, 654 third and fourth graders, and 435 fifth and sixth graders. There were 45 teachers in charge of the student camp. The study was divided into 3 groups: 1st and 2nd grade, 3rd and 4th grade, and 5th and 6th grade elementary school students, and the camp was conducted for 2 days (8 hours) considering their level.

Table. 1. Participant

| Category | 1-2 grade | 3-4 grade | 5-6 grade |
|----------|-----------|-----------|-----------|
| Student | 216 | 654 | 435 |
| Teacher | 3 | 22 | 20 |

2.2 Research Tool

The scale constructed by this research team has a total of 22 questions: 5 questions about interest in software and artificial intelligence, 5 questions about self-efficacy, 4 questions about outcome expectations, 4 questions about value perception, and 4 questions about career goals. The questions were modified and supplemented to fit the purpose of this study with reference to previous research so that they could be easily understood and answered by elementary school students [7]. Among the student survey questions, the interest element is as follows. I enjoy studying software and artificial intelligence. Software and artificial intelligence classes are fun. I am very interested in software and artificial intelligence. I like learning activities related to software and artificial intelligence. There is a lot of interesting content in the software and artificial intelligence class.

The questionnaire on self-efficacy factors is as follows. Learning about software and artificial intelligence is not difficult. I can learn software and artificial intelligence just as well as my

friends. Understand how software and artificial intelligence are used in our lives. I can excel at tasks related to software and artificial intelligence. You can use what you have learned about software and artificial intelligence.

The questionnaire on results expectation elements is as follows. Software and artificial intelligence will help my life in the future. Learning about software and artificial intelligence will help you get a better job. By learning software and artificial intelligence, I will be able to do what I am interested in. By learning software and artificial intelligence, you will be able to make the world a better place.

The questionnaire on value recognition elements is as follows. Software and artificial intelligence play an important role in developing our country. Software and artificial intelligence help our lives. Software and artificial intelligence have a lot to do with our lives. Software and artificial intelligence are essential to us.

The survey on career goals elements is as follows. I want to enter a department related to software and artificial intelligence in college. I would like to further study related to occupations that utilize software and artificial intelligence. I want to learn about jobs related to software and artificial intelligence. When I become an adult, I will choose a career that is highly related to software and artificial intelligence.

The teacher surveyed 12 questions as follows. The program operation time was appropriate (Q1). The number of participants in the program was appropriate (Q2). The program content was appropriate (Q3). The materials, teaching materials, and media used in the program were appropriate (Q4). The environment for the program was appropriately allocated (Q5). I am overall satisfied with the operation of this program (Q6). I think this program was helpful to students in terms of their studies, careers, and leisure (Q7). This program is helpful when students take software and artificial intelligence classes (Q8). The contents of this program can be applied to classes and used as class materials (Q9). In this program, you can learn about the content and direction of new artificial intelligence and software education (Q10). If I get the chance, I would like to apply for this type of program again (Q11). In the future, such programs need to be opened and operated continuously (Q12).

For analysis, the mean and standard deviation were calculated using the spss 25.0 program.

3. Result

As a result of the study, it was analyzed that both teachers and students were overall satisfied with the camp, with an average score of 4 or more (5-point Likert scale). In the case of teachers, it was analyzed that they sympathized with the content and purpose of the camp. In other words, it can be said that the original purpose of opening the camp, which is to spread artificial intelligence and software education to schools, has been achieved.

Table 2. Teacher Perceptions of Camp

| Catagory | 1-2 grade | 3-4 grade | 5-6 grade |
|----------|------------|------------|------------|
| Q1 | 4.67(0.58) | 4.86(0.35) | 4.30(1.13) |
| Q2 | 4.67(0.58) | 4.73(0.63) | 4.45(0.89) |
| Q3 | 4.67(0.58) | 4.91(0.29) | 4.70(0.66) |
| Q4 | 5.00(0.00) | 4.82(0.50) | 4.70(0.98) |
| Q5 | 5.00(0.00) | 4.86(0.35) | 4.55(0.94) |
| Q6 | 5.00(0.00) | 4.82(0.50) | 4.40(1.05) |
| Q7 | 4.67(0.58) | 4.82(0.39) | 4.50(1.15) |
| Q8 | 5.00(0.00) | 4.73(0.63) | 4.50(1.15) |
| Q9 | 5.00(0.00) | 4.77(0.61) | 4.60(0.94) |
| Q10 | 5.00(0.00) | 4.77(0.61) | 4.60(0.99) |
| Q11 | 5.00(0.00) | 4.91(0.43) | 4.30(1.26) |
| Q12 | 5.00(0.00) | 4.77(0.61) | 4.60(1.10) |

Through the camp, student perceptions were surveyed based on the following factors: Interest, Self-Efficacy, Results Expectation, Value Recognition, and Career Goals.

Table 3. Student Perceptions of Camp

| Catagory | 1-2 grade | 3-4 grade | 5-6 grade |
|---------------|----------------|----------------|----------------|
| Interest | 4.35 (0.70) | 4.28 (0.82) | 4.00 (0.88) |
| Self-Efficacy | 4.22 (0.72) | 4.13 (0.84) | 4.00 (0.85) |
| Results | 4.14 | 4.06 | 4.01 |
| Expectation | (0.79) | (0.90) | (0.88) |
| Value | 4.13 | 4.17 | 4.17 |
| Recognition | (0.77) | (0.82) | (0.82) |
| Career Goals | 3.83 (1.01) | 3.59 (1.10) | 3.54 (1.09) |

As a result of the study, it was analyzed that there was a positive perception, but there was little

change in perception at the level of career occupation. In addition, the positive impact was found to decrease as the grade increased, suggesting the need to apply software and artificial intelligence education at an early age.

4. Conclusions

Korea has been criticized for lacking educational investment and policies in software and artificial intelligence compared to other countries. Digital Sprout Camp spread software and artificial intelligence education throughout Korea with the aim of overcoming this situation. The purpose of this study was to derive policy implications for the elementary level of artificial intelligence and software education through the camp. First, although the short-term camp was helpful in cultivating positive perceptions, it was found to have no effect on changing students' perceptions in terms of future career paths and jobs. Therefore, in order to fully achieve the purpose of education to respond to a society that is changing due to software and artificial intelligence, it is necessary to find ways to provide regular education in public education. Second, since the effects of education are greater in lower grades than in upper grades, there is a need to come up with a plan to provide software and artificial intelligence education at an early age. Lastly, it is necessary to recall that national investment had this positive effect. Continuous investment is needed at the future level of the country. Future research needs to investigate parents' perceptions, and continuous follow-up studies of students and teachers are needed.

References

- [1] M. Bauer, B. Bruegge, G. Klinker et al., "Design of a component-based augmented reality framework," in *Proc. of International Symposium on Augmented Reality (ISAR)*, 2001.
- [2] Computer Science Teachers Association, "K-12 Computer Science Standards," revised 2017.
- [3] H.R. Kim, "The Research about Policy Background of Computer Science Education in UK School : Lesson from the UK," *Journal of The Korean Association of Information Education*, Vol 20, No. 2, pp. 207-218, 2016.
- [4] China State Council, "Notice of the State Council Issuing the New Generation of Artificial Intelligence Development Plan," 2017.
- [5] KERIS, "ICT4ED 2017: Overseas Software Education Trend Analysis and Implications," 2017.
- [6] Lee, J. H., Ha, J. S. "'AI education' to be introduced as a regular curriculum in 2025, on-site. reactions are mixed," News Article on December 06, 2020 from Chosun Ilbo (in Korean).
- [7] Ministry of Education, "Cultivating 1 million talents to become protagonists of the digital era," Press Release on August 22, 2022.
- [8] Korea Foundation for the Advancement of Science & Creativity, "Analysis of the effectiveness of SW·AI education camp," 2023.
- [9] Korea Foundation for the Advancement of Science & Creativity, "A Study on Development Strategies through Performance Analysis of the Digital New Software·AI Camp," 2023.

Position Estimation using Stereo Camera on AR Glass and 3D Reconstructed Map

Sukwoo Jung, Youn-Sung Lee, Yunju Lee, and KyungTaek Lee

Contents Convergence Research Center, Korea Electronics Technology Institute,
Seoul, South Korea

[e-mail: swjung@keti.re.kr]

*Corresponding author: Sukwoo Jung

Abstract

Recent studies in sensor position estimation have gained significant attention due to their broad applicability across various fields, including robotics, Virtual Reality (VR), and Augmented Reality (AR). The current state-of-the-art indoor tracking methods typically rely on technologies such as Wi-Fi or Bluetooth. In this research, we introduce an innovative approach that departs from traditional techniques, which typically rely solely on images or IMU (Inertial Measurement Unit) sensors for position estimation. Instead, we used pre-reconstructed 3D maps to significantly enhance the accuracy of the pose estimation. To achieve this, we utilize high-precision indoor maps obtained through the use of the Leica RTC360 laser scanner. These 3D reconstruction maps are used for training our system, allowing us to effectively track objects using data from camera sensors and accurately calculate positional coordinates within indoor environments.

Keywords: 3D reconstruction, pose estimation, AR glasses, feature matching, stereo camera

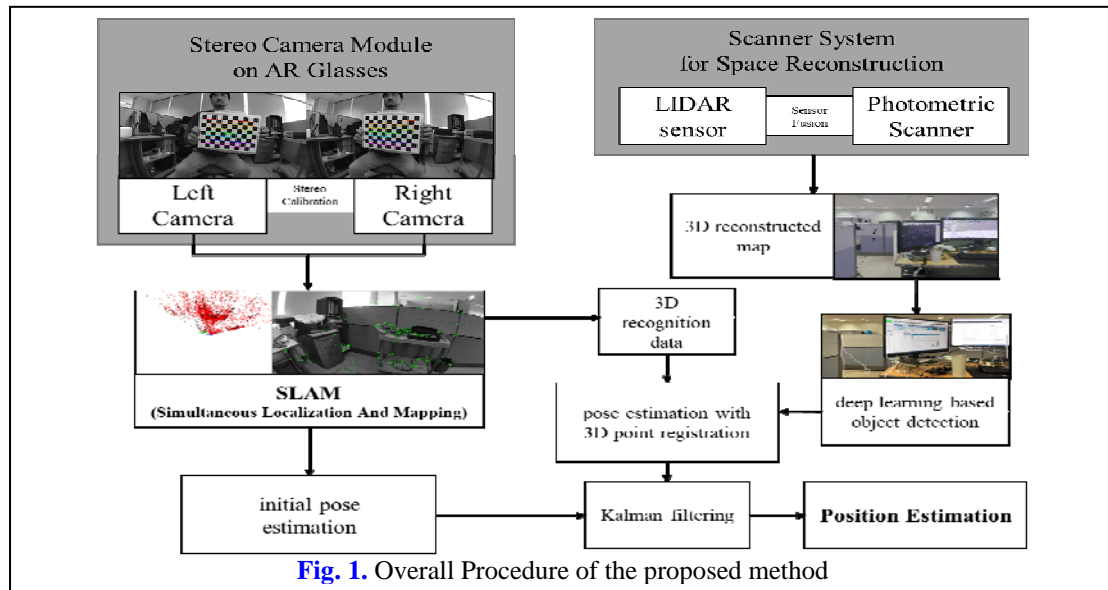
1. Introduction

In today's landscape, widely adopted position estimation systems primarily come in the form of Augmented Reality (AR) applications [1].

While the Global Positioning System (GPS) is the go-to choice for outdoor scenarios, it falls short in reliability when deployed indoors. On the other hand, cutting-edge indoor localization methods utilize the devices known as beacons. However, WiFi based localization demands the establishment of a dense network of access points, resulting in substantial maintenance costs, while offering only limited positional accuracy.

Various studies have explored methods for position tracking, employing either single or multiple sensors [2] such as Inertial Measurement Units (IMUs), or approaches based on the recognition of feature points [3] within images, including Simultaneous Localization and Mapping (SLAM). However, these methods often have limitations in terms of accuracy and real-time performance. 3D spatial reconstruction refers to the process of reconstructing the 3D shape and 2D texture of the objects or environments in 3D space [4].

In this paper, we present a method that combines 3D spatial reconstruction information with 3D recognition technique to enhance the accuracy and efficiency of estimating the position.



2. Proposed Method

The overall procedure of proposed method is vividly illustrated in Fig. 1. To begin the process, we obtained point cloud data by employing LIDAR scanner for 3D spatial mapping. These diverse multiple sets of 3D data were merged to create a unified reconstructed map [5]. From this map, we extracted various types of features points and generated templates for matching. Next, we collected data from stereo cameras attached to AR glasses, calibrated the two cameras using a chessboard panel [6], and then used SLAM with video from the stereo cameras to determine the sensor's position and the 3D information of the surrounding environment. Finally, we further utilized the Kalman Filtering method to find the optimal position by matching the pre-reconstructed 3D map with the surrounding environment. Finally, we utilized the Kalman Filtering method to find the optimal position by matching the pre-reconstructed 3D map with the surrounding environment. It's important to recognize the challenges posed by issues like significant occlusions in images or substantial changes in previously reconstructed maps, which may affect object recognition accuracy. To address these issues, we can leverage dynamic object recognition technology to distinguish moving objects and employ SLAM technology to dynamically perceive the environmental changes and update the 3D map.

3. Conclusions

In conclusion, this paper has introduced a novel position tracking method that combines 3D reconstruction information with 3D recognition technology, enhancing user experiences in various applications and contributing to future technological advancements. As we conclude this endeavor, it's crucial to emphasize the importance of continuous research and innovation in the field of 3D map-based position estimation to meet the evolving demands of technologies and applications.

References

- [1] G. Georg, E. Vonach, and H. Kaufmann. "HyMo Track: A mobile AR navigation system complex indoor environments," *Sensors*, 16.1, 17, 2015
- [2] B. Gaurav. "A survey of augmented reality navigation," *arXiv preprint arXiv:1708.05006*, 2017.
- [3] A. Ziad, Z. Shaghaghian, and W. Yan. "BIMxAR: BIM-empowered augmented reality for learning architectural representations," *arXiv preprint arXiv:2204.03207*, 2022.
- [4] S. Jung, et al. "3D reconstruction using 3D registration-based ToF-stereo fusion," *Sensors* 22.21, 8369, 2022.
- [5] S. Jung, et al. "Range image registration based on 2D synthetic images," *Computer-Aided Design* 94, 16-27, 2018.
- [6] S. Jung, S. Park, and K.T. Lee. "Pose tracking of moving sensor using monocular camera and IMU sensor," *TIIS*, 15.8, 3011-3024, 2021.

An Autonomous Control System Implementation for Maintaining Dissolved Oxygen Level in Aquaculture

Juhyoung Sung, Sungyoon Cho, Wongi Jeon, Yangseob Kim and Kiwon Kwon

Korea Electronics Technology Institute (KETI)

Seongnam-si, South Korea

[e-mail: {jh.sung, sycho, jeonwg, likant, kwonkw}@keti.re.kr]

*Corresponding author: Kiwon Kwon

Abstract

There are many indicators to be considered to manage the aquaculture environments. Among them, managing a dissolved oxygen (DO) level is significant because it is relevant to the fish welfare. However, the DO level is highly affected by various internal and external factors such as water temperature and feeding timing. Moreover, the appropriate DO level is also changed according to the fish rearing environments. Since the DO levels that are too high or too low are not suitable for survival conditions the DO level should be maintained within a certain range. Typically, the aquaculture manager directly intervenes to maintain the DO level. However, it is inefficient due to the nature of fish farms that must operate all the time. Moreover, inaccuracy is also inevitable according to the human error. In this paper, we present an autonomous control system implementation to manage the DO level to handle this problem. Our solution shows an overall system architecture and how the DO level can be controlled in conjunction with the actual water tank.

Keywords: Aquaculture, Autonomous Control, Dissolved Oxygen

1. Introduction

In aquaculture environments, the managers should take care of a number of factors to increase the productivity. Especially, an environmental indicator such as a dissolved oxygen (DO) level is one of the major factors since the DO level is closely related to the fish growth. However, the DO level is varying frequently according to the numerous conditions. For example, as the fishes in the tank become more active, the DO level highly decreases with feed supply. Since low DO is not suitable for survival environments, it is important to increase the dropped DO level again. In contrast, the excessively high DO level is also not appropriate since the movement of fishes gets to slow down. Traditionally, the manager manually control the

DO level. However, there are several limitations since aquaculture need to be operated 24 hours/7 days. Moreover, the human error also can occur. To deal with this kind of issues, autonomous management systems have been studied to minimize the human intervention [1]. In this paper, we go one step further and present a structure for linking with the aquaculture end point. Moreover, we verify the DO level through the actual water tank after configuring the test environment.

The rest of this paper is organized as follows. In section 2, we introduce a system model to manage the DO level in aquaculture. In section 3, we analyze the control result in conjunction with the real water tank before concluding this paper in section 4.

This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government(MSIT) (No. 2021-0-00225, Development of digital aqua twin core platform for optimal aquafarm design and operation)

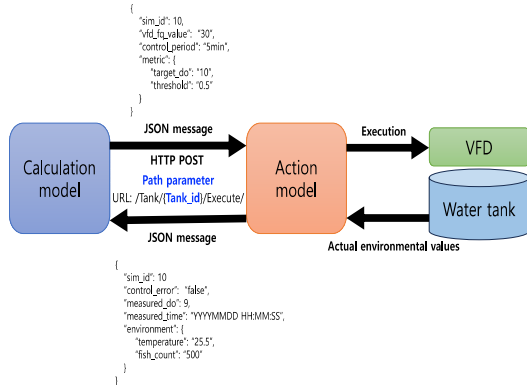


Fig. 1. Proposed DO level control system concept.

2. System Model

2.1 Basic considerations

There are several methods to control the DO level. Among them, we consider the method via a variable frequency drive (VFD). The VFD adjust the frequency of a motor which can affect the DO level in the water tank. As the value of the frequency increases, the water speed also increases which leads to increase the DO level. In contrast, when the frequency value decreases, the result becomes the opposite.

2.2 Control system implementation

As we mentioned in the previous subsection, the frequency value of VFD plays a key role in our control system. However, there are also several parameters other than the frequency value also to be considered which affect the DO level. Moreover, the manager also can configure the target tank with the desired DO level through the system. The system also receives the feedback results from the actual measured values and modifies the control value if it is required. The proposed control system consists of two models. The first model calculates the frequency value considering the environmental conditions in the water tank and transfers it with a control period. The second model triggers the VFD to adjust the frequency value according to the control information and returns the feedback results to the calculation model. The calculation model re-transfers the control value if the threshold condition is not satisfied. In order to support cross platform operation and real-time management, communication between the

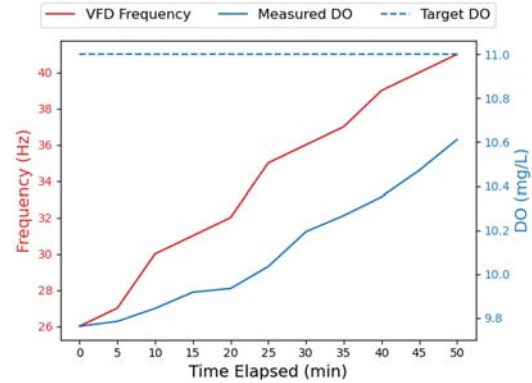


Fig. 2. Changes in frequency of VFD and DO level over time.

modules follows hypertext transfer protocol (HTTP) based representational state transfer application interface (REST API) [2]. Thus, all kinds of messages are expressed as javascript object notation (JSON). Fig. 1 illustrates the proposed control system implementation concept. Each model shares the common path parameters to indicate the target water tank and POST based HTTP method which can take advantage of security and containing various JSON key-value pairs.

3. Integrated Test Results

In order to demonstrate the operation of the proposed control system, we link the system to the actual water tank used in the real aquaculture industry. The manager configures the target DO level as 11 mg/L. Fig. 2 describes the variations of the frequency value of VFD and the DO level over time. As expected, the control system tries to increase the frequency value to meet the desired DO level as the calculation model and the action model interact. Since the threshold is 0.5 mg/L, the control has completed when the measured DO value reaches 10.61 mg/L. The frequency value remains as long as the DO level does not go out of the threshold range.

4. Conclusions

In this paper, we introduced the integrated control system to manage the DO level in the aquaculture industry. The system consists of the calculation model and the action model depending on the purpose. In addition, we also carried out verification on the implemented

system and confirmed that the system worked well as intended.

References

- [1] Z. Peng, Y. Chen, Z. Zhang, Q. Qiu, and X. Han, "Implementation of water quality management platform for aquaculture based on big data," *International Conference on Computer Information and Big Data Applications (CIBDA)*, pp. 70-74, 2020.
- [2] R. T. Fielding *et al.*, "Reflections on the REST architectural style and "principled design of the modern web architecture" (impact paper award)," in *Proc. 11th Joint Meeting on Foundations of Software Engineering (ESEC/FSE)*, pp. 4-14, 2017.

Homography-based Object Position Tracking Method using Fixed Camera

Youn-Sung Lee, Sukwoo Jung and Kyung-Taek Lee
Korea Electronics Technology Institute, Seoul, South Korea
[e-mail: yslee@keti.re.kr]
*Corresponding author: Youn-Sung Lee

Abstract

This paper presents the implementation of a homography-based method for tracking the position of objects using fixed cameras. The proposed method is implemented with a detection module, multi-process queues, and a GUI module. The primary advantage of this method is its capability to derive 3D position coordinates from 2D images captured by stationary cameras, ensuring both speed and real-time functionality.

Keywords: 3D reconstruction, pose estimation, navigation, feature matching, object detection

1. Introduction

Recent advancements in camera and AI technologies have led to the widespread utilization of object recognition across various domains. In manufacturing environments, there is an increasing demand for efficient methods to track the positions of workers and operating robots. Video-based object tracking methods commonly rely on hardware such as stereo cameras, Time-of-Flight (ToF) cameras, or depth sensors to measure the distance between the camera and objects[1,2]. This paper presents a method for tracking the position of objects using an RGB camera without the need for additional hardware like depth cameras or depth sensors.

2. Proposed Object Position Tracking Method

The proposed object position tracking method involves several steps: camera calibration, homography computation, object detection, and

2D-to-3D coordinate conversion processes. The camera calibration process calibrates the intrinsic and extrinsic parameters of a camera. This process includes rectifying distortion caused by the camera lens and establishing correspondence between the 2D image and the actual 3D scene. In the homography computation process, a homography matrix is calculated, representing the transformation between the 2D ground plane in the image captured by the camera and the 2D ground plane in the target space. The object detection process identifies the location of one or multiple objects within the camera-captured image, using bounding boxes to enclose the objects with rectangular shapes. The resulting bounding boxes indicate the positions of the objects within the 2D plane of the image. By applying the pre-computed homography matrix to the bottom center coordinates of the detected objects, the corresponding coordinates on the target plane are calculated. The 2D-to-3D coordinate conversion process transforms the 2D coordinates of objects in the camera-captured image into 2D coordinates of the ground plane in the target space. Assuming objects only move on

the ground, the 2D coordinates (x, y) of objects can be converted into 3D coordinates (x, y, z=0). This allows for tracking objects in a 3D space.

The proposed object position tracking method is implemented with a detection module, multi-process queues, and a GUI module as shown in Fig. 1, and its operation is as follows: The detection module receives real-time video data streaming from an IP camera through the RTSP protocol. The received video data is processed by a deep learning object detection model, specifically YOLO[3], to generate bounding boxes around objects, such as humans and robots. The generated bounding boxes are overlaid on the RGB image to create a visualization. The detection module extracts 2D coordinates from the generated bounding boxes and transforms them into 3D coordinates that can be mapped onto a 3D space. These 3D coordinates, along with the identifications of detected objects, are then transmitted to the GUI module through multi-process queues. The GUI module displays the visualization on the screen and represents the 3D coordinates in a 3D environment. Additionally, it provides information about the identifications of detected objects and their 3D coordinates. Fig. 2 shows the implementation of the GUI software for the proposed method.

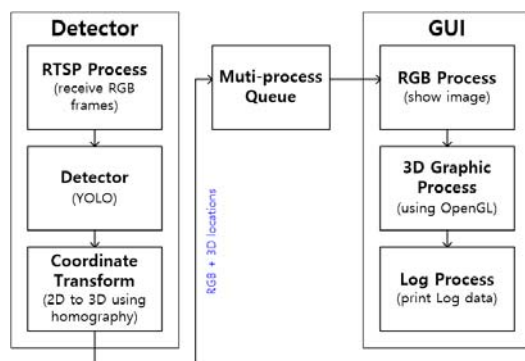


Fig. 1. Block diagram of the proposed object position tracking method

This method enables object position tracking using only cameras, eliminating the need for additional equipment such as stereo cameras, depth cameras, or depth sensors. It can also be visualized in virtual reality. However, it's important to note that there are still limitations, including challenges in object tracking in complex environments and reduced precision for

distant objects. Tracking object positions with fixed cameras using the proposed method holds significant potential for applications in computer vision and robotics, contributing to future research and development.

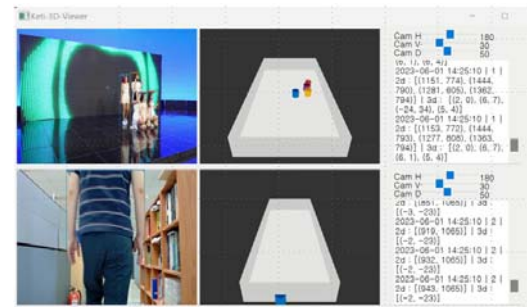


Fig. 2. GUI software of the proposed object position tracking method

3. Conclusions

In this paper, an object position tracking method using fixed cameras was proposed. This method is based on homography, which represents the transformation between the 2D ground plane of the camera-captured image and the 2D ground plane of the target space. The proposed method provides object identifications and their positional information. It also tracks the movement paths of objects. The key advantage of this method is its ability to extract 3D position coordinates from 2D images captured by fixed cameras, ensuring both speed and real-time functionality.

References

- [1] S. Jung, S. Park, KT Lee. "Pose tracking of moving sensor using monocular camera and IMU sensor," *KSII Transactions on Internet and Information Systems (TIIS)*, 15(8), 3011-3024, 2021.
- [2] S. Jung, YS Lee, KT Lee. "3D reconstruction using 3D registration-based ToF-stereo fusion," *Sensors*, 20(21), 8369, 2022.
- [3] R. Huang, J. Pedoeem, C. Chen. "YOLO-LITE: A Real-Time Object Detection Algorithm Optimized for Non-GPU Computers," *IEEE International Conference on Big Data*, 2018.

An Analysis of the Imported Ornamental Fish Quarantine Environment for CNN based Recognition of Ornamental Fish Species in Korea

Hyunhak Song¹, Sungyoon Cho¹, Yangsub Kim¹, WonGi Jeon¹, KyungWon Park¹,
Byoungchul Song¹, Jongho Paik² and Kiwon Kwon¹

¹Smart Network Research Center, Korea Electronics Technology Institute (KETI)
Seongnam-si, South Korea

[e-mail: {rainy930, sycho, likant, jeonwg, kwpark, songbc, kwonkw}@keti.re.kr]

²Department of Software Convergence, Seoul Women's University
Seoul, Republic of Korea

[e-mail: paikjh@swu.ac.kr]

*Corresponding author: Kiwon Kwon

Abstract

This paper analyzes the quarantine environment of imported ornamental fish in Korea in depth and studies the image recognition of ornamental fish in consideration of the complex characteristics in the environment. In this study, the unique elements of the quarantine environment were investigated and analyzed in detail. And based on the analyzed results, learning was performed by augmenting the learning dataset using image processing techniques, and the method of applying the conditions of the quarantine environment improved the accuracy of image recognition. The study conducted recognition experiments on a total of 50 species of ornamental fish to demonstrate that image augmentation methods similar to quarantine environments are effective.

Keywords: Ornamental fish, Quarantine, Augmentation, CNN, YOLO

1. Introduction

Convolutional Neural Network (CNN) is a key technology attracting attention in the modern image recognition field, showing high performance due to its unique structure and pattern recognition ability[1]. By effectively extracting and learning image features, CNNs realize accurate and efficient image classification and recognition in a variety of applications. At the same time, for effective learning of a CNN for a specific object, it is essential to construct a learning dataset with diversity in various environments. Data that takes diversity into account like this helps the

model learn robust features under various conditions and have high generalization ability[2]. This paper aims at robust image recognition for ornamental fish imported into Korea and analyzed several characteristics of the quarantine environment for ornamental fish in Korea. We then performed CNN learning and experiments using training image data based on the analyzed features and data that did not, and present comparative results.

2. Quarantine Environment Analysis

We actually visited domestic ornamental fish quarantine sites and analyzed environmental

characteristics. As shown in **Fig. 1.**, ornamental fish imported into Korea are individually packaged in plastic. Accordingly, the characteristics of ornamental fish change sensitively in images taken at quarantine sites depending on the medium of vinyl, lighting, and water. In the case of vinyl, if wrinkles occur, the shape of the ornamental fish may appear curved. Additionally, the unique color characteristics of ornamental fish may be expressed differently depending on the intensity and color of lighting. Water contained in plastic can distort the shape of aquarium fish and make them appear cloudy. In addition to these media, images taken while aquarium fish are moving may show shaking or various rotation angles.



Fig. 1. Conservation status of imported ornamental fish during quarantine process

3. Experiment

In this experiment, using the same CNN model, we compared the coronary recognition performance through each learning of the training image data with the quarantine environment analyzed in Chapter 2 and the data without. Version 4 of the YOLO model has been selected to improve processing time and awareness. Learning data with a quarantine environment was applied with single and compound conditions of rotation, inversion, blurring, distortion, and brightening using OpenCV library. Data that did not are augmented by applying only rotation and enlargement. As shown in **Fig. 2.**, a total of 50 ornamental fish classes were selected, and 2,000 image data were built for each class, and the YOLO model was learned 10,000 times with the option of batch=32 and subdivision=32. Using the finally extracted weight file, the figures are summarized in **Table 1.** as 100 test image data per class applied with a quarantine environment that was not used for learning. Here, 'A' means a model learned with an image applied with quarantine environmental conditions, and 'B' means a model learned with an image applied with some quarantine

environmental conditions.

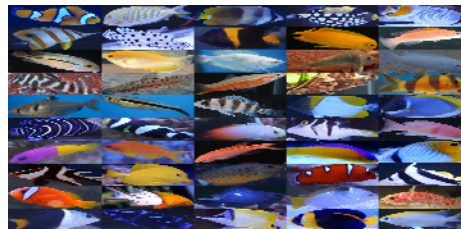


Fig. 2. Images of 50 species of ornamental fish trained from CNN model

Table 1. Images of 50 species of ornamental fish

| Type | A | B |
|------------------------|-------|-------|
| Inference Accuracy (%) | 89.11 | 97.24 |

4. Conclusions

This paper aims to recognize strong ornamental fish in the domestic quarantine environment, and analyzes the quarantine environment in detail and evaluates the learning performance of the artificial intelligence model through image augmentation according to conditions. Because it is difficult to collect images in an actual quarantine environment, we used open data such as Google and FishBase to augment and learn images to suit the quarantine environment, and achieved a higher recognition rate of approximately 8% than the existing method. This proves that the quarantine environmental conditions analyzed in this paper actually work meaningfully

References

- [1] J. G. A. Barbedo, "A Review on the Use of Computer Vision and Artificial Intelligence for Fish Recognition, Monitoring, and Management," *Fishes*, Vol. 7, No. 6., pp. 335, 2022.
- [2] C. Shorten, "A survey on Image Data Augmentation for Deep Learning," *J. Big Data*, Vol. 6, No. 1., 2019.
- [3] H. H. Song, "A study on Ornamental Fish Recognition in Quarantine Environment using CNN-based YOLOv4 and YOLOv5," *The 17th Asia Pacific International Conference(APIC)*, pp. 347-348, 2022.

Automatic Calibration of Multiple LIDAR using Plane in Vehicle Intersection Environment

Hyunhak Song, Byoungchul Song and Kyungwon Park

Korea Electronics Technology Institute (KETI)

Seongnam-si, South Korea

[e-mail: {rainy930, songbc, kwpark }@keti.re.kr]

*Corresponding author: KyungWon Park

Abstract

This paper proposes an automatic calibration method for the development of fully autonomous driving technology using multiple LiDAR sensors in a vehicular intersection environment. The proposed approach features rapid adaptation to potential initial calibration changes that may occur in LiDAR sensors installed at various positions in the intersection environment. To effectively address the physical changes in LiDAR sensors, the method suggests utilizing 3D plane points for automatic calibration.

Keywords: Automatic calibration, Iterative closest point, LiDAR, Plane

1. Introduction

Recent research has been focused on advancing fully autonomous driving technology, extending beyond vehicles to the integration of LiDAR in the surrounding environment[1-2]. Particularly, in the context of intersection environments, a method involving the attachment of LiDAR sensors to poles requires the simultaneous use of multiple sensors to ensure broad coverage. This necessitates the calibration of different LiDARs to account for their distinct characteristics.

However, this approach relies on using calibration values calculated at the initial installation position, making the LiDAR sensors vulnerable to physical changes induced by external factors, such as environmental fluctuations. In this paper, we propose a method to address these physical changes by utilizing a plane for automatic correction, allowing for adaptive calibration.

2. Automatic Calibration

The proposed methodology in this paper presupposes a scenario as illustrated in Fig. 1, wherein multiple LiDAR sensors are strategically positioned at various fixed locations within an intersection environment, accompanied by the application of inter-sensor calibration. Moreover, the automatic calibration process is executed following the steps outlined in sections 2.1 to 2.3.

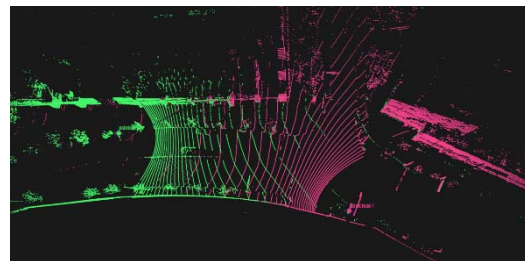


Fig. 1. 3D points of LiDAR Sensors Calibrated in Intersection Environment

This work was supported by Institute for Information & communications Technology Promotion(IITP) grant funded by the Korea government(MSIP) (No.2021-0-01314,Development of driving environment data stitching technology to provide data on shaded areas for autonomous vehicles)

2.1 Defining Plane in Intersection

In a vehicular intersection environment, a common feature is the plane of the vehicular road. In **Fig. 1**, after the initial calibration, a predefined planar region is established for the vehicle road or surrounding walls. The configuration of this planar region can be expressed by six coordinates representing the maximum and minimum values for x, y, and z coordinates.

2.2 Enhancing 3D Point Density

The density of 3D points in the common area varies depending on the position of the LiDAR. A low point density results in reduced accuracy in registration through section 2.3. Therefore, to enhance the density of 3D points in the common planar region defined in section 2.1, points within the plane are generated. The 3D points are randomly selected with x, y, z values satisfying the plane equation of $ax + by + cz + d = 0$ established for the common region in Section 2.1. **Fig. 2** illustrates the generation of numerous 3D points for each plane in the defined common area that has been altered by external factors.

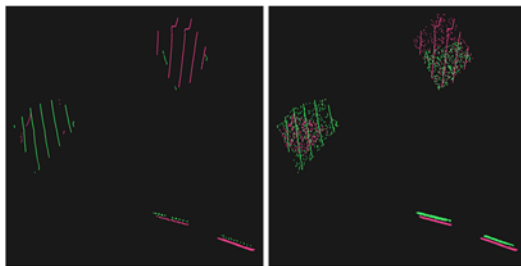


Fig. 2. Enhanced 3D Point Density in Common Planar Region (Left: 2.1, Right: 2.2)

2.3 Scan Matching for Point Registration

Using the 3D points from the overlapping common area generated in Section 2.2, we applied the scan matching algorithm to perform alignment between the two LiDARs. This scan matching algorithm employed the ICP (Iterative Closest Point) algorithm known for its fast computation and iterative process of finding optimal transformations by locating neighboring points. **Fig. 3** illustrates the results of applying the ICP algorithm from the Open3D library to different 3D points from two LiDARs affected by external environmental changes. The algorithm was configured with options of 10 iterations and a threshold of 1 for the computed transformation matrix.

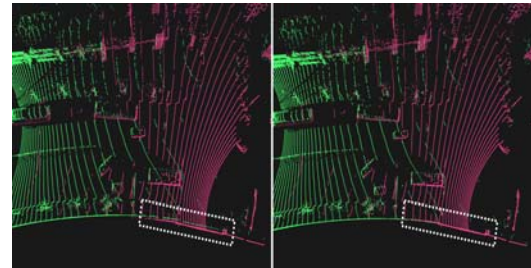


Fig. 3. Registration Results of Two LiDAR Sensors through sections 2.1~2.3 (Left: before, Right: after)

3. Analysis of Automatic Calibration

The proposed automatic calibration method enhances system performance with higher precision in the initial setup, facilitating effective calibration in environments utilizing multiple LiDARs. However, as the magnitude of physical changes increases, automatic alignment becomes challenging due to the non-overlapping common areas between the two LiDAR sensors. Additionally, variations in 3D point density based on LiDAR positions can impact accuracy, and convergence may pose challenges in specific scenarios.

4. Conclusions

This paper proposes a method for automatic calibration in response to initial calibration changes in multiple LiDAR sensors deployed in a vehicle intersection environment. The proposed method has the advantage of promptly adapting to slight rotational changes in LiDAR sensors. Future research will focus on improving the accuracy of the technique in response to rotational changes, performing automatic alignment through the detection of motion changes, and exploring various aspects such as providing an alarm service for detecting changes in initial calibration.

References

- [1] J. Jiao et al., "A novel dual-LiDAR calibration algorithm using planar surfaces," *IEEE Intelligent Vehicles Symposium*, pp. 1499-1504, 2019,
- [2] H. Song, "A study of HD-Map based Point Cloud Stitching for Multi LiDAR Sensor Calibration," *KSII FALL Conference*, 2022.

A Reconstruction of Korean Traditional Performance in Virtual Reality through 3D Pose Estimation across Multi-views

Taemin Hwang, Myoungjin Kim, Jaehyuk So, and Minjoon Kim

Department of Data Convergence Platform Research Center, KETI

Seongnam, South Korea

[e-mail: taemin.hwang, kmjketi, sojh0124, mjoon@keti.re.kr]

*Corresponding author: Minjoon Kim

Abstract

This research explores the intersection of cultural preservation and technology by reconstructing Korean traditional performances in virtual reality using 3D pose estimation across multiple camera views. With multiple edge devices, the proposed system can capture these performances in the open spaces. Through the proposed system, we introduce a comprehensive methodology for capturing, processing, and presenting these performances in VR, emphasizing the authenticity of traditional movements and postures. In addition, we illustrate the effectiveness of our approach and discuss its potential applications in the realms of education, tourism, and cultural exchange.

Keywords: 3D Pose Estimation, Edge Processing, Multi-View Reconstruction, Virtual Reality (VR)

1. Introduction

Human 3D pose estimation from video has found wide range applications such as sports analysis, surveillance, and virtual reality (VR). With rapid development of computer vision, particularly the advent of convolutional neural networks (CNN), various research have emerged to enhance this field [1]. Recently, to mitigate self-occlusion problem, various approaches using multiple camera views for 3D pose estimation have been introduced [2, 3].

Despite these remarkable advancements, there remains absence of a generalized methodology for creating 3D animation and presenting them through the VR domain. In this paper, we introduce a comprehensive methodology for capturing, processing, and presenting the

performance in VR through 3D pose estimation across multi-views.

2. Proposed System

2.1 Real-time 3D Pose Estimation

The proposed system is shown in Fig. 1. To mitigate self-occlusion and motion blur, we consider multiple view system. Each edge device transfer 2D pose to the central server. Then, the central server reconstructs 3D pose across multiple 2D poses [3].

2.2 Performance Reconstruction

In order to facilitate real-time operation, a pre-rigged 3D model was used instead of volumetric data. The central server merges the reconstructed 3D pose to the pre-rigged 3D model. Through

this process, the performance transforms into a coherent animation.

2.2 Cross Platform Accessibility

The proposed system provides the compatibility with a wide range of devices and platforms. To ensure that these cultural treasures can be experienced by as many individuals as possible, the reconstructed performance could be played in web or PC as well as HMD.

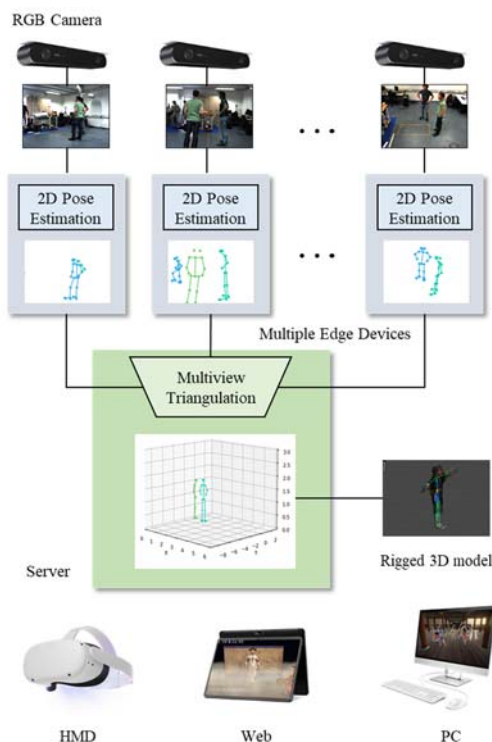


Fig. 1. Proposed System for the Performance Reconstruction

3. Experimental Results

To demonstrate the proposed system could reconstruct 3D performance, we capture Korean traditional performance with our prototype. As discussed earlier, the fusion process involves the integration of the reconstructed 3D pose data with a rigged 3D model. Consequently, the 3D model exhibits dynamic movements and follows the precise 3D pose. **Fig. 2.** Shows the reconstructed performance as viewed through HMD. We used Oculus 2 as HMD device.

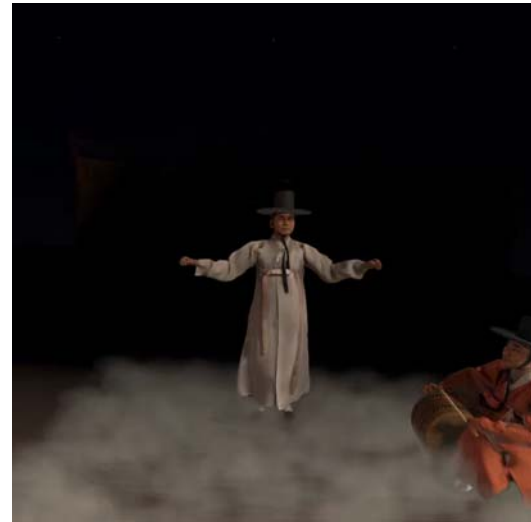


Fig. 2. Reconstructed performance as viewed through HMD

4. Conclusions

This paper introduces a methodology for the capture, processing, and presentation of Korean traditional performance in a virtual reality (VR). Through the implementation of prototype, we validate the efficacy of our approach, demonstrating its ability to reconstruct these performances in the VR environment.

References

- [1] L. Chen, H. Ai, R. Chen, Z. Zhuang, and S. Liu, "Cross-View Tracking for Multi-Human 3D Pose Estimation at Over 100 FPS," in *Proc. of IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, WA, USA, pp. 3276-3285, 2020.
- [2] E. Remelli, S. Han, S. Honari, P. Fua, and R. Wang, "Lightweight Multi-View 3D Pose Estimation Through Camera-Disentangled Representation," in *Proc. of IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, WA, USA, pp. 6039-6048, 2020.
- [3] T. Hwang, J. Kim, and M. Kim, "A Distributed Real-time 3D Pose Estimation Framework based on Asynchronous Multiviews," *KSII Transactions on Internet and Information Systems*, vol. 17, no. 2, pp. 559-575, Feb. 2023.

Spatial Information Scanning Based on a Mems Mirror for Mobile 3D Scan Device

Young Bo Shim* and Sung Hun Chae

Department of Data Convergence Platform Research Center, Korea Electronics Technology Institute (KETI)
Gyeonggi-do, South Korea
[e-mail: ybshim@keti.re.kr]

*Corresponding author: Young Bo Shim

Abstract

In this paper, a MEMS mirror including optics for mobile 3D scan device for spatial information acquisition is investigated. The efficient pulsed laser driver project uniform beam profile over the field of view area. A MEMS mirror and directional optics perform x-y raster scan. The performance of the proposed technique is demonstrated exploiting a time-to-digital converter to reconstruct 3D spatial information. As a result of demonstration, the proposed technique is capable of detect ~5 m distance within 30x30 degrees of field of view in mobile form.

Keywords: Spatial information, Mobile 3D scan

1. Introduction

Recently, there is a high demand for technology to acquire 3D spatial information using mobile devices [1]. Mobile devices such as ToF sensors, which have recently been commercialized, have begun to be installed in mobile phones due to the advantage of simple signal processing, but there is a limitation in that the measurable distance is limited.

Since typical ToF sensors utilize continuous wave (CW) laser illumination with low power due to the heat problem, the ToF sensors has relatively short range measurement capability to a few meters.

To overcome this limitation, there are attempts to miniaturize flash LIDAR. However the cost and complexity of the SPAD module and DSP system, which are the core of flash LIDAR, are high. On the other side, for a scanning LIDAR with beam-steering laser on the emitter and a wide-angle lens on the receiver, there is a problem that the highly sensitive light detection

element is easily saturated by environmental ambient light [2].

In this paper, a MEMS mirror including optics for mobile 3D scan device for spatial information acquisition is investigated. The efficient pulsed laser driver project uniform beam profile over the field of view area. A MEMS mirror and directional optics perform x-y raster scan. The performance of the proposed technique is demonstrated exploiting a time-to-digital converter to reconstruct 3D spatial information.

2. Z-axis depth measurement

Fig. 1 shows the Z-axis depth measurement of the proposed system. A micro lens array (MLA) is used as a beam homogenization element. The pulse timing of the illumination is measured with a yellow line.

An incident laser pulse is divided into many beamlets due to the pupils of small lenslets of the MLA. Since the beam centers of each beamlet has parallel propagation direction, each beamlet is projected at the same superimposed objective

surface by a final projection lens.

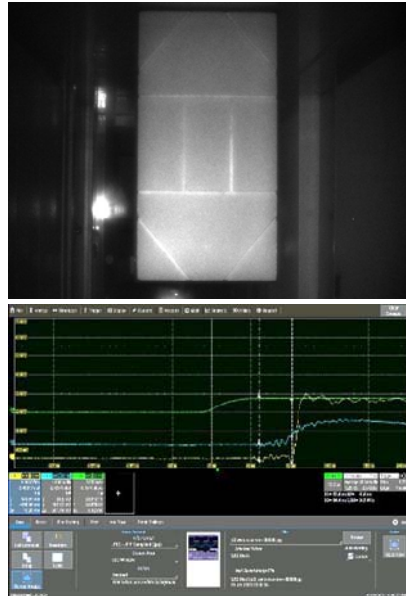


Fig. 1. Object depth measurement data

The return pulse received by SiPM through the clear aperture of a mems mirror is measured with a yellow line.

