AICCC & ADIP 2024

2024 7th Artificial Intelligence and Cloud Computing Conference (AICCC 2024)
2024 6th Asia Digital Image Processing Conference (ADIP 2024)

https://www.aiccc.net
https://adip.org

December 14-16, 2024 | Tokyo, Japan

Supported by

Indexed by

Table of contents

Welcome Message

On behalf of the organizing committee, it's our great pleasure and honor to welcome you all to attend the 2024 7th Artificial Intelligence and Cloud Computing Conference (AICCC 2024) and the 2024 6th Asia Digital Image Processing Conference (ADIP 2024), which will be held in Tokyo, Japan during December 14th to 16th, 2024.

We are gathered here at this significant event to explore and exchange the latest advancements, ideas, and research findings in the fields of artificial intelligence, cloud computing, and digital image processing. These two conferences have always served as important platforms for professionals, scholars, and industry leaders from around the world to come together, share their insights, and collaborate.

We're confident that over the two days you'll get the theoretical grounding, practical knowledge, and personal contacts that will help you build long-term, profitable and sustainable communication among researchers and practitioners working in a wide variety of scientific areas with a common interest in Digital Image Processing, Artificial Intelligence and Cloud Computing Conference.

We received submissions from all over the world, including Thailand, India, Japan, China, Czech Republic, Korea, Germany, United States, Pakistan, Finland, Sweden, Estonia, Vietnam, Philippines, UK, Switzerland and many other countries. We have obtained the support of reviewers worldwide. They have volunteered time in helping us review papers and giving constructive suggestions to these papers.

During the conference, there are four Keynote Speeches, four Invited Speeches and eleven sessions. We truly hope you will enjoy the conference and get what you expect from it.

Once again, we sincerely would like to thank all the authors as well as the technical program committee members and reviewers. The reviewers' high competence, enthusiasm, time and expertise knowledge, enabled us to prepare the high-quality final program and helped to make the conference become a successful event.

AICCC&ADIP Organizing Committee
December, 2024
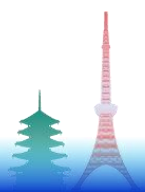
Conference Committees

| Conference Co-Chairs

Seiichi Ozawa, Kobe University, Japan
Wenbing Zhao, Cleveland State University, USA
Kiyoshi Hoshino, Meiji University, Professor Emeritus of University of Tsukuba, Japan
Chinthaka Premachandra, Shibaura Institute of Technology, Japan

| Program Co-Chairs

Kenji Suzuki, Institute of Integrated Research at Institute of Science Tokyo, Japan
Tomoaki Ohtsuki, Keio University, Japan
John P. Sullins, Sonoma State University, USA
Shuichi Onami, RIKEN Center for Biosystems Dynamics Research, Japan
Seokwon Yeom, Daegu University Gyeongsan, Korea
Chia-Hung Yeh, National Taiwan Normal University

| Steering Co-chairs

Yuji Sekiya, The University of Tokyo, Japan
Dmitriy K. Levonevskiy, Russian Academy of Sciences, Russia
Takeshi Saitoh, Kyushu Institute of Technology, Japan
Keisuke Kameyama, University of Tsukuba, Japan

| Chapter Chairs

Montathar Faraon, Kristianstad University, Sweden
Katsuhiro Honda, Osaka Metropolitan University, Japan
Rebeka Sultana, Tokyo University of Agriculture and Technology, Japan
Qiang (Shawn) Cheng, University of Kentucky College of Engineering, USA

| Publicity Chairs

Ting-Sheng Weng, NCYU
Azizi Bin Ab Aziz, Universiti Utara Malaysia, Malaysia
Tetsuya Shimamura, Saitama University, Japan
Ruiheng Zhang, Beijing Institute of Technology, China
Amir Akramin bin Shafie, International Islamic University Malaysia, Malaysia

4
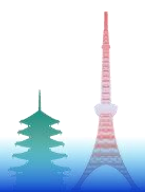
| Publication Chair

Hamimah binti Ujir, Universiti Malaysia Sarawak, Malaysia

| Technical Committee Chairs

Yulia Kumar, Kean University, USA
Huiyu Zhou, University of Leicester, UK
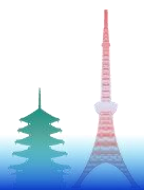
| Technical Program Committees