3. XY-axis raster scanning

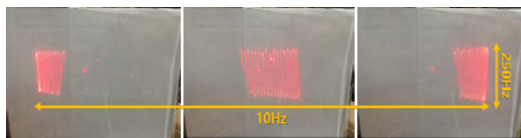


Fig. 2. Raster scanning trace

Fig. 2 shows the raster scan analysis of the x and y axes based on mems mirror up to 250 Hz. The number of scan line is measured to be ~100 which means high scan density.

4. Experimental performance analysis

Fig. 3 shows the experimental performance demonstration in a test lab environment. The flattened rectangular field of view is formed with 10 Hz refresh rate. Since the FoV has uniform distribution, the designed structure does not require any extra projection lens.



Fig. 3. Demonstration of the scan system

5. Conclusions

In this paper, a MEMS mirror including optics for mobile 3D scan device for spatial information acquisition is investigated. The efficient pulsed laser driver project uniform beam profile over the field of view area. A MEMS mirror and directional optics perform x-y raster scan. The performance of the proposed technique is demonstrated exploiting a time-to-digital converter to reconstruct 3D spatial information.

References

- [1] M.-González, M. Ángel, V. D. Pietra, and M. Piras, "Evaluation of Different LiDAR Technologies for the Documentation of Forgotten Cultural Heritage under Forest Environments," *Sensors*, Vol 22, No. 16, pp. 6314, 2022.
- [2] A. Incoronato, I. Cusini, K. Pasquinelli and F. Zappa, "Single-Shot Pulsed-LiDAR SPAD Sensor with on-chip Peak Detection for Background Rejection," *IEEE Journal of Selected Topics in Quantum Electronics*, Vol 28, No. 5, pp. 1-10, 2022.

Range Measurement Technique Based on VCSEL Source and SiPM Receiver

Young Bo Shim* and Sunghun Chae

Department of Data Convergence Platform Research Center, Korea Electronics Technology Institute (KETI)
Gyeonggi-do, South Korea
[e-mail: ybshim@keti.re.kr]

*Corresponding author: Young Bo Shim

Abstract

In this paper, a range measurement technique for mobile devices is investigated. A homogenized VCSEL beam illumination module and highly sensitive SiPM module is exploited as a light source and receiver, respectively. The fabricated optics for VCSEL source includes focusing lenses and beam homogenizer for flattened beam projection over large area. The reflection time and intensity are measured by SiPM. The demonstration shows various work error characteristics according to color and view angle.

Keywords: ToF sensor, VCSEL, Homogenizer

1. Introduction

Recently, with the development of cultural content technology, there are increasing attempts to acquire 3D information in real space and apply it to virtual spaces such as the metaverse [1].

To acquire 3D information in real world, high-performance equipment such as LIDAR scanners are mainly used. LIDAR is divided into two types: scanning LIDAR for ADAS systems and flash LIDAR for 3D spatial information. Since conventional flash LIDARs emit a single laser pulse beam across the LIDAR's entire field of view (FoV) at once, a strong laser power was required to compensate for beam divergence. Accordingly, the capacity of the laser amplifier should have been increased. This limitation is disadvantageous in terms of cost and miniaturization [2].

In this paper, an emission control technology capable of direct pulse control of 5 ns or less was implemented using an integrated laser array

element and a small multi-channel switch element and its performance is verified.

2. Direct-pulsed Range measurement

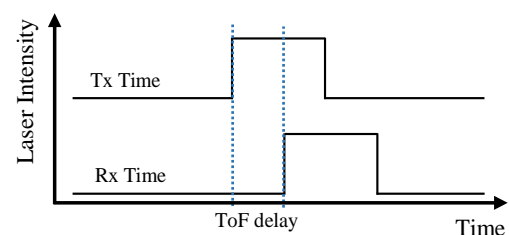


Fig. 1. Basic scheme of the direct-pulsed range measurement

Fig. 1 shows the basic scheme of the range measurement. The basic principle of the direct pulsed system is the arrival time measurement of the back scattered laser signal. A laser source emits a pulsed light and a detector receives the back scattered laser from an object.

3. VCSEL array source

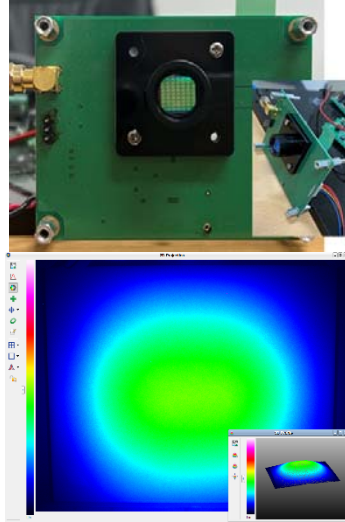


Fig. 2. Fabricated VCSEL laser source

Fig. 2 shows the fabricated module of proposed array source. The driver circuit is based on a multi-channel switch element unlike conventional capacitor-discharge drivers. The module has effective power and heat-generation characteristics since low driving voltage is directly modulated at multiple channel within 5 ns. Since the VCSEL can be mounted parallel to the surface, it occupies smaller space and emits laser perpendicular, as opposed to the conventional laser diodes. The fabricated optics for VCSEL illumination includes focusing lenses for divergence control. A multi-lens array is used as a beam homogenization element. The beam distribution is measured using a commercial CCD profiler.

Since the array source's light-emitting operation principle is based on parallel low voltages and the beam profile is even as well, there are no restrictions on multi-source configuration. The multi-source configuration is expected to enhance power and FoV.

4. Range measurement result

Fig. 3 shows the range measurement result. A commercial SiPM receiver is exploited as a back scattering detector of VCSEL source. The SiPM device is capable of detect single-photon. An IR bandpass filter is exploited to reject the ambient light and increase the signal-to-noise-ratio.

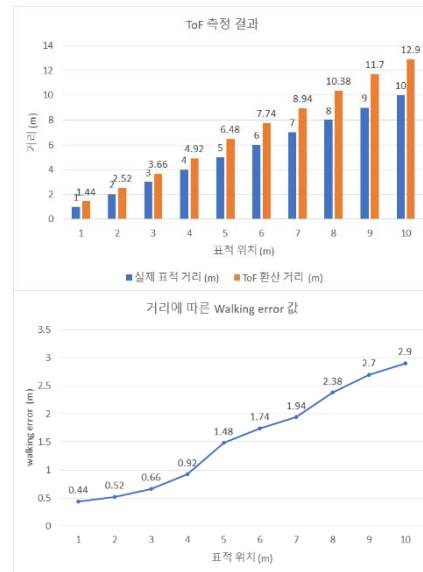


Fig. 3. Range measurement result based on VCSEL source and SiPM receiver

The result shows linear response characteristics. The walk-error can easily be compensated in the signal processor.

5. Conclusions

In this paper, an emission control technology capable of direct pulse control of 5 ns or less was implemented using and integrated laser array element and a small multi-channel switch element and its performance is verified.

Since the proposed technology is suitable for a multi-source structure, various modifications are possible, such as improving output and expanding the light emitting area, depending on the needs of the application field.

References

- [1] M.-González, M. Ángel, V. D. Pietra, and M. Piras, "Evaluation of Different LiDAR Technologies for the Documentation of Forgotten Cultural Heritage under Forest Environments," *Sensors*, Vol 22, No. 16, pp. 6314, 2022.
- [2] A. Tontini, L. Gasparini, and M. Perenzoni, "Numerical Model of SPAD-Based Direct Time-of-Flight Flash LIDAR CMOS Image Sensors," *Sensors*, Vol 20, No. 18, pp. 5203, 2020

Sunghun Chae* and Young Bo Shim

Department of Data Convergence Platform Research Center, Korea Electronics Technology Institute (KETI)
Gyeonggi-do, South Korea
[e-mail: shcahe@keti.re.kr, ybshim@keti.re.kr]
*Corresponding author: Sunghun Chae

This paper presents the real-time large-channel lighting control interface module is proposed. Based on the standard protocols Art-Net and DMX standards, we have implemented a miniaturized interface module that integrates existing consoles and DMX modules. And also, extension conventional standard protocol to scalable controlling non-standard devices. Interface module that can link and control more than hundreds of devices and is linked with generative art content to suit the scenario. The real-time control operation performance of a generative art content was verified through field demonstration.

1. Introduction

this paper, we propose an interface system based on a miniaturized interface module that replace a control consoles and DMX module. The proposed system verified the operation through real-time control through performance of generative art content by direct linkage with hundreds of lighting devices.

2.1 Architecture of Proposed System

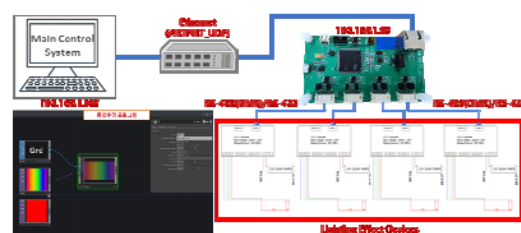


Fig. 1. Multi-channel lighting control interface system

The proposed system consists of an integrated operation program, interface module, and lighting device, as shown in Fig. 1. The Artnet protocol was applied to link between the integrated operation program and the interface module, and the UART interface was used for devices linkage based on DMX protocol.

2.2 Implement of Interface Module

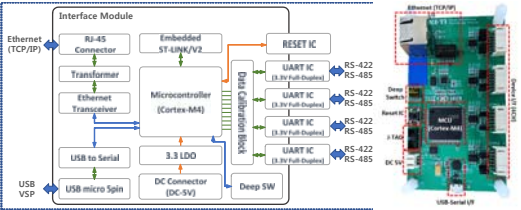


Fig. 2. Interface module block diagram and prototype

Table 1. Artnet to DMX data packet structure

| System <=> Interface Module (variable size of packets) <=> Effect Devices | | | | | | | |
|---|----------|---|---------|----------|----------|----------|--------|
| Name | ED | OpCode | PresVtr | Sequence | Physical | Universe | Length |
| Detail | 1 | 2 | 2 | 1 | 1 | 1 | 2-512 |
| Description Header | ED | "Art-Net" | | | | | |
| | OpCode | "Art-Net" (0x0000) (Packet Type) | | | | | |
| | PresVtr | Init (Version of the Art-Net protocol) | | | | | |
| | Sequence | Init-0x00 (Physical part that generated the packet) | | | | | |
| Art-Net Packet | Physical | Init-0x00 (Physical part that generated the packet) | | | | | |
| | Universe | Pres-Address | | | | | |
| | Length | Length(0x00) (Byte) : (Most Significant byte of length of data array) Length(0x01) (Byte) : (Least Significant byte of length of data array) | | | | | |
| Init | Detail | Init-0x00 | | | | | |

The proposed interface module consists of Cortex-M4-based MCU, Ethernet, USB, UART (RS-422/485) 4CH, deep switch, etc as shown in Fig. 2. It supports RS-485 communication for standard, and the RS-422 for non-standard devices [2]. In addition, a control protocol was developed according to the system to be applied through the MCU based control unit and designed to be selectively expandable. The designed data packet includes Art-Net header and DMX data as shown in Table 1. The data packet is only an example and can be adaptively changed depending on the implemented system.

2.3 Experimental Result

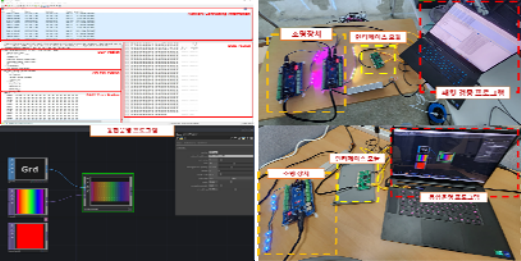


Fig. 3 Test environment to verify the interface module



Fig. 4. Performance verification with service contents

To verify the performance of the interface module, a test environment was established as shown in Fig. 3. After receiving the data packet from the program, each control packet was sent to the lighting devices through the output port of the interface module to confirm. Additionally, the interface module operation performance was verified through field demonstration through linking generative art content as shown in Fig. 4.

3. Conclusions

In this paper, interface system proposed to support multi-channel sensor and effect device control. The proposed system includes an interface module and the module connected devices that can independently control up to hundreds lighting devices by control commands from the integrated operation program. The Art-Net to DMX standard based protocol was implemented to enable control of various lighting devices, and real-time control was implemented in consideration of generative art content that require direct linkage with hundreds of lighting devices through linkage with actual devices and integrated operation programs. The operational performance of interface module and system was verified through field demonstration.

References

[1] K. Du and L. Lv, "The design of DMX512 and STM32 based intelligent interactive lighting control system," in *Proc. of 2014 11th China International Forum on Solid State Lighting (SSLCHINA)*, Vol 1, pp. 76-78, 2014

[2] S. Chae, Y. B. Shim, "Implementation of Interface System through DMX512 Standard Expansion," in *Proc. of The 8th International Conference on Consumer Electronics (ICCE) Asia*, pp. 339-340, 2023.

Emotion and Theme Recognition by combining multiple classifiers

Dalwon Jang, Hanyi Shin, Jaewon Lee, and JongSeol Lee

Communications and media R&D Division

Korea Electronics Technology Institute

[e-mail: dalwon@keti.re.kr, hannie12@keti.re.kr, jwlee0121@keti.re.kr, leejs@keti.re.kr]

*Corresponding author: Dalwon Jang

Abstract

This paper presents experimental results of recognizing emotion and theme in music. Tagging music with emotion/theme labels is realized with classification strategy. However, combination of multi-classifier can give the better results than using only one classifier. Here, we present results of various classifiers and result of average ensemble of multiple classifiers.

Keywords: Music theme, music emotion, music classification

1. Introduction

With the recent proliferation of digital music, there is increasing demand for management of music information [2]. For automatic extraction of useful information, music information retrieval research area has been developed in various ways. Some researches focus finding musical structure information such as melody or beat information [5]. However for music listener, tag or label information, that can explain entire song, is much useful. Thus, music classification problem, which extract tag of label with input of music recording, have been studied. Early researches focused on musical genre [2,3]. Nowadays music emotion and theme are also big concern [1,4]. Early researches tried developing new features and combining various features [2,3]. Recent music classification researches focus development of classifier. With only one classification, satisfactory results cannot be obtained. Thus, by combining the results of multiple classifiers final classification result is

determined[1,2].

This paper presents experimental results of classifiers with MTG-Jamendo dataset [4] by referring previous research [1]. Trying to simulate the study in [1], test results of various cases are obtained.

2. Classifiers used in experiments

As in [1], ResNet, Frequency-Aware ResNet (FA-ResNet)[6], VGG, and convolutional recurrent neural network (CRNN)[7] are used in our experiments. With the classifiers ensemble of classifiers are also tested.

As an input of classifiers, Mel-spectrogram, which is provided by the dataset [4] is used as it is. From 12000 Hz sampling rate, 512 sample is converted to 96 Mel bins with 256 hop size. We used 96 by 512 Mel-spectrogram, whose length is about 10.9 second.

3. Experimental results

We used MTG-Jamendo dataset, and from the dataset mood/theme tag information was used [4]. There are 59 tags, thus classification of 59 label is performed. Here, we compute Receiver operating characteristic area under curve (ROC-AUC) as a performance metric. Table 1. shows the results of various classifiers. As shown in the table, combining multiple classifiers gives the better results. Combining ResNet and CRNN gives slightly better result then CRNN does. Combining FA-ResNet and CRNN gives slightly better result then FA-ResNet, which leads to the best result of single classifier, does. What we did in the experiments is similar to that in [1], but use of Mel-spectrogram provided by dataset is a difference.

Table 1. Experimental results

| Classifiers | ROC-AUC |
|--------------------------|---------|
| ResNet | 0.6293 |
| FA-ResNet | 0.7074 |
| VGG | 0.6732 |
| CRNN | 0.7013 |
| Ensemble (ResNet+CRNN) | 0.7077 |
| Ensemble(FA-ResNet+CRNN) | 0.7121. |

4. Conclusions

In this paper, we present the experimental results of music emotion/theme recognition. Using MTG-Jamendo dataset, various classifiers with Mel-spectrogram are tested. And, ensemble averaging is also tested to enhance the performance. In our experiments, it is verified that combining classifiers leads to performance enhancement.

References

- [1] K. Koutini, S. Chowdhury, V. Haunschmid, H. Eghbal-Zadeh, and G. Widmer, "Emotion and Theme Recognition in Music with Frequency-Aware RF-Regularized CNNs," in MediaEval 2019 Workshop, Sophia Antipolis, France, 27-30 October 2019, vol.2670.
- [2] D.Jang, M.Jin, and C. D. Yoo, "Music genre classification using novel features and a weighted voting method," In Proceeding of the 2008 IEEE International Conference on Multimedia and Expo, pp.1377–1380, April 2008.
- [3] G. Tzanetakis and P. Cook, "Musical genre classification of audio signals," IEEE Trans. Speech Audio Process., vol.10, no.5, pp. 293-302, Sep. 2002.
- [4] D. Bogdanov, M. Won, P. Tovstogan, A. Porter, and X. Serra. 2019. The MTG-Jamendo dataset for automatic music tagging.
- [5] J. Lee, D. Jang, and K. Yoon, "Automatic melody extraction algorithm using a convolutional neural network," KSII Trans. Internet Inf. Syst., vol. 11, no. 12, pp. 6038–6053, 2017.
- [6] K. Koutini, H. Eghbal-zadeh, and G. Widmer, "Receptive-field-regularized CNN variants for acoustic scene classification," In Proceedings of the Detection and Classification of Acoustic Scenes and Events 2019 Workshop (DCASE2019).
- [7] K. Choi, G. Fazekas, M. Sandler, and K. Cho, "Convolutional recurrent neural networks for music classification," In Proc. Of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2392–2396, 2017.

Comparison of object recognition performance using Embedded AI SoC

Seho Park and SaetByul Yu

Contents Convergence Research Center, Korea Electronics Technology Institute
Seoul, South Korea

[e-mail: sehopark@keti.re.kr, sbyu@keti.re.kr]

*Corresponding author: Seho Park

Abstract

Most automobile manufacturers employ AI technology in the development of autonomous driving systems. However, these technologies are primarily applied to newly released vehicles. As a result, it has become challenging to integrate convenient advanced driver-assistance systems into existing vehicles. In this paper, we address this issue by developing our own low-power and compact AI SoC (System on a Chip) chipset. We also create an object recognition model suitable for rear-view cameras in automobiles. The object recognition model is implemented using three different lightweight deep learning networks. To evaluate the performance, the implemented AI object recognition model is tested on our self-developed AI SoC board, nVidia AGX board and Android tablet (Galaxy Tab S8). By comparing the results, we aim to assess the commercial viability of our self-developed AI SoC board.

Keywords: object recognition, AI SoC, Performance analysis, EfficientNet, MobileNet, Yolo, rear camera application

1. Introduction

Object detection in still or moving images is a fundamental and widely utilized technique in the fields of image processing and computer vision. This technology plays a crucial role in various domains, with applications extending to the integration of object detection into commercial products.

However, artificial intelligence-based object recognition demands significant computational power, typically implemented on high-performance devices equipped with GPUs. Consequently, this leads to challenges in terms of price competitiveness and maintenance during commercial implementations. To address such issues, various forms of AI System on Chip (SoC) chipsets are being developed. In this paper, we compare the performance of our self-developed AI SoC Board with nVidia AGX Board and Android Tablet when executing object recognition models. Additionally, this paper compares three lightweight models among object

classification and object recognition models. The deep learning networks used for performance comparison are Efficient Detect Lite0, MobileNet V2, and YoloV3 [1] [2].

2. Object recognition AI model

2.1 Dataset for Object Recognition

The scope for comparing object recognition models is aimed at integrating them into the rear cameras of vehicles. Therefore, the dataset used for training comprises images captured by rear cameras in vehicles. The datasets utilized for training include nu Image dataset, AI Hub dataset, and a dataset constructed in-house. Given the nature of rear cameras, there is no necessity to recognize very small objects at a long distance [3]. Consequently, objects smaller than 0.1% of the entire input image size were excluded from training datasets. A total of 83,551 images from rear cameras were utilized

for training, and the test set for performance evaluation consisted of 21,423 images

2.2 Deep Learning Network Models for Object Recognition

For the comparative validation tests on the terminal, lightweight deep learning networks, namely Efficient Detect Lite0, MobileNet V2, and Yolo V3, were selected. The EfficientDet AI model was generated using the modelmaker API provided by TensorFlow, and for lightweight deployment, it was converted to tflite for performance testing. The SSD MobileNet V2 model also utilized the Object Detection API provided by TensorFlow, with an input size set to 300 x 300. YoloV3 was trained using darknet, with batch-64 and an input size of 416 x 416.

3. Performance analysis results

3.1 Object detection performance

The model was trained using a self-constructed dataset for rear cameras to distinguish objects such as bus, car, cycle, pedestrian, traffic cone, and truck. The Average Precision (AP) values for each AI model are as follows. Due to the nearly identical performance of the three AI models in object recognition, the fastest object recognition module will be implemented in the product.

Table 1. Mean Average Precision of Object Detection Network

| | Efficient Detect | MobileNet V2 | Yolo V3 |
|---------------|---------------------|-----------------|------------|
| mAP(%) | 47.56 | 46.12 | 47.57 |

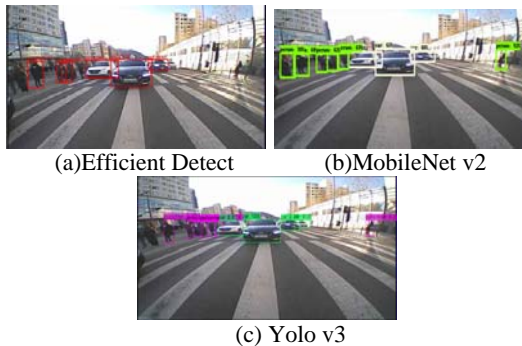


Fig. 1. Object Recognition Results from Rear Cameras for Each Model

3.2 Comparison of AI model processing speed

The table below presents the processing speed, measured in FPS, when deploying AI models applicable to rear cameras on various devices. The three object recognition models used for evaluation were converted into Tflite format. Based on the results of testing each AI model, it was concluded that YoloV3 exhibited slow performance on all devices, making it challenging for practical applications. In contrast, EfficientDet and MobileNet demonstrated sufficiently fast performance suitable for use on embedded boards.

Table 2. Comparison of Processing Speed (Frames Per Second) on Each Device

| | Efficient Detect | MobileNet v2 | Yolo v3 |
|-----------------------|---------------------|-----------------|-------------|
| nVidia AGX | 6.06 | 10.08 | 0.22 |
| Android Tablet | 14.83 | 26.97 | 4.59 |
| AI SoC Board | 39.98 | 53.08 | 0.60 |

4. Conclusions

In this paper, we developed an object recognition model applicable to rear cameras and deployed the developed model on general-purpose terminals and our self-implemented AI SoC. We compared the performance results across different devices. The object detection performance on our self-implemented AI SoC, tested with a lightweight object recognition module, demonstrated a relatively superior performance compared to other platforms. In the future, we plan to confirm its superiority not only in terms of performance but also in power consumption compared to commercial terminals. additional research is needed on Yolo models suitable for deployment on embedded boards.

References

- [1] F. Sajid, A. R. Javed, A. Basharat, N. Kryvinska, A. Afzal and M. Rizwan, "An Efficient Deep Learning Framework for Distracted Driver Detection," in IEEE Access, vol. 9, 2021.
- [2] Q. -C. Mao, H. -M. Sun, Y. -B. Liu and R. -S. Jia, "Mini-YOLOv3: Real-Time Object

- Detector for Embedded Applications," in *IEEE Access*, vol. 7, 2019.
- [3] K. Chang, B. -H. Oh and K. -S. Hong, "An implementation of smartphone-based driver assistance system using front and rear camera," *IEEE ICCE*, pp. 280-281, 2024.

Metric Learning for Gait Recognition with WiFi Sensing

Seho Park, Myeongseop Kim, and Taehyeon Kim[†]

Contents Convergence Research Center,
Korea Electronics Technology Institute,
Seoul, South Korea

[e-mail: sehopark@keti.re.kr, myeongseopkim@keti.re.kr, taehyeon.kim@keti.re.kr]

^{*}Corresponding author: Dr. Taehyeon Kim

Abstract

This paper presents an innovative method for gait recognition using Channel State Information (CSI) derived from WiFi signals, incorporating a neural network-based approach for signal analysis. The proposed method advances the field by using a metric learning strategy with triplet loss, which significantly refines the neural network's ability to recognize and differentiate between individual gait patterns. Our approach addresses the privacy limitations of conventional surveillance systems by providing a non-intrusive means of identification. Extensive experiments were conducted to compare the proposed method against benchmark techniques, revealing superior accuracy in both laboratory and cubic environments, with different numbers of participants. The results of our study demonstrate that the proposed CSI-based gait recognition method can serve as a powerful tool for enhancing indoor security measures and personal identification with high accuracy, without compromising individual privacy.

Keywords: WiFi sensing, channel state information, gait recognition, deep learning, metric learning

1. Introduction

With the full establishment of the ubiquitous wireless communication era, WiFi environments have been set up in many spaces indoors. During the development of WiFi wireless communication, the technique of measuring signal quality using channel state information (CSI) has evolved to improve communication quality [1, 2]. CSI serves as an indicator signal that reflects the state of signal quality in a WiFi wireless communication environment, influenced by space and obstacles through diffraction and reflection. This information can be utilized as an input for technologies aimed at

enhancing the quality of WiFi wireless communications.

Recently, the use of Channel State Information (CSI) in WiFi wireless communications has expanded beyond just enhancing signal quality. By integrating it with neural network-based data interpretation techniques, CSI is being repurposed for various applications such as detecting indoor human movements and identifying individuals [3]. This is because, as previously mentioned, CSI values change in response to the indoor spatial layout, as well as the movement or position of objects. Therefore, more precisely, in an indoor environment with WiFi wireless communication, human

movement will alter the signal characteristics of CSI. These changes can be observed in the coefficients of CSI's subcarriers. By inputting these variations into a neural network, desired inference results can be extracted.

In this paper, we introduce a technique that enhances the recognition performance of neural network through a metric learning strategy in the process of performing gait recognition using a neural network-based WiFi CSI analysis method. The proposed technique can directly address the critical privacy issues inherent in traditional CCTV or IoT home camera-based indoor security systems. Its high versatility stems from its ability to operate in any indoor environment where WiFi-based wireless communication is established.

2. Proposed Method

CSI data is composed of the number of antennas \times number of subcarriers \times number of packets. In this paper, to utilize Convolutional Neural Networks (CNN), CSI images of the size 'Number of Subcarriers \times Number of Packets' are structured with 'Number of Antennas' as channels for input into the CNN. This approach allows for intuitive use of CNN without the need for additional preprocessing and has the added advantage of enabling the use of fundamental image recognition models as initial models. The proposed WiFi CSI-based Gait Recognition Method utilizes triplet loss, which is one of the representative losses used in metric learning [4]. This loss is defined by the following formula:

$$L = \sum_{i=1}^N \max(0, d(a_i, p_i) - d(a_i, n_i) + m),$$

where a_i represents the anchor sample, p_i is the positive sample, n_i denotes the negative sample, and $d(x, y)$ means the Euclidean distance between two samples.

The fundamental idea of Triplet Loss is to use three related data points (known as a 'triplet') during the training process. This triplet typically consists of an "anchor," "positive," and "negative" sample. The anchor is the reference data point. The positive is a data point that belongs to the same category (or class) as the anchor. The negative is a data point from a different category than the anchor. Based on this,

the principle of triplet loss works to minimize the distance between the anchor and the positive, and maximize the distance between the anchor and the negative. Therefore, the Triplet Loss function calculates the difference in distances between the anchor and the positive and the anchor and the negative, guiding the learning model to ensure that the distance between the anchor and the positive sample is smaller than the distance to the negative sample by a "margin."

This approach enables effective separation of data from various classes in WiFi CSI Sensing-based gait recognition, and allows deep learning models to place similar items close together and different items further apart, enabling effective feature learning.

3. Experimental Results

Table 1. CSI Gait Recognition Results

| User Size | [5] | [6] | Ours | Loc |
|-----------|--------|--------|--------|-------|
| 2 | 86.64% | 93.41% | 97.56% | Lab |
| | 89.15% | 93.41% | 97.58% | Cubic |
| 5 | 80.35% | 89.12% | 96.73% | Lab |
| | 84.68% | 89.66% | 94.69% | Cubic |

This study presents the comparative analysis of gait recognition methods using WiFi CSI-based techniques. The evaluation was conducted across two different user group sizes and in two separate environments, namely a laboratory (Lab) and a cubic space (Cubic). The methods were compared based on the User Size, which refers to the number of participants in the experiment, with groups of 2 and 5. The performance of each method is represented as a percentage accuracy. The results clearly indicate that our proposed method consistently outperforms the existing methods [5] and [6] in terms of accuracy across different user sizes and environmental conditions. This improvement is significant and suggests that the proposed method can be effectively applied for gait recognition in various indoor settings, showcasing its robustness and adaptability.

4. Conclusion

Our study introduced a novel WiFi CSI-based Gait Recognition Method that leverages neural network analysis techniques to enhance gait recognition accuracy in indoor environments. By

employing a triplet loss metric learning strategy, our method achieved significant improvements in distinguishing between different gait patterns. Experimental results demonstrated that our proposed method consistently outperformed existing approaches across various user group sizes and environmental settings. Specifically, in a laboratory environment, our method attained accuracies up to 97.56% for two users and 96.73% for five users. Similarly, in cubic environments, our method reached accuracies of 97.58% and 94.69% for two and five users, respectively. These findings suggest that our approach not only offers a viable solution to the privacy concerns associated with traditional visual surveillance systems but also provides a highly accurate and robust framework for gait recognition using wireless signals. The versatility and effectiveness of our method hold promise for broader applications in security and user identification domains.

References

- [1] Lu, Lu, *et al.*, “An overview of massive MIMO: Benefits and challenges,” *IEEE journal of selected topics in signal processing*, 8.5, 742-758, 2014.
- [2] Armstrong, Jean, “OFDM for optical communications,” *Journal of lightwave technology*, 27.3, 189-204, 2009.
- [3] Yang, Jianfei, *et al.*, “SenseFi: A library and benchmark on deep-learning-empowered WiFi human sensing,” *Patterns*, 4.3, 2023.
- [4] Kaya, Mahmut *et al.*, “Deep metric learning: A survey,” *Symmetry*, 11.9, 1066, 2019.
- [5] J. Zhang *et al.*, “Gate-ID: WiFi-based human identification irrespective of walking directions in smart home,” *IEEE Internet Things J.*, vol. 8, no. 9, pp. 7610–7624, May 2021.
- [6] D. Wang *et al.*, “CSIID: WiFi-based human identification via deep learning,” in *Proc. 14th Int. Conf. Comput. Sci. Educ. (ICCSE)*, pp. 326–330, 2019.

Method of Operation and Control of Detection Equipment for Patrol Drone-Cop System

Sungwook Jung, Jong-Hong Park, Sungchan Choi, Wonseok Jung, and Il-yeup Ahn

Autonomous IoT Research Center,
Korea Electronics Technology Institute (KETI)
Seongnam, South Korea

[e-mail: {sungwook87, jonghong, csc, wonseok, iyahn}@keti.re.kr]

*Corresponding author: Il-yeup Ahn

Abstract

This paper presents a ground control system for operating Patrol Drone and a method for operating detection equipment to be mounted on Patrol Drone among drone cop systems to be operated to respond to illegal drones. This system is made through interworking with Mobius, an IoT platform based on oneM2M standards, and its performance was evaluated using a virtual data generator before it was actually installed. We briefly describe the structure of the system and present the possibility of utilizing the patrol drone-cop system through GUI design

Keywords: Ground Control System, Simulator, Illegal drone

1. Introduction

As various drones are closely related and expanded in real life in various forms, secondary problems that did not exist before arise [1]. Among them, research on methods to respond to illegal drones, which has recently become a major issue, has been actively conducted. This research team is upgrading operational S/W that can respond autonomously and quickly by developing a Drone Cop Control System (GCS) for anti-drone systems and a control system for various mission equipment. Among them, this paper presents a method for operating detection equipment mounted on the patrol drone cop system and a control method through GCS, and a method for securing stability and reliability through a simulator environment.

2. Software Structure

2.1 Interworking structure

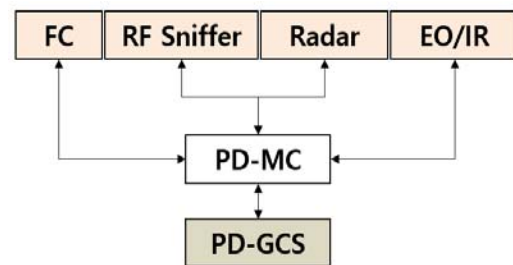


Fig. 1. Patrol drone connection structure

The patrol drone cop (PD) is equipped with flight controller, RF Sniffer, Radar, and EO/IR



Fig. 2. GCS for operating Patrol Drone

equipment, and each equipment is linked to the GCS through the drone's mission computer. Each device transmits current status information and Ack information to the GCS, which sends commands down to the devices. Since no actual equipment was installed, the test was conducted through a virtual data generator in this paper.

2.2 Detection equipment User Interface



Fig. 3. Radar UI

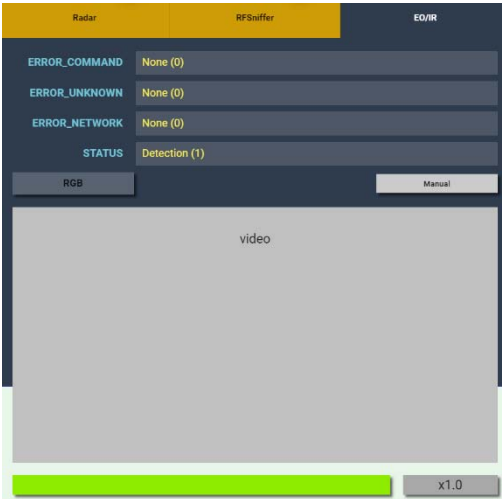


Fig. 5. EO/IR UI



Fig. 4. RF Sniffer UI

3. Conclusions

Among the drone-cop systems to be operated to respond to illegal drones, this paper presents a ground control system for operating the Patrol Drone and a method for operating detection equipment to be mounted on the Patrol Drone.

References

[1] S. Park et al., “Survey on Anti-Drone Systems: Components, Designs, and Challenges,” IEEE Access, vol. 9, pp. 42635-42659, 2021.

Task Offloading Strategy Based on Different Network Planning Approaches in Edge Computing Networks

**Md. Delowar Hossain, Tangina Sultana, Sharmen Akhter, Md. Imtiaz Hossain,
Md Nosin Ibna Mahbub, Ngo Thien Thu, Ga-Won Lee and Eui-Nam Huh**

Department of Computer Science and Engineering, Kyung Hee University
Yongin-si, South Korea

[E-mail : {delowar, tangina, sharmen, hossain.imtiaz, nimahbub, thu.ngo, gawon, johnhuh}
@khu.ac.kr]

*Corresponding author : Eui-Nam Huh

Abstract

Edge computing has emerged as a promising paradigm to meet the low-latency demand for highly reliable services in various applications. By decentralizing computing resources closer to the data source, edge computing minimizes latency, optimizes bandwidth usage, and enhances overall system reliability. However, one of the crucial aspects of edge computing is the efficient offloading of tasks from user devices to edge servers, which demands careful consideration of network planning and resource allocation. In addressing these challenges, this study explores the potential of different network planning approaches, particularly the M/M/1 and M/M/2 queuing models, within the context of edge computing networks. By conducting a thorough comparative analysis of these queuing models concerning task offloading strategies, this research aims to provide insights that can lead to the development of more efficient, adaptable, and responsive edge computing systems. In this study, we evaluated the performance of various task offloading strategies through the EdgeCloudSim simulator and incorporated three distinct applications to model a range of real-life scenarios. The results of the simulation demonstrated that our proposed edge (M/M/1) with cloud task offloading scheme performed better in comparison to baseline approaches across various scenarios. Our proposed scheme can significantly reduce service time by 88.4% compared to local computing (mobile). Furthermore, when contrasted with an edge offloading (M/M/1) scheme, it still achieves a substantial reduction of approximately 14%.

Keywords: Task offloading, M/M/1 and M/M/2 queuing models, mobile edge computing.

1. Introduction

In the modern world, mobile devices have gained immense popularity. The rapid proliferation of smart mobile devices and the emergence of

demanding services and applications have resulted in a substantial increase in data traffic. These applications are typically characterized by their high computational requirements and, as a consequence, result in substantial energy

This research was partly supported by the MSIT (Ministry of Science and ICT), Korea, under the ITRC (Information Technology Research Center) support program (IITP-2023-RS-2023-00258649) supervised by the IITP (Institute for Information & Communications Technology Planning & Evaluation) and the MSIT (Ministry of Science and ICT) and Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (No.RS-2023-00220631, Edge Cloud Reference Architecture Standardization for Low Latency and Lightweight Cloud Service)

consumption, which is a significant drawback associated with smart mobile devices [1]. In the past, the concept of mobile cloud computing (MCC) was introduced to address the aforementioned issues. MCC offers several advantages, such as extending battery life by offloading computationally intensive applications to the cloud and enhancing data storage capabilities for users [2]. However, despite the numerous benefits of MCC, it is not particularly well-suited for real-time applications and faces unavoidable challenges, including high latency and limitations in backhaul bandwidth. To mitigate the challenges mentioned above, a promising new concept called mobile edge computing (MEC) has emerged [3].

Mobile edge computing is a distributed computing paradigm that brings cloud computing capabilities closer to the edge of a mobile network, where mobile users and their devices are located. MEC deploys computing resources, such as servers or edge nodes, in close physical proximity to base stations, access points, or cell towers within the mobile network. Because of this close proximity, data is required to travel a shorter physical distance, resulting in reduced latency and improved response times. MEC serves as a bridge between cloud computing and mobile networks, offering a decentralized, responsive, and efficient computing environment [4]. On the other hand, one of the pivotal roles in optimizing the performance of edge computing systems is efficient task offloading. Task offloading in MEC is the process of transferring computation and data processing tasks from mobile devices, such as smartphones or IoT devices, to edge servers or nodes located at the edge of the mobile network. The development of task offloading strategies for edge computing networks has attracted a lot of research attention in the last few years [5, 6]. Using queuing theory, Liu et al. [5] introduced the MEC system's computational offloading strategy. Their research aimed to determine an ideal offloading probability for mobile users within the context of MEC systems. A game theory-based approach was proposed by Messous et al. [6] to solve the computationally demanding task offloading issue from unmanned aerial vehicles (UAVs) in the MEC paradigm.

Task offloading strategies play a pivotal role in determining whether tasks should be processed

locally at the edge, offloaded to nearby edge servers, or sent to centralized cloud infrastructure. The critical challenge lies in identifying the most suitable task offloading strategy, and this challenge is significantly influenced by the chosen network planning approach. However, the intricate relationship between different network planning models (particularly the M/M/1 and M/M/2 queuing models) and the design of task offloading strategies within edge computing remains largely unexplored. Furthermore, offloading tasks to edge servers may cause network congestion, particularly in places with high population densities or during periods of peak usage. Therefore, an efficient network planning model is necessary to prevent congestion and minimize latency. Therefore, our research aims to bridge this gap by exploring how these network planning approaches affect the execution of computational tasks in edge computing networks.

2. System Architecture and Network Model

2.1 Proposed Model

The task offloading architecture for edge computing networks is depicted in Fig. 1, which consists of three distinct layers: end users, edge servers, and cloud server. In the first layer, a huge number of end users are present, each having the option to either locally execute their tasks on their own devices or offload them to the nearby edge server. The second layer comprises various access points (AP), with each AP associated with a deployed edge server. The edge servers are responsible for receiving offloaded computation tasks from mobile users and processing them efficiently. They serve as a middle layer between mobile users and remote cloud server, reducing latency and providing faster responses. In the final layer, the core network is linked to a centralized cloud server responsible for delivering cloud computing services. This layer can handle more complex and resource-intensive tasks that may not be suitable for processing at the edge. As depicted in Fig. 1, we denote the set of end users as U , represented as $U = \{1, 2, 3, \dots, U\}$, and the set of access points as S , which serve to cover end users in edge computing networks, defined as $S =$

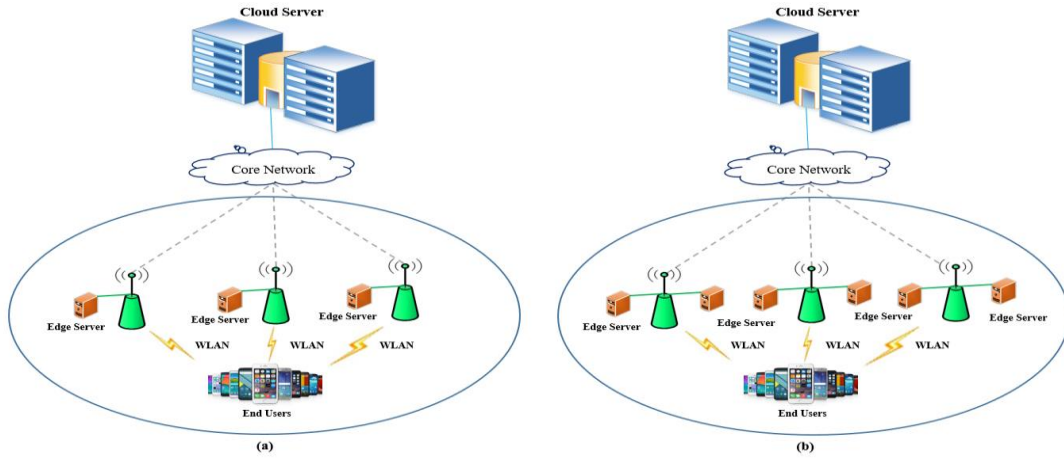


Fig. 1. Network planning-based edge computing networks: (a) M/M/1 queuing based offloading (b) M/M/2 queuing based offloading

$\{1, 2, 3, \dots, S\}$. Each computation task, denoted as T_i , for a specific user device i , is typically characterized by three essential parameters, as expressed below:

$$T_i = \{d_i, \omega_i, \psi_i^{\max}\} \quad (1)$$

For task T_i , d_i denotes the input data size required for computation. ω_i represents the desired number of CPU cycles required to fulfill the task T_i . Lastly, ψ_i^{\max} refers to the maximum allowable delay that the computation task can tolerate.

2.2 Task Offloading Based on the M/M/1 Queuing Planning Approach

The M/M/1 queuing model represents a single-server system where tasks arrive through a Poisson distribution and are serviced exponentially. The M/M/1 model is a mathematical framework for analyzing queuing systems, representing the average time a task spends in a queuing system before it is serviced. The architecture of the M/M/1 queuing planning approach is shown in Fig. 2.

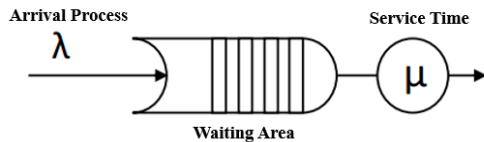


Fig. 2. M/M/1 queuing model

In an M/M/1 queue (single-server queue),

- Arrival Process (M): Arrivals to the queue follow a Poisson process, which means they occur randomly over time.
- Service Time (M): Service times are exponentially distributed, indicating the time

it takes to serve a user follows a specific statistical distribution.

- Single Server (1): There's one server in the queue capable of serving one user at a time.

Response time is a critical performance metric, as it measures the average time an entity spends in the system, including both queue waiting time and actual service time. An M/M/1 queue's response time \mathcal{R} can be written as follows:

$$\mathcal{R} = 1 / (\mu - \lambda) \quad (2)$$

where μ is the service rate (rate at which the server can complete tasks) and λ is the arrival rate. The response time affects both the arrival rate and the service rate. If the rate of arrivals is lower than the rate of services ($\lambda < \mu$), the system is stable, and the response time is finite. On the other hand, the system becomes unstable, and the response time approaches infinity, if the arrival rate exceeds the service rate ($\lambda > \mu$).

2.3 Task Offloading Based on the M/M/2 Queuing Planning Approach

The M/M/2 queuing model extends the single-server concept to a dual-server scenario. In queuing theory, the M/M/2 queue represents a queuing system with two servers where tasks arrive following a Poisson process and are served with exponentially distributed service times. There are two servers in the queue to serve tasks concurrently. This means that two different tasks can be served simultaneously if they arrive when both servers are available. This model allows us to explore the implications of parallel task processing on network performance. The architecture of the M/M/2 queuing planning

approach is shown in Fig. 3. The response time in an M/M/2 queue depends on the utilization of the servers. The response time \mathfrak{R} in an M/M/2 queue can be calculated using the formula:

$$\mathfrak{R} = 4\mu / (2\mu - \lambda)(2\mu + \lambda) \quad (3)$$

where μ represents the service rate for each of the two servers and λ is the total arrival rate.

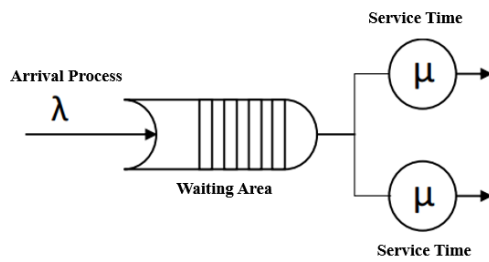


Fig. 3. M/M/2 Queuing Model

3. Performance Evaluation

In this section, we present the results of our study on task offloading strategies based on different network planning approaches in edge computing networks. We evaluated the performance of various task offloading strategies and their implications through the EdgeCloudSim simulator [7]. In this study, we have used three different applications in a simulation to model various real-life scenarios. The three applications used are health monitoring (HM), infotainment (I), and augmented reality (AR). Each of these applications has distinct characteristics that affect their sensitivity to latency, delay tolerance,

and computational demands. The key characteristic parameters for these applications, as well as other simulation parameters, are detailed in Tables 1 and 2, respectively [7]. In this simulation, a uniform computing capacity of 40 giga instructions per second (GIPS) has been assumed for each access point across various offloading schemes. For instance, when applying the M/M/1 queuing approach, an edge server computing capacity of 40 GIPS was utilized. In contrast, for the M/M/2 queuing approach, 20 GIPS of computing capacity were allocated to each edge server. The system model incorporates various performance metrics (response time, offloading success rate, etc.) to assess the effectiveness of task offloading strategies and network planning approaches. Moreover, to assess the performance of our proposed (edge with cloud offloading) scheme across different scenarios, we conducted a comparative analysis with two distinct benchmark offloading schemes. The two benchmark schemes, along with our proposed scheme, are outlined as follows:

- Local computing (mobile): In this scheme, mobile devices perform all processing locally, without offloading tasks to the edge or cloud.
- Edge offloading (M/M/1): Under this scheme, mobile users' tasks are offloaded to adjacent edge servers for processing.
- Edge (M/M/1) with cloud offloading: In this scheme, tasks from mobile users are first offloaded to edge servers and then, if necessary, to the cloud for processing.

Table 1. Application characteristics regarding three different applications

| Parameters | Application Types | | |
|------------------------------|-------------------|----------|----------|
| | AR | HM | IA |
| Usage percentage | 40 | 30 | 30 |
| Task Interarrival (sec) | 2 | 3 | 7 |
| Active/Idle Period (min) | 40/20 | 45/90 | 30/45 |
| Upload/Download Data (KB) | 1500/250 | 200/1250 | 250/1000 |
| Task length (GI) | 12 | 6 | 15 |
| VM utilization on Edge/Cloud | 8/0.8 | 4/0.4 | 10/1 |

Table 2. Simulation parameters

| Parameters | Value |
|--|-------------------------|
| Number of simulation round | 10 |
| Number of mobile devices | 10~100 |
| Network model | M/M/1 and M/M/2 |
| Number of cores per edge server | 4 (M/M/1) and 2 (M/M/2) |
| Total Capacity of edge server | 40 GIPS |
| WLAN Bandwidth | 100 Mbps |
| Number of locations | 3 |
| Mean waiting time in Type 1/2/3 Places | 8/5/2 minutes |

3.1 Simulation Results

Fig. 4 illustrates the comparison between edge offloading (M/M/1) and edge offloading (M/M/2) schemes in terms of the number of mobile users versus average service time under the same edge server capacity. As the number of mobile user's increases, the server's workload becomes heavier. This often leads to increased queuing delays and longer average service times. The average service time is directly influenced by the queuing delays, which can grow linearly with the number of users.

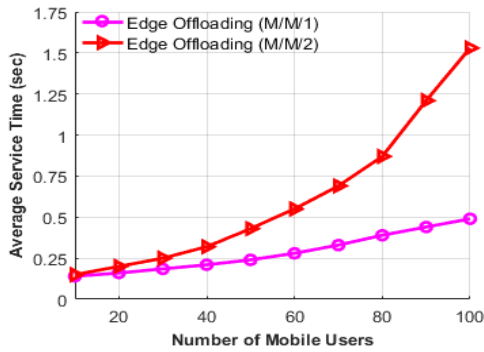


Fig. 4. The average service time in terms of varying quantities of mobile users

From the analysis in **Fig. 4**, we can observe that the edge offloading scheme using the M/M/1 queuing approach consistently offers a lower service time when compared to the edge offloading scheme using the M/M/2 queuing approach.

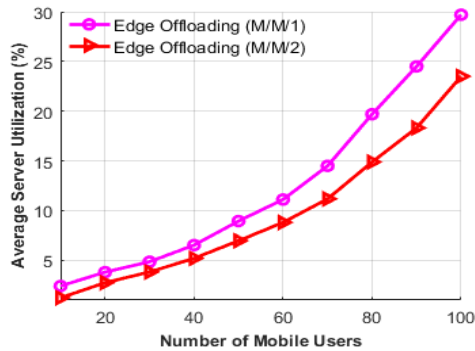


Fig. 5. The average server utilization for a varying number of mobile users

On the other hand, **Fig. 5** illustrates the relationship between average server utilization and the quantity of mobile users. The figure makes it clear that as the number of mobile users grows, the average server utilization also rises in

all scenarios. Notably, edge offloading using the M/M/1 queuing approach exhibits higher server utilization compared to the M/M/2 approach in all cases. In summary, when both schemes have the same edge server capacity, the edge offloading (M/M/1) scheme offers significant advantages in terms of response time and efficient server utilization as the number of mobile users increases.

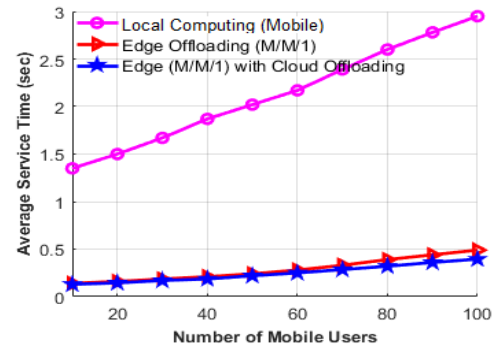


Fig. 6. Comparison of the average service time in relation to the number of mobile users

In order to validate the importance of our suggested edge (M/M/1) with cloud offloading strategy, we conducted an experiment to assess the average service time across different quantities of mobile users, as depicted in **Fig. 6**. Based on **Fig. 6**, it can be seen that when there are between 10 and 70 mobile users, the edge offloading and our proposed schemes will produce nearly similar results, while the local computing scheme exhibits longer service times. This is because of the limited computing capacity of mobile devices, which struggle to manage a high volume of tasks. The average service time rises in all scenarios as the number of mobile devices increases. However, in the case of our proposed approach, tasks are distributed between the edge server and the cloud. Consequently, the service time increase is comparatively restrained when contrasted with the local computing and edge offloading schemes. For instance, with 80 mobile users, the service times for local computing (mobile), edge offloading (M/M/1), and our proposed strategies are 2.6, 0.39, and 0.32 seconds, respectively. Consequently, our suggested approach is capable of reducing the average service time by around 88.4% in comparison to local computing (mobile) and 14% when contrasted with the edge offloading (M/M/1) scheme.

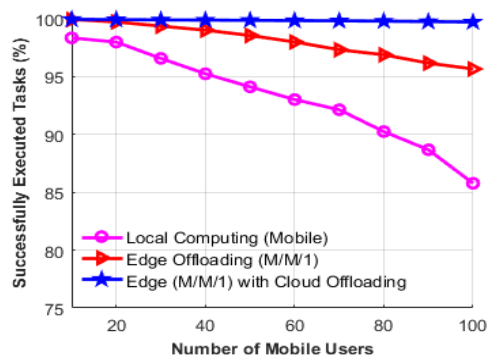


Fig. 7. The number of successfully executed tasks in terms of the number of mobile users

The rate of successful task completion is a critical metric for assessing the performance of task offloading. **Fig. 7** illustrates the success rate of completed tasks among varying numbers of mobile users. As depicted in **Fig. 7**, it is clear that all the discussed offloading techniques are capable of efficiently managing the majority of offloaded tasks when the system operates under a light workload. However, as the number of mobile users in the network increases, the success rate of offloaded tasks declines rapidly. For instance, in the case of the local computing (mobile) scheme, congestion is encountered when the number of mobile users reaches 20, while the edge offloading (M/M/1) scheme experiences congestion at 50 mobile users. Conversely, the edge (M/M/1) with cloud offloading scheme demonstrates a remarkable ability to handle up to 100 mobile users without congestion. This is attributed to the optimal task distribution between the edge and cloud servers in our proposed edge (M/M/1) with cloud offloading scheme, effectively mitigating congestion at the edge layer. Based on the simulation results, it is evident that our proposed scheme outperforms both the local computing (mobile) and edge offloading (M/M/1) schemes in managing offloaded tasks efficiently.

4. Conclusions

This paper thoroughly investigates and compares two prominent network planning approaches, namely the M/M/1 and M/M/2 queuing models, within the context of edge computing networks. Moreover, we develop different task offloading strategies by using these network planning approaches. The study affirmed that, when both

schemes have an identical edge server capacity, the M/M/1 approach consistently provides higher server utilization and lower service times than the M/M/2 approach across various scenarios. In addition to that, the effectiveness of the proposed edge (M/M/1) with cloud framework has been confirmed through testing with EdgeCloudSim simulators. It has demonstrated its capability to reduce service times and enhance the percentage of successfully completed offloaded tasks while maintaining the quality of service (QoS) for various edge applications.

References

- [1] T. Taleb, A. Ksentini, and R. Jantti, "Anything as a service for 5G mobile systems," *IEEE Network*, Vol 30, No. 6, pp. 84–91, 2016.
- [2] A.u.R. Khan, M. Othman, S.A. Madani, and S.U. Khan, "A survey of mobile cloud computing application models," *IEEE Communications Surveys & Tutorials*, Vol 16, No. 1, pp. 393–413, 2014.
- [3] P. Mach and P. Becvar, "Mobile edge computing: A survey on architecture and computation offloading," *IEEE Communications Surveys & Tutorials*, Vol 19, No. 3, pp. 1628–1656, 2017.
- [4] Y. Liu, M. Peng, G. Shou et al., "Toward Edge Intelligence: Multiaccess edge computing for 5G and Internet of Things," *IEEE Internet Things Journal*, Vol 7, No. 8, pp. 6722–6747, 2020.
- [5] L. Liu, Z. Chang, X. Guo, and T. Ristaniemi, "Multi-objective optimization for computation offloading in mobile-edge computing," in *Proc. of IEEE Symposium on Computers and Communications (ISCC)*, 2017, pp. 832–837.
- [6] M.-A. Messous, S.-M. Senouci, H. Sedjelmaci, and S. Cherkaoui, "A game theory based efficient computation offloading in an UAV network," *IEEE Transactions on Vehicular Technology*, Vol 68, No. 5, pp. 4964–4974, May 2019.
- [7] C. Sonmez, A. Ozgovde, and C. Ersoy, "EdgeCloudSim: An environment for performance evaluation of Edge Computing systems," *Transactions on Emerging Telecommunications Technologies*, Vol 29, No. 11, pp. 1–17, 2018.

A Research on the Visualization Services in BIM-Based Virtual Reality Environments

Kijung Ryu, Joonyoung Ahn and Dongho Kim

Department of IT Policy and Management, Graduated School of Soongsil University
Seoul, South Korea

[e-mail: yoo123111@gmail.com, corea003@gmail.com, dkim@gmail.com]

*Corresponding author: Dongho Kim

Abstract

As digital technology continues to advance rapidly, its impact on various economic activities is also growing. In particular, the construction industry is also undergoing digitalization, and digital technologies in the construction field, often referred to as 'ConTech,' are emerging with the solutions that can innovate the previously rigid economic structures in construction and architecture, bringing about new dynamism. Amid this digital transformation, the importance of technologies like BIM is increasing. BIM represents the physical and functional characteristics of a building in a digital form, integrating and sharing information to enable all stakeholders in a construction project to have access to the same information and understanding. This leads to the improved decision-making and reduction in errors throughout all stages, from design to construction and operation. Virtual reality has rapidly emerged as a significant technology in the construction industry, demonstrating relevance in supporting various construction activities. VR enables users to intuitively explore and analyze the interior of the buildings in a 3D environment. This research aims to study visualization services in a BIM-based virtual reality environment to promote the dissemination and utilization of this technology.

Keywords: Building Information Modeling, Virtual Reality, Facility Management, Data Visualization

1. Introduction

Recent advancements in the industrial sector have placed a spotlight on Virtual Reality (VR). The launch of various VR devices, the advancement of high-resolution displays, and the evolution of development software have accelerated the technological progress of virtual reality. These advancements in VR technologies are becoming popularized in our daily lives, offering users the opportunity to experience the boundary between virtual and reality at a relatively low cost. However, in the architectural industry, the application of VR is still limited to safety training on construction sites and promotional leasing, with its frequency of use being notably low[1]. Recently, Building Information Modeling (BIM) is being actively

introduced in the construction sector, and its adoption is on the rise due to its ability to solve various problems in advance[2]. The utilization of BIM data, as an approach in construction and facility management fields, is expanding not only during the construction process but also to the operation and maintenance phases of a building or facility[3]. There is an increasing demand to manage control data required for the maintenance of facilities (detailed information on building components and attributes) based on BIM technology and to maximize the value of the data created through BIM[4]. There is also a growing need for the visualization of BIM data and real-time collaboration features to minimize the daily maintenance inspection costs of various buildings, and the necessity for facility simulation and data visualization using VR is being emphasized[5]. Moreover, research is

ongoing to incorporate virtual reality in exploring and verifying the information contained in BIM[6], and international development cases have been observed. This study aims to describe the visualization service in a Virtual Reality environment based on Building Information Modeling (BIM).

2. BIM and Virtual Reality

2.1 Virtual Reality

Virtual Reality (VR) is an environment created using computer systems and various sensors, allowing users to have an experience similar to reality. Virtual reality employs computer technology to recreate specific environments or scenarios, allowing the person experiencing it to interact as if they were truly present in that setting. Additionally, it's recognized as a key technology playing a vital role in the fourth industrial revolution, emphasizing deep connections and interactions between humans and computers. Virtual reality operates with the user at its center, recognizing and understanding the surrounding environment through various human senses (visual, auditory, tactile, olfactory, gustatory)[7]. Such sensory information has been researched to provide a sense of realism genuinely felt through interaction with the virtual space.

2.2 BIM-based Virtual Reality

BIM-based Virtual Reality (VR) is being utilized as an innovative technology in the construction and facility management sectors. This technology integrates Building Information Modeling (BIM) with Virtual Reality (VR) capabilities to offer a new dimension of visualization and interaction for building and facility management. BIM is a method of creating and managing a digital 3D model of buildings and facilities, effectively managing information throughout its life cycle, including design, construction, and maintenance. The VR technology leverages this BIM data, enabling users to navigate and interact with buildings or facilities in a virtual space. BIM-based VR is employed for various purposes, such as safety training and simulations, visual reviews of design modifications, optimization for cost savings, and sustainability analysis.

2.2 BIM-based visualization service

Visualization services refer to the representation of data or information in forms such as graphics, images, charts, videos, etc., making it easy and intuitive for users to understand. It simplifies complex data or concepts, facilitating effective communication and is utilized in various domains. In the construction industry, employing visualization services aids stakeholders involved in a project to comprehend the model and its characteristics. Users can experience immersion, interacting with materials and objects, and through BIM models, they can easily identify structural inefficiencies, modeling errors, design issues, and the like. This provides the potential for more effective monitoring of construction activities and building maintenance.

3. BIM-based Virtual Reality Visualization Service

The Visualization Service of BIM-based virtual reality environments is a key technology in urban planning and management, particularly related to the construction of virtual spaces for urban infrastructure, visualization of digital twin data, and integrated operational management platforms. The aim of constructing a virtual space for urban infrastructure is to accurately reproduce the intricate structures and facilities of a city within a virtual 3D environment.

Fig. 1 depicts an example from Daewoo Engineering & Construction, which introduced a virtual reality model house called "MetaGallery"[8]. This was developed based on BIM and utilized the Unity Engine to create a realistic virtual space.



Fig. 1. Daewoo Engineering & Construction "Virtual reality model home using BIM"

Digital twin-based data visualization serves as a

bridge connecting the physical and digital worlds. A digital twin is a virtual replica of a physical object or system, enabling the collection and analysis of real-time data. Research on the integrated operational management platform for smart city services focuses on methods to manage various smart city services and technologies in a unified manner on a single platform.

Fig. 2 presents an example of utilizing visualization services in a BIM-based virtual reality environment[9].

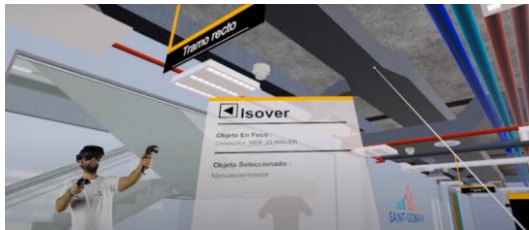


Fig. 2. Introduction to VT-Lab's BIM-based virtual reality environment visualization service

Such technology makes accessing BIM data more accessible and can enhance customer and user satisfaction. Especially, there's the advantage of experiencing movement within the model and being able to explore and use the model from any location.

4. Conclusions

Building Information Modeling (BIM) technology creates a 3D digital representation of all elements, spaces, systems and connections of a building. The main features and advantages of BIM-based VR visualization can be described as follows. First, real-time interaction is possible. Second, it provides effective support for construction-related decision-making. Third, it can provide simulations of building energy efficiency, lighting, and ventilation. Fourth, it can improve communication between building stakeholders. Fifth, it can be used in various education and training scenarios such as safety education and maintenance training at construction sites. Therefore, BIM-based VR visualization provides great help in improving decision-making and increasing efficiency

during the design, construction, and operation stages of buildings.

References

- [1] Jo, B., Cho, H.-H., & Son, C.-B., "An Analysis on the Utilization Status of VR and AR in the Construction Industry and Its Improvement Measures," *Journal of the Architectural Institute of Korea*, 37(1), 229–236, 2021.
<https://doi.org/10.5659/JAIK.2021.37.1.229>
- [2] Chen, L., & Luo, H., "A BIM-based construction quality management model and its applications," *Automation in Construction*, 46, 64–73, 2014.
- [3] Moon, S. W., Kim, S. D., & Park, M. K., "Application of a 3D graphic model for bridge maintenance," *Korean Journal of Construction Engineering and Management*, 12(2), 64–71, 2011.
- [4] Lim, J., You, S., & Kim, Y., "Monitoring System based on Bigdata Platform for Safety Management of Road Facilities," *Journal of advanced information technology and convergence*, 18(11), 139–151, 2020.
- [5] Leen-Seok Kang, Jin-Seok Moon, & Jung-Hee Kwon, "Simulation Function of Railway Station Project by Virtual Construction Technology," *Journal of the Korean Society for Railway*, 11(5), 433–440, 2008.
- [6] Lee, Jae-Sung, & Lee, Sang-Yun, "A Study of the Integrated Design Process with the Use of VR Technology - Focused on Differences between Architectural Design Process and the Integrated Design Process with the Use of BIM," *Journal of the Architectural Institute of Korea*, 39(7), 85–95, 2023.
- [7] YOUN JI-HYUN, "Simulation of Efficient Virtual Space using 3D Holography," *한국디자인포럼*, (19), 295–304, 2008.
- [8] Daewoo Engineering & Construction develops a 3D BIM model virtual model home,
<http://www.kharn.kr/news/article.html?no=19171>
- [9] <https://www.vt-lab.com/en/home/>

Dynamic Access Control Policy Based on Blockchain and Machine Learning to Enhance Security in IoT Smart Homes

Hafiz Adnan Hussain, Zulkefli Mansor, Zarina Shukur, and Uzma Jafar

Faculty of Information Science & Technology, Universiti Kebangsaan Malaysia, Selangor, Malaysia
[e-mail : hafizadnan87@gmail.com, kefflee@ukm.edu.my, zarinashukur@ukm.edu.my, uzmajafar@gmail.com]

*Corresponding author: Hafiz Adnan Hussain

Abstract

The rapid growth of Internet of Things (IoT) technology in smart homes has presented new challenges for security and access control that traditional models are unable to cope with them. To address this issue, a new dynamic access control policy framework has been introduced for IoT smart homes, which utilizes both Blockchain and Machine Learning (ML). The framework utilizes a decentralized architecture where access control policies are stored and verified on a Blockchain network, ensuring that they are immutable and transparent. Machine learning algorithms, such as Random Forest and Decision Trees, are incorporated to continuously analyze device and user behavior, updating access control policies dynamically as required. Initial results indicate a 98% success rate in detecting unauthorized access and a minimal 2% false positive rate, which significantly improves over the traditional method. However, the study was conducted in a controlled environment, and more real-world testing is needed.

Keywords: Security, Internet of Things (IoT), Data Access Control, Blockchain, Artificial Intelligence, Machine Learning

1. Introduction

The Internet of Things (IoT) has revolutionized the world of smart homes, with an estimated 27 billion devices projected to be connected by 2025. While these devices provide numerous conveniences and automate various tasks, they pose significant security challenges [1-3]. Traditional security models, which are often centralized, are not suitable for handling the dynamic and distributed nature of IoT smart homes [4, 5]. This security gap calls for innovative approaches to establish effective access control mechanisms. This research aims to address this issue by introducing a dynamic

access control policy that leverages the immutable and transparent nature of Blockchain technology, along with the predictive power of ML algorithms. The proposed framework aims to significantly improve security measures, reduce unauthorized access, and offer a scalable solution compatible with existing IoT devices in smart homes [6, 7].

Based on our research, we have designed a decentralized architecture for access control with the help of Blockchain. To enable real-time analysis and dynamic updating of access control policies, we have utilized Machine learning algorithms, specifically Random Forest and Decision Trees. Furthermore, we have evaluated the effectiveness of this framework through

simulation and real-world testing.

This paper is organized as follows: Section 2 reviews related work in IoT security, Blockchain, and machine learning. Section 3 outlines the methodology and the implementation details in Section 4. Section 5 presents the results, and Section 6 discusses their implications. Finally, Section 7 concludes the paper and outlines future research avenues.

2. Related Work

The security of IoT devices in smart homes has been researched extensively. However, traditional approaches focus mainly on centralized architectures and have scalability and reliability issues. Several studies have pointed out the limitations of using static, rule-based access control policies in smart homes. For instance, [8, 9] emphasized the need for dynamic solutions that adapt to changing conditions. Blockchain technology has emerged as a robust solution for enhancing security across various domains, including the Internet of Things (IoT). In a study by [10], a blockchain-based framework for IoT that focuses on providing secure and transparent access control mechanisms was proposed [11]. Machine learning algorithms have been utilized to analyze behavioral patterns and automatically update access control policies. [12] demonstrated a significant reduction in false positives and false negatives using machine learning algorithms. However, they have started exploring the synergy between Blockchain and machine learning for enhanced security, but their focus has yet to be on the smart homes [13]. It presents an excellent opportunity to propose a framework that combines the immutability and transparency of Blockchain with the predictive analytics of machine learning for IoT smart home security.

3. Methodology

3.1 Research Design

This study will use an experimental research design, including simulation and real-world testing. It will be divided into four phases:

1. Designing and developing a Blockchain-based architecture.
2. Integrating ML algorithms for dynamic policy updates.
3. Testing the framework's effectiveness through simulation.
4. Testing the framework in a controlled smart home environment.

3.2 Data Collection

Data will be collected from multiple IoT devices in a smart home setup. The dataset will include device IDs, timestamps, access requests, and user behavior patterns. All of this data will be securely stored on a Blockchain.

3.3 Blockchain Implementation

Ethereum will be used to implement the Blockchain layer. Its modular architecture makes it simpler to develop decentralized applications, as shown in Eq.1.

$$\text{Blockchain Layer: } \mathcal{B}_i = \{b_1, b_2, \dots, b_n\} \quad (1)$$

Where B_i represents a block containing access control policies and transaction data.

3.4 Machine Learning

A combination of Random Forest and Decision Trees will be utilized for real-time analysis, with the ML model dynamically updating access control policies based on collected data, as shown in Eq.2.

$$\text{ML Model: } M(x) = \text{Policy Update Decision} \quad (2)$$

Where x is the feature vector containing device and user behavior data.

3.5 Evaluation Metrics

The framework will be evaluated using the following metrics:

1. Security Efficacy: Measured by the number of unauthorized access attempts detected and prevented.
2. Efficiency: Measured by the system's response time for access requests.
3. Scalability: Evaluated by adding increasing devices and users to the system.

4. Implementation

4.1 Blockchain Architecture

The Blockchain layer serves as the backbone for secure and transparent access control. It is implemented using Ethereum, a public Blockchain that supports smart contracts. Ethereum's Solidity language defines smart contracts for access control policies and their enforcement.

4.2 Smart Contract Algorithm

The smart contract algorithm for access control can be represented in Eq.3.

$$SmartContract(T, P) = \begin{cases} GA & \text{if } H(T) \in P \\ DA & \text{otherwise} \end{cases} \quad (3)$$

Where T represents the transaction request, P is the set of predefined policies, GA represents Grant Access, DA represents Deny Access, and $H(T)$ is the hash of the transaction request.

4.3 Machine Learning Integration

Due to their efficiency and effectiveness in classification tasks, real-time analysis of user behavior and device interactions is performed using ML models, specifically Random Forest and Decision Trees.

4.4 Real-Time Policy Update Algorithm

The ML model analyzes historical data to predict potential security threats and update access control policies dynamically. The algorithm can be represented in Eq. 4, where D represents the dataset of historical user behavior and device interactions, M is the ML model, UP represents the Update Policy, MP represents the Maintain Policy, and PT represents Potential Threat.

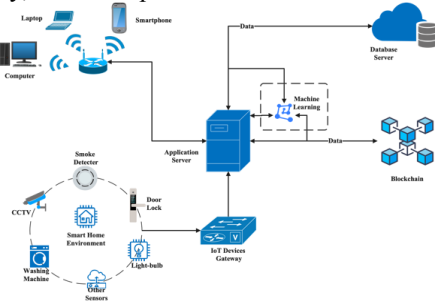


Fig. 1. Design Architecture of Dynamic Access Control Policy Based on Blockchain and Machine Learning for IoT

4.5 Integration and Testing

The Blockchain and ML layers are integrated using RESTful APIs. The system has been initially tested using a simulated smart home environment, and further testing in real-world conditions is planned.

5. Results

5.1 Simulation Testing

The framework underwent various scenarios in the simulation tests, including both authorized and unauthorized access attempts, to test its security efficacy.

1. Unauthorized Access Detection Rate: Our framework detected 98% of unauthorized access attempts.
2. False Positives: The system had a low false positive rate of only 2%, indicating high accuracy in distinguishing between legitimate and illegitimate access requests.

5.2 Simulation Testing

The security efficacy can be calculated using the Eq.5.

$$SE = \frac{\text{Number of Unauthorized Access Detected}}{\text{Total Number of Access Attempts}} \times 100 \quad (5)$$

Where SE is Security Efficacy.

5.3 Real-world Testing

The framework was deployed in a smart home equipped with various IoT devices in the real-world test environment. The results were promising:

1. Efficiency: The average response time for access requests was less than 300 ms, showcasing the system's efficiency.
2. Scalability: The system handled increased devices and users without significantly impacting performance.

The efficiency can be measured using Eq.6. The Radar Chart compares four security models across three key metrics: Unauthorized Access Detection Rate, False Positives, and Average Response Time. The chart shows that the Proposed Framework, represented in blue, outperforms the other models in all metrics. It

has the highest unauthorized access detection rate, the lowest rate of false positives, and the fastest response time. This visual representation supports the argument made in the paper that combining Blockchain and ML through a hybrid approach offers better performance in securing IoT smart homes than traditional or single-technology models.

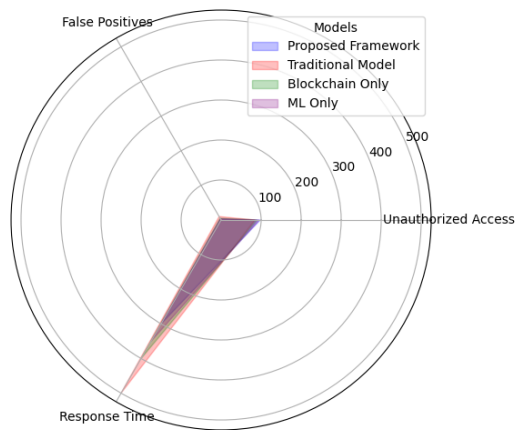


Fig. 2. Performance Comparison Across Models

6. Discussion

During simulation testing, the proposed framework was found to be highly effective in enhancing security, with an unauthorized access detection rate of 98% and only 2% false positives. The framework's efficiency and scalability were further demonstrated during real-world testing, confirming its practical application. This framework combines Blockchain's immutability and transparency with ML's predictive analytics, resulting in a robust, dynamic, and scalable solution for access control. It has significant implications for IoT security, specifically in smart homes. However, it's worth noting that real-world testing was conducted in a controlled environment, and results may vary in more complex setups. Additionally, the ML model may need retraining periodically to adapt to new threats. For future research, we can extend the framework to support more complex IoT ecosystems beyond smart homes and investigate the integration of more advanced ML algorithms to refine access control policies.

7. Conclusion

As smart homes become more sophisticated, security solutions must also become more robust, scalable, and adaptable. In this paper, we present a novel framework for dynamic access control in IoT-enabled smart homes. Our approach integrates both Blockchain technology and Machine Learning (ML) to ensure an efficient and secure system. We use Ethereum's smart contract execution capability to ensure a decentralized, transparent, and immutable process to access control. Additionally, we have incorporated ML algorithms to offer real-time analysis and dynamic updates to access control policies. Our simulation and real-world testing confirm the effectiveness of our proposed framework. Our system has a high rate of unauthorized access detection and low false positives, demonstrating significant promise in enhancing smart home security. Furthermore, our framework is efficient, scalable, and has the potential for real-world applications. Our research has limitations that offer avenues for future work. Some of the potential research directions include extending the framework to more complex IoT ecosystems and investigating the integration of advanced ML algorithms. Our research contributes to the ongoing efforts to secure IoT smart homes and presents a comprehensive solution that could be the cornerstone for future developments.

8. References

- [1] U. Jafar, M. J. Ab Aziz, Z. Shukur, and H. A. Hussain, "A Cost-efficient and Scalable Framework for E-Voting System based on Ethereum Blockchain," in *2022 International Conference on Cyber Resilience (ICCR)*, 2022: IEEE, pp. 1-6.
- [2] U. Jafar, M. J. Ab Aziz, Z. Shukur, and H. A. Hussain, "A Systematic Literature Review and Meta-Analysis on Scalable Blockchain-Based Electronic Voting Systems," *Sensors*, vol. 22, no. 19, p. 7585, 2022.
- [3] H. A. Hussain, Z. Mansor, and Z. Shukur, "Comprehensive survey and research directions on blockchain iot access control," *International Journal*

- of Advanced Computer Science and Applications*, vol. 12, no. 5, 2021.
- [4] U. Jafar, M. J. A. Aziz, and Z. Shukur, "Blockchain for electronic voting system—review and open research challenges," *Sensors*, vol. 21, no. 17, p. 5874, 2021.
- [5] K. Wüst and A. Gervais, "Do you need a blockchain?," in *2018 crypto valley conference on blockchain technology (CVCBT)*, 2018: IEEE, pp. 45-54.
- [6] Y. Liu, F. R. Yu, X. Li, H. Ji, and V. C. Leung, "Blockchain and machine learning for communications and networking systems," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 2, pp. 1392-1431, 2020.
- [7] U. Jafar and M. J. A. Aziz, "A state of the art survey and research directions on blockchain based electronic voting system," in *Advances in Cyber Security: Second International Conference, ACeS 2020, Penang, Malaysia, December 8-9, 2020, Revised Selected Papers 2*, 2021: Springer, pp. 248-266.
- [8] H. A. Hussain, Z. Mansor, Z. Shukur, and U. Jafar, "Ether-IoT: A Realtime Lightweight and Scalable Blockchain-Enabled Cache Algorithm for IoT Access Control," *Computers, Materials & Continua*, vol. 75, no. 2, 2023.
- [9] M. Conti, E. S. Kumar, C. Lal, and S. Ruj, "A survey on security and privacy issues of bitcoin," *IEEE communications surveys & tutorials*, vol. 20, no. 4, pp. 3416-3452, 2018.
- [10] I. Makhdoom, M. Abolhasan, J. Lipman, R. P. Liu, and W. Ni, "Anatomy of threats to the internet of things," *IEEE communications surveys & tutorials*, vol. 21, no. 2, pp. 1636-1675, 2018.
- [11] M. A. Ferrag, M. Derdour, M. Mukherjee, A. Derhab, L. Maglaras, and H. Janicke, "Blockchain technologies for the internet of things: Research issues and challenges," *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 2188-2204, 2018.
- [12] I. H. Sarker, A. I. Khan, Y. B. Abushark, and F. Alsolami, "Internet of things (iot) security intelligence: a comprehensive overview, machine learning solutions and research directions," *Mobile Networks and Applications*, vol. 28, no. 1, pp. 296-312, 2023.
- [13] R. Vinayakumar, M. Alazab, K. Soman, P. Poornachandran, A. Al-Nemrat, and S. Venkatraman, "Deep learning approach for intelligent intrusion detection system," *Ieee Access*, vol. 7, pp. 41525-41550, 2019.

A Study On Adaptation Tasks Between Suitable DBMS For Cloud Migration

Joonyoung Ahn, Changik Oh, Sugil Hwang and Dongho Kim

Department of IT Policy and Management, Graduated School of Soongsil University
Seoul, South Korea

[e-mail: corea003@gmail.com, lowerright@gmail.com, soogool@gmail.com, dkim@ssu.ac.kr]

*Corresponding author: Dongho Kim

Abstract

This study aims to conduct the research work for the ongoing rapid transition of the current legacy-based system to a cloud-based system. When transitioning from a legacy environment to a cloud-based one, it is important to anticipate potential issues such as system performance degradation and excessive scale-out due to the failure of initial CPU usage estimation in the DBMS. To address these problems, we intend to verify suitable queries using TPC-H BMT test.

As part of the preliminary research, we will categorize and summarize the types and content of TPC's BMT (Business Metric Test) and define a testing methodology that can be applied to common tasks during the activation test.

The ultimate goal of this study is to provide suitable business tasks for each query in TPC-H for the legacy environment's DBMS. We conclude the paper by initially estimating the query execution times for each DBMS in the legacy environment and confirming their characteristics..

Keywords: DBMS, Cloud migration, TPC-H

1. Introduction

After the enforcement of the Cloud Computing Act in South Korea in September 2015, a total of 1,961 tasks from 467 public sector organizations were transitioned to cloud computing. In the private sector, support was provided for the use of cloud computing by 8,000 companies. Furthermore, based on the 3rd Basic Plan in 2021, there are plans to expand the use of internal business systems in government agencies, and support for the use of cloud computing is also being planned for over 10,000 systems in the private sector. [1] However, due to insufficient technical capabilities, a lack of personnel, and inadequate research and development, as well as inefficient operations in preparation for demand, major tasks are being operated without transitioning to global corporate data centers or transitioning from legacy formats from the past. [2]

While various guides such as the Practical Guide for Cloud Adoption in the Private Sector[3], the Guidelines for the Use of Private Clouds in Government Public Agencies[2], and the Self-Diagnosis and Type-Specific Adoption Guidelines for Cloud Service Introduction in the Public Sector[4] provide specific information regarding the transition, the current cloud migration process is underway without detailed instructions on capacity calculation and usage guidelines for DBMS.

In the event of issues arising within the DBMS during cloud migration, Auto Scaling can be triggered to address and resolve them. However, in cases of fundamental design errors, problems such as service slowdown and service interruptions can occur, and in severe situations, it may necessitate a complete reconstruction.[5]

Therefore, this paper aims to conduct BMT (Benchmark Test) on various heterogeneous

DBMS that can be used in legacy systems to measure the performance of DBMS. It also seeks to analyze which DBMS is suitable for specific tasks by reflecting the characteristics of each query in the Power Test.

2. Related Research

The Transaction Processing Performance Council (TPC) was established in August 1988 in the United States and provides standards for measuring transaction processing performance. TPC has a total of 18 measurement methods, and currently, 10 of them are in active use. [6]

2.1 Active TPC Benchmarks Types

Table 1. Active TPC Benchmarks Types

| BenchMarks | Fields | Notes |
|------------|---|-------|
| TPC-C | OLTP-Based Basic Test | |
| TPC-DS | Industry Standard Benchmark Test for Measuring the Performance of Decision Support Solutions. | |
| TPC-E | TPC-C and Financial Market Interaction Test. | |
| TPC-H | Benchmark Test for Decision Basic Business System Test. | |

Table 1 summarizes four TPC measurement methods out of the total ten methods available for assessing TPC.

2.2 TPC-H Model

TPC-H is an improved testing method over TPC-D, a DBMS performance measurement model for decision support systems, by incorporating Power Tests and Throughput Tests. [6]

Korea's performance measurement testing certifying authority, TTA, and a full member of TPC, also utilizes TTA's DBMS performance BMT. In this paper, the testing is conducted using the Power Test model.

The Power Test utilizes 22 queries, with each query defined as shown in **Table 2** below. [6]

Table 2. TPC-H PowerTest Definition

| Query NO | Execution Sequence | Query Content | Query propensity |
|----------|--------------------|--|------------------|
| 1 | 17 | Price Summary Report | Update |
| 2 | 2 | Selection of Minimum Cost Supplier | Update |
| 3 | 11 | Determining Shipment Prioritization | Insert |
| 4 | 14 | Order Priority Confirmation | Insert |
| 5 | 20 | Domestic Supplier Transaction Volume | Writing |
| 6 | 5 | Revenue Change Prediction | Insert |
| 7 | 21 | Measuring the Value/Quantity of Products | Update |
| 8 | 8 | Domestic Market Share | Insert |
| 9 | 3 | Profit Measurement by Product Type | Update |
| 10 | 18 | Returns Merchandise Reporting | Insert |
| 11 | 15 | Identification of Critical Inventory | Writing |
| 12 | 22 | Shipment Status and Order Priority | Update |
| 13 | 10 | Customer Distribution by Order Quantity | Update |
| 14 | 1 | Effect of Promotions | Update |
| 15 | 16 | Outstanding Suppliers | Writing |
| 16 | 13 | Supplier Relationships | Update |
| 17 | 6 | Analysis of Small-Order Revenue | Update |
| 18 | 7 | Top Customers | Update |
| 19 | 19 | Discounted Sales Revenue | Update |
| 20 | 4 | Promotion Potential | Writing |
| 21 | 9 | Supplier Analysis by Delivery Timeframe | Writing |
| 22 | 12 | Sales Opportunity Analysis | Update |

3. TPC-H Test

3.1 Test Environment

Tests were conducted using commercial DBMS products from companies O and T, and open-source DBMS products from companies P, A, and M, for a total of 5 different products. The tests were conducted under the same environment on a single device, and the specifications of the device are summarized as shown in Table 3.

Table 3. Test Machine Specifications

| Division | Details 1 | Details 2 | Notes |
|----------|--|---------------------------|-------|
| CPU | Intel(R) Xeon(R) Silver 4216 CPU @ 2.10GHz | 16 Cores 32 Threads | |
| MB | ASUS WS C621E SAGE | | |
| RAM | 128 GB | 64GB * 2EA (2400 MT/s) | |
| STORAGE | 1.40 TB | NVMe SSD | |

The virtualization environment for the test equipment was configured with 1 socket and 100GB of memory, and three different types of testing were performed with 8 vCores, 12 vCores, and 16 vCores.

3.2 Test Results

The test results were collected for only 3 out of the 5 products. For A and M products, which operate in a Single Core mode, three queries timed out, and therefore, these results were excluded.

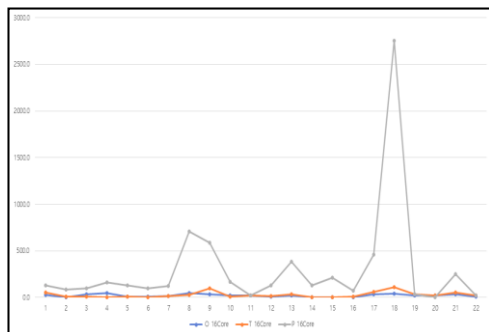


Fig. 1. Power Test Result of 16vCores

Reviewing the power test results, it is evident that two products exhibit similar performance, while one product shows a significant

performance difference in specific queries. This is notably more pronounced in 16vCores compared to 8vCores and 12vCores.

Due to the nature of the job, for queries 3 and 4, which involve frequent updates, O company's product outperformed, while for queries 9, 17, and 18, which involve frequent inserts, T company's product performed better.

4. Conclusions

In this paper, compared the performance of DBMS products currently used in various countries using internationally recognized benchmark tests, aiming to identify which products excel in performance for each query. The analysis of the nature of tasks to be used during the transition to the cloud can serve as foundational data for selecting the most suitable DBMS products for the tasks.

In the future, it will be necessary to conduct these tests in a cloud environment to verify if they yield consistent results with legacy environments. Additionally, further research could focus on deriving conversion factors between commercial and open-source products.

References

- [1] "3rd Cloud Computing Master Plan," *The 14th Information and Communication Strategy Committee*, pp. 1-5, 2021. 09.
- [2] "Government Agency Cloud Business Environment Adoption Guide" *Ministry of the Interior*, pp. 11-13, 2016.11.
- [3] "Practical Guidelines for Cloud Adoption in the Private Sector - "Focusing on Public-Private Cloud," *Korea Communication Commission*, pp. 6-8, 2012.12.
- [4] Jungyup Kim., "Suitability Assessment Guideline for Adopting Cloud Computing in the Public Sector," *Telecommunications Technology Association*, TTAK.KO-10.0799, pp. 13-18, 2016.12.
- [5] Chiha Hwang, Jiwon Yang, "Cloud X Security Practical Guide : Custom Guide for Secure Cloud Use, Design, and Establishing Security Governance" *Freerack*, pp. 24-36, 2020.11.03.
- [6] <https://www.tpc.org/information/benchmark5.asp>

AI Companion for Children: A Service Architecture to aid Social-Emotional Development

Seunghye Han, Soo Jung Chang and Keon Chul Park

Urban Computing & Innovation Lab., Advanced Institute of Convergence Technology, Korea

[e-mail: sjchang1186@snu.ac.kr, rejeena@snu.ac.kr, parkkc07@snu.ac.kr]

*Corresponding author: Keon Chul Park

Abstract

The purpose of this study is to propose a service architecture for AI(Artificial Intelligence) companion technology that supports the healthy socio-emotional development of infants. As AI technology becomes more pervasive in society, AI services that provide various conveniences are appearing in various parts of life. However, despite these advantages, there are also growing concerns about new social side effects such as bias and abuse of technology. In particular, young children are in a period of social and emotional development, and the misuse of technology can harm their emotional health, requiring special attention. This study identifies essential elements to support healthy socio-emotional development of infants in the process of interaction and social exchange and presents a service flow and architecture of AI companion technology for infants based on them. Through interaction with an AI companion implemented based on social characteristics, infants can learn how to interact socially, improve their social-emotional competence, and receive emotional support for healthy development and growth.

Keywords: AI Companion for Young Children, Service Architecture, Social Companion, Social Features

1. Introduction

AI companions for young children provide services such as life monitoring, health care, therapy, education and learning, play and leisure, and provide benefits such as blocking harmful information, reducing media exposure, play-based learning, and providing customized content. Through social interaction with an AI companion, infants can improve their socio-emotional competence and receive emotional support for healthy development and growth. AI companions can also fulfill a variety of roles as social companions for infants. Rogge (2023) categorized the roles of AI companions as

secretaries and servants, friends and partners, caregivers and guardians, teachers and coaches, and children and pets[1], and Rasouli (2022) categorized them as therapists and social mediators, screening and diagnostic and adjunctive therapeutic media, conversation buddies, peers and interactive playmates, and coaches or instructors[2]. In particular, a companion technology that interacts with infants by flexibly recognizing imperfect speech and natural language is expected to enhance the accessibility of AI services for relatively young children.

However, despite these benefits, there are also growing concerns about new social side effects

of technology, such as bias and abuse. In particular, the social-emotional development of infants requires special attention as the misuse of technology can harm their emotional health. This study identifies essential elements to support healthy socio-emotional development of infants in the process of interaction and social exchange, and presents the service flow and architecture of AI companion technology for infants based on these elements.

2. Theoretical Foundation

2.1 A Conceptual Model of AI Companion

An AI companion is a technology that exchanges and effectively processes more complex information in human natural language and conversations to understand the content and intent of the user's utterances and provide appropriate responses and actions to help the user [3].

An AI companion generally consists of a conversation management module, a knowledge learning module, a user modeling module, and an inference engine and knowledge structure [4]. The dialogue management module analyzes the conversation and provides appropriate responses, the knowledge learning module acquires utterances and information not processed by the dialogue management module and generates new responses, and the user modeling module analyzes the intent of the utterance and updates the inference engine to enable user-adapted response inference. The inference engine analyzes utterances to determine user intent, and the knowledge structure stores and manages conversation scenarios and response scripts that correspond to the utterances (Fig. 1).

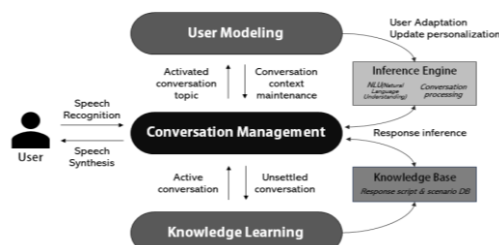


Fig. 1. Basic structure of AI companion

2.2 Social Traits for a AI Companion

Benyon & Mival (2007) categorized the relational characteristics of AI companions as utility, form, personality and trust, and social attitude [5]. Rogge (2023) categorized the social characteristics of AI companions into verbal and non-verbal communication quality, user- and context-adaptive interaction quality, engagement, personality building, and appearance [6]. This study aims to extend the criteria presented in related studies and define the social characteristics of AI companions for young children that support infants' daily life and promote their socio-emotional development based on the aforementioned discussion on the benefits and risks of AI companions for infants as follows (Table 1).

Table 1. Social Features of Artificial Intelligence Companions for Children

| Features | Description |
|---|---|
| Inclusive and ethical conversation technology | Active, inclusive, and ethical conversation techniques considering speech characteristics of young children |
| Child-friendly appearance and interaction | Secure, appealing appearance and user-adaptable, continuous multi-sensory interaction design |
| Emotional response and support | Expressing emotions and providing social-emotional responses to provide bond and mental support |
| Healthy relationship building | Creating a mutually moral and ethical relationship that does not infringe on the freedom and rights of young children |
| Motivation and persuasion techniques | Supporting the positive development and independence of young children by encouraging and inducing them to perform their daily tasks and to acquire social skills |
| Personalization and continuity | Building a two-way learning model that provides continuous and long-term experiences and grows with young children |

3. A Service Architecture of AI Companion for Children

Based on the functional elements and social characteristics of an AI companion defined earlier, this study derived a service architecture for AI companion technology that supports the healthy socio-emotional development of children. The core service elements of AI companion

technology for young children consist of 1) developmental analysis system, 2) developmental support system, and 3) interaction reasoning system.

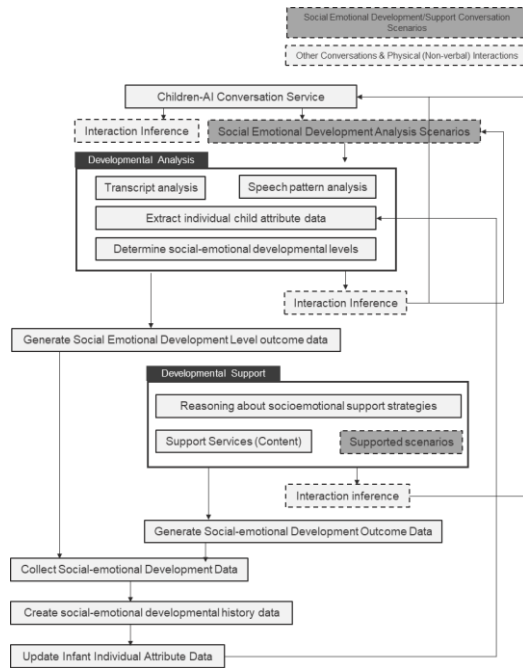


Fig. 2. A Service Architecture of AI Companion for Children

4. Conclusions

This study identifies essential elements to support healthy socio-emotional development of infants in the process of interaction and social exchange, and proposes a service flow and architecture for AI companion technology for young children based on these elements.

A common user perception was found that the companion quickly loses interest as the child grows, resulting in a narrow scope of use and target audience. This is a factor that can lead to poor usability, cost, etc., which can be a barrier to user acceptance. In terms of personalization and continuity, it is necessary to aim for a long-term usage model that grows with the infant through interactive learning. Therefore, an AI companion for young children needs to secure long-term memory of usage information, user modeling, customized content, and dialogue scenarios that respond to the growth of infants to build an adaptive service model that can continue to interact with infants without losing interest.

In addition, AI companions for young children are not expected to completely replace the role of a caregiver or friend, but rather to serve as an excellent mediator, preserving the parental sphere of intervention. The companion will be able to connect infants with parents and teachers through a dedicated application that controls and tracks various information. Personalized and tailored socio-emotional support content can ease the burden of parenting and support the healthy development of children through appropriate advice and supplementary guidance content. Using this research as a foundation, we will continue to conduct follow-up research on the architecture of an AI companion for young children, service development, and preliminary user studies.

References

- [1] A. Rogge, "Defining, Designing and Distinguishing Artificial Companions: A Systematic Literature Review", *International Journal of Social Robotics*, pp.1-23, August 2023.
- [2] S. Rasouli, G. Gupta, E. Nilsen, and K. Dautenhahn, "Potential applications of social robots in robot-assisted interventions for social anxiety", *International Journal of Social Robotics*, Vol. 14, pp.1-32, January 2022.
- [3] J. F. Allen, D. K. Byron, M. Dzikovska, G. Ferguson, L. Galescu, and A. Stent, "Toward conversational human-computer interaction", *AI magazine*, Vol.22, No.4, pp.27-27, December 2001.
- [4] J. H. Hong, "Intelligent Conversation Agent Using Compound Artificial Intelligence Modules", Master's thesis, Yonsei University, Seoul, 2004
- [5] D. Benyon, and O. Mival, "Introducing the Companions project: Intelligent, persistent, personalised interfaces to the internet", *In Proceedings of HCI 2007 The 21st British HCI Group Annual Conference University of Lancaster*, UK, pp.1-2, September 2007.
- [6] A. Rogge, "Defining, Designing and Distinguishing Artificial Companions: A Systematic Literature Review", *International Journal of Social Robotics*, pp.1-23, August 2023.

A study on college students' telemedicine acceptance

Semi Han and Sangmin Lee

School of AI Healthcare, CHA University
Department of Business Administration and Data Science, CHA University
Pocheon, South Korea
[e-mail: mm32@cha.ac.kr, slee@cha.ac.kr]
*Corresponding author: Sangmin Lee

Abstract

This study investigated a previously understudied group of college students who are potential users of telemedicine. In this study, we applied the uses and gratifications theory as a conceptual framework, specifically focusing on the media aspects of telemedicine. Our goal was to delve into the motivations and personal characteristics of college students that influence their acceptance of telemedicine and to uncover the mechanisms through which these factors contribute to a stronger intention to accept telemedicine. To empirically test our research model, we designed an online survey tailored for college students. The survey aims to unveil specific values and benefits that telemedicine should offer to resonate with this demographic. Beyond identifying these values, our study will also shed light on the role of influential individuals and personal characteristics that play a pivotal role in enhancing the intention to embrace telemedicine. The findings from this research endeavor will contribute valuable insights into tailoring telemedicine services to the unique needs and preferences of college students.

Keywords: Telemedicine, Acceptance, College students, Motivation

1. Introduction

The coronavirus disease (COVID-19) has brought many changes to various industries, particularly the healthcare system. For example, the need to take care of patient's health remotely and the demand for tools for care at home became more robust than before [1]. It brings increased use of telemedicine, which diagnoses and treats patients by medical professionals remotely. Accordingly, several studies, including telemedicine definition, medical effects, and policy, have been studied [2]. Based on these excellent earlier studies, we attempt to understand telemedicine from the user's point of view, in particular, college students. As telemedicine has not only medical service characteristics but also media service characteristics, we employ the Uses and Gratifications theory, which is the most powerful

and widely used theory to understand media use behavior. Additionally, we use theories about user characteristics and efficacy that influence the use of healthcare services. Based on these theories, this study aims to explore college students' motivations for telemedicine and the mechanisms involved in how the motivations lead to stronger acceptance intention of telemedicine.

2. Research Framework and Hypotheses

The uses and gratifications theory is one of the most widely used theories in media research to understand how and why individuals use the media. It suggests that media use behavior is goal-directed and purposeful based on users' needs, wants, or expectations. Uses and gratifications research is based on the

fundamental premise that media users are active individuals, and the motivation and behavior of media users must be first understood [3]. This premise is also applicable to telemedicine. In other words, it is crucial to grasp the motivations of potential users using the uses and gratifications theory, especially considering the user-centric nature of telemedicine compared to traditional medical care. Understanding these motivations enhances the likelihood of active acceptance and use of telemedicine. Our research model builds on existing literature, outlining motivations and user characteristics impacting telemedicine acceptance. Subsequently, we investigated how these factors influence college students' intention to accept telemedicine.

2.1 Motivations

It has been proven that when telemedicine is useful and helpful in managing the user's health, the willingness to accept it is high [4]. Telemedicine offers users the advantage of accessing medical services without any time or location constraints. This allows users to conveniently receive medical assistance anytime and anywhere, potentially motivating them to embrace telehealth. [5]. This convenience benefit of telemedicine allows users to receive medical services without visiting medical centers and consequently reduces medical expenses. The benefit of saving on medical expenses would also motivate the willingness to accept telemedicine [6]. In addition, the perceptions and opinions of others about telemedicine influence potential users' willingness to accept it [7]. Others here are those who influence the behavior and opinions of potential users. Given the characteristics of college students with a strong sense of community [8], social influence is highly likely to motivate their telemedicine acceptance behavior.

2.2 User characteristics

Health Technology Self-Efficacy (HTSE) is the self-efficacy toward healthcare technology. Since self-efficacy is situationally specific and varies according to various situations [9], it is necessary to understand specific self-efficacy tailored to the specific domain. In this study, HTSE is a specific self-efficacy in the telemedicine domain. HTSE for telemedicine is

defined as a belief in one's ability to use it without difficulty [10]. HTSE includes not only belief in one's ability to use healthcare technology well but also emotional state in one's health status [11]. In other words, the user's feelings about health affect HTSE. Therefore, this study assumes that users' health status and concerns affect HTSE. Poor health or high health anxiety generally increases the demand for traditional medical services [12]. However, some studies show that individuals with better health or lower health anxiety are more willing to accept healthcare using technology [11]. Additionally, other research illustrates that these health-related factors do not affect the acceptance behavior of healthcare using technology [12]. Despite such contradictory results, the effect of health-related factors on the telemedicine acceptance intention of college students has not been studied. Therefore, it is necessary to examine the effect of health-related factors on it. The hypotheses regarding health status and health anxiety were established based on the literature about traditional medical services.

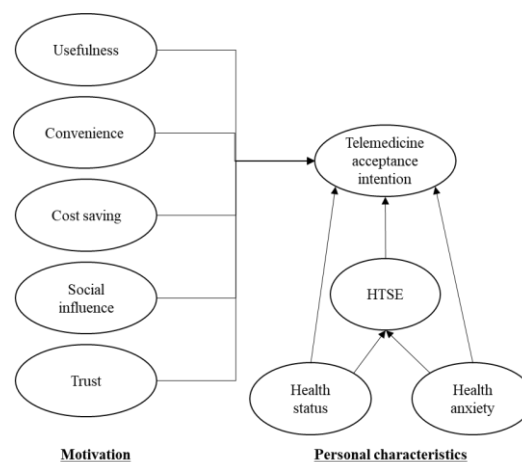


Fig. 1. Research model and hypotheses

3. Material and Methods

The subject of this study is college students with no experience in telemedicine services. We will recruit the respondents from the online community for college students. The measurement items in this study were developed based on the existing scales whose validity and reliability were verified. Some questions have been partially modified to reflect the

characteristics of telemedicine. Usefulness is the perception that telemedicine helps manage users' health more effectively than in-person care, and it was measured by five items [4]. Convenience is the expectation that users can use telemedicine more efficiently without time and spatial limitations. A total of six items were measured for convenience motivation [5]. Cost-saving motivation was measured by three items. Social influence is the degree to which the behaviors and perceptions of acquaintances had on college students' acceptance intentions of telemedicine, and it was measured by five items. Next, the degree of trust in telemedicine providers, such as the accuracy of medical activities and reliable information provided, was measured by four items [7]. HTSE and health anxiety were measured by each of the four items using the existing scales [10]. Subjective health status was measured by evaluating the respondent's health status on a 5-point scale. Telemedicine acceptance intention was measured using four items, including use plan and preference for telemedicine.

4. Conclusions

In this study, we present the values and benefits that telemedicine should provide to increase college students' intention to accept it. In addition, other factors that have a positive influence on the acceptance of telemedicine, such as influential people and personal characteristics, can be suggested. Based on these results, we can propose strategic ways to improve college students' willingness to accept telemedicine.

References

- [1] S. Ajami, M. Mohammadi, "Telemedicine against CoVID-19 crisis," *Health Policy Technology*, Vol 9, No.3, pp. 277-278, 2020.
- [2] W. Zhang, D. He, G. Wang, et al. "Analyzing national telemedicine policies in China from the perspective of policy instrument (1997-2020)," *International Journal of Medical Informatics*, No.166, pp.104854. 2022.
- [3] A. Rubin, "Audience activity and media use," *Communications Monographs*, Vol. 60, No.1, pp. 98-105. 1993.
- [4] S. Kamal, M. Shafiq, P. Kakria, "Investigating acceptance of telemedicine services through an extended technology acceptance model (TAM)," *Technology in Society*, No. 60, pp. 101212. 2020.
- [5] X. Zhang, B. Zaman, "Adoption mechanism of telemedicine in underdeveloped country," *Health Informatics Journal*, Vol. 26, No.2, pp.1088-1103. 2020.
- [6] A. Grustam, J. Severens, J. van Nijnatten, et al. "Cost-effectiveness of telehealth interventions for chronic heart failure patients: a literature review," *International Journal of Technology Assessment in Health Care*, Vol. 30, No.1, pp. 59-68. 2014.
- [7] D. Wu, H. Gu, S. Gu, et al. "Individual motivation and social influence: A study of telemedicine adoption in China based on social cognitive theory," *Health Policy and Technology*, Vol. 10, No.3, pp. 100525. 2021.
- [8] E. Cicognani, C. Pirini, C. Keyes, et al. "Social participation, sense of community and social well being: A study on American, Italian and Iranian college students," *Social Indicators Research*, No.89, pp. 97-112. 2008.
- [9] A. Bandura, "Social foundations of thought and action: A social cognitive theory," Englewood Cliffs, NJ: Prentice-Hall: 1986.
- [10] M. Rahman, M. Ko, J. Warren, et al. "Healthcare Technology Self-Efficacy (HTSE) and its influence on individual attitude: An empirical study," *Computers in Human Behavior*. No.58, pp. 12-24. 2016.
- [11] C. Anderson, R. Agarwal, "The digitization of healthcare: boundary risks, emotion, and consumer willingness to disclose personal health information," *Information Systems Research*, Vol. 22, No.3, pp. 469-490. 2011.
- [12] S. Lee, A. Rothbard, S. Choi, "Effects of comorbid health conditions on healthcare expenditures among people with severe mental illness," *Journal of Mental Health*, Vol. 25, No.4, pp. 291-296. 2016.

A Study On The Access Devices For The Activation Of Web 3.0

Joonyoung Ahn and Dongho Kim

Department of IT Policy and Management, Graduated School of Soongsil University
Seoul, South Korea

[e-mail: corea003@gmail.com, dkim@ssu.ac.kr]

*Corresponding author: Dongho Kim

Abstract

This study aims to investigate individuals' web access environments for the activation of Web 3.0. This study analyzes the issues contributing to the problems for the activation of Web 3.0 and confirms that improving the access environment at the OS level rather than the platform, along with ensuring easy accessibility based on the access devices, is a crucial factor.

As a preliminary study, this research summarizes the process of the evolution from Web 1.0 to Web 3.0, and elaborates on the essential technologies of Web 3.0, including the Metaverse, Blockchain, and LLM (Large Language Model). Additionally, it provides information regarding the current status of OS development by system software development companies. In conclusion, it suggests research directions for the improvement of the accessibility in Web 3.0, focusing on the operating environment and associated technological factors related to access devices.

Keywords: Web3.0, Access Devices, OS, Blockchain, Metaverse, LLM

1. Introduction

Currently, the term 'WEB3.0' has been frequently mentioned post-COVID, but there's been more discussion about it than concrete progress in terms of its precise definition, requisite technologies, and environments needed for its advancement.[1] Other than being defined as an intelligent and personalized web, WEB3.0 remains mostly a list of speculative terms—cryptocurrency, Bitcoin, blockchain, decentralization, ICO, anti-monopoly, data ownership—without much content on specific use cases.[1][2]

Thus, this paper compares Web 1.0, 2.0, and 3.0, investigating siloed technologies that constitute the platform known as the foundation of WEB 3.0.

It analyzes the case of mobile OS that accelerated the technology in WEB 2.0. It suggests the integration of OS and internet access gateways as a means to unify siloed technologies. It also provides recommendations for devices enabling this integration.

2. Related Research

Table 1 summarizes the forms of usage, software, and device changes during the evolution of the WEB. [3]

Table 1. Web Evolution History

| | WEB 1.0 | WEB 2.0 | WEB 3.0 | Notes |
|--------------------|-----------------|----------------|-----------|-------|
| Structure | 1D | 2D | 3D | |
| Form | Potal | Platform | Metaverse | |
| Communication Type | Uni Directional | Bi Directional | N:N | |
| Software | Web | App | Bot | |
| Device | PC | PC, Smartphone | Wearable | |

Intelligent, Customized, Productivity Enhancement, and Ownership Assurance are the elemental technologies of WEB3.0.

2.1 Blockchain

Blockchain is a technology that securely stores

target data in a distributed data storage environment called 'blocks,' formed by countless small-scale data interconnected in a chain-like structure based on a peer-to-peer (P2P) method. This makes it impossible for anyone to arbitrarily modify the data and allows anyone to view the outcome of changes.[4]

In Web 3.0, there has been recognition of usability in terms of proving ownership of generated content and transaction histories.[3]

2.2 Metaverse

Metaverse refers to a network of 3D virtual worlds focusing on social and economic connections. Recently, its scope has expanded to not just encompass simple virtual worlds but also a mix of present and virtual spaces.[5][6]

Web3.0 is drawing attention with its decentralized iterative concept.[6][7]

2.3 AI (Artificial Intelligence)

AI refers to intelligence in machines or software, and the development of machine learning and deep learning since 2012 has enhanced its usability. [8]

Philosophical issues concerning human usability and personality based on machine intelligence have come to the forefront. [8]

2.4 LLM (Large Language Model)

LLM are a type of language model that performs general language understanding and generation. During training, they learn numerous parameters using vast amounts of data and adopt a form of deep learning that consumes computational resources between training and operation.[9]

Recently, there has been an evolution towards acquiring specific knowledge about the inherent syntax, semantics, and existence within human language corpora. [9][10]

3. Suggestions for Siloed Integrated Technologies

The technologies mentioned in the preliminary research have reached a level where commercialization as elements of Web3.0 poses no issue, and some are already in commercial use. However, their current use is in the form of platforms based on Web2.0, making their application in the comprehensive usage of

Web3.0 challenging.[3] Therefore, propose two suggestions to integrate siloed technologies.

3.1 Suggestions for Integrated Software Platform Technologies

The emergence of smartphones significantly contributed to the activation of Web 2.0. While Microsoft's Windows created the best environment on PCs, it didn't generate the acceleration from Web 1.0 to 2.0. Apple's iOS and Safari, alongside Google's Android and Chromium, spurred the acceleration to Web 2.0 through the realization of ubiquity, becoming a catalyst that allows envisioning Web 3.0.[3]



Fig. 1. Integrated Software Environment for Performing Various Tasks

It should present a new Metaverse environment distinct from traditional web and app environments. The currently platformized Metaverse environment should integrate into the operating system, allowing access to the Metaverse environment through the web-exclusive gateway simultaneous with OS booting, enabling various tasks. Subsequently, leveraging AI and LLM for enhanced productivity in content creation for external communication becomes crucial. At this stage, blockchain-based ownership proof for all generated content becomes the focal point.

In essence, it is proposed that the current segregated OS, WEB access gateways, and Metaverse platforms should converge into a singular entity.

3.2 Suggestions for Devices for WEB 3.0 Access

WEB 3.0 second proposition is a device running an integrated OS, web access gateway, and Metaverse platform as one.

The current web operates through PCs, smartphones, and some embedded devices, with a clear limitation in its proposed web access.[11] Therefore, it's suggested that the utilization of WEB 3.0 would be possible when unrestricted

web access is achieved through wearable devices.



Fig. 2. Apple MR Headset and Samsung Smart Ring

Currently, companies like Apple, Google, and Samsung have made significant technological advancements in wearable devices. However, unresolved issues persist, such as battery life, fatigue from extended usage, and display limitations due to miniaturization.

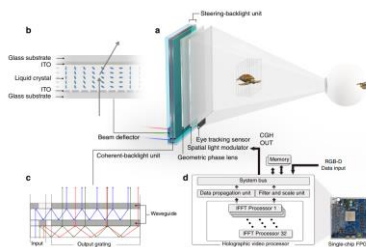


Fig. 3. Slim-panel holographic video display

Recently, Samsung Electronics introduced an advanced technology through the 'Slim-panel holographic video display' paper.[12]

4. Conclusions

This paper provides an analysis of the evolution of WEB3.0, defines the elemental technologies, outlines their applications, and presents two suggestions for activation. WEB3.0 holds a completely different communication method and productivity from the existing WEB1.0 and WEB2.0, becoming the focal point of change by ensuring ownership rights for content and its dynamism.

This suggests the integration of the OS, WEB Access Gateway, and the Metaverse platform, and suggests a device suitable for the integrated environment.

The current technology has found direction in the Software, but the Device section requires significant improvement in elemental technology, which necessitates ongoing research for enhancement.

References

[1] Jihyun Lee, Euna Park, Haeduk Song, "An

Exploratory Study on the Capabilities of E-Learning Course Designers in the Web 3.0" *The Korean Journal of Human Resource Development Quarterly*, Vol.16, No.1, pp. 143-168, 2014.

- [2] Posner, Eric A., and E Glen Weyl, *Radical Markets, "Uprooting Capitalism and Democracy for a Just Society"* Princeton University Press, pp. 43-51, 2017.
- [3] Yongtae Kim, "Web 3.0 Metaverse : The Business Rules Transformed by NFTs and ARGs" *Yeonamsa Publishing Co.*, pp. 50-122, 2022.
- [4] Iansiti, Marco; Lakhani, Karim R. "The Truth About Blockchain" *Harvard Business Review*, 2017.2.18
- [5] Newton, Casey, "Mark Zuckerberg is betting Facebook's future on the metaverse" *The Verge*, 2021.10.25
- [6] Ritterbusch, Georg David; Teichmann, Malte Rolf, "Defining the Metaverse: A Systematic Literature Review" *IEEE Access*. 11: 12368–12377. 2023.2.2.
- [7] Rajan, Amala; Nassiri, Nasser; Akre, Vishwesh; Ravikumar, Rejitha; Nabeel, Amal; Buti, Maryam; Salah, Fatima, "Virtual Reality Gaming Addiction" *2018 Fifth HCT Information Technology Trends (ITT)* ISBN 978-1-5386-7147-4, pp. 358–363. 2018.11.28.
- [8] Russell Stuart J, Norvig, Peter, "Artificial Intelligence: A Modern Approach" *Prentice Hall*, pp. 26–48, 2020.
- [9] Goled, Shraddha, "Self-Supervised Learning Vs Semi-Supervised Learning: How They Differ" *Mystery Vault*, 2021.5.7.
- [10] Manning, Christopher D., "Human Language Understanding & Reasoning" *Daedalus*, pp. 123-138, 2022.
- [11] Svetlana Kim, Man-Soo Chung, YongIk Yoon, "Mobility Process Orchestration for Multi-Device Environment" *KIPS Tr. Comp. and Comm. Sys*, Vol.4, No.3, pp. 91-96, 2015.5.
- [12] Jungkwuen An, Kanghee Won, Young Kim, Jong-Young Hong, Hojung Kim, Yongkyu Kim, Hoon Song, Chilsung Choi, Yunhee Kim, Juwon Seo, Alexander Morozov, Hyunsik Park, Sunghoon Hong, Sungwoo Hwang, Kichul Kim, Hong-Seok Lee, "Slim-panel holographic video display" *Nature Communications*, Vol 11, pp. 1-7, 2020.11.

Development of Elementary Information Education Program based on SDGs

Jaekwoun Shim, Youngkwon Bae and Wooyeol Kim

Dept. of Computer Education, Daegu National University of Education
Daegu, South Korea

[e-mail: jkshim@dnue.ac.kr, bae@dnue.ac.kr, jhon@dnue.ac.kr]

*Corresponding author: Wooyeol Kim

Abstract

According to the OECD's report on developing literacy skills in a digital world, the digital skills of Korean students were analyzed as below average. To improve these problems, the 2022 revised curriculum aims to strengthen basic knowledge and capabilities to actively prepare for future changes. Therefore, basic digital knowledge appeared in the curriculum, and information education was expanded to become the basis for learning all subjects and lifelong learning. In addition, the educational method changed from simple memorization-centered education to self-directed exploratory learning by providing learning linked to the students' lives. In this study, we develop a program to experience information education to improve students' digital capabilities and revitalize information education in the field. Specifically, for this purpose, an information education program linked to the SDGs was developed and applied to fourth grade elementary school students.

Keywords: Elementary, Information Curriculum, Information Education Program, Sustainability Development Goals

1. Introduction

In Korea, social demand for education on artificial intelligence and software is increasing [1]. Additionally, in the information society, everyone is expected to have a minimum level of digital literacy and perform any job or task. From a welfare perspective, it is very meaningful to educate students starting from elementary school so that everyone can have these minimum digital skills. The Ministry of Education and the Ministry of Science and ICT are making various efforts to promote information education for elementary school students. Among them, an educational project was carried out to form a network of students, teachers, parents, local communities, and companies with the purpose of revitalizing information education [2]. The main purpose of the project was to apply information education to school settings, and the main tasks

were educational programs for elementary school students, textbook development, and teacher training. In order to create a good educational program, it is necessary to consider the level of students and conduct education that includes future-oriented values. In addition, in order to achieve in-depth learning linked to students' lives, a shift from an educational method centered on simple memorization to in-depth learning based on exploration and concepts is required.

This study developed an educational program by the Information curriculum of Korea Association Information Education Society with the topics of SDGs. The six types of educational programs developed considered the areas of computing systems, data, algorithms and programming, artificial intelligence, and digital culture. The developed program will be applied to 4th grade elementary school students, and the possibility of

the educational program will be confirmed through interviews.

2. Related research

2.1 Sustainable Development Goals

Sustainable development, adopted by the United Nations in 2015, is a goal to end global poverty, protect the planet, and ensure peace and prosperity for all people by 2030 [3]. The goal was decided through the following conference. Sendai Framework for Disaster Risk Reduction, Addis Ababa Action Agenda on Financing for Development, Paris Agreement on Climate Change and Transforming our world: the 2030 Agenda for Sustainable Development with its 17 SDGs was adopted at the UN Sustainable Development Summit in New York in September 2015. Now, the annual High-level Political Forum on Sustainable Development serves as the central UN platform for the follow-up and review of the SDGs.

With the slogan ‘no one will be left behind’, it sets out 17 goals and 169 detailed directions for humanity in five areas: people, earth, prosperity, peace, and partnership. It is presented as a goal. Each sustainable development goal is closely linked to each other, so actions toward one goal have an organic impact on achieving other goals, and social, economic, and environmental sustainability are balanced through development [4].



Fig. 1. SDGs

2.2 Elementary Information Curriculum

Elementary information education curriculum to improve elementary school students' computational thinking was developed by the Korean Association of Information Education

Society [5]. The nature of information education is to recognize changes in the digital world caused by data and information in a society defined by artificial intelligence, to explore the social value of information, and to develop real-world and This is a subject that cultivates the ability and attitude to solve problems in various academic fields. The goal of information education is to focus on developing the basic skills necessary for rational decision-making and problem solving using computing technology, including artificial intelligence, as digital democratic citizens based on experience solving real-world problems based on computational thinking. The five detailed goals are as follows. (1) Understand computing systems that operate through the organic connection of hardware and software, and develop the ability to use information and communication technologies to generate and transmit data in information handling systems. (2) Understand the meaning and types of data processed by computers, and develop the attitude of using data to solve real-life problems based on experience collecting and visualizing various data. (3) Understand the meaning and importance of algorithms based on procedural thinking, and develop an attitude of discovering and putting into practice the value of cooperation and sharing in the process of designing algorithms to solve problems in everyday life and implementing them in programming languages. (4) Through experience, understand the operating principles of artificial intelligence and its impact on our society, and develop the attitude and ability to explore the artificial intelligence development process based on artificial intelligence ethics. (5) As a digital democratic citizen, develop the ability to understand social changes with digital ethics and explore the connection with career paths. The educational content consisted of five areas: computing systems, data, algorithms and programming, artificial intelligence, and digital culture. The content system consists of core ideas, knowledge and understanding, processes and skills, values and attitudes, covering seven grades from kindergarten to sixth grade, and presents content elements and achievement standards for each grade. For example, the detailed content structure of the artificial intelligence area for 4th grade elementary school students is as follows.

Table. 1. AI area for 4th grade of elementary

| Catagory | Explanation |
|-------------------------|--|
| Core Idea | · Artificial intelligence is a program or system created by imitating human intelligence. · Artificial intelligence affects various areas of life. |
| Knowledge Understanding | · Meaning of artificial intelligence · Principles of Artificial Intelligence |
| Process Function | · Experience artificial intelligence in daily life · Explore the artificial intelligence development process · Attitude to understand the impact of artificial intelligence on society |
| Values Attitude | · Attitude to understand and practice artificial intelligence ethics |

spike, mekimake, etc.) to be covered in the educational program are selected.

Table. 2. Overview of Information Edu-Programs

| Catagory | Education Goal | SGDs |
|-------------------------|---|------------|
| Artificial Intelligence | · Utilizing artificial intelligence image recognition technology | 14, 15 |
| Artificial Intelligence | · Understand and apply machine learning principles | 15 |
| Algorithm | · Solving Problems with Robot Programming | 13, 14 |
| Computing System | · Rescue endangered animals using robots | 13, 14, 15 |
| Computing System | · Understand sensors and take environmental issues quizzes | 13 |
| Computing System | · Creating and sharing promotional media related to nature conservation | 13, 14, 15 |

3. Method

3.1 Research Subject

The study provided camp activities that any 4th grade elementary school student, and approximately 500 students participated.

3.2 Research Process

An information education experience program was developed for the purpose of improving elementary school students' digital capabilities and computational thinking and revitalizing information education. Development was carried out through the following four steps. Selection of sustainable development goals (step 1), Selection of elementary information curriculum areas and achievement standards (step 2), Elementary Information curriculum content elements and platform selection (step 3), Development of information education experience program activities (step 4)

Select one from 17 SDGs and allow elementary school students to experience problem solving using information technology. Select one or two information education achievement standards related to SDG problem solving. Based on the information educational area and achievement standards, content elements and platforms (entry,

Materialize the educational program by planning program topics, goals, and activities based on the designed achievement standards and content elements. At this time, check whether the goals and activities of the education program are related to the SDGs.

4. Result

4.1 Education Programs

Through this study, 6 types of information education programs for fourth grade elementary school students were developed. The developed program was developed so that it could be used during the first period (40 minutes) of elementary school class time.

Table. 3. Example of Information Edu-Programs

| Program | Education Contents | SGDs |
|---------|---|--------|
| 1 | · Learn about animals at risk and find ways to protect them · Understand artificial intelligence and learn image recognition methods · Find the names of endangered animals and learn about crisis situations | 14, 15 |

| Program | Education Contents | SDGs |
|---------|---|------------|
| 4 | using image recognition artificial intelligence. | |
| | ·Share animal protection measures | |
| | · Protecting climate crisis and plants and animals using robots | 14, 15 |
| | · Find problems related to marine debris, climate crisis, and animal extinction | |
| 6 | · Eliminate the causes associated with each problem using a robot | |
| | ·Producing media related to climate crisis and animal protection | 13, 14, 15 |
| | ·Create and share images and videos related to the climate crisis and animal protection | |

4.2 Research Process

The 6 types of programs developed by providing a camp to 4th grade elementary school students (approximately 500 students) were applied.



Fig. 2. Camp for 4th grade elementary students

After application, some students were interviewed. As a result of the interview, they said, “It was good to learn about SDGs issues,” “It was interesting to solve problems with robots,” and “Artificial intelligence technology is helpful in solving SDGs issues.”

4. Conclusions

This study developed an educational program for the purpose of elementary students' digital capabilities and expanding digital literacy and information education in the 2022 revised curriculum. The education program was

produced by the SDGs and the information education curriculum developed by the Korean Association of Information Education Society. There are 6 educational programs developed, which were applied to 4th grade elementary students and received positive evaluation. In the future, there is a need to confirm whether the capabilities related to digital literacy of students who studied the 6 developed programs have improved. It is expected that if various types of information education experience programs are developed in the future, it will contribute to revitalizing information education in elementary schools.

References

- [1] Korea Education and Research Information Service, "Digital education global trends," Vol 20, 2021.
- [2] Ministry of Education, 2022 revised elementary, middle, high school and special education curriculum confirmed and announced, 2022.12.22.
- [3] United Nations, Sustainable Development Goals, 2015.
- [4] United Nations, The Sustainable Development Goals Report Special edition, 2023.
- [5] Korean Association of Information Education Society, Elementary School Information Curriculum, 2023.

The U.S. Pre-service Teachers' Perspectives and Analysis of NCTM Illuminations: Interactives

Yong Joon Park

Department of Teaching and Learning, Indiana State University
Terre Haute, USA
[e-mail: air1027@gmail.com]

Abstract

This study aimed to describe how and what kinds of NCTM Interactives pre-service teachers selected and analyzed. Pre-service teachers of Elementary Education analyzed the math games utilizing both the National Council of Teachers of Mathematics (NCTM) Content and Process Standards and the state standard of Mathematics Education in Indiana, USA. Participants of the study consist of 100 who took a course related to Kindergarten Curriculum. The data of the study, in which case study method was used, was collected with document analysis. As a result of the study, the pre-service teachers selected most of the games focusing on the NCTM Content Standards of Number and Operations and Number Sense and Computation (Indiana Standard), especially for encouraging students to practice counting skills fluently such as “one to one correspondence,” and the relationship between numerals and quantities. Most pre-service teachers think that NCTM Illumination games that they selected would be motivational, effective, and developmentally appropriate.

Keywords: Pre-service Teachers, NCTM Illuminations, Mathematics Education, Web-based Games

1. Introduction

Froebel, the founder of the kindergarten system, stated that young children learn through play and hands-on manipulatives are important. Today's kindergartners also like to use and learn more about the tech-based tools (Park & Yang, 2016; 2018). Vogel et al. (2006) stated that computer gaming and interactive simulation for learning are diverse. Most young children are skillful or confident to use the technology tools and like to play educational games in a classroom or in a computer lab at school (Park & Yang, 2018). Park & Kim (2016) found that even hard of hearing or deaf children could play the game with the proper assistance and language competency such as American Sign Language. Park (2012; 2013) found that the experimental group with a

combination of hands-on manipulatives and computer math games performed better for number and operations than the control group using only hands-on learning materials. Thus, it is important for teacher educators to provide teacher candidates with the proper educational resources of early mathematics and let the teacher candidates be ready to work with Today's kindergartners.

The National Council of Teachers of Mathematics (NCTM) and NCTM Illuminations

The National Council of Teachers of Mathematics (NCTM) stated that NCTM Illuminations (NCTM Illuminations, n.d.) are one of the most popular resources that have been used in a classroom setting (NCTM, 2000). Akcay (2017) stated that NCTM Illuminations

site and the National Library of Virtual Manipulatives (NLVM) site (National Library of Virtual Manipulatives, 2015) were frequently selected by pre-service teachers, especially by elementary education major students in the case study. Akcay found that both NCTM Illuminations and NLVM are the most popular web-based virtual manipulatives tools among pre-service teacher's selection of various technology tools to support mathematical concepts through different representations. However, there are very few articles on how pre-service teachers scientifically select and analyze the NCTM Interactives. Thus, in this case study, the purpose of study is to describe how and what kinds of NCTM Interactives pre-service teachers select and analyze. The main research question is how the teacher candidates think about their selected NCTM math game through their observations and analysis.

2. Methods

Participants and Sites

100 pre-service teachers participated in this study. Most of them are juniors majoring in Elementary Education in a teacher's college, Midwest, USA. 95% of the participants are female and only 5% of them are male. Most of them are Caucasians, one Asian, and one African American student. The sites are (1) a college classroom setting at a university and (2) a Kindergarten classroom setting in a local public school.

Data Collection and Analysis

I chose the case study design because it involves "detailed, in-depth data collection involving multiple sources of information rich in context" (Creswell, 2013, p. 61). Yin (2003) recommends six types of information: documents, archival records, interviews, direct observations, participant observations and physical artifacts. I collected the participants' documents that consist of an observation reflection paper and a rating scale of Math Game Evaluation. It was one of the major assignments of a required elementary

education method course offered each semester. The data was collected to the end of each semester for one Academic year and analyzed. I also kept my field notes when I observed the participants in a kindergarten classroom setting for six weeks. I reviewed, summarized, and analyzed pre-service teachers' documents and my own observations.

Procedures and Instrument

For the data collection procedures, I designed the activity procedures with the template, *Math Game Evaluation*. Here is the procedure of using the Math Game Evaluation template per semester. First, in a college classroom setting, the participants reviewed and discussed the NCTM content and process standards and state standards of Mathematics Education for Kindergartners. Next, the participants explored the instrument, NCTM Illuminations Interactives (i.e., math games), selected one of the math games, analyzed it using the templates such as the rating scale of Math Game Evaluation and Pre-game Analysis. Second, the participants spent three weeks interacting and communicating with Kindergartners in a classroom setting and observed a focal Kindergartener's playing the selected NCTM math game and wrote their observation reflection with the Post-game Analysis. And, lastly, the participants shared their experiences including the math game evaluation and analysis with other participants in a college classroom setting.

3. Findings

There is no significant difference between the pre- and post- math game analysis. The most popular game was "Octa's Rescue" that the preservice teachers selected and studied about from NCTM Illuminations. Many preservice teachers believe that their selected NCTM math game is aligned well with the state and national standards of mathematical contents and process. They think that their selected math game was motivational for their focal child who played the game through their observation and the game was developmentally appropriate for Kindergartners and the effective learning tool for practicing various math skills and knowledge such as counting skills, one-to-one correspondence, and numeral and quantity relationships.

Survey of Math Game Evaluation

Most preservice teachers completed the rating scale and rated their selected NCTM math games highly. The selected math games are as follows:

1. Encourages higher levels of thinking/critical thinking.
2. Stresses state or national math standards.
3. Can be used for math skill building.
4. Active feedback and appropriate assessment are provided.

4. Conclusions

As a result of the study, the pre-service teachers selected most of the games focusing on the NCTM Content Standards of Number and Operations (NCTM, 2000) and Number Sense and Computation (Indiana Department of Education, 2020), especially for encouraging students to practice counting skills fluently such as “one to one correspondence,” and the relationship between numerals and quantities. Most pre-service teachers think that NCTM Illumination games that they selected would be effective and developmentally appropriate for Kindergarteners if a teacher carefully selects the game with the prior assessment and utilizes it in a right moment and place. The participants told me that they seemed like a teacher researcher (Belanger, 1996) for this intervention process such as observing a focal child playing the game at school.

References

- [1] Akay, A. O. (2017). Instructional Technologies and Pre-Service Mathematics Teachers' Selection of Technology. *Journal of Education and Practice*, 8(7): 163-173.
- [2] Belanger, J. (1992). Teacher as researcher: Roles and expectations. *An annotated bibliography*. ED 342 751.
- [3] Creswell, J. W. (2013). *Qualitative inquiry & research design: Choosing among five approaches* (3rd ed.). Thousand Oaks, CA: Sage.
- [4] Indiana Department of Education (IDOE) (2020). Indiana Academic Standards Mathematics: Kindergarten. Retrieved from <https://www.in.gov/doe/files/Kindergarten-Math-Standards-Dec-2020-R.pdf>
- [5] NCTM (2000). *Principles and standards for school mathematics*. Reston, VA: NCTM.
- [6] NCTM Illuminations. (n.d.). NCTM Illuminations Interactives (PreK to the 2nd grade). Retrieved from https://illuminations.nctm.org/Search.aspx?view=search&type=ac&st=n_a_g_m_d&gr=Pre-K-2
- [7] National Library of Virtual Manipulatives. (2015). Virtual Manipulatives. Utah State University. Retrieved from <http://nlvm.usu.edu/en/nav/stieinfo.html>
- [8] Yin, R. K. (2003). *Case study research: Design and methods* (3rd ed.). Thousand Oaks, CA: Sage.

Development of Automated English Speech Evaluation System Using Machine Learning

Woochun Jun

Department of Computer Education, Seoul National University of University
Seoul, Korea
[e-mail: wocjun@snue.ac.kr]

Abstract

The status of English as a global language can be said to be absolute, and no one can deny it. Currently, various private institutions along with public educational institutions are holding many English speech contest. Evaluating many candidates or applicants for a short period of time in English speech or speech contests is very tricky and also difficult to evaluate objectively. In this study, an automated English evaluation system is developed. The proposed system was developed to learn the English oratory test abilities of English experts based on a linear regression method. As a result of applying this system to the English speech of general students, it showed very excellent performance. The results of this work are expected to be very helpful for automating English speech evaluation in the future.

Keywords: English Speech Evaluation, Automated Evaluation, Linear Regression, Machine Learning

1. Introduction

Its influence of English as a global language and official language is absolute, and its influence is expanding day by day. English is becoming a must-learned foreign language, even in countries that are globally also hostile to Britain and the United States. In the context of the "global village" where our world becomes one, English learning in the future is becoming essential worldwide.

The purpose of this work is to develop an automatic scoring system for English speech. Unlike the existing automatic scoring system, the system proposed in this study has the advantage of learning the examination patterns and methods of English reviewers, learning them by machine learning, and then applying them to the examination immediately. This system is expected to be used widely in the future for its

high accuracy, simple, and high implementation.

2. Related Works

Existing studies related to automated English speaking evaluation are as follows.

In the study of [1], they examined the feasibility of an automatic scoring system (ASS) in the context of domestic English-speaking education. They reviewed the scope, test items, evaluation criteria, scoring methods, and reporting methods of six overseas English-speaking tests using ASS.

In the study of [2], an effective method for developing an automated test system for English speaking using an automated evaluation system for English essay writing was proposed. This paper examines the development of recent English essay automation tests based on two

categories: 1) the analysis of automated English essay test systems and 2) the analysis of the scoring process of automated essay scoring tools.

In the study of [3], a tool for automatically scoring Koreans' free speech in English was implemented using a voice recognizer, and its performance was verified using the correlation with the evaluator's manual scoring score.

As we have seen above, it is difficult to automatically score English speech and it is not easy to increase the accuracy of scoring. In this study, a system for learning and automatically scoring the English evaluation of English experts is implemented through machine learning.

3. The Proposed System

3.1 Methodology

In this work, a multilinear regression analysis model was implemented because there were 23 independent variables. Multilinear regression models find variables w_i (weight) and bias(b) that best represent the data in the equation and

$$y = \sum w_i \times x_i + b$$

for given feature x_i and presentation score y

3.2. Implementation

In this paper, the TensorFlow machine learning framework was used to implement the speech score prediction model, and the version is TensorFlow-gpu 1.15. Model learning was conducted on a single machine and is equipped with an AMD 12-core processor, 128GB RAM, and NVIDIA GeForce RTX 2080 Ti (11GB RAM) GPU.

In this paper, we used speech data from 30 high school students and one English professor for learning and evaluating the speech presentation score prediction model. Excluding missing data, 22 data were finally obtained, and for model learning, verification, and performance evaluation, it was randomly separated into 18 training data, two validation data, and two evaluation data.

4. Experimental Results

We evaluated the performance of the proposed speech presentation score prediction model with evaluation data. For reproducibility and statistical verification, two evaluation data were randomly selected and fixed. The correct answers to the published scores of the two selected evaluation data are 75 and 80 points, respectively.

The results of averaging the errors and predictions after learning the model 5 times from the optimal hyperparameter are as follows.

-Average error in verification data: 2.088
 -Test Data Prediction Average: 70.76 points (75 points correct), 77.85 points (80 points correct)
 -Average prediction of verification data: 70.21 points (70 points correct), 66.13 points (65 points correct)

It can be seen that almost similar predictions are made for test data and verification data.

5. Conclusions

In this paper, we have developed an automatic scoring system for English speech based on multilinear regression. In other words, it automatically scores speech recordings by learning the learning patterns and scoring styles of English experts. Although for little data, the system has made fairly accurate predictions. In the future, this system can be used widely in the field of English speech scoring.

References

- [1] Lee, Jinwha, et al. "Analysis of English automated speaking scoring tests", *English Teaching*, Vol. 78, No. 2, pp. 223-244, 2023.
- [2] Lee, Donghan & Ryu, Do-Hyung. "A plan for development of an automated English-speaking testing system using automated English essay evaluation systems", *STEM Journal*, Vol. 14, No. 3, pp. 193-210, 2013.
- [3] Kwon, Ohwook, "Speech recognition-based fluency evaluation for English speaking education", *Ministry of Education*, Report 1345261260, 2018.

A survey on the perception of prospective teachers after learning EPL using the tinkering method

Byoung Rae HAN

Department of Computer Education, Chinju National University of Education
Jinju, South Korea
[e-mail: raehan@cue.ac.kr]

Abstract

Software education is a process of solving problems by directly creating programs. Teachers should provide opportunities for children to recognize problems and various approaches. Tinkering is a useful educational method for creating various solution for problem solving. Prospective teachers do not have experience creating programs, but when they become teachers, they must provide programming education in which children develop software solutions.

This study aims to apply programming education using tinkering educational methods to prospective teachers and investigate their thoughts. As a result of investigating prospective teachers' perceptions of programming education, the following concepts emerged as results: "improvement of children's problem-solving skills, sense of achievement, motivation to learn, various approaches to creating solutions, students' trial and error, and changes in the teacher's role." It was confirmed that the prospective teachers' perception on the tinkering educational method were positively perceived.

Keywords: Software education, Prospective teachers education, Tinkering in programming education, coding education, computational thinking

1. Introduction

Current prospective teachers who must teach programming to elementary school students have no experience with programming in their own education. After graduating from a university of education, prospective teachers are required to implement SW-AI education. However, the number of credits allocated to prospective teachers in university of education is insufficient [1, 2]. While increasing the number of credits in university of education would be a good way for prospective teachers to learn about SW-AI education, it is not easy to revise the University's curriculum.

In SW education, learners continuously produce

software. In the process of learning, students recognize problems and sequentially create tools to solve them. The essence of SW education is to help students learn the computational thinking and use computing power so that students who will live with computers in the future can communicate smoothly with computers.

For prospective teachers to effectively perform SW education in elementary schools, first, they must change their perception. Second, prospective teachers need to learn programming education skills. Third, they need to actively conduct classes when they become in-service teachers.

In this study, I would like to apply program creation education that adopts the tinkering

method to prospective teachers and investigate the perceptions of prospective teachers who have experienced the training program.

2. Theoretical background

2.1 EPL and programming education for prospective

EPL was developed for the purpose of programming education and is useful for beginners in programming, and is used to develop algorithmic thinking and problem-solving skills [3, 4]. Byoung-rae Han et al.(2015) proposed a programming education method for elementary school students applying the tinkering method[5]. Prospective teachers cannot teach programming well by learning the theories of programming education.

It is important for prospective teachers to experience tinkering method programming education. Programming education incorporating tinkering can be effective when using a block-type programming language.

2.2 Tinkering and programming education

Tinkering is a method of using the real thing first before rules or manuals are established. This means touching it first and trying it to operate. Tinkering is a term that means welding, fill in for something, fiddling around, and fiddling, it is also used in fields other than programming [6].

Tinkering can be applied to programming education because the program developing environment has changed, such as the availability of block-based programming languages [5].

Prospective teachers can understand programming education methods that apply tinkering and experience education programs.

3. Tinkering programming education instructional design and progress

The Elementary Computing II (Elementary Computing Education Theory) course taken by prospective teachers consists of 2 credits. Among the 15 weeks of classes, programming education classes were conducted for 4 weeks on

programming education. The education program contents for the 4 weeks are structured as shown in <Table 1> below.

As second-year university students, the prospective teachers had experience of learning the basics of ICT use for one semester in the second semester of their first year. The study was conducted on 55 students in two classes (30 students in class A, 25 students in class B). In weeks 1, 2, and 3, prospective teachers repeated the process of understanding the characteristics of educational programs and creating programs to solve problems. Although the instructor did not explain the characteristics of program commands in detail, instructor provided hints to students who were having difficulty creating program [5]. In the 14th week of the semester, an exhibition was held to present the program they had planned.

Table 1. Programming education plan applying tinkering method

| class schedule | Topics and Activities |
|---|--|
| Week 1) EPL overview, Scratch features, understanding tinkering methods | <ul style="list-style-type: none"> - Sign up for Code.org site, Sign up for Scratch site - Write a program using 2 commands (cat running game) - Program modification and re-game implementation - Peer review of various methods of program modification |
| Week 2) Scratch problem solving project (solve the mission) | <ul style="list-style-type: none"> - Move the cat(sprite) using the arrow keys on the keyboard - Compare different movement methods of sprites according to direction key input - Move the character sprite to the destination point - Check collisions between sprites. Add another monsters. - Upgrade to 2 player game |
| Week 3) Write a Scratch program production plan | <ul style="list-style-type: none"> - Understand how to write a program plan through the process of repeated projects (understanding problems and creating solutions) - Creating a program using the message input/output function (quiz program) |
| Week 14) Scratch production | <ul style="list-style-type: none"> - Holding an exhibition of programs created based on the software production plan |

| | |
|--------------------|--|
| results exhibition | - Provide feedback through peer-to-peer evaluation |
|--------------------|--|

4. Discussion of prospective teachers' responses

I collected the opinions of class participants to find out the thoughts of prospective teachers who experienced classes that applied tinkering-based programming education methods.

To find out about prospective teachers' perceptions of programming education methods that apply tinkering, I asked prospective teachers to answer the following questions.

The prospective teachers' perceptions of the questions are listed based on keywords as shown in <Table 2>.

Table 2. Prospective teachers' perception of programming education using tinkering.

| Class | Prospective teacher's opinion |
|---------|--|
| Class A | Student learning occurs without direct instruction from the teacher. I reconsidered teacher's role It's a great way to learn a new field Provide children with trial and error experiences and cultivate their creative abilities Can increase children's participation in class Creativity is boosted |
| Class B | Children create their own creativity, imagination, and a sense of accomplishment during the production process. Develops children's thinking, problem-solving, and creativity. Develops children's intuition, creativity, and cooperation. A teaching method that fosters children's self-direction ability and creativity Children learn on their own how new command blocks work. |

Through 3 weeks of software production classes and 1 week of software exhibitions, I was able to find that prospective teachers were mainly aware

of project learning, self-directed learning, and creating various solutions.

From the response results, I am able to confirm that the prospective teachers understood teaching and learning that students' activities are the center and teachers' role is promoting students' activities. In addition, many responses were made that prospective teachers had the opportunity to think on their own terms by coding to solve problems.

In addition, positive contents such as creativity, thinking ability, problem-solving ability, cooperation, communication, cooperation, creative solution, sharing with others, and broadening of thinking were presented.

This experience changes prospective teachers' perception, and they can apply it to their classroom when a teacher in the future.

References

- [1] B. R. Han, "Case Study of Prospective Computer Science Teacher Education through Computer Science Assignment Exhibition," *The KAIE Research Journal*, Vol 4, No. 3, pp. 35-40, 2013
- [2] Y. S. Jeong, "Improvement of a Composition of the Computer Education Curriculum for the Universities of Education," *Journal of The Korean Association of information Education*, Vol 16, No. 1, pp. 1-10, 2005.
- [3] S. Papert, *Mind, Storms: Children, Computers, and Powerful Ideas*. New York, 1980.
- [4] S. K. Han et al, "The Development of the Game Addiction Remedy Program based on Scratch Programming," *Journal of The Korean Association of information Education*, Vol 14, No. 1, pp. 61-68, 2010.
- [5] B. R. Han, J. M. Gu, "A Proposal of Teaching Method for Programming with Tinkering in Elementary Class," *The KAIE Research Journal*, Vol 6, No. 3, pp. 53-59, 2015
- [6] Nolgong, *No effort, fun wins, storytree*, 2013

Q-Learning-based Access Priority Control in Internet of Medical Things

Jung-Hyok Kwon¹, Jaehoon Park², and Eui-Jik Kim^{3*}

¹ Smart Computing Laboratory, Hallym University
Chuncheon, South Korea

² School of Nano Convergence Technology, Hallym University
Chuncheon, South Korea

³ Division of Software, Hallym University
Chuncheon, South Korea

[e-mail: jhkwon@hallym.ac.kr, jaehoonpark@hallym.ac.kr, ejkim32@hallym.ac.kr]

*Corresponding author: Eui-Jik Kim

Abstract

This paper presents a Q-Learning-based access priority control, abbreviated QL-APC, in the Internet of Medical Things (IoMT), which aims to ensure urgent data transmission. The QL-APC uses Q-learning to adaptively adjust the contention window (CW) of urgent traffic considering the acknowledgment (ACK) reception. To this end, each sensor device maintains a Q-table dedicated to urgent traffic, which contains the Q-values for different CWs and Actions (i.e., decrease, keep, increase). The Q-values are updated based on the ACK reception, and one of the actions is selected considering the updated Q-value and the current CW. The sensor device determines the CW for urgent traffic using the current CW and the selected action. To verify the superiority of QL-APC, we conducted an experimental simulation. The results showed that the QL-APC obtains better performance compared to the existing IEEE 802.15.6 medium access control (MAC).

Keywords: Access priority control, IEEE 802.15.6, Internet of medical things, Q-learning, urgent traffic

1. Introduction

Recently, as interest in personal healthcare has increased, demand for the Internet of Medical Things (IoMT) has rapidly increased [1]. In IoMT, a hub and sensor devices form a wireless body area network (WBAN) generally based on the IEEE 802.15.6 standard to collect medical data [2]. The standard supports eight traffic classes with different user priorities (UPs) for prioritizing critical and delay-sensitive data. Specifically, for each UP, the minimum and maximum contention windows (CW_{min} and

CW_{max}) are differentiated, thereby providing different access priorities for each traffic class.

However, under the existing standard, urgent traffic may encounter frequent transmission failures and significant delays, particularly when the number of sensor devices increases. This is because it uses a carrier sense multiple access with collision avoidance (CSMA/CA)-based channel access mechanism, where the CW is initialized to CW_{min} on the first transmission attempt and is doubled only if the number of transmission failures is even.

To address the problem, in this paper, we propose a Q-learning-based access priority

control that adaptively selects CW considering the reception of acknowledgment (ACK) to urgent traffic. The detailed design of the proposed scheme is presented in Sect. 2. In Sect. 3, the experiment results are given. Finally, the Sect. 4 concludes this paper.

2. Proposed Scheme

Before transmitting urgent traffic, the sensor device adaptively selects CW considering the result of ACK reception. To this end, each sensor device uses a Q-agent maintaining Q-table and Q-function. **Table 1** shows the Q-table dedicated to urgent traffic (i.e., UP=7). In the table, different Q-values are listed depending on the state and action. The state denotes CW in the range of 1 to 4, and the action includes Decrease, Keep, and Increase.

Table 1. Q-table for urgent traffic (UP=7)

| State (s) | Action (a) | | |
|---------------|----------------|----------|----------|
| | Decrease | Keep | Increase |
| CW=1 | $Q(0,0)$ | $Q(0,1)$ | $Q(0,2)$ |
| CW=2 | $Q(1,0)$ | $Q(1,1)$ | $Q(1,2)$ |
| CW=3 | $Q(2,0)$ | $Q(2,1)$ | $Q(2,2)$ |
| CW=4 | $Q(3,0)$ | $Q(3,1)$ | $Q(3,2)$ |

The Q-values in the table are updated using the Q-function. The update is executed after every transmission regardless of traffic type. The Q-function is given by Eq. (1).

$$Q'(s_t, a_t) \leftarrow (1 - \alpha) \cdot Q(s_t, a_t) + \alpha [r_t + \gamma \cdot \max_a Q(s_{t+1}, a)] \quad (1)$$

where $Q'(s_t, a_t)$ is the updated Q-value for the t -th transmission, α is the learning rate that is determined in the range of $[0, 1]$, the $Q(s_t, a_t)$ is the old Q-value for the t -th transmission, and $\max_a Q(s_{t+1}, a)$ is the estimate of optimal future Q-value. In addition, r_t and γ are the reward and the discount factor for the t -th transmission, respectively. In Eq. (1), r_t is determined based on the result of ACK reception. r_t can be obtained through Eq. (2).

$$r_t = \begin{cases} 1, & \text{ACK} \\ -1, & \text{no ACK} \end{cases} \quad (2)$$

With the updated Q-table, the sensor device selects one of the actions to determine the next state. Specifically, it selects the largest Q-value among the Q-values in the current state.

3. Performance Evaluation

Fig. 1 shows the average number of collisions for urgent traffic when the number of sensor devices changes. In the simulation, each sensor device randomly selects the traffic type. In the figure, the proposed scheme obtains 10.76% fewer number of collisions compared to the existing IEEE 802.15.6, on average. This is because the Q-agent in our proposed scheme finely adjusts the CW taking into account network conditions.

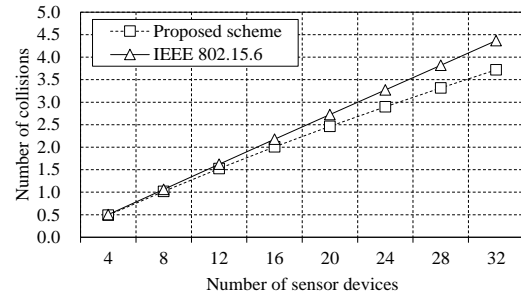


Fig. 1. Number of collisions.

4. Conclusions

In this paper, we propose a Q-learning-based access priority control in IoMT. The proposed scheme reduces the number of collisions of urgent traffic by adaptively adjusting CW. The simulation results showed that our proposed scheme has 10.76% fewer number of collisions than the existing IEEE 802.15.6 standard.

References

- [1] A. Ghubaish, T. Salman, M. Zolanvari, D. Unal, A. Al-Ali, and R. Jain, "Recent Advances in the Internet-of-Medical-Things (IoMT) Systems Security," *IEEE Internet of Things Journal*, Vol. 8, No. 11, pp. 8707–8718, 2021.
- [2] IEEE Standard for Local and Metropolitan Area Networks—Part 15.6: Wireless Body Area Networks.

The Spillover Effects of US Economic Policy Uncertainty on China's Economy: Monetary Policy Uncertainty vs. Fiscal Policy Uncertainty

Chi GONG¹, Shang WU², Hong JIN³ and Xiaoye Lu¹

¹ Department of Global Business, Kyonggi University, Suwon, South Korea

² School of Economics, Sichuan University, Sichuan, China

³ China University of Political Science and Law, Beijing, China

*Corresponding author: Xiaoye Lu

Abstract

In this paper, we examine the shock effects of two types of the US policy-related uncertainty on China's economy, monetary policy uncertainty (Henceforth MPU) and fiscal policy uncertainty (Henceforth FPU). We find that both types of US policy uncertainties will increase the fluctuation in China's economy, and create an aggregate negative effect on the China's economy. The spillover effects of MPU are much stronger than those of the FPU, while the latter persist in much longer periods. In addition, we find that the real exchange rate is more sensitive to the uncertainty shocks of MPU, which is different from the prediction of the Dornbusch theory.

Keywords: monetary policy uncertainty, fiscal policy uncertainty, international spillover effects, China, SVAR

1. Introduction

Friedman (1968), Rodrik (1991), Higgs (1997) and Hassett and Metcalf (1999), among others, consider the detrimental economic effects of monetary, fiscal, and regulatory policy uncertainty. Baker et al. (2011) develop economic policy uncertainty (henceforth EPU) and show that the impact of policy uncertainty on an economy increases the fluctuations in many economy sectors. Compared with other uncertainty indicators of an economy, such as CCI, VIX, VP etc., the EPU index is more accurate and more closely related to others indicators (Bloom, 2013). A large literature shows that exploring the spillover effects of the uncertainty in an economy plays an important role in predicting the economic trend of related

economies and that of global economy, examples including Bernanke (2008), Caballero (2008); Obstfeld & Rogoff (2009), Blanchard & Milesi-Ferretti (2009).

There has been a large literature exploring the international transmission of EPU of an economy. Bernal, Gnabo & Guilmin (2017) examine the impact of economic policy uncertainty on risk spillovers within the core economies of the Eurozone. Colombo (2013) quantifies the possible spillovers going from the US to the Euro area economies and finds a negative and significant reaction of Euro area price and quantity indicators. Ballia et al. (2017) find that countries with higher vulnerability tend to experience stronger shocks from EPUs. Gauvin (2014) examines the extent to which EPUs in advanced countries spill over

to emerging market economies (EMEs) via gross portfolio bond and equity flows. He argues that the impact of EPU on portfolio equity flows differ markedly depending on whether EPUs originate from the United States (US) or from the European Union. Han, Qi & Yin (2016) claim that EPUs of the US economy, compared with those from other economies, appear to have the strongest spillover shocks on the fluctuation of China economy.

China and the US, as the world's largest developing country and the largest developed country respectively, both play an important role in affecting the economic development in many other economies. Since China's economic opening and reforms in 1978, the economic ties between China and the US has become closer and closer. The study on the spillover effects of economic conditions or EPUs of the US economy may deepen our understanding of the history of economic development in China during the past decades of years.

Most studies on the international transmission of economic policies are based on the framework of Mundell-Fleming-Dornbusch Mode or New Open Economy Macroeconomics. They focus on two main international transmission channels, the trade channel and the financial channel (Obstfeld and Rogoff, 1995; Corsetti and Pesenti, 2001; Glick and Hutchison, 1990; Rosa, 2011; Kim, 2001). According to the empirical evidence from many researchers, including Kaminsky and Reinhart (1999), Forbes (2012), and Forbes and Warnock (2012), cross-country trade and finance serve as significant contagion channels for several types of economic crises during the 1980s-2000s.

In this paper, we are going to explore the impact of the US EPU shocks on China's economy. The EPU is classified into the monetary policy uncertainty (henceforth MPU) and the fiscal policy uncertainty (henceforth FPU) in the US economy. We will study how the above two types of EPUs affect China's major economic indicators, through the trade channel and the financial channel. Furthermore, we are going to examine the similarities and differences between the impacts of MPU and FPU.

The rest of this article is organized as follows.

Section 2 briefly reports on data and methodology. Section 3 discusses the empirical results, and Section 4 concludes.

2.Data and Methodology

In this chapter, in order to analyze the spillover effects of US policy uncertainty on China, we introduce the following p^{th} order VAR model:

$$Y_t = C + A_1 Y_{t-1} + A_2 Y_{t-2} + \dots + A_p Y_{t-p} + \varepsilon_t \quad (1)$$

where C is a $k \times 1$ vector of constants, A_i is a time-invariant $k \times k$ matrix and ε_t is a $k \times 1$ vector of error terms. Vector $Y_t = [\text{MPU or FPU}, \text{CPI}, \text{EXE}, \text{CIBOR}, \text{RGDP}, \text{FER}, \text{Con}, \text{Inv}]$, where,

CPI: Consumer Price Index

EXE: Exchange Rates

CIBOR: China Inter-Bank Offered Rate

RGDP: Real GDP

FER: Foreign Exchange Reserves

Con: Consumption

Inv: Investment

The above variables are all important in identifying monetary policy of the central bank. It is clear that the monetary authority regulates the exchange rates through foreign exchange transactions. Moreover, EPUs will lead to changes in the flow of international capital, thus affecting other countries' consumption and investment levels. The changes in consumption and investment can also reveal the structure of an economy, which in turn reflects the overall stability of the economy. Therefore, consumption and investment, two fundamental driving forces of an economy, are also introduced in the model.

In this paper, we use the monthly data from January 2000 to June 2017, which mainly comes from the policy uncertainty website, IFS database and CQER1 data website.

Fig. 1 shows that the degree of MPU and that of FPU are different. Especially after 2008, FPU becomes more unstable, compared with MPU. One of the main focuses of this paper is to examine the differences between the spillover effects from MPU and FPU.

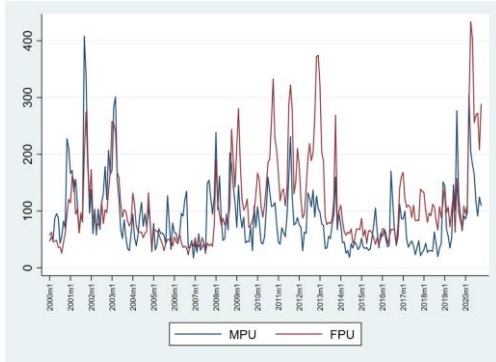


Fig. 1. the trend of US MPU and FPU index.

In this paper, we define the structural effects among the variables in our SVAR model as below.

1. The changes in a country's economic policy will directly affect its domestic economy, then spread to other economies by trade and financial channels after some periods. Therefore, we assume that there is some time lag, at least one period, in the spillover effects of the MPU and FPU of the US on China's Economy.

2. The MPU and FPU of the US can be triggered immediately by changes in both domestic and international economic environments in the current period. Hence, China's economic fluctuation and economic policies will have immediate impacts on the EPUs in the US in current period.

3. It is obvious, by definition, that CPI will affect real GDP and real exchange rate.

4. Foreign reserve will be affected by exchange rate, consumption, and inter-bank offered rate.

5. CPI will be affected by current consumption. However, current consumption is affected by the CPI_{t-1}, real GDP_t and current inter-bank offered rate.

6. Real exchange rate will be affected by CPI, foreign reserve, and inter-bank offered rate at current period.

7. Inter-bank offered rate is affected by current circumstance, such as foreign reserve, consumption and investment. While investment was affected by real GDP and inter-bank offered rate.

Given the above structural effects, we build a SVAR model (AB-model) as below:

$$\begin{bmatrix} 1 & g_{12} & g_{13} & g_{14} & g_{15} & g_{16} & g_{17} & g_{18} \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & g_{32} & 1 & g_{34} & 0 & g_{36} & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & g_{46} & 0 & 0 \\ 0 & g_{52} & 0 & 0 & 1 & g_{56} & 0 & 0 \\ 0 & 0 & g_{63} & g_{64} & 0 & 1 & g_{67} & 0 \\ 0 & 0 & 0 & g_{74} & g_{75} & 0 & 1 & 0 \\ 0 & 0 & 0 & g_{84} & g_{85} & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u_{1t}^{mpu} \\ u_{2t}^{cpi} \\ u_{3t}^{crepo} \\ u_{4t}^{rgdp} \\ u_{5t}^{fer} \\ u_{6t}^{u5t} \\ u_{7t}^{inv} \\ u_{8t}^{inv} \end{bmatrix} = \begin{bmatrix} e_{1t}^{mpu} \\ e_{2t}^{cpi} \\ e_{3t}^{crepo} \\ e_{4t}^{rgdp} \\ e_{5t}^{fer} \\ e_{6t}^{u5t} \\ e_{7t}^{inv} \\ e_{8t}^{inv} \end{bmatrix} \quad (2)$$

3. Results

Fig. 2 shows that both the MPU and the FPU have negative spillover effects on China's economy in the aspects of price and quantity. The impact of the MPU lasts not as long as that of the FPU, but it is stronger and more volatile. On the contrary, the impact of the FPU is relatively gentle and less volatile, its duration is longer.

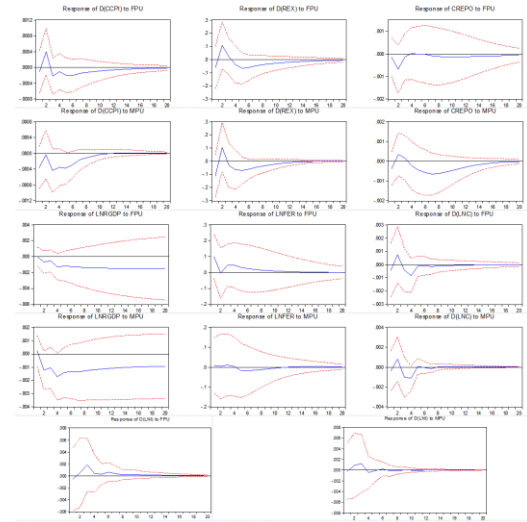


Fig. 2. the results of impulse-response function.

2. Inflation. The above analysis shows that an increase in the MPU or the FPU in the US may serve to suppress the inflation in China, which is consistent with underlying economic theories and transmission channels. According to the Real Options Theory², when the MPU or FPU increases, consumers and companies will delay their consumption and investment, so the aggregate domestic demand will decrease correspondingly. Moreover, foreign direct investment and the demand from other economies will decrease, due to the spillover effects through international economic transmission channels. If the aggregate domestic demand in China remains the same, a decrease in the demand of the US will lead to a drop in the

demand of the export products in this economics. The manufacturers in China will shift their target markets to the domestic, and will adopt various measures, such as price reductions, to reduce their inventory and raise their capital turnover. As a result, the inflation rate in China might decline due to lower price levels.

In the long run, the inhibiting impact on inflation of EPU is weakened according to (a) and (b), which describe the different response of inflation to EPU before and after the crisis. Before the crisis, a decline in inflation lasts longer and the impact of EPU still exists after 15 months. While after the crisis, the inflation increases in the first four periods and the impact of EPU disappear by the large from the 9th period. In addition, the impact of economic policy uncertainty on inflation from restraining to promoting. Since mild inflation will promote economic growth, this may indicate that the market is more confident on China's economy and treat it as a safe haven due to its good performance in 2008.

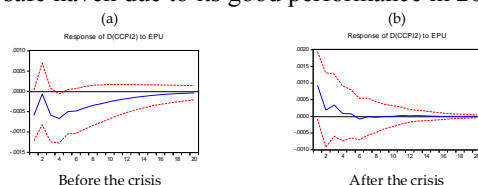


Fig. 3. Pre-crisis and post-crisis response of CPI.

3.Exchange rate. In the short run, the MPU and the FPU in the US may intensify the fluctuation in the real exchange rate between US Dollars and Chinese Yuan. Nonetheless, in the medium and long run, their spillover effects will reduce the fluctuation. The underlying mechanism is that in the short period, following an increase of the MPU in the US, both domestic investors in the US and investors in other markets become more prudent, thus inhibiting both the outflow from and the inflow of capital into the US market, thus reducing the fluctuation of exchange rate between US dollars and Chinese yuan. In the following period, after domestic and foreign investors affirm their evaluation that the MPU has increased, domestic investor will raise their investments in other markets, while at the same time, the foreign investors will divert their capital flow from the US to China. This will intensify the fluctuation in the exchange rate between US Dollars and Chinese Yuan. As we can see from **Fig. 3**, the fluctuation rises in the first two periods, and peaks at the second period.

Then the fluctuation starts to decline and then slowly converged to zero, as the investors finish adjusting the directions of capital flow³.

An increase in the FPU means a rise in the uncertainty of government expenditures in the following periods. Nonetheless, government expenditures in the current period was decided in the previous periods, hence they are not affected by the increase in FPU in current period. Therefore, there is a lag in the impact of the FPU, which means the market is entitled to more time to adjust to the changes in government fiscal policies. Consequently, the impact of an increase in the FPU lasts longer periods, but it leads to less dramatic fluctuations in the exchange rate between US Dollars and Chinese Yuan, compared with the MPU.

This implies that if the Chinese economy is in a downward phase, the uncertainty of U.S. policy will exacerbate the rate of exchange rate devaluation in the short term, but it will help China to stabilize the exchange rate in the medium and long term. This partly resembles the impulse response functions estimated by Clarida and Gali (1994), Eichenbaum and Evans (1995) and Grilli and Roubini (1994).

Compare the response of real exchange rate to EPU before and after the crisis, we can see that exchange rates have moved more widely and for a longer period of time in the short term, but more stable in the long term.

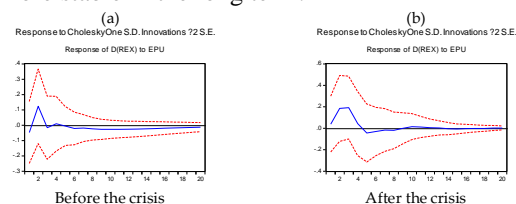


Fig. 4. Pre-crisis and post-crisis response of real exchange rates.

4.Interest rates. Our findings show that both the MPU and the FPU in the US have negative cumulative spillover effects on the real interest rate in China. Nonetheless, as monetary and fiscal policies affect macroeconomic variables through different channels, the macroeconomic variables varies in their response to these two types of policies. Since US Dollars is acting an important role in world economy as the international currency of free settlement, the MPU will make the investors more cautious in the short term, including those from China,

which will cause a temporary liquidity squeeze. The global liquidity crunch will raise the interest rate in the first two periods and begins to fall gradually from the third period as showed in the Fig. 2. However, the FPU will bring uncertainty to aggregate demand of US, and investors tend to shift their investment to abroad, which leads to an increase in liquidity in China, causing the interest rate drop and rebound rapidly in the second period. And it will fall again after reaching the peak in the fifth period. In general, both MPU and FPU will cause the decline of interest rates, but the two fluctuate in different ways in the early stage due to the unlike influence mechanisms.

Before the crisis, the reaction of interest rate to EPU was more like its response to MPU. While after the crisis, interest rates reacted to the EPU more like it did to the FPU and fluctuated more wildly. But the spillover effect of EPU lasts shorter after the crisis as shown in Fig. 5, where interest rate become stable in the 12th period.

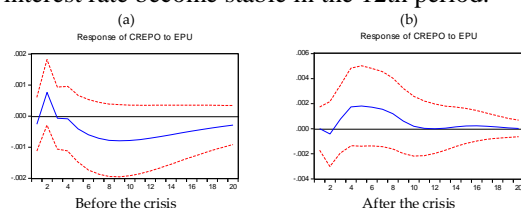


Fig. 5. Pre-crisis and post-crisis response of CREPO.

5. GDP. MPU and FPU have similar effects on output. Considering the role of US economy on global economy, its shocks will pull down China's growth. The result of impulse-response function shows that the reaction of china is similar to the result of "wait-and-see" dynamics(Bloom, 2009). Furthermore, Fig. 2 shows that at the beginning of US economic uncertainty shocks, the fundamentals of China's economy are relatively stable and the lags of the shocks will impel the real GDP slightly increased. But this won't last long, after one period, the response inverses to negative and declined to the lowest point at the 4th period, then slowly returning to a path of balanced growth. The whole process is similar to the process of "overshoot".

From the impact of uncertainty shocks on output, the response of the MPU impulse function is consistent with the "wait and see" reaction pattern in Bloom (2009). It can be seen from the

picture that in the early stages of uncertainty in the US economic policy, China's economic fundamentals are relatively stable, resulting in a slight increase in China's actual GDP. Due to the transmission of uncertainties, the growth only manage to sustain one period before the negative impact takes place. Basically, it falls to the lowest point in the fourth period, then the output shows a slow return to the balanced growth path, which is similar to the phenomenon of "overshooting".

In comparison of Fig. 6. (a) and (b), we can identify that after the crisis, the response of GDP to EPU has decreased markedly, and the affected time has significantly shortened. This may reflect the impact of China carrying out a series of internal economic reforms to strengthen the coping mechanism against external economic shocks and uncertainties after the financial crisis.

6.Foreign exchange reserve. The impact of foreign exchange reserves is similar to the impact of interest rates. In the short term, due to the relative stability of China's economic fundamentals, the uncertainty of U.S. monetary policy will increase China's foreign exchange reserves, but soon this will affect China's overall macroeconomic fundamentals through the transmission effect of China's export-oriented economy, and ultimately lead to a decrease in foreign exchange reserves. Judging from the cumulative effect, the uncertainty of US monetary policy will reduce China's foreign exchange reserves in the medium and long term. The uncertainty of the US fiscal policy is slightly different. In a short term, it is similar to that of the MPU, but this impact will be quickly stabilized, and then gradually converge. Judging from the cumulative effect, the uncertainty of US fiscal policy will increase China's foreign exchange reserves in the medium and long term. Because the uncertainty of fiscal policy will directly affect consumption and investment, it is easy to form a big impact on its own market mechanism. Under this condition, the spillover effect of fiscal policy uncertainty will not be as strong as the uncertainty of monetary policy. Instead, it remains in the United States to a greater extent, and triggers a consistent trend in the overall impact and a different cumulative effect on China's foreign exchange reserves.

Compared with before and after the crisis, we can also observe that the volatility of foreign exchange reserves is also larger, but the duration

of the shock is shorter as well.

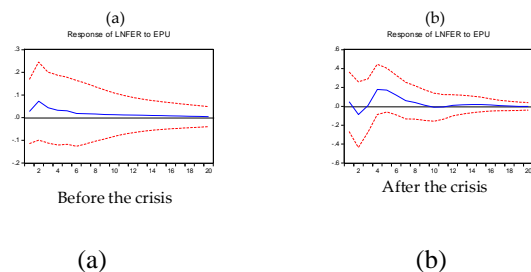


Fig. 7. Pre-crisis and post-crisis response of FER.

4. Conclusions

In this study, the SVAR model was used to investigate the spillover effects of US economic policies on China economy, and the differences in the impact of MPU and FPU were compared. The impulse response function showed us the details. In short, FPU and MPU will increase China's economy fluctuation and have an aggregate negative affects on the Chinese economy, but there are differences in intensity and persistence. Specifically,

- 1.The increase in MPU and FPU will help to suppress inflation in China, which is in line with Real Options Theory.
- 2.MPU and FPU intensify real exchange rate fluctuations in the short-term, but reduce the fluctuations in the medium and long-term.
- 3.Both the MPU and the FPU in the US have negative cumulative spillover effects on the real interest rate in China, but the fluctuations of the two in the previous period were different.
- 4.From the impact of uncertainty shocks on output, the impact of MPU and FPU on output is very similar, and the response of the MPU impulse function is consistent with the "wait and see" reaction pattern in Bloom.
- 5.The impact of MPU increases China's foreign exchange reserves in short run and decreases in long run. The short-term impact of FPU is similar to that of MPU, but the medium and long-term will also increase foreign exchange reserves.

In conclusion, both MPU and FPU will increase the volatility of China's economy and leave with an overall negative impact, suppressing the inflation and interest rate. In terms of scope, the spillover effects of MPU on interest rates, exchange rates and inflation are much stronger than those of the FPU, whereas the latter's

impact lasts longer on consumption, investment and output. Among the indicators examined in this paper, we find that the real exchange rate is the most sensitive to FPU and MPU. And in comparison of the results before and after crisis, China became more sensitive to US economic policy uncertainty, but after a short but sharp period of adjustment, China's economy can reach a new stable state more quickly, since China has become more experienced and capable to deal with external shocks by consciously deepening its domestic economic reform and optimizing the trade structure to disperse the risk.

References

- [1] 1.Ben S. Bernanke, 2007. "Monetary policy under uncertainty," Speech 327, Board of Governors of the Federal Reserve System (U.S.).
- [2] 2.Ciuriak, Dan, Ali Dadkhah, and Jingliang Xiao. 2017. "Brexit Trade Impacts: Alternative Scenarios," GTAP Resource #5252, presented at the GTAP Annual Conference, West Lafayette, 6-8 June, 2017.
- [3] 3.Rosa C. The High-frequency Response of Exchange Rates to Monetary Policy Actions and Statements[J]. Journal of Banking & Finance, 2011, 35(2):478-489.
- [4] 4.Kim S. International Transmission of U.S. Monetary Policy Shocks: Evidence from VAR's [J]. Journal of Monetary Economics, 2001, 48(2):339-372.
- [5] 5.Glick R, Hutchison M. Financial Liberalization in the Pacific Basin: Implications for Real Interest Rate Linkages [J]. Journal of the Japanese & International Economies, 1990, 4(1):36-48.
- [6] 6.Obstfeld, Maurice & Rogoff, Kenneth, 1995. "Exchange Rate Dynamics Redux," Journal of Political Economy, University of Chicago Press, vol. 103(3), pages 624-660, June.
- [7] 7.Corsetti, Giancarlo, and Paolo Pesenti [2001]. "Welfare and Macroeconomic Interdependence." Quarterly Journal of Economics, 116 (2), 421-446.
- [8] 8.Friedman, & Milton. (1968). Role of monetary policy, vol. 58, no. 1.
- [9] 9.Rodrik, D. . (1991). Policy uncertainty and private investment in developing

- countries. *Journal of Development Economics*.
- [10] 10. Kevin, A., Hassett, Gilbert, E., & Metcalf. (1999). Investment with uncertain tax policy: does random tax policy discourage investment. *Economic Journal*.
- [11] 11. Higgs, & Robert. (1997). Regime uncertainty. *Independent Review*.
- [12] 12. Baker, S. R., Bloom, N., & Davis, S. J. . (2015). Measuring economic policy uncertainty. *CEP Discussion Papers*.
- [13] 13. Bloom Nicholas. (2013). Has Economic Policy Uncertainty Slowed Down the World Economy?. *World Development Report*(2014).
- [14] 14. Caballero, R. J., & Engel, E. M. R. A. . (2008). Missing aggregate dynamics: on the slow convergence of lumpy adjustment models. *Cowles Foundation Discussion Papers*.
- [15] 15. Milesiferretti, G., & Blanchard, O. . (2009). Global imbalances. *Economist*, 393(8654), 117.
- [16] 16. Bernal, O., Gnabo, J. Y., & Guilmin, Grégory. (2016). Economic policy uncertainty and risk spillovers in the eurozone. *Journal of International Money & Finance*, 65, 24-45.
- [17] 17. Colombo, V. . (2013). Economic policy uncertainty in the us: does it matter for the euro area?. *Economics Letters*, 121(1), 39-42.
- [18] 18. Gauvin, L., McLoughlin, C., & Reinhardt, D. . (2014). Policy uncertainty spillovers to emerging markets - evidence from capital flows. *Working papers*.
- [19] 19. Han, L., Qi, M., & Yin, L. . (2016). Macroeconomic policy uncertainty shocks on the chinese economy: a gvar analysis: *applied economics*: vol 48, no 51. *Applied Economics*.
- [20] 20. Obstfeld, M., & Rogoff, K. . (2009). Global imbalances and the financial crisis: products of common causes. *CEPR Discussion Papers*, 131-172.
- [21] 21. Reinhart, C. M., & Kaminsky, G. L. . (1999). The twin crises: the causes of banking and balance-of-payments problems. *American Economic Review*, 89.
- [22] 22. Forbes, K. . (2012). The "big c": identifying contagion. *nber working papers*.
- [23] 23. Forbes, K. J., & Warnock, F. E. . (2011). Capital flow waves: surges, stops, flights and retranchment. *SSRN Electronic Journal*, 88(2), 235-251.
- [24] 24. Balli, F., Uddin, G. S., Mudassar, H., & Yoon, S. M. . (2017). Cross-country determinants of economic policy uncertainty spillovers. *Economics Letters*, 156(jul.), 179-183.
- [25] 25. Chun Chang, Kaiji Chen, Daniel F. Waggoner, *Trends and Cycles in China's Macroeconomy*, NBER Macroeconomics Annual 2015, Volume 30
- [26] 26. Nouriel Roubini, Vittorio Grilli, (1995). *Liquidity Models in Open Economies: Theory and Empirical Evidence*, NBER Working Paper No. 5313
- [27] 27. Ludovic Gauvin, Cameron McLoughlin, Dennis Reinhardt. (2014). *Policy Uncertainty Spillovers to Emerging Markets – Evidence from Capital Flows*. Bank of England Working Paper No. 512

A Study on the Improvement of Pet Registration System Using Zero-Knowledge Proof

Joonyoung Ahn, Kijung Ryu, Yunhoo Na and Dongho Kim

Department of IT Policy and Management, Graduated School of Soongsil University
Seoul, South Korea

[e-mail: corea003@gmail.com, yoo123111@gmail.com, kelleya99@gmail.com, dkim@ssu.ac.kr]

*Corresponding author: Dongho Kim

Abstract

This study analyzes the issues related to the distribution, registration, and management of pets in the environment of rapidly growing pet industry in South Korea. The solutions are based on Zero-Knowledge-Proof from the birth of pets through the entire distribution process until their end of life.

A preliminary research studies the current animal registration system and distribution process, and analyze the issues. The ultimate target of the research is to propose alternatives to address those problems using Zero-Knowledge-Proof.

Keywords: Pet Registration, Zero-Knowledge Proof

1. Introduction

During the COVID-19 pandemic, it's estimated that there are 6.02 million households and 1,306 individuals engaged in raising pets in South Korea.[1] The Animal Protection Act was revised in 1991, and in 2014, animal registration became mandatory, leading to computerization for identifying animal identities and locating lost animals. In 2021, there has been effective progress in preventing and penalizing abuse, strengthening penalties against abusive acts.[2] As the number of pet owners increases, so does the number of animals being abandoned or lost. Issues related to inhumane breeding practices and safety concerns, such as aggressive dogs, have emerged. Although pet registration has been mandatory since 2014, the effectiveness has been limited due to lack of information updates, damage to microchips, and external chip loss. Consequently, problems persist, including illegal trading, abandonment, and abuse. Moreover, incidents of personal information leakage of registered pet owners from the system have been reported.[1][3]

Therefore, this paper aims to review the overall issues surrounding the animal registration system, focusing on the entire lifecycle of pets, from production to disposal. It seeks to explore the use of a blockchain technology called Zero-knowledge Proof for clear proof of ownership, transaction verification, and the protection of owners' personal information.

2. Related Research

2.1 Pet Registration System

The Ministry of Agriculture, Food and Rural Affairs has been enforcing a mandatory pet registration system for owners of dogs aged over two months since 2014.[4] This system involves implanting an identification device, either wireless or non-wireless, in the animal's body, which has proven highly effective in identifying and confirming ownership of pets, aiding in finding owners in cases of abandonment or loss.[2][4]

2.2 Distribution Process of Pets

Pets are produced through two methods: specialized breeding and owner-initiated breeding. The distribution of pets occurs as depicted in Fig. 1.[1]

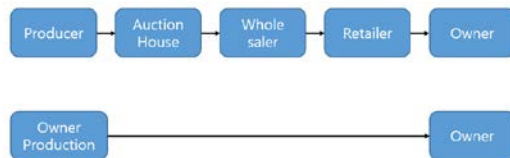


Fig. 1. The Distribution Process of Pets

2.3 BlockChain

Blockchain is a technology that securely stores target data in a distributed data storage environment called 'blocks,' formed by countless small-scale data interconnected in a chain-like structure based on a peer-to-peer (P2P) method. This makes it impossible for anyone to arbitrarily modify the data and allows anyone to view the outcome of changes. Utilizing blockchain enables tracking transaction histories and provides clear proof of ownership.[5]

2.4 Zero-knowledge proof

Zero-knowledge proof is a method by which one party (the prover) can prove to another party (the verifier) that a given statement is true, while avoiding conveying to the verifier any information beyond the mere fact of the statement's truth.[6] The intuition underlying zero-knowledge proofs is that it is trivial to prove the possession of certain information by simply revealing it; the challenge is to prove this possession without revealing the information, or any aspect of it whatsoever.[7] In light of the fact that one should be able to generate a proof of some statement only when in possession of certain secret information connected to the statement, the verifier, even after having become convinced of the statement's truth, should nonetheless remain unable to prove the statement to third parties.[8]

3. Proposal for Improving the Pet Registration System Using Zero-Knowledge-Proof

3.1 Pet Registration System Using Blockchain

Fig. 2 depicts the current pet registration system and the point of registration for pets. The issue lies in the complete absence of ownership history registration from the production of pets to the retail stage.[3]

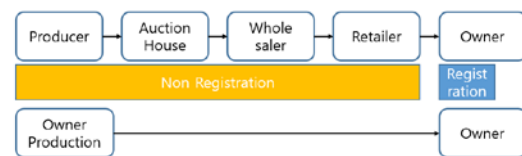


Fig. 2. Pet Registration Point

DNA-based pet registration from Producers or Owners is proposed. Registration details would include information about the producer, the parents of the produced pet, and the DNA of the pet, akin to human birth registration. Subsequently, this registration would facilitate tracking transaction information.

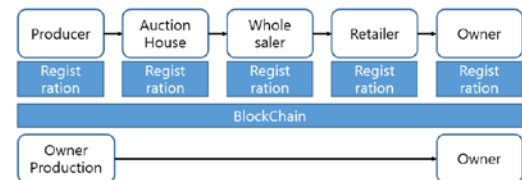


Fig. 3. Proposal for Pet Registration Point

Even auction houses should conduct pet ownership registration using blockchain technology and should only auction pets that have been registered on the blockchain. Subsequently, wholesalers and retailers should also undergo the registration process on the blockchain.

This involves registering owners' information regarding ownership history on the blockchain, enabling privacy protection. It allows for the tracking of pet ownership and transaction history. Particularly, it prevents the trade of illegally produced animals, enabling the tracking of pet lineage and health information.

3.2 Protection and Verification of Owner Information Using Zero-Knowledge Proof

Applying Zero-knowledge proof to pet registration provides a great solution for identifying owners of abandoned or lost pets. The current method of inserting and managing microchips in animals doesn't allow for ownership identification if the microchip is lost or damaged, making it impossible to trace ownership of lost or abandoned pets, especially if a new owner maliciously damages the chip.

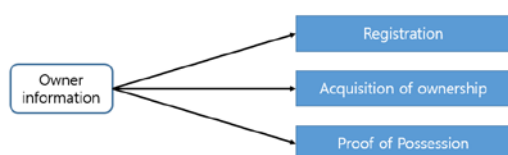


Fig. 4. Primary Use Cases for Ownership Proof

In the usual registration process, personal information needs to be disclosed during registration, and when reclaiming ownership of an abandoned or lost pet, the owner must provide proof of their personal information. In such cases, the exposure of personal information becomes unavoidable. Therefore, protecting the personal information of pet owners becomes another crucial application.

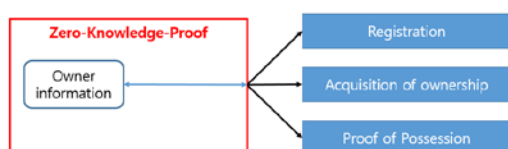


Fig. 5. Protection of Owner Information Using Zero-Knowledge Proof

The use of Zero-Knowledge Proof for pet registration allows completion of the registration process without exposing the owner's information. Moreover, veterinary treatment for pets can be carried out without verifying the owner's information, and changes in ownership or alterations can also be tracked. Additionally, the technology's significant application lies in providing proof of pet ownership without any exposure of the owner's personal information.

4. Conclusions

The pet registration system is an essential legal framework in the rapidly growing domain of pet ownership. However, the current application of this system lacks proactive measures in handling issues such as owner's personal information leaks and situations involving abandonment or theft of pets. This paper proposes the registration of pet production and distribution stages using blockchain technology and further suggests protection using Zero-Knowledge proof for owner information and proof of ownership.

Applying this technology provides a convenient way for proof of ownership and transaction verification. It also offers a solution to counter issues such as hacking in the system.

Subsequent research following this paper could involve developing the application architecture of blockchain systems and executing protection proof using Zero-Knowledge proof to demonstrate improvements in information security.

References

- [1] "Comprehensive Animal Welfare Plan 2020-2024", *The Ministry of Agriculture, Food and Rural Affairs*, pp. 1-5, 2020.
- [2] "Ways to Strengthen Animal Welfare", *The Ministry of Agriculture, Food and Rural Affairs*, pp. 11-13, 2022.
- [3] "2020 Pet Protection and Welfare Survey Results", *The Ministry of Agriculture, Food and Rural Affairs*, pp. 3-27, 2021.
- [4] "Ways to Strengthen Animal Welfare", *The Ministry of Agriculture, Food and Rural Affairs*, pp. 3-4, 2022.
- [5] Iansiti, Marco; Lakhani, Karim R. "The Truth About Blockchain" *Harvard Business Review*, 2017.2.18.
- [6] Valentin Mulder, Alain Mermoud, Vincent Lenders, Bernhard Tellenbach, "Zero-Knowledge Proof" *Trends in Data Protection and Encryption Technologies*, Cham: Springer Nature Switzerland, pp. 25-30.
- [7] Goldreich, Oded, "Foundations of Cryptography" *Cambridge University Press*. pp. 184. 2021.
- [8] Goldreich, Oded, "Foundations of Cryptography" *Cambridge University Press*. pp. 247. 2021.

Thin Graph Convolution Network in Recommendation Systems

Sang-min Lee, Kim Heechan, Jeon Janggun and Namgi Kim

Department of Computer Science, Kyonggi University
Suwon, South Korea

[e-mail: d9249@kyonggi.ac.kr, train1212@kyonggi.ac.kr, 201812285@kyonggi.ac.kr,
ngkim@kyonggi.ac.kr]

*Corresponding author: Namgi Kim

Abstract

Recommender systems help users quickly find the content they are looking for on the Internet, which is overflowing with information. A popular algorithm for efficiently implementing such RISs is the graph convolutional network (GCN) algorithm. In this work, we improve the structure of LightGCN, which is often used as a backbone network in research on GCN-based recommender systems, to show better recommendation performance, and demonstrate through experiments that the proposed ThinGCN model with WF methodology has improved accuracy over the existing LightGCN model.

Keywords: Graph Convolution Network, learning method, Recommendation Systems

1. Introduction

Recommendation systems are a key tool in providing users with the information and content they want in a world of information overload. Recommendation systems analyze a user's past behavior, preferences, and interests to effectively present the most relevant information or products. Recent studies on these recommendation systems utilize convolutional networks (GCNs) to improve their performance, and in subsequent studies, the LightGCN model is often used as a backbone network, and improvements to the LightGCN model can advance the overall recommendation system research methodology. Therefore, in this paper, we propose an improvement method for combining layers of the LightGCN model.

2. Proposal methods

2.1 LightGCN Layer Combination

Equation 1 is the propagation rule for LightGCN, where \mathcal{N}_i and \mathcal{N}_u are the connected nodes of user u and item i , respectively. Then, a symmetric normalization term ($=1/\sqrt{|\mathcal{N}_u|}\sqrt{|\mathcal{N}_i|}$) is performed to obtain the next embedding layer, and a weighted sum of Equation 2 is performed to obtain the final user-item embedding, and the recommendation score is calculated using the inner product of these embeddings.

$$e_u^{(k+1)} = \sum_{i \in \mathcal{N}_u} \frac{1}{\sqrt{|\mathcal{N}_u|}\sqrt{|\mathcal{N}_i|}} e_i^{(k)};$$
$$e_i^{(k+1)} = \sum_{u \in \mathcal{N}_i} \frac{1}{\sqrt{|\mathcal{N}_i|}\sqrt{|\mathcal{N}_u|}} e_u^{(k)} \quad (1)$$

$$e_u = \sum_{k=0}^n \alpha_k e_u^{(k)}; e_i = \sum_{k=0}^n \alpha_k e_i^{(k)} \quad (2)$$

However, this propagation rule captures the loss of embedding values, and to overcome this, the Weight Forwarding(WF) methodology and the

weighted summation of embeddings used by it can be omitted.

2.2 Proposed ThinGCN

The proposed WF methodology scalarly multiplies the previous embedding by its weight and passes it to the next layer, as shown in Equation 3. In the process, the information from the previous layer is more strongly conveyed, allowing us to make recommendations using only the last layer's embedding, as shown in Equation 4.

$$e_u^{(k+1)} = w_k \sum_{i \in \mathcal{N}_u} \frac{1}{\sqrt{|\mathcal{N}_u|} \sqrt{|\mathcal{N}_i|}} e_i^{(k)};$$

$$e_i^{(k+1)} = w_k \sum_{u \in \mathcal{N}_i} \frac{1}{\sqrt{|\mathcal{N}_i|} \sqrt{|\mathcal{N}_u|}} e_u^{(k)} \quad (3)$$

$$e'_u = e_u^{(k)}; e'_i = e_i^{(k)} \quad (4)$$

3. Experiments

As experimental data, we used Yelp2018, Douban-book, and MovieLens-1M, which are frequently used in existing research. The composition of the dataset is shown in **Table 1**, and the experimental results of the existing method and the proposed initialization method are shown in **Table 2**.

Table 1. Experiments data

| | Yelp2018 | |
|-------------|--------------|-----------|
| | Train | Train |
| User | 31,668 | 31,668 |
| Item | 38,048 | 38,048 |
| Interaction | 1,237,259 | 1,237,259 |
| Sparsity | 99.90% | 99.90% |
| | Douban-book | |
| | Train | Train |
| User | 12,638 | 10,882 |
| Item | 22,222 | 19,075 |
| Interaction | 478,730 | 119,690 |
| Sparsity | 99.83% | 99.94% |
| | MovieLens-1M | |
| | Train | Train |
| User | 6,038 | 5,989 |
| Item | 3,492 | 3,190 |
| Interaction | 460,359 | 114,922 |
| Sparsity | 97.82% | 99.40% |

Table 2. Experiments result

| | Yelp2018 | | |
|----------|--------------|---------|---------|
| | Weight | recall | NDCG |
| LightGCN | - | 0.05949 | 0.04887 |
| ThinGCN | 8 | 0.06873 | 0.5656 |
| | Douban-book | | |
| | Weight | Recall | NDCG |
| LightGCN | - | 0.14783 | 0.12508 |
| ThinGCN | 16 | 0.16701 | 0.14632 |
| | MovieLens-1M | | |
| | Weight | Recall | NDCG |
| LightGCN | - | 0.27196 | 0.30274 |
| ThinGCN | 128 | 0.27748 | 0.30830 |

4. Conclusions

This paper, we proposed a ThinGCN model using WF methodology, and experimental results show that both recall and NDCG show performance improvement, which proves the superiority of ThinGCN utilizing only the last layer.

Acknowledgments

This research was supported by the National Research Foundation of Korea (NRF) Basic Science Research Program funded by the Ministry of Education (No. 2020R1A6A1A03040583) and this research was supported by the MSIT(Ministry of Science and ICT), Korea, under the National Program for Excellence in SW supervised by the IITP(Institute of Information & communications Technology Planning & Evaluation) (2021-0-01393) and this work was supported by Korea Institute for Advancement of Technology(KIAT) grant funded by the Korea Government(MOTIE) (P0008691, HRD Program for Industrial Innovation)

References

- [1] X. He, K. Deng, X. Wang, Y. Li, Y. Zhang *et al.*, "LightGCN: Simplifying and Powering Graph Convolution Network for Recommendation," in *Proc. Special Interest Group in Information Retrieval (SIGIR)*, China, pp. 639–648, 2020.

A Study On Conversion Vector Between Suitable DBMS Operated As A Cloud Service

Joonyoung Ahn, Changik Oh, Sugil Hwang and Dongho Kim

Department of IT Policy and Management, Graduated School of Soongsil University
Seoul, South Korea

[e-mail: corea003@gmail.com, lowerright@gmail.com, soogool@gmail.com, dkim@ssu.ac.kr]

*Corresponding author: Dongho Kim

Abstract

This study is conducted as a preliminary research to reduce the system load when operating in cloud systems after transitioning from the existing legacy systems. While the initially configured cloud systems can instantly respond to sudden loads thanks to its Auto-Scaling, this responsiveness may lead to situations such as session interruptions and service disruptions during the scaling process. However, by understanding the differences between the existing system and the transitioning system and applying the conversion vector, it is possible to prepare for such situations. The purpose of this paper is to execute TPC-H on DBMS with internationally established test specifications to calculate conversion vector. Summarizing the Power Test of TPC-H, conversion vector are calculated after conducting tests on one commercial product and one open-source product. Following that, a second round of TPC-H testing is conducted in the cloud, and the validation of conversion vector is carried out.

Keywords: DBMS, Cloud Service, TPC-H, Conversion Vector

1. Introduction

Currently, in South Korea, starting with the opening of the National Information Resources Management Institute's Daegu Center in December 2022, there is a plan for a comprehensive migration of public sector systems to private public clouds by 2025.[1] However, the preparation is insufficient, and confusion is expected as a result. In particular, in the case of DBMS, there is a lack of standardized conversion coefficients.[2][3] Especially in the nature of the cloud, while stabilizing through Auto Scaling in cases of system overload is possible to an extent, through this research, the aim is to describe the research content with the purpose of early stabilization without subjecting the system to various risk factors until a more precise DBMS sizing is achieved for stabilization.[3]

2. Related Research

The Transaction Processing Performance

Council (TPC) was established in the United States in August 1988 as a committee that sets standards for measuring transaction processing performance. TPC holds a total of 18 measurement methods, with 10 of them currently in active use.[4]

2.1 Active TPC Benchmarks Types

Table 1. Active TPC Benchmarks Types

| BenchMarks | Fields | Notes |
|------------|---|-------|
| TPC-C | OLTP-Based Basic Test | |
| TPC-DS | Industry Standard Benchmark Test for Measuring the Performance of Decision Support Solutions. | |
| TPC-E | TPC-C and Financial Market Interaction Test. | |
| TPC-H | Benchmark Test for Decision Basic Business System Test. | |

Table 1 summarizes four TPC measurement methods out of the total ten methods available for assessing TPC.

2.2. TPC-H Throughput Numerical Quantity

The results of the throughput test are used to compute TPC-H Throughput at the chosen database size. It is defined as the ratio of the total number of queries executed over the length of the measurement interval, and must be computed as:[4]

$$\text{TPC-H Throughput@Size} = \frac{(S * 22 * 3600)}{T_s * SF} \quad (1)[4]$$

TPC-H Throughput@Size are Queries per hour * Scale-Factor, reported to one digit after the decimal point, rounded to the nearest 0.1.[4]

3. TPC-H Test

3.1 Test Environment

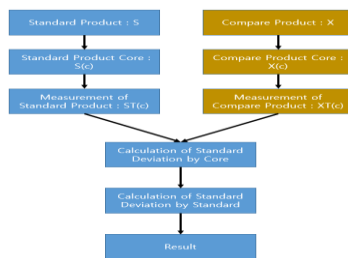


Fig. 1. Test Flow of conversion vector between suitable DBMS

Tests were conducted using commercial DBMS products from companies O and T, and open-source DBMS products from P, A, and M, for a total of 5 different products. The tests were conducted under the same environment on a single device.

Fig. 1 represents the flow for calculating the Conversion Vector for different products, based on the reference product, which is O company's product.

3.2 Test Results

The test results cover only 3 out of 5 products. Among them, A and M products, operating on a Single Core system, encountered Time Outs in 3 queries, leading to their exclusion from the results.

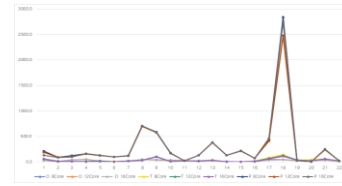


Fig. 2. Test Results

Queries 8, 9, 17, and 18 of product P exhibited abnormal results, consequently, they were excluded from the calculation and comparison during the assessment.

The calculated Conversion vector excluding the 4 queries showing peak values in the test results is as shown in Table 2.

Table 2. Conversion Vector

| Division | O Product | T Product | P Product | Notes |
|----------|-----------|-----------|-----------|-------|
| 8 vCore | 1 | 1.023903 | 6.154799 | |
| 12 vCore | 1 | 1.122751 | 6.527803 | |
| 16 vCore | 1 | 1.051428 | 7.26987 | |

4. Conclusions

This paper conducted a foundational study for cloud migration, focusing on the research of interconversion coefficients for heterogeneous DBMS. It was possible to measure the interconversion coefficients between different systems and identify factors in the measured values that could prevent issues during conversion in advance.

References

- [1] "3rd Cloud Computing Master Plan," *The 14th Information and Communication Strategy Committee*, pp. 3-11, 2021. 09.
- [2] "Practical Guidelines for Cloud Adoption in the Private Sector - "Focusing on Public-Private Cloud," Korea Communication Commission, pp. 23-29, 2012.12.
- [3] Jungyup Kim., "Suitability Assessment Guideline for Adopting Cloud Computing in the Public Sector," *Telecommunications Technology*
- [4] <https://www.tpc.org/information/benchmarks5.asp>

Egocentric Human Pose Estimation for VR Interactions Using RGB-D Data

Seongmin Baek¹ and Yejin Kim²

¹ Intelligent Realistic Content Lab, Electronics and Telecommunications Research Institute
Daejeon, Republic of Korea
[e-mail: baeksm@etri.re.kr]

² School of Games, Hongik University
Sejong, Republic of Korea
[e-mail: yejkim@hongik.ac.kr]

*Corresponding author: Yejin Kim

Abstract

Recently, virtual reality (VR) technology has evolved into a metaverse where individuals share information about their real-world activities through head mounted displays (HMDs). For smooth interaction between users, the position and rotation information of the hands and head of a user are recognized by the camera sensor and controller of an HMD while the rest of the body is inferred from the inverse kinematics. However, this can lead to inaccurate pose estimation and reduce the user's immersion in the virtual space. In this paper, we introduce a method to estimate user pose from egocentric (inside-in) data acquired from multiple RGB-D sensors installed on an HMD. In the proposed method, we calibrated multiple RGB images and depth data to the same coordinate system and estimated user pose in 3D using a deep neural network (DNN). Compared to external sensor-based systems using outside-in data, our method generates accurate user poses in 3D and provides better freedom of movement for VR interactions.

Keywords: Human pose estimation, VR interaction, RGB-D sensor, viewpoint transformation, DNN

1. Introduction

Recently, virtual reality (VR) technology has evolved into a metaverse where individuals share information about their real-world activities using head-mounted displays (HMDs). For smooth interaction between users, the position and rotation information of the hands and head of a user are recognized by the camera sensor and controller of an HMD while the rest of the body is inferred from the inverse kinematics [1]. However, this type of camera-based acquisition

is difficult to recognize the user's lower body. It requires an external tracker applied to the feet [2], or simulates the lower body in accordance with the upper body motion [3]. Moreover, the user's lack of experience in sense of ownership (SoO) or sense of agency (SoA) can reduce immersion in the virtual space [4].

In this paper, we introduce a method to estimate user pose from egocentric (inside-in) data acquired from multiple RGB-D sensors installed on an HMD. In the proposed method, we calibrated multiple RGB images and depth data to the same coordinate system and estimated user

pose in 3D using a deep neural network (DNN). Compared to external sensor-based systems using outside-in data, our method generates accurate user poses in 3D and provides better freedom of movement for VR interactions.

2. System Overview

2.1 Configuration of RGB-D Sensors

Commercial RGB-D sensors have a narrow field of view since they are designed to be installed externally, which makes it difficult to see the user's entire body when the sensor is placed on an HMD. For this problem, we attached multiple RGB-D sensors to the frame and mounted them on the HMD as shown in Figure 1. The main sensor was installed in a horizontal orientation to view the upper torso while the left and right sub-sensors were installed in a vertical orientation to view the left and right lower torso, respectively.

2.2 Sensor Calibration

As shown in Fig. 1, RGB and depth data from each sensor are represented in different coordinate systems. We unified them into one coordinate system through coordinate system conversion, which takes at least 3 images from different positions by using the checkerboard in the overlapping part of the main sensor and sub-sensor images. A total of $3N$, where N is the number of extracted images, positions in 3D were estimated from each image. The rotation (\mathbf{R}) and translation (\mathbf{T}) matrices were calculated using the iterative closest point (ICP) algorithm, respectively. By rotating and translating the incoming sub-sensor data to the main sensor, they are aligned in the same coordinate system as shown in Fig. 1.



Fig. 1. RGB-D sensors on the HMD (left) and their calibrated result (right).

2.3 Pose Estimation Using DNN

Using the egocentric (inside-in) viewpoint, the lower parts of the body (i.e., wrist and ankle joints) are difficult to track owing to the occlusion problem. For a full body motion, we applied a multilayered perceptron (MLP) method to find the hand and foot positions from the visible joint data. We used motion capture data as the training set which consist of walking, hand, and foot waving motion (i.e., a total of 40,000 frames). Since the motion data varies depending on the user's height, we converted the data into joint vectors which were centered on a local axis of the pelvis joint. For MLP, we used 3 layers with a total of 11 joint vectors as an output pose, which were multiplied by the inverse matrix for the local axis and fed back into MLP as input.

3. Experimental Results

For RGB-D sensors, we used the Intel RealSense D435i model [5] owing to its size, weight, angle of view, and minimum recognition distance. Fig. 2 shows the pose estimation process which generates the skeletal joint positions using DNN and articulated ICP methods. Since the foot and hand joints were not clearly represented in the reconstructed skeleton, we used the data reconstruction method using MLP. We compared the estimated results to an external sensor-based system that uses multiple MS Azure Kinects [6]. Comparing a total of 3,000 frames of various user poses, we achieved less than 5% in the positional differences of joint positions.

4. Conclusions

In this study, we introduced a 3D pose estimation method that uses multiple RGB-D sensors on an HMD. The proposed method generates a wide-angle view, transforms the viewpoint, and applies a deep learning framework for the pose estimation. To restore the occluded lower body data, we introduced an DNN-based method that uses MLP. Compared to the previous system using multiple external sensors, the proposed method achieved similar accuracy in 3D pose estimation and provided better flexibility in freedom of movement for VR interactions.

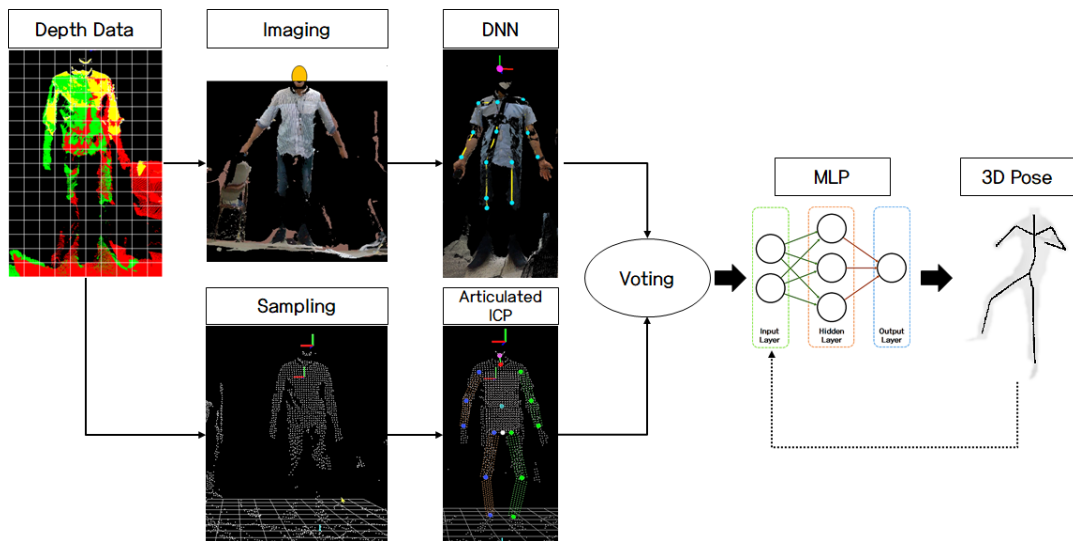


Fig. 2. 3D pose estimation using DNN.

Currently, we are improving the configuration of the sensors attached on an HMD for more comfortable wear and optimization of the estimation process for performance. In addition, we plan to apply the proposed method to the user interactions in a metaverse and analyze the user experiences by conducting user tests.

References

- [1] Demotion. Available online: <https://blog.deepmotion.com> (accessed on 30 October 2023).
- [2] S. Merker, S. Pastel, D. Bürger, A. Schwadtke, and K. Witte, "Measurement Accuracy of the HTC VIVE Tracker 3.0 Compared to Vicon System for Generating Valid Positional Feedback in Virtual Reality", *Sensors*, Vol 23, No 17, 7371, 2023.
- [3] A. Winkler, J. Won, and Y. Ye, "QuestSim: Human Motion Tracking from Sparse Sensors with Simulated Avatars," in *Proc. of SIGGRAPH Asia*, 2022.
- [4] N. Braun, S. Debener, N. Spsychala, E. Bongartz, P. Sörös, H.H.O. Müller, and A. Philipsen, "The Senses of Agency and Ownership: A Review," *Front. Psychol.*, Vol 9, 2018.
- [5] Intel RealSense. Available online: <https://www.intelrealsense.com> (accessed on 30 October 2023).
- [6] MS Azure Kinects. Available online: <https://azure.microsoft.com> (accessed on 30 October 2023).

Over-smoothing in Light GCN doesn't happen much

Sangmin Lee, Heechan Kim, Janggun Jeon and Namgi Kim

Department of Computer Science, Kyonggi University
Suwon, South Korea

[e-mail: d9249@kyonggi.ac.kr, train1312@kyonggi.ac.kr, 201812285@kyonggi.ac.kr,
ngkim@kyonggi.ac.kr]

*Corresponding author: Namgi Kim

Abstract

Recommendation systems help people make choices on the Internet, where a large amount of content and information is generated on various platforms. To efficiently implement such a recommendation system, Graph Convolution Network (GCN) is a widely used algorithm. In this paper, we evaluated the performance of LightGCN, one of the GCN models, by increasing the number of layers from 1 to 10. In our experiments, we found that LightGCN suffers from over-smoothing when the number of layers exceed 3, but in reality, it does not have a significant impact and can be stacked more than 3 layers.

Keywords: Graph Neural Network, Graph Convolution Network, Recommendation System, Deep Learning

1. Introduction

There's something called the paradox of choice. It refers to the phenomenon that when people are given too many options, it makes it harder for them to make a decision and they end up making a poorer choice than if they had fewer options, or even giving up on the choice altogether. This is true on the internet today, where there's a lot of content and information available on a variety of platforms. Therefore, there is a need for a recommendation system that filters and delivers the content or information that users want. These recommendation systems are a type of information filtering technology that recommends suitable content or information to users based on their preferences and past behavior. Many companies, such as Netflix, Amazon, and YouTube, are using recommendation systems to reduce user frustration and increase sales. Recent studies on these recommendation systems utilize GCN to improve their performance, and subsequent

studies often use the LightGCN model[1], and improvement of the LightGCN model[1] is equivalent to improvement of the recommendation system method. Therefore, in this paper, we actually evaluated the performance of LightGCN model to improve LightGCN model. Specifically, we experimented to see if LightGCN actually suffers from over-smoothing when stacking more than 3-layers.

2. Related Work

2.1 LightGCN[1]

LightGCN can only learn the embedding parameters of the first layer. It then combines the representations from the k-layers into a simple weighted sum. In the LightGCN paper, they said that as the number of layers increases, the over-smoothing problem occurs, so they use layer combinations. Therefore, in this paper, we increased the number of layers to 10 and

performed a performance evaluation to see if the over-smoothing problem was significant in practice.

3. Results

The experimental datasets use commonly used Last.fm[2] Datasets. We also experimented with stacking up to 10 layers of LightGCN to see how performance changes, as shown in **Table 1**.

Table 1. Experiments data

| Last.fm | | | |
|--------------|-------|--------|--------|
| | Epoch | recall | NDCG |
| LightGCN_L1 | 500 | 0.0517 | 0.0357 |
| LightGCN_L2 | 500 | 0.0451 | 0.0348 |
| LightGCN_L3 | 500 | 0.0992 | 0.0812 |
| LightGCN_L4 | 500 | 0.1039 | 0.0866 |
| LightGCN_L5 | 500 | 0.0932 | 0.0812 |
| LightGCN_L6 | 500 | 0.1205 | 0.1003 |
| LightGCN_L7 | 500 | 0.111 | 0.0913 |
| LightGCN_L8 | 500 | 0.1156 | 0.0964 |
| LightGCN_L9 | 500 | 0.1061 | 0.0853 |
| LightGCN_L10 | 500 | 0.1147 | 0.091 |

In **Table 1**, you can see that there is not much performance difference beyond layer 3.

4. Conclusions

In this paper, we experimented with LightGCN by stacking up to 10 layers to see if there is an over-smoothing problem. As a results of the experiment, we found that the over-smoothing problem is not so big even if we stack more than 3 layers. Therefore, in future research, we plan to use the fact that the over-smoothing problem is not so big to improve the performance of LightGCN.

Acknowledgments

This research was supported by the National Research Foundation of Korea (NRF) Basic Science Research Program funded by the Ministry of Education (No. 2020R1A6A1A03040583) and this research was supported by the MSIT(Ministry of Science and ICT), Korea, under the National Program for Excellence in SW supervised by the IITP(Institute of Information & communications Technology Planning & Evaluation) (2021-0-01393) and this work was supported by Korea Institute for Advancement of Technology(KIAT) grant funded by the Korea Government(MOTIE) (P0008691, HRD Program for Industrial Innovation)

References

- [1] X. He, K. Deng, X. Wang, Y. Li, Y. Zhang *et al.*, “LightGCN: Simplifying and Powering Graph Convolution Network for Recommendation,” in *Proc. Special Interest Group in Information Retrieval (SIGIR)*, China, pp. 639–648, 2020.
- [2] T. B. Mahieux, D. P. W. Ellis, B. Whitman, and P. Lamere, “The Million Song Datasets,” in *Proc. ISMIR*, Miami, Florida, pp. 591-596. 2011.

Analyzing the Financial Effects of Smart Factory Support Program using Propensity Score Matching

Sangun Park and Tai-Woo Chang

Department of Industrial and Management Engineering, Kyonggi University
Suwon, South Korea

[e-mail: supark@kgu.ac.kr, keenbee@kgu.ac.kr]

*Corresponding author: Tai-Woo Chang

Abstract

Smart factory is a human-centered, high-tech intelligent factory that integrates all production processes. The Korean government has deployed about 30,000 smart factories by 2022, highlighting the necessity for a professional and systematic assessment of the effectiveness of support at manufacturing sites. Consequently, this study analyzes the impact of the smart factory support program on the financial performance of the supported company in order to provide reliable data for future smart factory policy development and institutional improvement. In this study, the financial performance of supported and non-supported companies was compared to analyze the impact of the government's smart factory support program. To prevent inaccurate results due to confounding variables, propensity score matching (PSM) was employed to mitigate selection bias for supported companies, allowing for more accurate comparisons. The comparison revealed differences between supported and non-supported companies in the average CAGR from 2018 to 2021, and annual sales growths in 2019 and 2021.

Keywords: Financial effect analysis, Propensity score matching, Smart factory, Governmental program

1. Introduction

Smart factory is a human-centered, high-tech intelligent factory that integrates all production processes from product planning and design, manufacturing and processing, distribution and sales with information and communication technology to produce customized products at minimal cost and time [1].

Through the deployment of smart factories, companies can actively respond to rapidly changing external environments and advanced customer requirements by aiming to intelligently, flexibly, optimally, and efficiently manage their operations.

The Korean government has established and promoted a smart factory policy since 2014 to innovate manufacturing and has deployed more than 30,000 smart factories by 2022.

As the number of smart factories built each year increases, there is a need for professional and systematic measurement, analysis, and diagnosis of the effectiveness of support at production and manufacturing sites.

Through this study, we aim to provide reliable data for future smart factory policy establishment and institutional improvement by analyzing the performance of smart factory support program from the perspective of economic policy support effects.

In order to analyze the impact of the smart factory support program on the financial performance of the company, it is necessary to evaluate whether the companies that implement a smart factory with the support program actually improved their financial performance, such as sales, compared to the companies that did not [2].

However, it is problematic to simply compare the financial performance of supported and non-supported firms. In order to analyze the impact of support for a company on its financial performance, it is possible to compare it under the assumption that all other conditions are equal. This requires the use of a randomized controlled environment, which is not feasible. If there are confounders that affect both the company's support and financial performance, it is not possible to make an accurate comparison.

Propensity score matching (PSM)-based performance analysis reduces selection bias for applicants, allowing for more accurate performance comparisons [3].

Therefore, this study aims to more accurately compare the financial performance of companies that received support from the smart factory program and those that did not by using PSM.

2. Research Method with PSM

2.1 PSM Analysis Ideas

PSM-based performance analysis reduces selection bias for applicants, allowing for more accurate performance comparisons [4]. By calculating propensity scores for supported and non-supported companies and matching companies with the closest propensity scores, the effect of confounding variables can be minimized. By analyzing the matched experimental group (supported companies) and the control group (non-supported companies), it is possible to more accurately analyze the effect of support on financial performance while the effect of confounding variables is minimized.

2.2 PSM Analysis Procedure

PSM analysis consists of data preprocessing, variable generation, support effect analysis before PSM, PSM performance, and support effect analysis after PSM as shown in Fig. 1.

In the data preprocessing stage, basic preprocessing such as missing value removal and data transformation and merging is performed, and in the variable generation stage, target variables and confounding variables are created for performance comparison, and outlier removal is performed based on the generated variable values.

In addition, the effect of the support effect analysis through the PSM performance is identified by analyzing the support effect before and after the PSM performance, respectively.

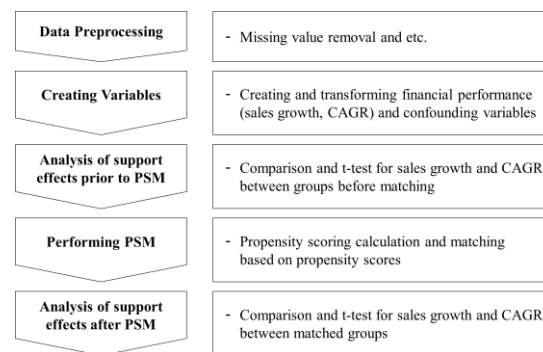


Fig. 1. PSM analysis procedure

3. Results of the Financial Effects Analysis using PSM

Fig. 2 shows the propensity score distribution before and after PSM matching for supported and non-supported companies. Raw Treated represents the propensity score distribution of applied companies before matching, and Raw Control represents the propensity score distribution of non-applicable companies before matching. Before PSM matching, the propensity score distributions of supported and non-supported companies are different, which means that the comparison may be distorted by confounding variables.

In the figure, Matched Treated represents the propensity score distribution of supported companies after PSM matching, and Matched Control represents the propensity score distribution of non-supported companies after PSM matching. After PSM, you can see from the graph that the propensity scores of supported and non-supported companies are distributed almost similarly. This means that the influence of confounding variables on newly extracted

groups of support and non-support can be minimized.

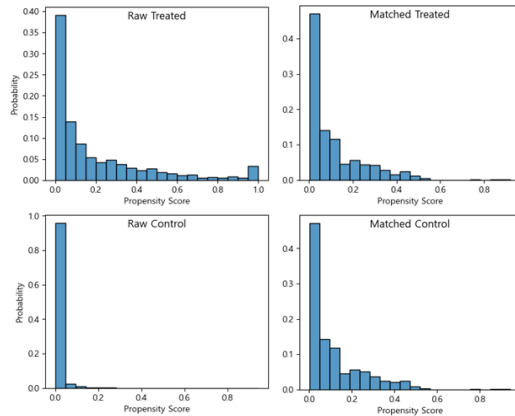


Fig. 2. The propensity score distribution before/after PSM

As a result of t-test on the annual sales growth rate and CAGR of supported and non-supported companies before PSM, it was found that there was no significant difference in all, as shown in **Table 1**. The ‘Result’ column in the table represents the results of whether there is a difference between the groups of supported and non-supported companies. This shows that government support for smart factories had no effect on the company’s sales.

| | t | p-value | Result |
|------------------------|--------|---------|--------|
| Sales growth in 2019 | -0.529 | 0.597 | No |
| Sales growth in 2020 | 0.724 | 0.469 | No |
| Sales growth in 2021 | 0.148 | 0.883 | No |
| Avg. CAGR in 2018~2021 | -1.830 | 0.067 | No |

Table 1. t-test results between supported vs. non-supported before PSM

However, as shown in **Table 2**, as a result of the t-test for supported/non-supported companies after PSM, it was confirmed that there was a difference in sales growth rate and CAGR except for 2020. This is in contrast to the previous t-test results before PSM, which found no differences. These results show that incorrect analysis caused by confounding variables can be prevented using PSM, and consequently show that the government’s support for smart factories is effective.

| | t | p-value | Result |
|------------------------|--------|---------|--------|
| Sales growth in 2019 | -4.544 | 0.000 | Yes** |
| Sales growth in 2020 | -1.634 | 0.102 | No |
| Sales growth in 2021 | -2.046 | 0.041 | Yes* |
| Avg. CAGR in 2018~2021 | -5.976 | 0.000 | Yes** |

Notes: *Significant at the $p < 0.05$ level,

** Significant at the $p < 0.001$ level

Table 2. t-test results between supported vs. non-supported after PSM

4. Conclusions

In this study, the financial performance of supported and non-supported companies was compared to analyze the impact of the government's smart factory support program. To prevent inaccurate results due to confounding variables, PSM was performed to reduce selection bias for supported companies, allowing for more accurate comparisons. As a result of the comparison, it was found that there was a difference between supported and non-supported companies in the average CAGR from 2018 to 2021, and annual sales growths in 2019 and 2021.

References

- [1] E. Hozdić, “Smart factory for industry 4.0: A review”, *International Journal of Modern Manufacturing Technologies*, Vol 7, No 1, pp.28-35, 2015.
- [2] J.J.Heckman, H. Ichimura, and P.E. Todd, “Matching as econometric evaluation estimator: evidence from evaluating a job training programme”, *Review of Economic Studies*, Vol 65, No 2, pp.605~654, 1997.
- [3] Z. Zhao, “Using matching to estimate treatment effects: data requirements, matching metrics, and Monte Carlo evidence”, *The Review of Economics and Statistics*, Vol 86, No 1, pp.91-107, 2004.
- [4] P. R. Rosenbaum and D. B. Rubin, The Central Role of the Propensity Score in Observational Studies for Causal Effects, *Biometrika*, Vol 70, No 1, pp. 41-55, 1983.

Comparing Scalable Multiclass Classification Models to Rate Amazon Books

Hsiu-Ping Lin, *Min-goo Kang, Jongwook Woo

Department of Information Systems, California State University, Los Angeles

*Department of Computer Information Systems, Hanshin University, Korea

{hlin54, jwoo5}@calstatela.edu, kangmg@hs.ac.kr

Abstract

This paper aims to predict the book's ratings using the large-scale dataset of Amazon books. The dataset is too large to handle with the traditional approach, so we adopt the Big Data platform, PySpark, in the Amazon EMR cluster that processes large-scale data. We build prediction models as multiclass classifiers using various PySpark algorithms – Logistic Regression (LR), Decision Tree (DT), Random Forest (RF), and Multi-Layer Perceptron Neural Network (MLP-NN). So, we can predict book ratings in five and binary classes. We compared Multiclass Classifiers' accuracies, especially Precision, which has the smallest False Positive. As a result, we observed that the LR model exhibits the highest precision in binary and multiclass classification models and boasts the fastest computing time. Besides, MLP-NN demonstrates good performance, although it requires more time for model construction than LR. RF and DT show lower performance and demand more computing time than LR and MLP-NN. Furthermore, binary classification models have better performance in precision than multiclass ones.

Keywords: Big Data, Spark, Scalable Computing, Neural Network, Multiclass Classification, Rating

1. Introduction

Amazon has emerged as one of the world's most successful e-commerce enterprises. Its book-selling business has been producing data related to curated lists of bestsellers, popular books, and a recommendation system. Reviews text of the data are generated for the respective products when customers provide feedback and ratings on the website. It would be helpful to have a prediction model to recommend the right books to customers. The challenge for this predictive analysis is that first, the data size is too big to use the traditional computing systems, and second, review data requires text processing and vectorization, which requires more computing time. So, this paper employs Big Data platform to build PySpark Machine Learning Models to predict book ratings efficiently using the

Amazon book review dataset from Kaggle. The PySpark Big Data cluster comprises multiple nodes for distributed parallel computing, which also attain scalability.

2. Related Work

Chen et al. explore machine learning models to predict Amazon product sentiment analysis. It aims to create an accurate model that can handle diverse sentiments expressed in reviews. The study evaluates NLP models (Light GBM, CatBoost, deep learning). The article discusses potential applications for sentiment analysis, concluding with results, key findings, and recommendations for future research [1]. Norinder et al. combine deep learning and conformal prediction for accurate sentiment analysis of Amazon product reviews across 12

categories. The paper emphasizes the approach's generalizability and its ability to handle imbalanced sentiment classes in Amazon reviews [2]. Our paper built LR, DT, RF, and MLP-NN models to predict ratings of Amazon Books in Big Data platform. We compare the computing time and accuracy of the models.

3. Data and Architecture

The dataset, obtained from Kaggle [3], comprises two files: *Books_rating.csv* and *Books_data.csv*, with a size of 2.9 GB. *Books_rating.csv* is a substantial file, containing information on 3 million book reviews, including user details for each review, along with columns: [ID, Title, Price, Profile name, Review Summary, Review Text, Review Helpfulness, Review Class]. *Books_data.csv* provides additional information, featuring details on genres, authors, cover designs, and descriptions, with columns: [Title, Description, Authors, Publisher, Categories].

It is not easy to handle data of 2.9 GB using traditional systems. So, we adopt the Big Data platform within the Amazon EMR cluster to process large-scale datasets. Table 1 shows the specifications of the cluster with Hadoop and Spark services.

Table 1. EMR cluster Specification

| | |
|------------------------|--------------------|
| Number of nodes | 3 |
| Hadoop Cluster Version | 3.2.1 |
| Spark | 3.1.1 |
| CPU speed | 1,995 GHz, 8 cores |
| Memory | 536.4 G |

4. Machine Learning Models

It is vital to recommend customers with higher-rated books. Our objective is to predict the rating classes of Amazon books using various data features.

We constructed prediction models as binary classifiers employing various PySpark algorithms, including Logistic Regression (LR), Random Forest (RF), Decision Tree (DT), and Multi-Layer Perceptron Neural Network (MLP-NN). Achieving the accuracy of each model, the smallest number of False Positives is

more important than minimizing False Negatives. In other words, *Precision* takes precedence over *Recall*. Amazon book ratings range from 1 to 5, with 5 being the highest score. In this structure, binary classification is also possible. In binary classification, we have assigned Classes 1, 2, and 3 to the label 0, and Classes 4 and 5 to the label 1.

To address missing values in the numeric features, we replaced them with their respective average values. For the text feature, we removed the corresponding rows. To handle the various data types of features, we processed the numeric features using *MinMaxScaler*, while the text feature underwent several NLP preprocessing steps, including *Tokenizer*, *StopWordsRemover*, *CountVectorizer*, and *IDF*. Additionally, we utilized *CrossValidation* techniques to generalize the models with hyperparameter tuning.

Table 2 represents each model's computation time (CT), Area Under Curve (AUC), Precision, and Recall.

Table 2. Performance of Binary Classification

| Algorithm | CT (min) | AUC | Precision | Recall |
|-----------|----------|-----|-----------|--------|
| LR | 16 | .87 | .86 | .86 |
| RF | 486 | .56 | .84 | .80 |
| DT | 756 | .64 | .83 | .82 |
| MLP-NN | 222 | .86 | .85 | .86 |

The study aims to identify a model that can recommend higher-rated books to customers, focusing on achieving highest *Precision* with the shortest computing time. **Fig. 1** represents the confusion matrix of LR. **Table 2** reveals that LR performed the highest *Precision*.

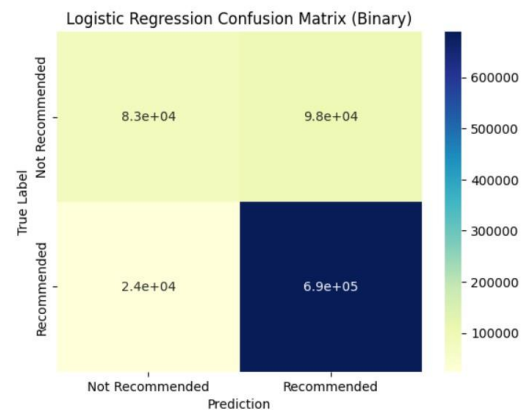


Fig 1. Confusion Matrix of Binary LR

LR model demonstrates the highest *AUC* (87%) and *Precision* (86%). It also requires a significantly shorter time (16 min) than other models. Beyond traditional machine learning models, the MLP-NN shows a high *AUC* (86%) and *Precision* (85%); however, it requires 14 times much longer running time than LR. As RF and DT exhibit relatively acceptable *Precision* and *Recall*, their *AUC* scores are 56% and 64%, respectively, which is notably unreliable compared to LR and MLP-NN.

Table 3. Performance of Multiclass Classification

| Algorithm | LR | RF | DT | MLP-NN |
|----------------------|-----|------|------|--------|
| Computing Time (min) | 19 | 492 | 1458 | 204 |
| Precision | | | | |
| Class 1 | .59 | .93 | .65 | .55 |
| Class 2 | .42 | - | .46 | .38 |
| Class 3 | .46 | - | .62 | .44 |
| Class 4 | .47 | .62 | .44 | .45 |
| Class 5 | .71 | .60 | .63 | .68 |
| Recall | | | | |
| Class 1 | .43 | .005 | .12 | .39 |
| Class 2 | .18 | 0 | .05 | .11 |
| Class 3 | .23 | 0 | .05 | .18 |
| Class 4 | .21 | .003 | .11 | .14 |
| Class 5 | .94 | .999 | .97 | .95 |

By adhering to the labels in the initial dataset, which range from score 1 to score 5, we have also presented the performance of multi-classification for each algorithm in Table 3. We observe that multi-classification performs less effectively than binary classification due to five classes, as expected.

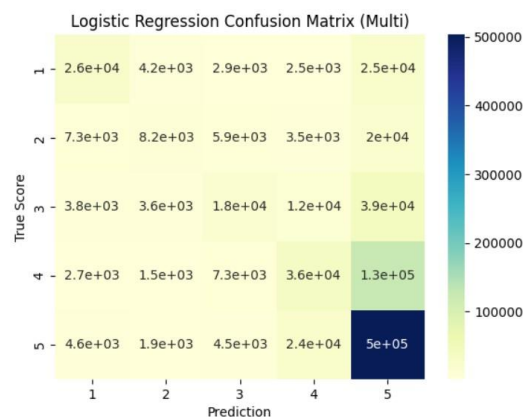


Fig 2. Confusion Matrix of Multiclass LR

LR outperforms others in *Precision* for Class 5 (71%), and has the shortest running time (19 minutes), as shown in Table 3. Although the MLP-NN demonstrates close *Precision* (68%), it requires 11 times longer running time (204 minutes) than LR. Additionally, RF predicts Class 1 with the highest *Precision* (93%), which is highly unfavorable to recommend a book.

5. Conclusion

The paper is to build models that can accurately predict Amazon book ratings and recommend a book to customers. The Amazon book data set is large-scale data, 2.9 GB, which does not allow the traditional systems to process complicated machine learning computation, so we adopt a Big Data platform, Spark cluster, using Amazon EMR service.

We developed binary and multiclass classification models - LR, RF, DT, and MLP-NN - for predicting Amazon book ratings. We compared the models with the accuracy and computing time as they process a large-scale data set. LR demonstrated the highest *Precision* and the shortest computing time. MLP-NN model also performed close *Precision* to LR, but necessitated 10 – 14 times longer computation time than LR.

References

- [1] A. Chen, J. Walsh, M. MacDonald, N. Chu, R. Ahmed, and S. Rao, “Amazon Review Rating Prediction with NLP”, 2021. Retrieved from <https://medium.com/data-science-lab-spring-2021/amazon-review-rating-prediction-with-nlp-28a4acdd4352>
- [2] U. Norinder, & P. Norinder, “Predicting Amazon customer reviews with deep confidence using deep learning and conformal prediction”, Journal of Management Analytics, Volume 9 (1), 2022.
- [3] Mohamed, B. Amazon Books Reviews, 2023. Retrieved from <https://www.kaggle.com/datasets/mohamedbakhet/amazon-books-reviews>

ViT-based Classification with Digital Signal Processing and Image Transformation of Time Series Data

Gyu-Il Kim¹, Seo-El Lee², Ji-Won Baek³, Hyun Yoo⁴ and Kyungyong Chung^{5*}

^{1,3}Department of Computer Science, Kyonggi University, Suwon-si, South Korea

²Department of Public Safety Bigdata, Kyonggi University, Suwon-si, South Korea

⁴Contents Convergence Software Research Institute, Kyonggi University, Suwon-si, South Korea

⁵Division of AI Computer Science and Engineering, Kyonggi University, Suwon-si, South Korea

[e-mail: ran6491@kyonggi.ac.kr, lse_1031@kyonggi.ac.kr, hy1233hh@naver.com, rhpa0916@gmail.com, dragonhci@gmail.com]

*Corresponding author: Kyungyong Chung

Abstract

This paper proposes a novel approach for analyzing and classifying defects in concrete bridge decks using time-series data through DSP(Digital Signal Processing) and image transformation, utilizing the Swin Transformer model. DSP is employed to reduce noise and emphasize significant signals in the data, while image transformation reinterprets the temporal characteristics of time-series data in a spatial context. The transformed images are classified using the Swin Transformer, a process that anticipates high accuracy and efficiency. This research offers a new perspective on time-series data analysis and image-based classification, exploring the potential application of the proposed methodology in various fields.

Keywords: Time series, Image transformation, Class classification, Digital signal processing, Transformer

1. Introduction

In today's data-centric society, time-series data plays a crucial role in various fields such as finance, healthcare, and meteorology. This data encompasses complex and diverse patterns that reflect changes in information over time. Traditional methods of time-series data analysis primarily rely on linear models, moving averages, and autoregressive models. However, these methods have limitations in adequately interpreting and utilizing nonlinear patterns and complex interactions within the data. They are particularly vulnerable to issues such as high-frequency noise, outliers, and non-stationarity, and face challenges in

accurately modeling data when changes over time are complex. Therefore, there is a need for new approaches to deeply analyze and utilize the complexity and diversity of time-series data [1, 2]. This study proposes a novel approach that analyzes time-series data through Digital Signal Processing (DSP), applies an image transformation algorithm to convert it into a visual form, and then classifies it using the ViT(Vision Transformer) model. DSP is useful in reducing noise and emphasizing important signals in time-series data, while image transformation expands the dimension of analysis by converting the temporal characteristics of this data into spatial features. Additionally, ViT, as a state-of-the-art deep learning model adept at processing image data, is

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(No. 2020R1A6A1A03040583)

expected to demonstrate exceptional performance in classifying time-series data transformed into images.

This study proposes a methodology that combines DSP and image transformation algorithms to enhance the characteristics of time-series data and effectively classify it using the ViT model. Through this, it is expected to open new horizons in the analysis of time-series data and contribute to research and applications in this field.

2. Related Works

2.1 The Role and Importance of Digital Signal Processing

Digital Signal Processing (DSP) plays a pivotal role in the analysis of time-series data. DSP is used to remove noise from data and emphasize important signals. Especially in the case of time-series data, it is effective in clearly separating and analyzing periodic patterns, trends, and seasonality. This is crucial in various applications such as predictive modeling, anomaly detection, and trend analysis [3]. The significance of DSP lies also in enhancing the quality of data. Through signal processing, it removes unnecessary or potentially misleading information from the data, thus allowing a more focused analysis on genuinely meaningful information. For example, techniques such as filtering, smoothing, and deconvolution are crucial in facilitating the understanding of time-series data and increasing the accuracy of analytical results. Moreover, DSP plays a role in simplifying complex time-series data into a more manageable and analyzable form. This allows for more efficient processing and analysis of the data while maintaining its unique characteristics. Therefore, DSP can be seen as an essential step in time-series data analysis for gaining higher-level insights.

2.2 Image Transformation Algorithms

The process of converting time-series data into images opens a new dimension in data analysis. This transformation reinterprets the temporal characteristics of time-series data into a spatial form, allowing for more intuitive and effective analysis of complex patterns and relationships. For example, techniques such as

RP(Recurrence Plots), GAF(Gramian Angular Fields), and MTF(Markov Transition Fields) transform time-series data into 2D images, visualizing the hidden structures and dynamic characteristics of the data.

Image transformation can be used as input for deep learning models, especially those optimized for processing image data like CNN(Convolutional Neural Network) or ViT. These models can deeply learn the complex patterns and relationships inherent in time-series data using transformed images. Additionally, this process contributes to dimensionality reduction and feature extraction, facilitating efficient analysis and improved classification performance.

3. ViT-based Classification with Digital Signal Processing and Image Transformation of Time Series Data

This study is structured in three stages. In the first stage, data collection and preprocessing are conducted to process the data. In the second stage, the preprocessed data undergoes Digital Signal Processing and the application of an image transformation algorithm. In the third stage, class classification is performed using the Swin Transformer model, based on the ViT approach. Fig. 1 shows the process of ViT-based classification with DSP(Digital Signal Processing) and the image transformation of time series data.

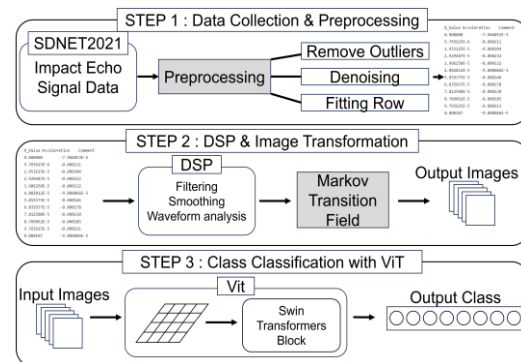


Fig. 1. Process of ViT-based Classification with Digital Signal Processing and Image Transformation of Time Series Data

3.1 Data collection and preprocessing

In this study, Impact Echo signal data from the SDNET2021 dataset is used. This data is

employed for the detection and assessment of defects in concrete bridge decks. The dataset contains sensor observations over time, which can be utilized for the detection and classification of defects in concrete bridge decks. Data preprocessing involves ensuring consistency in the data by aligning the number of rows according to different experiments, as the quantity predicted varies with each experiment. Additionally, noise and outliers are removed.

3.2 Applying Digital Signal Processing and transforming images

In this study, Digital Signal Processing techniques are applied to the Impact Echo signal data from the SDNET2021 dataset. The primary objective of DSP is to emphasize important signals and remove noise from the data. To achieve this, techniques such as filtering, smoothing, and waveform analysis are utilized. The appropriate method must be chosen based on the data's context. After DSP processing of the data, the time-series data is transformed into images. This process involves using techniques like RP, GAF, and MTF to convert time-series data into 2D images. Such transformation reinterprets the temporal characteristics of the data into a spatial form, allowing for more intuitive analysis of complex patterns and relationships. The DSP application and image transformation methods proposed in this study allow for the adjustment of DSP techniques and image transformation approaches according to different data types and situations. This enhances the flexibility and applicability of the research.

3.3 ViT-based Classification with DPS and Image Transformation

In this study, the Impact Echo signal data transformed into images is classified using the Swin Transformer model. The Swin Transformer is a deep learning model that exhibits excellent performance in image classification, object detection, and segmentation tasks. This model divides an image into multiple patches and learns the overall structure of the image through these patches. In this research, the Swin Transformer is employed to classify defect types in concrete bridge decks. The model is capable of accurately classifying areas with defects by learning the location, size, and type of defects from the transformed images. This process boasts high

accuracy and efficiency, making it suitable for classifying defects in complex concrete structures.

While this methodology is tailored to the specific dataset and context of this study, considering the flexibility and scalability of the Swin Transformer, it can also be applied to various types of image data.

4. Conclusions

This research proposes a novel approach for analyzing time-series data using DSP and image transformation, and classifying defects in concrete bridge decks using the Swin Transformer model. DSP processing plays a crucial role in emphasizing important signals and removing noise from the data, while image transformation contributes to reinterpreting the complex patterns of time-series data into a spatial form. The Swin Transformer effectively classified defects in the transformed images, promising high accuracy and efficiency.

The study provides a new perspective on time-series data analysis and image-based classification, laying the foundation for exploring potential applications in various fields. Additionally, the proposed methodology demonstrates flexibility, potentially applicable to other datasets and scenarios. Future research will focus on developing DSP analysis and image transformation methods tailored to specific data, and selecting the appropriate classification model.

References

- [1] J. Friedl, B. Zimmer, L. Perkhof, J. Zenisek, P. Hofer, and H. C. Jetter, "An empirical study of task-specific limitations of the overview+ detail technique for interactive time series analysis," *Procedia Computer Science*, Vol. 180, pp. 628-638, 2021.
- [2] Smart Policing Big Data Platform, <https://www.bigdata-policing.kr/>.
- [3] S. A. Pavan Kumar, P. K. Bora, "Time series analysis and signal processing," in *Proc. of National Conference on Computational Intelligence and Signal Processing (CISP)*, 2012.

NLP Models Classifying Helpful Ratings in OpenTable Dataset

Sunyoung Lee, *Hsiu-Ping Lin, Jihyun Park, Eugene Lim, *Jongwook Woo

Dongguk Business School, Dongguk University
Seoul, South Korea

*Department of Information Systems, California State University, Los Angeles
[e-mail: enhasu19@dongguk.edu, {zzhxxxn, euginoo}@gmail.com, {hlin54, jwoo5}@calstatela.edu]

*Corresponding author: Jongwook Woo

Abstract

This paper aims to build models to predict the helpfulness of reviews by analyzing review data from OpenTable, a restaurant reservation application spanning 14 years (2009 to 2023). Our goal is to accurately classify the helpfulness of the review column of the dataset. We adopted a DistilBERT Huggingface model, an example of LLM (Large Language Model), as a distilled version of the BERT base model for the Natural Language Process (NLP) to classify sentences. We built an NLP classification model by fine-tuning the pre-trained DistilBERT model of Huggingface with the OpenTable data. We also made traditional machine learning models using Logistic Regression (LR), Decision Tree (DT), Random Forest (RF), and Gradient Boost Tree (GBT) algorithms to compare the performance of the models. We observed that the fine-tuned DistilBERT model has 4 - 14 % better accurate Precision and AUC than the traditional models. The traditional models' computing time is 90 - 1,600 seconds, while the fine-tuned DistilBERT model's is 490 seconds. In summary, the fine-tuned DistilBERT model has the best performance among the models. This analysis allows us to identify consumers' (satisfied or dissatisfied) attitudes toward OpenTable's app service and pinpoint areas for improvement. Moreover, from a marketing perspective, featuring the predicted helpfulness review data at the top of the application can enhance the experience for existing customers and provide valuable insights to potential customers, which will gain more potential customers.

Keywords: LLP, Huggingface, Scalable Computing, OpenTable, BERT, NLP, Classification

1. Introduction

Online review systems have influenced the decision-making of potential customers [1]. In particular, classifying and predicting helpful reviews among many online reviews significantly impact customer acquisition [2]. This study aims to build models to predict reviews' helpfulness, comparing the performance of the traditional Machine Learning (ML) and Deep Learning (DL) Large Language (LLM) Models.

2. Related Work

Yu et al. built category classification models with The THUCNews dataset, combining neural network models such as word2vec, BiGRU, ELMo, BERT, CNN, and RNN. It shows that BERT-BiGRU model has the highest accuracy [4].

Garrido-Merchan et al. built sentiment prediction models with the traditional ML models and BERT with the datasets: IMDB, Tweets, News, Hotel Reviews. It concludes that BERT has the highest accuracy [5].

Our paper re-trains and builds the **DistilBERT** model with OpenTable data. The paper classifies helpfulness ratings in accuracy and computing time, comparing the traditional ML models and DistilBERT.

3. Background

3.1 Review of OpenTable Dataset

This paper collected online review data from **OpenTable**, a global restaurant reservation application, to build a model to predict the helpfulness of reviews. The 46,392 sample data was collected from the OpenTable's app from September 2009 to February 2023, about 14 years. The data includes the user's nickname, review content, star rating, helpfulness, and date and time.

3.2 Data Clustering Analysis

Topic modeling was performed to determine the frequency of co-occurrence of words. The consistency score (0.3952) and the inflection point were selected for 9 topics [3].

Table 1 presents how consumers rated the app with the 9 topics. Topics 1 and 6 provide an objective assessment of the quality of the app. It lists features that are important to them: ["booking", "new places", "recommendations"]. Topics 2 and 4 evaluate the UI and UX aspects of the app, particularly mentioning: ["adding application", "convenience"], indicating that consumers value functional convenience in using apps. Topic 3 evaluates the restaurant experience, focusing on the experience of a good restaurant, which is important and often leads to evaluating the app based on the reviewers' detailed restaurant experience. Topics 5 and 8 evaluate the app experience in terms of hedonic aspects. Emotional words such as "love" and "wonderful" are predominant, suggesting that users are highly satisfied when using the app. Topics 7 and 9 evaluate the accuracy and reliability of the app's content and system.

Table 1. Top 10 terms of co-occurrence of words.

| Topic | Words | % |
|-------|--|------|
| 1 | table, reserve, place, new, friendly, travel, easier, available, recommend, want | 13.6 |

| | | |
|---|--|------|
| 2 | make, make reservation, simple, easy make, need, options, cancel, problem, handy, make easy | 12.8 |
| 3 | time, way, experience, reliable, best, wish, dinner, reward, restaurant use, great experience | 11.6 |
| 4 | quick, use app, helpful, quick easy, service, try, app easy, nice, navigate, years | 11.4 |
| 5 | love, love app, book, super, issue, super easy, help, easy book, know, wonderful | 11.3 |
| 6 | great app, excellent, point, opentable, use opentable, use opentable, accurate, food, excellent app, website, add | 10.7 |
| 7 | useful, din, efficient, date, review, user, amaze, fast, look, time | 10.3 |
| 8 | good, awesome, opentable, thank, good app, email, let, love opentable, book, love | 9.2 |
| 9 | convenient, work, phone, easy convenient, work great, great restaurant, people, app work, convenient easy, pasword | 9.2 |

4. NLP Rating Classifying Models

Spark is a distributed parallel computing platform that supports machine learning APIs. It also provides NLP APIs such as Tokenizer, StopWordsRemover, and HashTF.

LLM in AI DL have received highlights lately, which can process, understand, and generate natural language. ChatGPT and BERT are the most popular LLM models in NLP. ChatGPT is built by Open AI that can generate texts. Google introduced BERT in 2018 that is suitable for text classification and Q&A.

4.1 ML and LLP Models

We built classification models to predict Helpful labels in the Spark platform - **LR**, **DT**, **RF**, **GBT** - with the users' review text from OpenTable data. We adopt Hashing Term Frequency for the models and set the parameters as {maxDepth:5, numTrees:3, maxIter: 20, regParam: 0.01}.

In addition, **BERT** can be used in text classification as well. It is a bidirectional transformer pre-trained model with raw texts

without human labeling on the *BookCorpus* and *English Wikipedia* datasets, combining next-sentence prediction and mask language modeling objectives. **DistilBERT** model is the small and fast version of BERT based on the BERT model. It is pre-trained on the same datasets as BERT. The model can determine a mask in a sentence to fill in. We train and build a text classification model on **DistilBERT** with Auto Tokenizer to predict Helpful labels.

4.2 Experimental Result

We built and executed the traditional machine learning models in single node **PySpark** with Spark 3.5.0, and **DistilBert** model in **Google Colab** as *e2-standard-4* with CPU speed 2.25 GHz, 4 vCPU, and 16 GB RAM.

Table 2. Performance Comparison of the Models

| | Precision | AUC | Computing Time: $\log(sec)$ |
|------------|-----------|-------|-----------------------------|
| LR | 0.682 | 0.644 | 201 |
| DistilBERT | 0.712 | 0.705 | 493 |
| RF | 0.667 | 0.524 | 634 |
| DT | 0.622 | 0.593 | 709 |
| GBT | 0.637 | 0.616 | 3,141 |

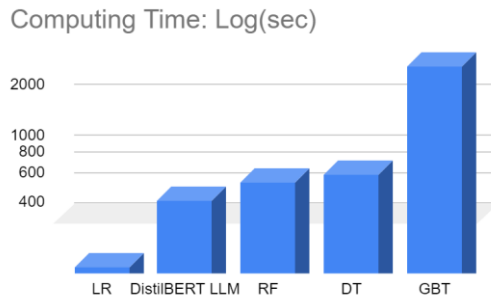


Fig. 1. Computing Time of the Models

We observed performance results of the models that predict the Helpful label ratings in **Table 2**. The precision and AUC of the DistilBERT LLM model are 71.2 % and 70.5 %, respectively. So, it has 4 - 14 % better Precision and AUC than the traditional models. LR has the shortest computation time to build the model. And, the next shortest time is the DistilBERT model, as shown in **Table 2** and **Fig. 1**.

5. Conclusions

This paper compares the performance of the traditional classification models and the DistilBERT LLM model to classify the review text of the OpenTable dataset. We observe that the DistilBERT LLM model shows the highest accuracy and the second fastest computing time compared to the traditional ML models to predict the *Helpfulness* label with the feature of the users' comments in the OpenTable data set as an NLP.

From a marketing perspective, helpful reviews significantly influence the decisions of potential customers. Therefore, developing a strategy to place helpful reviews at the top of your app using accurate and fast predictive models is critical.

References

- [1] Guo, Y., Barnes, S. J., & Jia, Q. , "Mining meaning from online ratings and reviews: Tourist satisfaction analysis using latent dirichlet allocation", *Tourism Management*, 59, 467–483, 2017.
- [2] Zhan, Mengmeng, et al. "The role of explained actions and reactions in the helpfulness of online reviews." *Electronic Commerce Research*, 1-30, 2023.
- [3] O'Callaghan, D., Greene, D., Carthy, J., & Cunningham, P., "An analysis of the coherence of descriptors in topic modeling", *Expert Systems with Applications*, 42(13), 5645–5657, 2015.
- [4] Q. Yu, Z. Wang, K. Jiang, "Research on Text Classification Based on BERTBiGRU Model ,", *Journal of Physics: Conference Series*, doi:10.1088/1742-6596/1746/1/012019, 2019.
- [5] E. C. Garrido-Merchán, R. Gozalo-Brizuela, S. González-Carvajal, "Comparing BERT against traditional machine learning text classification ,", *Journal of Computational and Cognitive Engineering*, <https://doi.org/10.47852/bonviewJCCE3202838>, 2023.
- [6] DistilBERT base model (Oct 2023), Retrieved from <https://huggingface.co/distilbert-base-uncased>

Study on Classification of mobile application based on mobile log data

Kyung Han You¹, Byung Hoon So¹, Minhee Lee¹, Jaehee Cho²

¹Department of Media & Communication Studies, Jeonbuk National University
Jeonbuk, Korea, Republic of
[e-mail: kyunghanyou@gmail.com]

² College of Media, Art, Science, Sogang University, Seoul, Korea, Republic of
[e-mail: jcho76@sogang.ac.kr]

*Corresponding author Jaehee Cho

Abstract

The present study aims to explore an alternative classification method to complement the existing genre-based classification method to accurately identify users' smartphone application usage patterns. To this end, by drawing on prior research on app repertoires, this study created a classification system consisting of 12 different categories for app usage optimization. With a focus on the time of usage, including weekdays, weekends, and overall usage, the present research employed cluster analysis to identify unique characteristics within each category. This approach not only reduces the risk of subjective categorization, which can be influenced by business biases when assigning apps to existing genres, but also relies on user data for precise measurement.

As a result, this classification system is expected to be much more effective in accurately identifying and classifying future apps. It takes into account the user's lifestyle and the specific context in which the app is used, even when comparing similar services.

Keywords: mobile application, clustering, application repertoire, mobile log data, mobile media use

1. Introduction

The conventional method for comprehending mobile application usage behavior is the "repertoire approach", which involves analyzing television, channels, and programs as indicators of usage behavior patterns. Specifically, people create their own collection of applications by choosing and downloading them from application markets. They then build their own media environment and use these applications in different ways in their daily lives. The study of this process of constructing an application repertoire is crucial for media studies that aim to comprehend how people use mobile media on smartphones. The objective of this study is to

examine how media users manage their collection of mobile applications for accessing media. This will be achieved by gathering individual mobile application usage data for the month of August 2023 and evaluating the app repertoire according to usage within different mobile application categories.

This study highlights the ineffectiveness of existing analytical tools in classifying app usage patterns based on lifestyle, specifically in recognizing recent trends across distinct genres. Furthermore, the classification of an application's genre is determined by the application provider and the operator of the Google Store. However, this classification can be inaccurate or fail to account for the fact that applications within the same genre can vary

depending on their purpose and context of use. Hence, the objective of this study is to surpass the current genre classification and categorize novel application usage categories that align with usage patterns, based on real user application usage data.

2. Research Question

This study is an exploratory study to identify the application repertoire of mobile users by utilizing the media repertoire approach, and the following two research questions are proposed.

Research question 1: How are applications categorized based on the usage of mobile applications?

Research question 2: What are the characteristics of each type of application according to the usage of mobile applications?

3. Methodology

This study gathered log data from 596 users, documenting their activity of a mobile application over the month of August 2023. The collected data includes the name of the application, the total amount of time it was used, the amount of time it was used on weekdays, the amount of time it was used on weekends, and demographic variables such as gender and age.

3.1. Analytical Plan

In this work, we performed a two-stage cluster analysis using the three usage times of the gathered data as variables. Subsequently, we determined the ideal number of clusters using K-means cluster analysis.

Initially, we analyzed the clustering schedule using a two-stage cluster analysis method, specifically employing the furthest item approach and the Pearson correlation coefficient. Our findings revealed a significant decrease in the coefficient value while considering 12 clusters. Subsequently, the researchers analyzed the dendrogram and saw that the initial stage formed seven distinct clusters. The researchers determined that utilizing seven and twelve

clusters yielded excellent results. They employed K-means cluster analysis to compare the clusters obtained from both configurations. Our analysis revealed that the 7 clusters were larger than expected, consisting of 156 individual clusters. Additionally, the 12 clusters were separated into 3 distinct clusters with specific characteristics. Hence, we deduced that the 12 clusters precisely mirrored the attributes of the individual populations.

4. Result

4.1 Characteristics of Each Group

4.1.1. Video content produced within the country / Services provided within the country

This app is specifically designed for domestic video content and local services. It features Naver Serieson SERIESON, Official Burger King BURGERKINGKOREA, SSGPAY, SmartThings, NS Home Shopping, and GoogleDuo. CJ Logistics offers home delivery services, among others.

4.1.2. Service for Local Membership and Convenience

The second category consists of service applications that prioritize convenience and are provided by local operators, such as CoupangEats, Lotte Mart GO, Seezn, KT Membership, and others.

4.1.3. Finance in the Business, Public Service, and Retail sectors

According to Korea Research Mobile, the third type is a hybrid of public and private activities, encompassing business, public service, and retail finance applications. It comprises 32 applications, including HanaOneQ, Kiwoom, KakaoPay, Kakao Bank, Government 24, JamLive, Woori Bank, and AppFree.

4.1.4. 2030 Academic Services

The fourth category comprises applications utilized by young adults, particularly university students, for academic pursuits and part-time

employment. This category includes Korea Scholarship Foundation, Welcome Digital Bank, Welcome Savings Bank, WorkNet, Liv, ZOOMCloudMeetings, MicrosoftWord, MicrosoftExcel, KBStar, BillLetter, and others.

4.1.5. Economic Job Search in 2030

The fifth category of application usage is specifically designed for the economic activities and employment search of young individuals. It comprises Hyundai Merchant Marine, Hana Card, Polaris Office, TOSS, Cloudberry, KakaoBus, employment Korea, and AllOneBank.

4.1.6. Food, Culture, and Leisure

The sixth category comprises programs designed for daily culture and leisure activities, such as Home & Shopping, Happy Order, Together Heart, Yes24 Bookstore, Bae-min, MEGABOX, Home Tax, Express Bus Service, wavve, U Mobiletv, TikTok, and others.

4.1.7. Emerging forms of media, music, culture, and cost-effective travel options.

The seventh category comprises applications that cater to the lifestyle and culture of young individuals. Examples include Palago, TVING, KAKAOWEBTOON, SmarterSubway, Upbit, Aladdin Internet Bookstore, Samsung Video Library, Beauty Point, Yogiyo, MUSINSA, Moarak, Melon, and others. These apps offer advantageous features for cost-effective travel, entertainment, and cultural experiences.

4.1.8. Consumer Preferences

The eighth category consists of widely used applications that are popular among a big user base. This category includes Homeplus, HOGANGNONO, Honeyscreens, Happy Point, Happy Screen, and Happy Money.

Haruna Weather2, Hana Members, Hana Money, Pocket CU, Panelock, TMON, Twitter, Coupang Play, and other applications are being used.

4.1.9. Services provided on an hourly basis

The ninth category refers to an application that caters to a non-traditional work schedule and effectively utilizes the resulting free time. It is closely associated with mobility, telecommuting, and side jobs. Examples of such applications include Haypole, Finng, Paycoin, PanelNow, TouchNM Vaccine, Courier Finder, Coinone, Korail Talk, KakaoTalk, Kakao Subway, Navigation 3.0, KakaoT, and Interpark Tour.

4.1.10. Home shopping/beauty oriented

The tenth category of application usage primarily involves shopping for beauty products. This includes applications like Hyundai Hmall, LOTTE Home Shopping, CJ OnStyle, and B612. These applications are predominantly used by middle-aged women in the office. Additionally, applications like Kakao Nav, TMAP, and GooglePlay Movies are used for mobility purposes.

4.1.11. Services provided in the workplace

The eleventh category of program usage is limited to NHpay and Google Docs, and is specifically tailored for certain office professionals as a specialized service consumption type.

4.1.12. Management of workplace and wealth

The twelfth category of application utilization encompasses workplace services and wealth management, which includes Korea Securities Investment, Hana Financial Investment, Hana OneCupro, and Kiwoom Securities. The list includes Youngwoong Moon S, Subscription Home, Securities Plus, Shinhan Financial Investment Alpha, Samsung Securities mPOP, Mirae Asset Securities m.Stock, Mobile Fax, Mobile Securities Tree Account Opening, HRDNet, and Health Status Self-Diagnosis Ministry of Education.

5. Conclusion

Thus far, we have categorized 12 application usage types that have been optimized, taking into account the period of usage (weekdays, weekends, and overall). Additionally, we have determined the distinctive features of each type through the implementation of cluster analysis. This classification system not only mitigates the potential for subjective categorization, which can be influenced by business biases when assigning apps to existing genres, but it also relies on user data for measurement.

Consequently, it is anticipated to be significantly more effective in accurately identifying the range of apps available in the future. This is because it can take into account the user's lifestyle and the specific context in which the app is used, even when comparing similar services.

References

- [1] Choi, Y. S., Park, J. H., & Lee, H. E., "Factors Affecting Smartphone Application's Functional Composition and Repertoires," *Journal of Social Science Review*, 163-198, 2015.
- [2] Kwon, S. E., Jang, S. I., & Hwangbo, H., "Mobile App Analytics using Media Repertoire Approach," *Journal of Society for e-Business Studies*, 26, 133-154, 2021.
- [3] Sim, M. S., "A Critical Review of Research on Multimodal Media Use," *Journal of Media Information Research*, 40-73, 2010.

An Operator Task Model for Constructing Human-in-the-loop Simulation in Manufacturing Systems

Byeonkuk Kim, Gibaek Kim, Hyeonsu Lee, and Dongmin Shin

Department of Industrial and Management Engineering, Hanyang University
Ansan, South Korea

[e-mail: kbbk960312@hanyang.ac.kr, vxz45@hanyang.ac.kr, criticaller@hanyang.ac.kr,
dmshin@hanyang.ac.kr]

*Corresponding author: Dongmin Shin

Abstract

Humans are playing critical roles in even highly automated manufacturing systems. For seamless collaboration between manufacturing resources and humans, computer simulations are considered an effective tool for designing and analyzing human-machine manufacturing systems. This paper presents a framework to incorporate human's behavior into a manufacturing simulation. A hierarchical model of human tasks is constructed so that human actions can fit the human-machine simulation framework. A simulation model framework is presented to illustrate the potential benefits of the simulation.

Keywords: Human-machine manufacturing, human-machine simulation, human task model, smart manufacturing

1. Introduction

Manufacturing systems are evolving toward automated and autonomous systems, resulting in a new paradigm of smart manufacturing. As smart manufacturing systems are rapidly spreading, simulation has become important than ever as crucial element in digital twins. A discrete event simulation (DES) is one of the promising techniques in designing and analyzing complex systems. Simulation has widely been used in manufacturing system design, operation, and analysis [1].

In spite of widespread of simulation techniques, it has been observed that human components in simulation models have not been considered than they should be. This is attributed partly to the

trend that most of simulations tend to focus on technical aspects such as automated processes.

Human's behavior should be considered in constructing a simulation event calendar. In this regard, atomic elements of human tasks within manufacturing systems needs to be identified and included in an event calendar. This requires a formal model of human tasks in manufacturing systems [2, 3].

This paper presents a formal model of human task in the context of manufacturing systems focusing on material handling tasks in an effort to plug-in human activities into computer simulations. A hierarchical model of tasks is suggested and several issues in terms of human task modeling are discussed. An simplified example from the on-going research is provided to illustrate the possibility of integrating human

manufacturing tasks into a simulation model.

2. Human-Machine Simulation Framework

Discrete event simulation models a real-world system by decomposing it into a set of logically separate objects. The objects evolve through time by make transitions in response to a series of events. Whenever an event occurs over time, each event is assigned a logical time, called, a timestamp. The result of an event is passed to one or more other objects to be further processed, or one event can generate new events.

Manual tasks can be perceived as an effect of them on environment. In other word, the outcome of human manual task can be observed by sensors embedded within a system of interest. These human tasks are considered events to a simulation model, triggering system state transitions. Fig. 1. Shows the flow of a human-machine simulation which takes human tasks as simulation events.

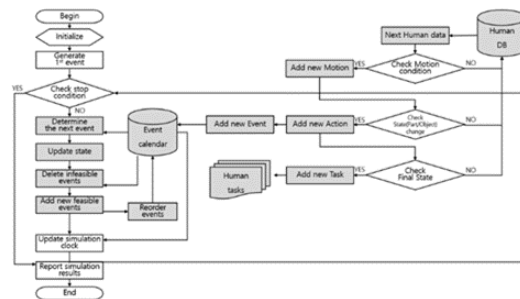


Fig. 1. Flow of a human-machine simulation

3. Hierarchical Human Task Model

3.1 Motion

In modeling a human task, motion is considered an atomic component which consists of a sequence of human body postures. It is defined as a change in human body postures. Several parameters can be associated with a motion such as distance, velocity, acceleration, and time while a change in coordinates of human body joints are primary factors. The motion of body can be observed by means of a variety of sensors, mainly motion capture sensors and position sensors embedded within an environment.

3.2 Action and Task

Action is defined a series of more than one consecutive motions. Action is mathematically described in terms of location and part parameters. Location means a set of positions where an action is taken whereas part is an object that is subject to an action. Task is composed of a series of consecutive actions which are associated with a specific part regardless of changes in location. Fig. 2 shows the relation among task, action, and motion.

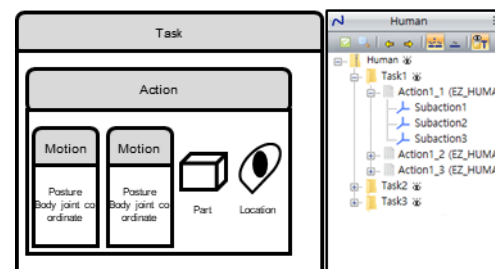


Fig. 2. Hierarchical structure of human task

4. Constructing Simulation Events from Human Behaviors

For a shop floor where human and machine work together to achieve a given goal, a human-machine simulation can be constructed as shown in Fig. 3. The human task model is incorporated into the simulation environment by defining tasks, actions, motions, and relevant parameters as properties of objects in a simulation model.

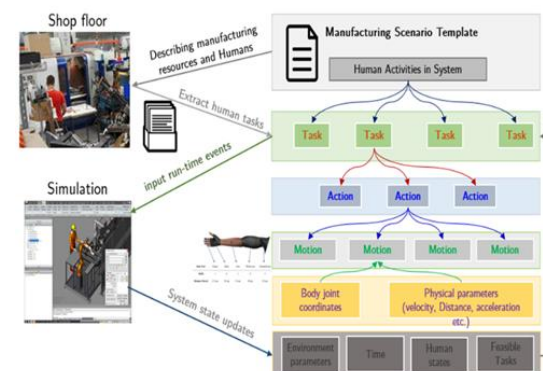


Fig. 3. Human-machine simulation framework

5. Conclusions

In this paper, the on-going research is briefly presented with a simulation framework and a hierarchical human task model. Specifically, a manufacturing shop floor where an human operator performs manufacturing task is considered a target system for the simulation.

Future works include validation of the human task model by comparing simulated results with real-world data or observations to ensure that the simulation reasonably reflects the behavior of human agents in the system.

References

- [1] L. Wang, R. Gao, J. Váncza, J. Krüger et al., “Symbiotic human-robot collaborative assembly,” *CIRP Annals*, Vol. 68, No. 2, pp. 701-726, 2019.
- [2] D. Shin, R. Wysk, and L. Rothrock, “A formal control-theoretic model of a human-automation interactive manufacturing system control,” *Int. J. Prod. Res.*, Vol. 44, No. 20, pp. 4273–4295, 2006.
- [3] M. Li, J. Wei, X. Zheng, and M. L. Bolton, “A formal machine-learning approach to generating human-machine interfaces from task models,” *IEEE Trans. Hum. Mach. Syst.*, Vol. 47, No. 6, pp. 822–833, 2017.

A Study of Stochastic Simulation Algorithm for Bond Pricing Model

Kisoeb Park and Moonseong Kim*

Department of IT Convergence Software, Seoul Theological University
Bucheon, South Korea

[e-mail: {moonseong, kisoeb}@stu.ac.kr]

*Corresponding author: Moonseong Kim

Abstract

This paper proposes a stochastic simulation algorithm to predict the pricing of bond using actual data set of Gilts which are bond is issued by the United Kingdom government. We obtain the pricing of bond through the proposed prediction algorithm which is generated random numbers by the random number generation scheme. In empirical computer simulations, we show that the lower value of limit of error for the pricing of bond under the proposed prediction algorithm corresponds to sharper estimates. It is used in order to develop the prediction algorithm for the term structure model of the interest rate, which takes place in various financial modelling.

Keywords: Stochastic Simulation Algorithm, Bond Pricing Model, Random Number Generation

1. Introduction

In this paper, we propose stochastic simulation algorithm for the pricing of bond model. Derivative of interest rate such as caps, swaptions, bond options, and captions securities have become more and more popular for a long time. A lot of scholar make an effort to find a standard model for the term structure models of interest rates. Basically, there are general equilibrium approach and arbitrage-free approach in the theoretical foundation for the interest rate model on continuous time. The pioneers of equilibrium approach could be regarded as Vasicek [1], Brennan and Schwartz [2], and Cox-Ingersoll-Ross [3]. On the other hand, the arbitrage-free approach to modelling the term structure of interest rate has its genesis in Ho and Lee model [4], the extended model of the Hull and White [5], and Heath, Jarrow, and Morton (HJM) have developed a new class of stochastic interest rate model [6]. In particular,

the HJM is widely accepted as the most general methodology for term structure models. Moreover, the jump-diffusion models [7] are particularly important in pricing and hedging with financial derivatives, since ignoring jumps in financial prices will cause pricing and hedging rates. More generally, infinitely divisible distributions and Lévy processes have been the topic of strong research and applications in late years [8].

We study a stochastic simulation algorithm to predict the bond price using actual data set of bond. Our goal in this paper is to predict more accurate and simple of bond price than previous studies through a prediction algorithm. We will be used an actual data set of gilts which are bonds issued by the United Kingdom government.

2. Stochastic Simulation Algorithm for the Pricing of Bond

In this section we propose a stochastic simulation algorithm to predict the bond price using actual data set of bond. Our goal is to predict more accurate of bond price than previous studies through a dynamic prediction algorithm. We will be used an actual data set of Gilts which are bonds issued by the United Kingdom government. This paper proposes a dynamic prediction algorithm to predict the bond option prices. The proposed algorithm is based on term structure model of the interest rate, which takes place in financial modelling. Fig. 1 shows yield curves for 266 actual data set of a two-year, five-year, ten-year and thirty-year gilts.

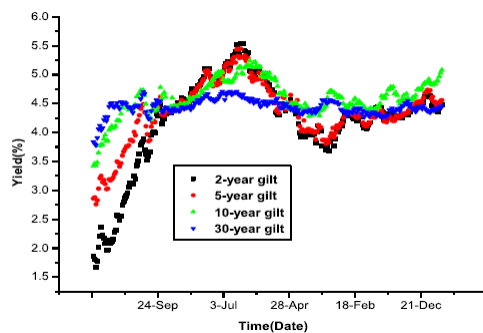


Fig. 1. Yield Curve of Gilts

We briefly explain about the procedure of stochastic simulation algorithm is as follows. We find the cumulative distribution functions (CDF) of actual data using the natural cubic spline (NCS) method [8], which is the frequently used numerical methods for interpolation. We obtain the inverse function of CDF using the inverse-transform technique then we generate arbitrary number from the uniform distribution (UD) [0, 1] and use the inverse function of CDF. Therefore, we can random number generation from previous steps. Therefore, we get the prices of bond through the proposed prediction algorithm. We now explain about the process of the stochastic simulation algorithm.

- I. Compute the CDF $F(x)$ that depends on the random variable x using the NCS interpolation.
- II. Setting the CDF on each interval of x .

III. Express the inverse function $F^{-1}(R)$ of the CDF $F(x)$.

IV. Generate random numbers R_0, R_1, \dots, R_n in UD [0,1] and compute the desired random variable x_i ($i = 0, 1, \dots, n$) using the inverse function F^{-1} .

Therefore, we predict bond prices through the proposed prediction algorithm.

3. Conclusions

In this paper, the stochastic simulation algorithm for Gilts issued by the United Kingdom government can be used to predict bond prices more efficiently and accurately than conventional methods. We can apply the proposed algorithm in order to predict the pricing of the more complicated financial derivatives.

References

- [1] O. Vasicek, "An Equilibrium Characterization of the Term Structure," *Journal of Financial Economics*, Vol 5, pp. 177-188, 1977.
- [2] M. Brennan, E. Schwartz, "A Continuous Time Approach to the Pricing of Interest Rates," *Journal of Banking and Finance*, Vol 3, pp. 133-155, 1979.
- [3] J. Cox, J. Ingersoll, S. Ross, "A Theory of the Term Structure of Interest Rate," *Econometrica*, Vol 53, pp. 385-407, 1985.
- [4] T.. Ho, S. Lee, "Term Structure Movements and Pricing Interest Rate Contingent Claims," *Journal of Finance*, Vol 41, pp. 1011-1028, 1986.
- [5] J. Hull, A. White, "Pricing Interest Rate Derivative Securities," *Review of Financial Studies*, Vol 3, pp. 573-592, 1990.
- [6] D. Heath, R. Jarrow, A. Morton, "Bond Pricing and the Term Structure of Interest Rates," *Econometrica*, Vol 60, pp. 77-105, 1992.
- [7] C. Peng, S. Olivier, "Linear-Quadratic Jump-Diffusion Modeling," *Mathematical Finance*, Vol 17, pp. 575-598, 2007.
- [8] M. Jeannin, M. Pistorius, "A transform approach to calculate prices and greeks of barrier options driven by a class of Lévy processes," Working paper, 2008.

Digital Transformation and its Implications for Technological Innovation in China's Cultural Industry: A Panel Data Analysis

Qilong Shi¹

¹ Anshan Normal University, 43 Ping 'an Street, Tiedong District, Anshan City, Liaoning Province, China; dalong1210@mail.asnc.edu.cn

Abstract

With the rapid advancement of digital technology, the cultural industry has emerged as a pivotal application domain. Concurrently, the establishment of industrial parks and the scale effects of industry agglomeration are perceived as vital strategies for catalyzing the growth of the cultural sector. Drawing upon panel data from 31 provinces in China spanning 2013-2021, this study delves into the interrelations among digitalization, cultural industry agglomeration, and technological innovation, especially after controlling for potential endogeneity concerns. The findings suggest that digitalization substantially augments technological innovation within the cultural industry, particularly in the realm of high-tech patent development where its impact is most pronounced. Moreover, in this digital era, as knowledge sharing and technology spillover effects among firms transcend geographical boundaries, the influence of cultural industry agglomeration on fostering technological innovation appears to diminish.

Keywords: Cultural Industry; Digitalization; Industry Agglomeration; Technological Innovation

1. Introduction

In the midst of a new technological revolution, digital technologies like the Internet, big data, cloud computing, and AI are reshaping global innovation and economic structures, with the cultural sector emerging as a key player in economic resurgence. Integrating digital advancements with the cultural sector for novel growth and global competitiveness is crucial. Digitization's potential to drive innovation in the cultural industry and its impact on cultural enterprises' capabilities remain key questions. The rise of cultural industry parks and agglomeration trends prompt inquiries about their alignment with digitization and innovation enhancement. Addressing these challenges requires theoretical and empirical analysis to guide future policy and development strategies.

2. Methods

In an era marked by technological advancement, digital technologies like big data and AI are transforming the cultural sector. This shift impacts all aspects of the cultural industry, from creation to consumption, and introduces new product formats and business models. Scholars note that digitization enhances the cultural sector's quality and innovation. However, empirical studies are scarce, and most focus on the theoretical mechanisms of how digitization drives innovation in the cultural industry.

This paper examines the effects of digital development on different types of patents and how it influences the relationship between cultural industry agglomeration and innovation. Findings suggest that digitalization boosts innovation across patent types and weakens the positive impact of geographical agglomeration

on innovation. The study highlights the need for further empirical research to understand the evolving dynamics of digitization in the cultural industry.

Based on this, the paper proposes:

Hypothesis 1: Digital development promotes technological innovation in the cultural industry by optimizing macro-control efficiency, advancing industrial structural upgrading, and reshaping the production value chain of the cultural industry.

Hypothesis 2: Digital development will weaken the promotional effect of cultural industry agglomeration on technological innovation in the cultural industry.

3. Model

To test whether Hypothesis 1 proposed in the theoretical analysis section holds true, and to explore the promoting effect of digital development on technological innovation in cultural industry enterprises, this paper employs the Fixed Effects Panel Model for the research. The baseline model is:

$$\ln Patent_{it} = \alpha_0 + \alpha_1 \ln Digit_{it} + \sum_j \alpha_j \times Control_{it} + r_t + \mu_i + \varepsilon_{it} \quad (1)$$

In this model, $\ln Patent_{it}$ serves as the dependent variable, where subscripts i and t respectively represent the i -th province and the t -th year. This paper selects the logarithm form of the number of authorized cultural patents to characterize the technological innovation level of the cultural industry; $\ln Digit_{it}$ represents the digital level of province i in year t , represented by the number of computers used per 100 people in each province to indicate the digital development level of the provinces. α_1 reflects whether digital development will have a promotional effect on technological innovation in the cultural industry. According to the theoretical analysis above, the expected regression coefficient α_1 is significantly positive, thus confirming that digital development indeed has a noticeable propelling effect on innovation in the cultural industry; $Control$ is a vector of control variables which, according to relevant literature, influences the quantity of patents and consists of a series of economic variables, including per capita GDP, higher education prevalence, degree of openness to the outside world, government support, and R&D investment; r_t represents the time fixed effects; μ_i denotes the cross-sectional effects of each province, which do not change over time, used to control for the impact on the

number of authorized cultural patents due to the potential special differences of provincial entities; ε_{it} is the stochastic disturbance term.

Moreover, to test whether digital development may weaken the geographical cultural industry agglomeration's promotional effect on industrial innovation, this paper, building upon the baseline model, incorporates the dummy variable representing the cultural industry specialization agglomeration degree, through the legal entity location entropy, and its interaction term with the level of digital development for regression.

The specific model form is as follows:

$$\ln Patent_{it} = \beta_0 + \beta_1 \ln Digit_{it} + \beta_2 D_1 + \beta_3 D_1 \times \ln Digit_{it} + \sum_j \beta_j \times Control_{it} + r_t + \mu_i + \varepsilon_{it} \quad (2)$$

Wherein, D_1 is a dummy variable representing the location entropy of legal entity units, indicative of the degree of cultural industry specialization and agglomeration. Location entropy, originally proposed by Hagerstrand, is typically utilized to quantify the agglomeration or specialization level of a specific industry within a region. The calculation equation can be expressed as follows:

$$LQ_{ij} = \frac{E_{ij} / \sum_j E_{ij}}{\sum_i E_{ij} / \sum_i \sum_j E_{ij}} \quad (3)$$

Where, LQ_{ij} represents the location quotient of industry i in province j , E_{ij} denotes the relevant indicators (such as industrial added value, employment, number of legal entities, etc.) of industry i in province j , $\sum_j E_{ij}$ signifies the relevant indicators of all industries in province j , $\sum_i \sum_j E_{ij}$ indicates the relevant indicators of industry i nationwide, and $\sum_i \sum_j E_{ij}$ represents the relevant indicators for all industries across the country. The numerical value of LQ_{ij} measures the level of industrial agglomeration; the larger the LQ_{ij} value, the higher the level of industrial agglomeration in industry i of province j . Generally speaking, if LQ_{ij} is greater than 1, it suggests that industry i in province j possesses a comparative advantage nationally; if LQ_{ij} is less than 1, it indicates a comparative disadvantage; if LQ_{ij} equals 1, it signifies that industry i in province j is at the national average level.

In this study, the dummy variable D_1 , measured by the size of the legal entity location quotient, is utilized as the representative indicator of the degree of specialized agglomeration in the cultural industry. The dummy variable D_1 takes

the value of 1 when the legal entity location quotient is greater than or equal to 1, and 0 otherwise. $D_1 \times \ln Digit_{it}$ represents the interaction term between the industrial agglomeration dummy variable and the level of digital development, with its coefficient β_3 reflecting whether the level of digital development will weaken the promotional effect of cultural industry agglomeration on industrial innovation, resonating with Hypothesis 2 proposed in the theoretical analysis section of this article. Based on previous analyses, it is anticipated that β_3 should be significantly negative, thereby indicating that, with the advancement of digitization, the enhancing impact of cultural industry agglomeration on industrial technological innovation has markedly declined. The meanings of the other variables remain consistent with those in Model (1).

4. Data

This study analyzes panel data from 31 Chinese provinces (excluding Hong Kong, Macao, and Taiwan) from 2013-2021 using a two-way fixed effects panel model. It examines digital development's impact on cultural industry innovation and how this relationship evolves under digital contexts. Data sources include various statistical yearbooks, focusing on the cultural, scientific, and technological sectors. The dependent variable is the logarithmic number of cultural patents, including grants for inventions, utility models, and design patents, indicating technological innovation levels in the cultural industry. The core explanatory variable is the logarithm of the number of computers used per 100 people in enterprises, reflecting regional digital development levels. Control variables include per capita GDP, higher education prevalence, openness to the outside world, government support intensity, and R&D investment. These factors significantly influence local cultural industry innovation. R&D investment is quantified as the sum of internal expenditures on applied research and experimental development. The variable setup, statistical results, and indicator calculations are detailed in Table 1.

Table 1. Variable Definitions and Descriptive Statistic

| Variable | Definitions | Mean | S.D. |
|---------------|---|---------|---------|
| <i>Patent</i> | Sum of invention, utility model, and design patent authorizations for | 3559.00 | 7127.00 |

| | cultural enterprises | | |
|----------------|---|------------|------------|
| <i>Digit</i> | Number of computers used per 100 employees in enterprises | 26.20 | 10.76 |
| <i>Gdp</i> | Actual GDP divided by end-of-year total population | 52096.00 | 24718.00 |
| <i>FDI</i> | Total investment of foreign-invested enterprises in RMB relative to GDP | 0.40 | 0.38 |
| <i>Gov</i> | Local general fiscal expenditure on culture, sports, and media per capital | 292.90 | 242.60 |
| <i>Edu</i> | Number of students enrolled in higher education institutions relative to end-of-year total population | 2609.00 | 769.20 |
| <i>R&D</i> | Total internal expenditure on R&D for applied research and experimental development category | 4978000.00 | 5800000.00 |

5. Results

5.1. Basic regression

Table 2. Impact of Digitalization Level on Technological Innovation in the Cultural Industry

| | lnPatent | |
|----------------------|----------------------------------|-------------------------------|
| | (1) Without control variables | (2) With control variables |
| <i>lnDigit</i> | 1.863*** (0.279) | 0.843*** (0.259) |
| <i>lnGdp</i> | | 0.458 (0.889) |
| <i>FDI</i> | | -0.151 (0.333) |
| <i>lnGov</i> | | 0.673** (0.318) |
| <i>lnEdu</i> | | -1.199 (0.822) |
| <i>lnR&D</i> | | 0.645** (0.256) |
| <i>Year</i> | YES | YES |
| <i>Province</i> | YES | YES |
| <i>Observations</i> | 217 | 217 |
| <i>R²</i> | 0.425 | 0.575 |

Note: Values in parentheses represent robust standard errors clustered at the provincial level; ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively. The R^2 values presented in the table refer to within-group R^2 . The same applies below."

Hypothesis 2: Digital development will weaken the promotional effect of cultural industry

agglomeration on technological innovation in the cultural industry.

5.2. Cultural Industry Agglomeration and Technological Innovation in the Context of Digitalization

Extending the initial analysis, the study examined if digitalization's growth reduces the positive impact of geographical agglomeration in the cultural industry on innovation. Table 3 presents equation (2)'s results. The positive coefficient for dummy variable D1 indicates regions with high cultural industry agglomeration still lead in innovation. However, the interaction term between D1 and digitalization level is significantly negative, showing increased digitalization lessens the agglomeration's stimulative effect on innovation, confirming Hypothesis 2.

Table 3. Regression Results Incorporating the Interaction Term between Digitalization Level and Cultural Industry Agglomeration

| | lnPatent |
|--------------|---------------------|
| lnDigit | 0.848*** (0.282) |
| D1 | 2.423** (0.907) |
| D1 ×lnDigit | -0.759** (0.296) |
| Control | YES |
| Year | YES |
| Province | YES |
| Observations | 217 |
| R^2 | 0.601 |

To enhance the study's scientific accuracy, we conducted robustness checks in three areas:

We replaced the core explanatory variable with the internet broadband users to total population ratio. Though not a direct measure of industrial digitalization, it reflects digital infrastructure prevalence. Results confirm this proxy still positively influences cultural industry innovation, supporting digitalization's role in driving innovation.

For cultural industry specialization, we used the ratio of regional to national cultural business revenues, reflecting cultural industry clustering. Table 5 shows that with this measure, digital

development's interaction term remains significantly negative, indicating digital advancement reduces the positive impact of cultural industry clustering on innovation.

To address outliers, we used winsorization, trimming extreme values in our regression data. The positive relationship between digital development and cultural industry innovation persisted, even after further trimming at 2%, 5%, and 10%, consistently supporting digitalization's influence on cultural industry innovation.

5.3. Heterogeneity analysis

This study not only explores digital development's general impact on cultural industry innovation via patents but also examines its varied effects on R&D for different patent types, with results detailed in Table 4.

Table 4. Impact of Digital Development on R&D of Different Types of Patents in the Cultural Industry.

| | (1) lnPatent 1 | (2) lnPatent 2 | (3) lnPatent 3 |
|--------------|---------------------|--------------------|-------------------|
| lnDigit | 1.341*** (0.244) | 0.779** (0.326) | 0.707* (0.400) |
| Control | YES | YES | YES |
| Year | YES | YES | YES |
| Province | YES | YES | YES |
| Observations | 214 | 217 | 217 |
| R^2 | 0.785 | 0.691 | 0.337 |

6. Discussion

Per the 'China Science and Technology Yearbook,' patents are classified into inventions, utility models, and designs. Invention patents cover new technical solutions, utility models focus on practical shape or structure innovations, and design patents on aesthetic and novel applications. Invention patents have the highest technical content.

Table 4 shows regression results for each patent type. Digital development's promotive effect is strongest on invention patents and weakest on design patents, indicating digitalization impacts higher technical content patents more within the cultural industry.

This study, using China's 31 province data from 2013 to 2021, explores digitalization, industrial agglomeration, and innovation in the cultural industry, considering endogeneity. Findings show digital development boosts cultural sector innovation but reduces the impact of geographical agglomeration due to increasing

industrial cloudification. Digitalization has varied effects on different patent types, favoring those with higher technical content.

For future development, China should integrate digital technology more with the cultural industry. Companies can use digital tools for cultural product innovation and regional resources for online platforms, moving beyond traditional geographical limits. This approach enhances innovation and efficiency in the cultural industry. Additionally, China should use digital technology for the innovative integration of the cultural industry with sectors like intelligent manufacturing and smart tourism, expanding the cultural industry's scope and promoting cross-industry collaboration.

References

- [1] 1.Jiang Xiaojian. Technology and Culture in the Digital Age [J]. China Social Sciences, 2021, (8).
- [2] 2.Gu Jiang, Chen Xin, Guo Xinru, et al. Logic Framework and Strategic Path for Improving the Modern Cultural Industry System during the 14th Five-Year Plan [J]. Management World, 2021, 37(3).
- [3] 3.Li Hui. Theoretical Mechanism, Practical Basis, and Policy Choices for Big Data to Promote China's High-Quality Economic Development [J]. Economist, 2019, (3).
- [4] 4.Wang Bin, Cai Hongbo. Connotation, Definition, and International Comparison of Digital Content Industry [J]. Finance and Trade Economics, 2010, (2).
- [5] 5.Liu Jing, Hui Ning, Nan Shijing. Non-linear Research on Data Empowerment Driving Cultural Industry Innovation Efficiency: An Empirical Test Based on the STR Model [J]. Economic and Management Research, 2020, 41(7).
- [6] 6.Zhou Jin, Wang Tingxin. Research on the Integration Development Mode and Path of Urban Cultural Tourism in the Digital Economy [J]. Jiangsu Social Sciences, 2021, (5).
- [7] 7.Zhang Sen, Wen Jun, Liu Hong. Exploration of Digital Economic Innovation: A Comprehensive Perspective [J]. Economist, 2020, (2).
- [8] 8.Li Wenjun, Li Qiaoming. Development Trends and Promotion Strategies for the Digital Creative Industry during the 14th Five-Year Plan [J]. Economic Horizon, 2021, (2).
- [9] 9.Marshall A. Principles of Economics [M]. London: Macmillan, 1920.
- [10] 10.Florida R, Tinagli I. Europe in the Creative Age [R]. Carnegie Mellon Software Industry Center, 2004.
- [11] 11.Boix-Domenech R, Soler-Marco V. Creative Service Industries and Regional Productivity [J]. Papers in Regional Science, 2017, (2).
- [12] 12.Lei Hongzhen, Li Yun. Spatio-temporal Differences and Influencing Factors of Cultural Industry Development Efficiency [J]. Contemporary Economic Management, 2020, 42(6).
- [13] 13.Wang Wenji, Gu Jiang. Measurement and Spatio-temporal Evolution of Cultural Industry Efficiency in the Yangtze River Economic Belt [J]. Nanjing Social Sciences, 2021, (6).
- [14] 14.Wang Ran, Yan Bo, Deng Weigen. The Impact and Mechanism of FDI on China's Industrial Independent Innovation Capability: From the Perspective of Industrial Association [J]. China Industrial Economy, 2010, (11).
- [15] 15.Guo Guofeng, Wen Junwei, Sun Baoying. Analysis of Factors Affecting Technological Innovation Capability: Based on Panel Data Empirical Study of the Six Provinces in Central China [J]. Quantitative Economic Technical Economic Research, 2007, (9).
- [16] 16.Sun Yanming, Chen Simiao. Spatio-temporal Evolution Pattern and Driving Factors of Green Technology Innovation Efficiency in the Yangtze River Delta [J]. Geographical Research, 2021, 40(10).
- [17] 17.Duan Dezong, Du Debin, Chen Ying, et al. Pattern and Influencing Factors of Urban Innovation Technology Transfer in China [J]. Journal of Geography, 2018, 73(4).

Does the Brain's Pre-existing Information Guide Our Explore Gaze?

Ji-Chan Yun*, Yeong-Woo Lim*, Junsik Lee*, Kyunghee Kim,
Jae min Kim**, Sanghoon Lee** and Do-Hyung Park***

* Graduate School of Business IT, Kookmin University, Seoul, Republic of Korea

** HYGGE Inc., Gyeonggi-do, Republic of Korea

[e-mail: chanyosh@kookmin.ac.kr, duddn7244@naver.com, jscx9303@gmail.com,
claire@neochair.com, jkim@neochair.com, shlee@neochair.com, dohyungpark@kookmin.ac.kr]

*Corresponding author: Do-Hyung Park

Abstract

In today's rapidly evolving technological landscape, the e-commerce market's importance cannot be understated. This has become particularly evident as societal shifts influenced by unforeseen events, such as the COVID-19 pandemic, have emphasized individualism more than ever. The emergence of "nano-society" as a trending keyword in 2021 is a testament to this shift. Beyond this, as society traverses the realms of postmodernism, there's an unmistakable accentuation on personal identity and subjectivity. This progression implies that businesses and marketers need a more nuanced and refined approach to customization than previously employed. In such a transformative environment, distinguishing oneself becomes beneficial and essential. For brands and e-commerce platforms, recognizing and responding to subtle, often subconscious, cues from potential customers can be the key to success. It's in these under-the-radar signals that valuable insights often lie. One such potent tool that's been increasingly acknowledged for its efficacy is gaze analysis. Marketers can gain deep insights into user preferences, interests, and potential purchase intent by tracking where, how long, and how often an individual's gaze lands on specific areas. This study leverages eye-tracking technology to delve into the realm of gaze analysis. Specifically, it seeks to discern how different product marketing strategies can influence and shift a user's gaze. Moreover, it aims to understand the implications of these shifts – what happens when a product's marketing strategy resonates well with a viewer versus when it misses the mark? In conclusion, we hold significant implications for sellers in the e-commerce arena. By understanding and leveraging the insights derived from gaze analysis, they can refine their sales strategies, ensuring higher engagement and, ultimately, better sales outcomes.

Keywords: eye-tracking, customization, marketing strategy, e-commerce

1. Introduction

With the imposition of restrictions on face-to-face interactions during the pandemic, the e-commerce market witnessed an unprecedented acceleration, positioning itself as an indispensable conduit for consumer transactions[1]. This surge in e-commerce continues to reverberate even as we emerge from the shadow of the pandemic, with online

shopping ingrained as a ubiquitous facet of daily life[2]. Within the fiercely competitive realm of online shopping, capturing consumers' visual attention has become a pivotal aspect of the industry[3].

In response to the evolving dynamics of this digital market, 'eye-tracking' has emerged as a crucial research technique, providing empirical validation of how consumers visually interact with products and how such interactions

contribute to determining purchasing behaviors[3,4,5].

Our investigation, centered around a type of experience goods known as 'chairs,' aims to elucidate subtle distinctions in consumers' visual attention between office chairs and gaming chairs. Additionally, this study aims to illuminate the intricate interplay between visual attention and consumer emotions, as well as its cascading effects on product evaluations.

2. Method

2.1 Overview of Eye-Tracking Experiment

In this experiment, we utilized an eye-tracking methodology to observe and quantify participants' eye movements and gaze patterns in near real-time, as outlined by Jeong et al. [6].

Areas of interest (AOIs) corresponding to different elements of the stimulus were designated to define specific regions of interest in the visual stimulus.

2.2 Research Question

The purpose of this study is to compare and analyze the factors that consumers make cognitive judgments and the points that their brains pay attention to when they want to buy a chair product online. Therefore, we propose the following two research questions.

RQ1. Where do consumers' eyes stay when purchasing a chair product?

RQ2. What factors do consumers consider important when purchasing a chair product?

2.3 Experiment Design and Method

This study was conducted on undergraduate and graduate students at a private university in South Korea. The total number of participants was 45.

The experimental stimuli used in this study were an office chair and a gaming chair. The chairs were selected from those currently sold in online shopping malls, focusing on products that do not reflect unusual designs. The experimental stimuli consisted of simultaneous front, side, back, and diagonal views of a single chair on a screen and

were observed for approximately 15 seconds. The office and gaming chairs were presented to participants in a randomized order.

Design evaluations for each chair were collected via a questionnaire after the stimulus observation. The survey questions were constructed using the Emotional Adjective Evaluation Vocabulary Scale, developed to explore participants' digital-based visual design and emotional evaluation criteria for chairs[7]. Using 32 adjective polarization scales developed in previous studies, we modified and restructured them to fit this study and finally organized 15 emotional adjective evaluation vocabulary into antonym pairs. The results are summarized in Table 1. We also added a question measuring attitudes toward the chairs in the survey to explore how each adjective evaluation vocabulary affects attitudes.

Table 1. Emotional Adjectives Evaluation Vocabulary Polarization Scale

| | | |
|--------------------|------------------------|----------------------------|
| Unique/Common | Cool/Warm | Sophisticated/Sloppy |
| Unusual/Familiar | Innovative/Traditional | Modern/Classic |
| Sturdy/Fragile | Practical/Ostentatious | Individualistic/Typical |
| Variable/Fixed | Efficient/Wasteful | Convenient/Inconvenient |
| Interesting/Boring | Simple/Complex | Functional/Design-oriented |

45 participants were randomly assigned to two groups and evaluated both office and gaming chairs with only a difference in order. Based on the 90 evaluations, we performed factor analysis and regression analysis. Based on the results obtained by presenting the evaluations of the 15 selected emotional words to the participants after observing the stimuli, factor analysis is performed, and the difference in the influence on chair satisfaction based on the newly formed factors is examined through regression analysis.

3. Results

We used eye-tracking methodology to track consumer gaze patterns when exploring products for purchase online, aiming to identify crucial points of influence during the consumer's decision-making process. The study revealed visual attention differences between office chairs and gaming chairs, shedding light on the distinct visual preferences within these two product categories. In addition, to understand the

emotional considerations of consumers when making chair-purchasing decisions, we asked participants to respond to a bipolar emotional scale. This process identified four key emotional factors deemed significant in the chair-buying context. Finally, through regression analysis, we examined the impact of these emotional factors on overall satisfaction with the purchased products.

The eye-tracking analysis revealed distinct visual attention patterns for gaming chairs compared to office chairs. Participants tended to seek unique cues in gaming chairs, such as brand symbols, while focusing on generic attributes like cushions in office chairs. The dispersion of gaze points was higher for office chairs, indicating more extensive and less focused visual exploration than gaming chairs. Gaze paths did not significantly differ based on product category. Participants consistently focused on the overall chair structure before directing attention to the wheels. Notably, participants evaluated all chair bodies before shifting focus to the wheels, suggesting a holistic visual processing approach. Fixation scores emphasized a concentration of gaze on the side view for both office and gaming chairs. Office chairs exhibited pronounced gaze fixation at the back, while gaming chairs showed minimal attention to this area. The higher gaze occupancy of the diagonal view for gaming chairs suggests that unique features, like backrest cushions, influenced participants' visual exploration, highlighting the importance of specific chair attributes in differential gaze patterns.

References

- [1] Gu, S., Ślusarczyk, B., Hajizada, S., Kovalyova, I., & Sakhbieva, A. (2021). Impact of the covid-19 pandemic on online consumer purchasing behavior. *Journal of Theoretical and Applied Electronic Commerce Research*, 16(6), 2263-2281.
- [2] Kawasaki, T., Wakashima, H., & Shibasaki, R. (2022). The use of e-commerce and the COVID-19 outbreak: A panel data analysis in Japan. *Transport Policy*, 115, 88-100.
- [3] Wedel, M., & Pieters, R. (2008). Eye tracking for visual marketing. *Foundations and Trends® in Marketing*, 1(4), 231-320.
- [4] Wedel, M., & Pieters, R. (2008). A review of eye-tracking research in marketing. *Review of marketing research*, 123-147.
- [5] Białowas, S., & Szyszka, A. (2019). Eye-tracking in marketing research. *Managing Economic Innovations-Methods and Instruments*, 1(69), 91-104.
- [6] Jeong, N. H., Lee, J., Yun, J. C., Park, D. H., & Park, S. B. (2023). Does wearing facial masks increase perceived facial attractiveness? An eye-tracking experiment. *Frontiers in Psychology*, 14, 1141319.
- [7] Han, Myoung-Heum, "Study on Emotional Evaluation of Furniture Design - Focusing on Folding Chair -", *Journal of the Korean Institute of Interior Design*, 30(3), 38-45, 2021.

Effect of Fan Systems for Lowering Leaf Temperature of Korean Ginseng

Byungsoon Kim, Myung Hwan Jang, and Ilkyeun Ra

Dept. of Computer Education, Andong National University, Gyungbook, South Korea

Punggi Ginseng Research Institute, Gyungbook, South Korea

Dept. of Computer Science and Engineering, University of Colorado Denver, Colorado, United States

[e-mail: bsgim@anu.ac.kr, hwanmj@korea.kr, ilkyeun.ra@ucdenver.edu]

*Corresponding author: Byungsoon Kim

Abstract

Due to climate change, elevated temperatures and recurring heatwaves have resulted in an annual increase in heat injury for ginseng, a shade-loving plant, due to heat stress. We studied how fan activation affects leaf temperature in Korean ginseng, comparing control and experimental groups with fans activated above 30°C. The results demonstrated a significant decrease in both ambient and leaf temperatures due to fan usage, as shown by lower median values in box plot analyses. The decrease, validated by Mann-Whitney U test outcomes, emphasizes the efficiency of fans in reducing leaf temperature.

Keywords: Korean ginseng, leaf temperature, fan, Mann-Whitney U test

1. Introduction

Effective heat stress mitigation in ginseng cultivation is critical, as high temperatures can significantly affect plant growth and yield. Therefore, it is essential to develop strategies to minimize heat damage to maintain the health and productivity of this temperature-sensitive crop. Agricultural practices are continually improved to optimize plant growth conditions, with ambient temperature being a crucial factor. High temperatures, particularly in the summer, can cause heat stress in plants, affecting their health and growth. In the cultivation of ginseng, it is crucial to maintain ideal leaf temperatures to ensure the plant's health and yield. However, high ambient temperatures pose a significant challenge. This study assesses the impact of activating fans on air and leaf temperatures in a ginseng cultivation environment. The study

compares a group with fans activated above 30 °C to a control group without fans.

2. Materials and Methods

We conducted the field experiments at Punggi Ginseng Research Institute, Gyungsangbuk-do, at the coordinates of latitude 36.4833° N, longitude 128.3230° E, and an altitude of 180 m above sea level. The experiment had two distinct groups, a control group and an experimental group, which is illustrated in [Fig. 1](#). Each group consists of two ridges, with each ridge measuring 24 m x 90 cm and planted with three rows of 4-year-old native strain ginseng. Three fans, set to activate at temperatures above 30°C, were installed on the experimental group.

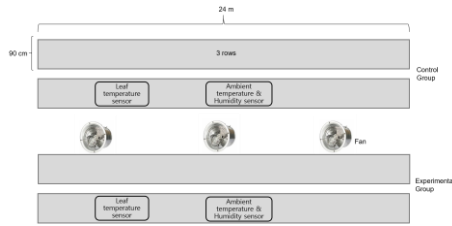


Fig. 1. Experimental testbed layout

The leaf temperature was measured using the LT-1T (Implexx Sense, Australia), and ambient temperature and humidity were measured using the AM2315C (ASAIR, China) sensors. To establish a network of sensors, we used LoRa (Long Range) technology. The measured data was transmitted and stored on a private server via LTE (Long Term Evolution), as shown in [Fig. 2](#) [1]. The data collection period extended from August 12 to September 10, 2023.

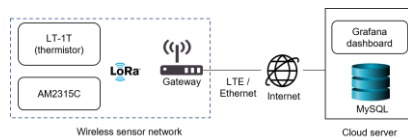


Fig. 2. Data collection system

3. Results and Discussion

[Fig. 3](#) shows box plots comparing the ambient and leaf temperatures, as well as the relative humidity, of the experimental and control groups under activated and non-activated fan conditions.

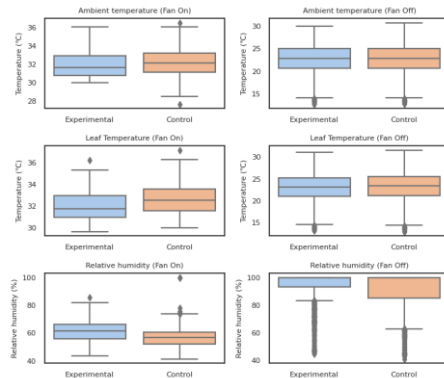


Fig. 3. Box plot of ambient temperature, leaf temperature and relative humidity

Under activated fan conditions, the ambient and leaf temperatures show a significant decrease compared to non-activated fan conditions. This

is evident from the lower median values in the activated fan scenarios, indicating that the use of fans effectively lowers the ambient and leaf temperatures. The range and spread of the data in the box plot also suggest that while there is variability in temperature, the overall effect of the fans is to create a cooler ambient environment. In activated fan conditions, the experimental group had slightly higher median humidity than the control group. Fans can alter the microclimate by affecting air movement and evaporation rates, which in turn can influence humidity levels.

Table 1. Mann-Whitney U test results

| Condition | Variable | U-statistics | p-value |
|---------------|---------------------|--------------|------------------------|
| Activated fan | Ambient temperature | 51,901.5 | 0.00031 |
| | Leaf temperature | 44,474.5 | 1.83×10^{-10} |
| | Relative humidity | 83,402.5 | 4.85×10^{-16} |

We conducted the Mann-Whitney U test [2] on the three variables as shown in [Table 1](#). For air temperature, leaf temperature, and relative humidity, the p -values are all less than the significance level of 0.05. We can see that there is a significant difference in the distributions of these variables between the fan-activated control group and the fan-activated experimental group.

4. Conclusions

Our study has proven that activating fans effectively reduces leaf temperatures, thereby mitigating high temperature damage in ginseng. Upon activation in the experimental group, a significant reduction in both ambient and leaf temperatures was observed.

References

- [1] B. Kim, "Assessing accuracy over warm-up time of Lepton 3.5 thermal imaging for measuring leaf temperature of crops," *J. of Applied Horticulture*, Vol 25, No. 1, pp. 39-42, 2023.
- [2] H.B. Mann & D.R. Whitney, "On a test of weather one of two random variables is stochastically larger than the other," *Annals of Mathematical Statistics*, Vol 18, pp.50-60, 1947.

A Compound Gateway-Activity Discovery Approach in Process Mining

Thanh-Hai Nguyen¹, Jin-Hyung Lee¹, Batt Chaya¹, Eun-Bee Cho¹, Sang-Eun Ahn², Kyung-Hee Sun², Kyoung-Sook Kim², Dinh-Lam Pham², Kwanghoon Pio Kim^{1,2*}

¹ Data and Process Engineering Research Lab.

Department of Computer Science, Graduate School

KYONGGI UNIVERSITY

154-42, Kwangkyosan-ro, Youngtong-gu, Suwon-si, Gyeonggi-do, 16227, Republic of KOREA

e-Mail: {nguyenthanhhai, batt, ebcho, kwang}@kgu.ac.kr

² Contents Convergence Software Research Institute

KYONGGI UNIVERSITY

154-42, Kwangkyosan-ro, Youngtong-gu, Suwon-si, Gyeonggi-do, 16227, Republic of KOREA

e-Mail: {sangeun1804, kh_sun, khmjmc, phamdinhlam, kwang}@kgu.ac.kr

*Corresponding author: Kwanghoon Pio Kim

Abstract

In this paper, we focus on exploring compound gateway-activity discovery possibility of the SICN-oriented process mining algorithm and its implemented system. The compound gateway-activity discovering issue is the most challengeable issue in the process mining literature. We had developed the process mining algorithm, which is named as *p-Algorithm* and is able to rediscover a structured information control net process model from an IEEE-XES-formatted dataset of process enactment event logs. However, the algorithm has a functional limitation in discovering the so-called compound gateway-activities that are formed in a multi-level and complex control-flow structure of split-join types of gateway-activity transitions such as exclusive, parallel, and iterative gateway-activity transitions. This paper, therefore, tries to functionally upgrade the *p-Algorithm* to properly deal with the challengeable issue through conceiving a novel approach of rediscovering such compound gateway-activities from a dataset of process enactment event logs formatted in the IEEE-XES standardized and extensible event stream format and fulfilling a couple of verification and validation experiments, as well.

Keywords: process enactment events log; process mining; compound gateway-activity rediscovery; process intelligence; process-related knowledge discovery

1. Introduction

In recent, we had developed the process rediscovery algorithm, which was named as “*p-Algorithm* [1],” and implemented its system for rediscovering a process model from a dataset of process execution event logs. In principle, the *p-Algorithm* is able to rediscover all the primitive process patterns, such as sequential, disjunctive (exclusive-OR transition), conjunctive (parallel-

AND transition), and iterative (LOOP transition) control-flow gateway-activities. However, the algorithm has a functional limitation in discovering the so-called compound gateway-activities that are formed in a multi-level and complex control-flow structure of split-join types of gateway-activity transitions such as exclusive-OR, parallel-AND, and iterative-LOOP gateway-activity transitions, which is the most challengeable research issue in the process

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT, Ministry of Science and ICT), Republic of Korea (Grant No. NRF-2022R1A2C2093002).

discovery and mining literature [1][2][3][4]. In this paper, therefore, we try to functionally upgrade the p -Algorithm to properly deal with the challengeable issue through conceiving a novel approach of rediscovering such compound gateway-activities from a dataset of process enactment event logs formatted in the IEEE-XES standardized and extensible event stream format [5] and fulfilling a couple of verification and validation experiments, as well.

2. Structural Formations of Compound Gateway-Activities

The eventual goal of the process mining approaches supported by the information control net process modeling methodology [7][8] is to

discover the complete structural formations of compound gateway-activities with keeping the sound and safe process requirements of matched-pairing and proper-nesting properties. Therefore, in this section, we describe the conceptual definition of the compound control-flow gateway-activities. In forming a structured information control net process model, the compound gateway-activity implies a structural formation of multi-level and complex control-flow gateway-activity (abbreviated as ML-CCFG) combinations of the primitive split-join gateway-activities. Fig. 1 depicts the six split-join (open-close) structural formations of compound control-flow gateway-activities, including the split-join pairings of Parallel-AND, Exclusive-OR, and Iterative-LOOP compound gateway-activities.

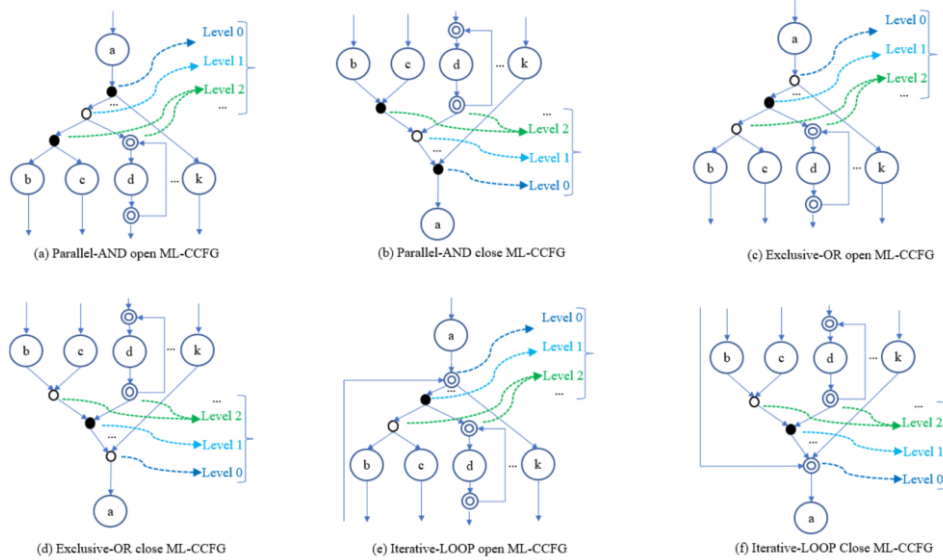


Fig. 1. Six Open-Close Structural Formations of Compound Control-Flow Gateway-Activities

3. Compound Gateway-Activity Discovery Algorithm and Experiment

In this section, we try to expand and upgrade the operational functionality and feasibility of the p -Algorithm [1] so as to properly deal with the six open-close structural formations of the compound control-flow gateway-activities with multi-level and complex combinational formations described in the previous section. Additionally, we carry out a couple of experiments to verify the upgraded p -Algorithm and its implemented process mining system. We

carried out a series of experiments on three process event log datasets [6][7][8] to evaluate the operational correctness and effectiveness of the proposed upgraded p -Algorithm.

Fig. 2 and Fig. 3 are the discovered process models discovered from the second dataset of the Review Example Large.xes by using the original p -Algorithm and the upgraded p -Algorithm, respectively. As you can see, the upgraded p -Algorithm can discover one of the structural formations of ML-CCFGs designed in the paper. Conclusively, we verified and proved the operational correctness and functional effectiveness of the upgraded p -Algorithm.

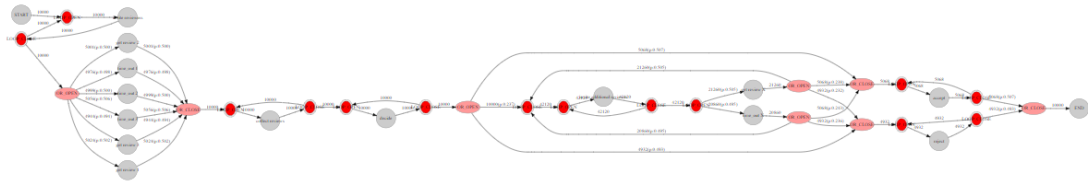


Fig. 1. ρ -Algorithm: The Discovered Gateway-Activities in the ICN-Process Model from the Review Example Large Dataset

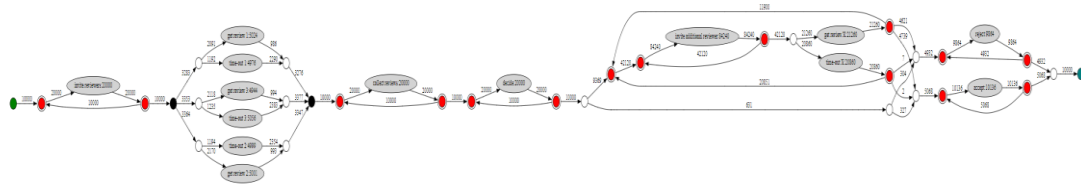


Fig. 2. Upgraded ρ -Algorithm: The Discovered Multi-Level Compound Control-Flow Gateway-Activities in the ICN-Process Model from the Review Example Large Dataset

4. Conclusions

In this paper, we proposed an approach of discovering compound control-flow gateway-activities that are characterized by the structural formations of multiple level and complex combinational gateway-activities. Especially, its implemented algorithm and system is extended and upgraded from the original ρ -Algorithm. Our experimental verification with the datasets shows that the upgraded ρ -Algorithm can effectively discovers the underlying information control net process model from process event log datasets. Surely, our work in this paper makes a valuable contribution to the progressive evolvement in the field of process mining.

Acknowledgement

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT, Ministry of Science and ICT), Republic of Korea (Grant No. NRF-2022R1A2C2093002).

References

- [1] K. S. Kim, D. L. Pham, and K. P. Kim, “ ρ -Algorithm: A SICN-Oriented Process Mining Framework,” *IEEE Access*, vol. 9, pp. 139 852–139 875, 2021.
- [2] W. M. Van der Aalst and A. J. Weijters, “Process mining: A research agenda,” *Computers in Industry*, vol. 53, no. 3, pp. 231–244, 2004.
- [3] W. Van Der Aalst, T. Weijters, and L. Maruster, “Workflow mining: Discovering process models from event logs,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 16, no. 9, pp. 1128–1142, 2004.
- [4] K. P. Kim and C. A. Ellis, “ σ -algorithm: Structured workflow process mining through amalgamating temporal workcases,” in *PacificAsia Conference on Knowledge Discovery and Data Mining*, 2007.
- [5] S. J. van Zelst, B. F. van Dongen, W. M. Van der Aalst, and H. M. Verbeek, “Discovering workflow nets using integer linear programming,” *Computing*, vol. 100, no. 5, pp. 529–556, 2018.
- [6] C. W. Günther and W. M. P. van der Aalst, “Fuzzy mining – adaptive process simplification based on multi-perspective metrics,” in *Business Process Management*, G. Alonso, P. Dadam, and M. Rosemann, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, pp. 328–343.
- [7] X. Wang, Z. Cao, R. Zhan, M. Bai, Q. Ma, and G. Li, “Density-based Outlier Detection in Multi-dimensional Datasets,” *KSII Transactions on Internet and Information Systems*, vol. 16, no. 12, pp. 3815–3835, 2022.
- [8] S. Liu, H. Qiao, L. Yuan, Y. Yuan, and J. Liu, “Research on data augmentation algorithm for time series based on deep learning,” *KSII Transactions on Internet and Information Systems*, vol. 17, no. 6, pp. 1530–1544, 2023.

Chat-GPT application for AICC of the Bank

Sudan Prasad Uprety*, **Yoosin Kim****

* AICC Solution Lab, Kathmandu, Nepal

** RTData Lab, Jeju, South Korea

[e-mail: sudan-u@gmail.com, iamyoosin@gmail.com]

*Corresponding author: Yoosin Kim

Abstract

Speech understanding and Dialog Systems are now a global trend and with the advancement of artificial intelligence and machine learning technic, has drawn attention from both academic and business communities. This paper explores the integration of ChatGPT, a cutting-edge language model developed by OpenAI, with the Artificial Intelligence Contact Center (AICC) system in the banking sector. The synergy between ChatGPT and AICC has the potential to revolutionize customer interaction by providing personalized, efficient, and human-like responses to customer queries. This study delves into the technical implementation, benefits, and challenges of incorporating ChatGPT within the AICC framework to enhance the overall customer experience in banking.

Keywords: NLP, AICC Dialog Systems, Joint Modeling, Chatbot

1. Introduction

In recent years, artificial intelligence has made significant strides, and ChatGPT stands as a testament to the advancements in natural language processing. Developed by OpenAI, ChatGPT is a state-of-the-art language model built on the GPT-3.5 architecture. After existence of ChatGPT, organizations consider to apply it as their business functions. Especially, the banking industry is undergoing a digital transformation, and customer interaction plays a pivotal role in this evolution[1]. The integration of ChatGPT with AICC aims to elevate the quality of customer communication by leveraging advanced natural language processing capabilities.

2. AICC Integration with ChatGPT

ChatGPT, a sophisticated language model created by OpenAI known as GPT-3.5, assist people by generating human-like text based on

the input what people provides. Whether people needs help with information, creative writing, programming assistance, or just wants to have a conversation, ChatGPT can engage with whom. However, ChatGPT 3.5 doesn't have personal experiences or emotions, and it's knowledge is limited to what was available up until January 2022)[2].

AICC refers to a contact center or customer service operation that leverages artificial intelligence (AI) technologies to enhance various aspects of customer interactions and support. The integration of AI into contact centers aims to improve efficiency, provide personalized customer experiences, and streamline communication processes. Key components and functionalities associated with AI Contact Center include Chatbots and Virtual Assistants, Natural Language Processing(NLP), Automated Speech Recognition(ASR), etc.

● ChatGPT and AICC Integration

This section explores the technical aspects of integrating ChatGPT with the AICC system. It

examines how ChatGPT's language generation capabilities can enhance the AICC's ability to understand and respond to customer queries effectively. Additionally, the paper discusses the potential use of ChatGPT for automating routine customer interactions and improving response times.

- **Personalized Customer Engagement**

One of the key advantages of incorporating ChatGPT into AICC is the ability to provide personalized customer interactions. This section investigates how the model can analyze customer data and transaction histories to offer tailored recommendations, assistance, and information.

- **Efficiency Gains**

By automating and streamlining customer interactions, the integration of ChatGPT with AICC has the potential to significantly improve operational efficiency. This section examines case studies and simulations to quantify the efficiency gains achieved through the implementation of this innovative approach.

- **Challenges and Mitigations**

While the integration of ChatGPT with AICC presents numerous benefits, challenges such as ethical considerations, potential biases, and model limitations must be addressed. This section outlines strategies and best practices for mitigating these challenges to ensure responsible and fair use of the technology.

- **Customer Trust and Security**

Maintaining customer trust and ensuring the security of sensitive information are paramount in the banking sector. The paper discusses how the integration of ChatGPT with AICC aligns with data protection regulations, and it explores strategies for building and maintaining customer trust in AI-driven interactions.

The integration of ChatGPT with AICC in the banking sector holds significant promise for revolutionizing customer interactions. By combining advanced natural language processing with automated communication systems, banks can deliver a more personalized, efficient, and engaging experience for their customers. This paper contributes to the understanding and implementation of this transformative technology in the context of banking customer service.

References

- [1] K. Nam, Z. Lee, B. G. Lee, "How Internet has Reshaped the User Experience of Banking Service?," *KSII Transactions on Internet and Information Systems*, vol. 10, no. 2, pp. 684-702, 2016. DOI: 10.3837/tiis.2016.02.014
- [2] OpenAI and ChatGPT service website. <http://www.openai.com>
<http://chat.openai.com>

3. Conclusion

The paper concludes with a discussion on the future directions of AICC and ChatGPT integration in the banking sector. It explores potential advancements, additional use cases, and ongoing research initiatives to further optimize and expand the capabilities of this innovative approach.

Application of Block-Chain Technology in Digital Music Copyright Management: A Case Study of VNT Chain Platform

Mingsheng Wang¹

¹ Anshan Normal University, 43 Ping 'an Street, Tiedong District, Anshan City, Liaoning Province, China; wangmingsheng521@hotmail.com

Abstract

In this paper, we design and construct a digital music copyright management system using the VNT Chain blockchain platform, which utilizes blockchain technology to provide proof of copyright and evidence storage, Shazam algorithm to provide proof of originality for music copyright, and smart contracts to ensure secure and reliable transactions. The system includes six functional modules: user management, copyright registration, copyright trading, infringement monitoring, evidence storage, and music ecology, covering the main aspects of copyright management. We use blockchain, InterPlanetary File System (IPFS), and MySQL as storage engines, depending on the different requirements of business data. Experimental results show that the registration time for each piece of music increased by approximately 1.9 seconds and the average feature fingerprint data for each piece of music stored on IPFS consumed approximately 8 MB, which meets the expected system performance requirements.

Keywords: Music Copyright; Blockchain; Shazam Algorithm; Music Ecology; InterPlanetary File System (IPFS)

1. Introduction

In the digital era, digital product copyright issues have emerged, particularly in the music industry disseminated online. Digital copyright involves creators' rights to store, copy, and distribute digital works online. Copyright can be acquired automatically or through registration, as per Chinese Copyright Law. The Chinese digital music industry faces challenges like copyright identification, infringement monitoring, evidence collection, and royalty distribution. Despite blockchain technology attempts, challenges remain.

This paper explores the VNT Chain platform's potential in managing digital music copyright in China. It discusses industry challenges, reviews blockchain research in digital copyright, and

introduces the VNT Chain platform's capabilities.

Researchers have examined blockchain for digital copyright management challenges, like infringement and royalty distribution. However, limitations persist, including scalability and regulatory compliance, and many studies overlook aspects like various digital work types and the full copyright service process.

The VNT Chain platform offers a blockchain-based solution for these challenges, encompassing all aspects of copyright services. The paper proposes a blockchain-based system focusing on the music industry, using the Shazam algorithm for identifying and preventing pirated works. It covers six functional modules and leverages Token economics on the VNT Chain public chain to incentivize creators and

address funding shortages. The system aims to regulate digital music copyright, stimulate creativity, and enhance the industry's prosperity..

2. Methods

2.1 Key technical

Blockchain's rise, exemplified by VNT Chain, offers secure data handling, complemented by IPFS for peer-to-peer file sharing. Smart contracts and the Shazam algorithm, a key audio identification tool, are also integral. This paper proposes 'digital DNA' for music file identification, using unique features like the Mel Frequency Cepstrum Coefficient (MFCC). Techniques like Echo print, Chroma print, and Landmark, particularly the Shazam algorithm, are explored for creating unique fingerprints. The Shazam algorithm uses feature extraction and FFT hashing for fast, reliable music file identification. We discuss managing music files in a decentralized environment using these technologies.

2.2 System requirements

As shown in Fig. 1, The blockchain-based digital music copyright management system authenticates music works and proves originality using the Shazam algorithm. It features smart contracts for secure transactions and authentic infringement evidence, plus a Token economy for music ecosystem growth. The system has six modules: user management, copyright registration, transactions, infringement monitoring, evidence preservation, and music ecosystem. User management handles login, registration, and personal info. Copyright registration allows users to register and download their work. The transaction module facilitates copyright sales and purchases. Infringement monitoring identifies pirated music online, and evidence preservation stores infringement files on-chain for legal protection. The music ecosystem module promotes industry prosperity, allows music promotion for revenue sharing, and supports crowdfunding for new music creation..

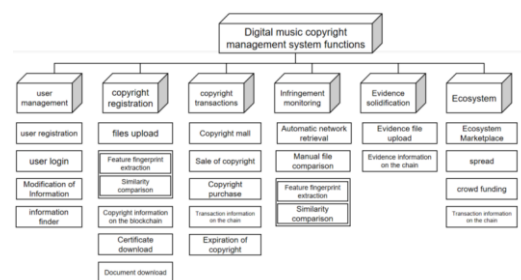


Fig. 1. Digital music copyright management system functions

3. System design

3.1 Design of System Architecture

The blockchain-based digital music copyright management system's architecture comprises three layers: front-end, back-end, and underlying storage. The storage layer includes blockchain platforms, IPFS, and databases. Blockchain ensures tamper-proof, traceable data ideal for judicial evidence but requires digital currency for storage, thus used for crucial data only. IPFS offers distributed file storage with security features but has privacy risks due to accessible content via hash addresses. MySQL provides easy operation and access but has higher security risks due to its centralized nature.

The back-end bridges storage and front-end, handling business logic, both smart contract and non-smart contract implementations. Smart contracts cover user management, copyright registration, transactions, evidence solidification, and music ecosystem. Non-smart contract tasks include infringement monitoring and Shazam algorithm execution.

The front-end, using Web with Vue.js and Element UI frameworks, provides user services via an interactive interface and connects to the back-end through Axios..

3.2 functional module

This article introduces a blockchain-based system for music copyright protection with six modules:

1. User Management: Manages user information, records personal data on the VNT Chain blockchain, and requires wallet registration on the VNT Chain website.
2. Copyright Registration: Verifies and registers music ownership, comparing tag information and user data, with the music file's

hash value stored as proof of ownership on the blockchain.

3. Copyright Transaction: Facilitates secure and transparent copyright buying and selling via smart contracts on the VNT Chain platform.

4. Infringement Monitoring: Detects copyright infringements by analyzing music file metadata, with automatic and manual monitoring options.

5. Evidence Solidification: Stores digital evidence of copyright disputes securely on IPFS, with hash values recorded on the blockchain for authenticity.

6. Music Ecosystem: Builds a vibrant music industry, offering services like music streaming, discovery, and artist promotion, and includes music promotion, crowdfunding, and a marketplace.

The system ensures music file authenticity and ownership, monitors and prevents copyright infringements, and provides a comprehensive range of music-related services. Smart contracts written in C language using vntlib.h header file are used for most modules, enhancing security and reliability.

4. 4. Performance Testing and Analysis

To assess its performance, the system was tested on a setup with an Intel Core i5-7300HQ CPU, 16GB RAM, Windows 10, and VMware hosting Ubuntu 16.04. Focusing on the copyright registration module, the most resource-intensive due to its use of the Shazam algorithm, the system operates on the VNT Chain test network.

For performance testing, a 10-second audio segment is used for similarity comparison, based on the Shazam algorithm's high confidence level for such clips. The system's database requires feature fingerprint data from entire music files.

In the test, 50 internet-downloaded music files were registered individually. 45 files were successfully registered, indicating a 90% success rate. The five failed files, numbers 10, 22, 23, 25, and 27, only failed post-similarity comparison, reflecting a 100% accuracy rate in this step. Their failure was due to insufficient memory allocation for extracting feature fingerprints from their approximately 5-minute durations. This issue could be addressed by modifying system configurations to increase virtual

memory and processing feature fingerprint data in batches.

Fig. 2 shows the registration time for each of the 45 successfully registered music files.

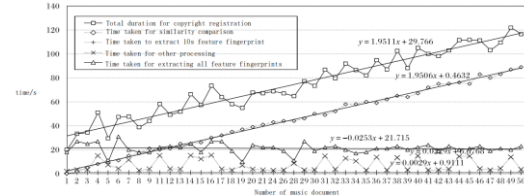


Fig. 2. Time spent on copyright registration

In **Fig. 4**, the slope for total registration time increases by about 1.95 seconds per music file, mainly due to longer similarity comparison times. Extracting feature fingerprints from 10-second segments takes about 1 second, while it takes roughly 22 seconds for whole files. The similarity comparison, nearly matching the total time increase, is the key factor in registration time growth. Additional time is needed for checking music file formats and converting non-compliant files, contributing to variations in processing times.

Fig. 3 further details the factors increasing similarity comparison time.

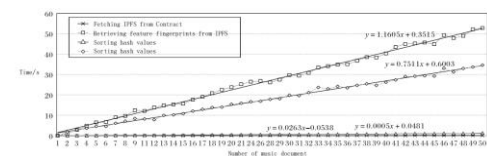
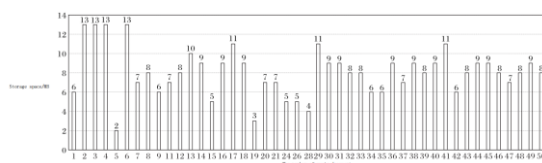


Fig. 3. Factors of copyright registration time increase

Analyzing **Fig. 5**, the combined slopes of the four lines roughly equal 1.95, matching the similarity comparison time slope from **Fig. 4**. The time increases caused by fetching IPFS addresses of feature fingerprints and sorting tens of thousands of hash values per fingerprint are relatively minor. Hash value matching uses a multi-process and binary search approach, with its logarithmic time complexity minimizing time increases as data grows. The most time-consuming aspect is retrieving feature fingerprint data from IPFS, taking about 1.1 seconds per music file's hash address. Future optimizations in data storage and transmission are planned to reduce registration time and improve system efficiency.



Of the 45 successfully registered music files, their feature fingerprint data on IPFS consumes about 8 MB per file, as shown in Figure 6, with the number of hash values per fingerprint displayed in Fig. 7. This storage size, though 2-3 times larger than the original music files, follows the Shazam algorithm's space-for-time trade-off, enhancing similarity comparison speed. The proposed system, compared to previous ones in Table 1, offers more comprehensive functionality, delivering end-to-end copyright services for efficient music copyright management.

With the improvement of people's living standards, an increasing number of individuals are inclined towards listening to music or creating their own. However, numerous issues in the existing music copyright industry pose significant obstacles to the development of the music industry. The emergence of blockchain technology provides a promising solution for music copyright management. In this study, we address the problems in the current music copyright industry and develop a digital music copyright management system based on blockchain technology. The system encompasses six major functional modules that can meet the

- [1] National Copyright Administration of the People's Republic of China. Copyright law of the People's Republic of China [EB/OL]. [2020-05-01]. <http://www.ncac.gov.cn/chinacopyright/contents/479/17542.html>.
- [2] Savelyev, A. Copyright in the blockchain era: promises and challenges. *Computer Law and Security Review*, 2018, 34(3): 550-561.
- [3] Zhang, B. Q., & Li, L. Balance of digital copyright interests based on blockchain technology. *China Publishing Journal*, 2019(11): 22-25.
- [4] Nakamoto, S. Bitcoin: a peer-to-peer electronic cash system [R/OL]. [2020-05-01]. <https://bitcoin.org/bitcoin.pdf>.
- [5] Chen, W. L., & Zheng, Z. B. Blockchain data analysis: a review of status, trends and challenges. *Journal of Computer Research and Development*, 2018, 55(9): 1853-1870.
- [6] Li, C., Dai, B. R., Wang, H. J., et al. Digital copyright protection and trading system based on blockchain technology. *Modern Computer*, 2018(29): 80-84.
- [7] Li, L., Zhou, S. Q., Liu, Q., et al. Blockchain-based digital copyright trading system. *Chinese Journal of Network and Information Security*, 2018, 4(7): 22-29.
- [8] Zhao, G. F., He, Y., Zhou, J. H. Digital copyright registration technology based on blockchain. *Information Technology and Network Security*, 2019, 38(4): 79-83.
- [9] Ayekum, K. O. B. O., Xia, Q., Liu, Y., et al. Digital media copyright and content

- protection using IPFS and blockchain. In Proceedings of the 2019 International Conference on Image and Graphics, LNCS 11903. Cham: Springer, 2019: 266-277.
- [11] Chen, Z. H., Li, Q., Gan, J., et al. VC Chain: an alliance audio-video copyright blockchain system. *Computer Engineering and Science*, 2019, 41(11): 1939-1948.
- [12] Pan, Y., Chen, Q., & Zhang, X. (2019). Research on the Application of Blockchain Technology in Copyright Protection. In 2019 IEEE International Conference on Intelligence and Security Informatics (ISI) (pp. 45-48). IEEE.
- [13] Zhang, X., Zhang, W., Xu, X., & Huang, J. (2020). A Blockchain-Based Copyright Protection Framework for Digital Media Assets. *IEEE Access*, 8, 27274-27284.
- [14] Zheng, Z., Xie, S., Dai, H. N., Chen, X., & Wang, H. (2017). An overview of blockchain technology: Architecture, consensus, and future trends. *IEEE International Congress on Big Data*, 557-564.
- [15] Kuo, T. T., & Ohno-Machado, L. (2017). ModelChain: Decentralized Privacy-Preserving Healthcare Predictive Modeling Framework on Private Blockchain Networks. *JMIR Medical Informatics*, 5(4), e46.
- [16] Yao, Y., Hu, C., Zhao, L., Fan, Y., & Yuan, Y. (2018). A Blockchain-Based Copyright Protection Framework for Digital Assets. In 2018 IEEE International Conference on Big Data and Smart Computing (BigComp) (pp. 609-613). IEEE.
- [17] Kshetri, N., & Voas, J. (2018). Can Blockchain Strengthen the Internet of Things? *IT Professional*, 20(3), 68-72.
- [18] Cachin, C. (2016). Architecture of the Hyperledger Blockchain Fabric. *Proceedings of the 2nd Workshop on Scalable and Resilient Infrastructures for Distributed Ledgers*, 1-6.
- [19] Swan, M. (2015). *Blockchain: Blueprint for a New Economy*. O'Reilly Media.
- [20] Zheng, Z., Xie, S., Ning, H., & Hao, F. (2018). An Overview of Consensus Algorithms Used in Blockchain. In 2017 IEEE International Congress on Big Data (BigData Congress) (pp. 380-387). IEEE.
- [21] Li, M., Jiang, P., Li, Y., Zhang, L., & Cai, Z. (2017). A Survey on the Security of Blockchain Systems. *Future Generation Computer Systems*, 107, 841-853.
- [22] Swan, M. (2017). *Blockchain: The Complete Guide to Understanding Blockchain Technology*. CreateSpace Independent Publishing Platform.
- [23] Dagher, G. G., Mohler, J., Milojkovic, M., Marella, P. B., & Fouad, M. M. (2018). Ancile: Privacy-preserving framework for access control and interoperability of electronic health records using blockchain technology. *Sustainable Cities and Society*, 39, 283-297.
- [24] Wang, H., & Wang, X. (2017). CyberMiles: A Next-Generation Blockchain for Smart Contracts Optimization. *arXiv preprint arXiv:1705.00647*.
- [25] Di Francesco Maesa, D., Di Francesco, M., & Kosta, E. (2018). Blockchain-Based Access Control for Secure Energy Trading in Industrial IoT. *IEEE Transactions on Industrial Informatics*, 14(8), 3690-3699.
- [26] Nakamoto, S. (2008). *Bitcoin: A Peer-to-Peer Electronic Cash System*. Retrieved from <https://bitcoin.org/bitcoin.pdf>
- [27] Tapscott, D., & Tapscott, A. (2016). *Blockchain Revolution: How the Technology Behind Bitcoin Is Changing Money, Business, and the World*. Penguin.
- [28] Swan, M. (2018). Tokens and Digital Securities: Blockchain Technology and the SEC. *Journal of Business, Entrepreneurship and the Law*, 11(1), 47-73.
- [29] Chen, Y., Sun, X., & Xu, J. (2019). A Decentralized Architecture for Music Copyright Management Based on Blockchain Technology. In 2019 18th International Symposium on Distributed Computing and Applications for Business Engineering and Science (DCABES) (pp. 56-60). IEEE.
- [30] Christidis, K., & Devetsikiotis, M. (2016). Blockchains and Smart Contracts for the Internet of Things. *IEEE Access*, 4, 2292-2303.
- [31] Drescher, D. (2017). *Blockchain Basics: A Non-Technical Introduction in 25 Steps*. Apress.
- [32] Ben S. Bernanke, 2007. "Monetary policy under uncertainty," Speech 327, Board of Governors of the Federal Reserve System (U.S.).