Abdel-Badeeh M. Salem, Ain Shams University, Egypt
Abdus Salam Siddique, Arkansas Tech University, USA
Ahmad Al Smadi, Zarqa University, Jordan
Ahmad Termimi Ab Ghani, Universiti Malaysia Terengganu, Malaysia
Akchay Srivastava, Apple, USA
Alaa Sheta, Southern Connecticut State University, USA
Alexandre, A., And Veyrie, Ecole Nationale de l'Aviation Civile, University of Toulouse, France
Amine Khaldi, Universite Kasdi Merbah Ouargla, Algeria
Amro Al-Said Ahmad, Keele University, UK
Angela Maria Alves, University of Campinas (Unicamp) / Renato Archer Information Technology Center, Brazil
Archit Yajnik, Sikkim Manipal Institute of Technology, India
Ashish Shenoy, Google Inc, United States
Beligiannnis Grigorios, University of Patras, Greece
Chang-Hsing Lee, Ming Chi University of Technology
Chansiri Singhtaun, Kasetsart University, Bangkok, Thailand
Cheng Siong Chin, Newcastle University in Singapore, Singapore
Cheng-Hsien Hsieh, Singapore University of Social Sciences, Singapore
Chien-Chih Wang, Ming Chi University of Technology
Ching-Chun Chang, National Institute of Informatics, Japan
Choo Peng Tan, Multimedia University, Malaysia
Chuan-Ming Liu, NTUT (Taipei Tech)
Costantin Volosencu, Politehnica University of Timisoara, Romania
CRISELLE CENTENO, Pamantasan ng Lungsod ng Maynila, Philippines
Criselle J. Centeno, MIT, Philippines
David Bong, Universiti Malaysia Sarawak, Malaysia
Dickson K.W. Chiu, The University of Hong Kong, China
Elmer R. Magsino, De La Salle University, Philippines
Firkhan Ali Bin Hamid Ali, Universiti Tun Hussein Onn Malaysia, Malaysia
Frank Werner, Otto-von-Guericke-University, Germany
Fusaomi Nagata, Sanyo-Onoda City University, Japan
Gabriel Gomes de Oliveira, University of Campinas (Unicamp) / Renato Archer
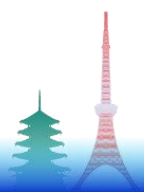
Information Technology Center, Brazil
Hamimah binti Ujir, Universiti Malaysia Sarawak, Malaysia
Hang Peng, Tongji University, China
Hector Rafael Morano Okuno, Tecnologico de Monterrey, Mexico
Herwig Unger, FernUniversität in Hagen, Germany
Igor Kotenko, St. Petersburg Institute for Informatics and Automation of the
Russian Academy of Sciences, Russia
Ilija Basicevic, University of Novi Sad, Serbia
Iouliia Skliarova, University of Aveiro, Portugal
Isao Nishihara, Toyama Prefectural University, Japan
J.M. Corchado, University of Salamanca, Spain
Jan Kubicek, VSB-Technical University of Ostrava, Czech Republic
Janelle Kyra A. Sagum, Polytechnic University of the Philippines, Philippines
Janusz Kacprzyk, Polish Academy of Sciences, Poland
Junfeng Qu, Clayton State University, USA
Kameyama Keisuke, University of Tsukuba, Japan
Khondker Shajadul Hasan, University of Houston-Clear Lake, USA
Koorosh Gharehbaghi, RMIT University, Australia
Lee Chien Sing, Sunway University, Malaysia
Li Yu, University of Electronic Science&Technology of China, China
Liang-Bi Chen, NPU
Linqiang Ge, Columbus State University, USA
Ljiljana Trajkovic, Simon Fraser University, Canada
Maki Habib, The American University in Cairo, Egypt
Malik Zawwar Hussain, University of the Punjab, Pakistan
Malinka Ivanova, Technical University of Sofia, Bulgaria
Man Fung LO, The University of Hong Kong, Hong Kong SAR, China
Marcin Paprzycki, Polish Academy of Sciences, Poland
Mariofanna Milanova, University of Arkansas, USA
MARK ANTHONY MERCADO, Pamantasan ng Lungsod ng Maynila, Philippines
Maytham Alabbas, University of Basrah, Iraq
Michael B. Dela Fuente MSGITS, Polytechnic University of the Philippines,
Philippines
Mobyen Uddin Ahmed, School of Innovation Design and Engineering,
Mälardalen University, Västerås, Sweden
Montathar Faraon, Kristianstad University, Sweden
Nittaya Kerdprasop, Suranaree University of Technology, Thailand
Noorlin Mohd Ali, Universiti Malaysia Pahang, Malaysia
Olarik Surinta, Mahasarakham University, Thailand
Paolo Bellavista, PUniversity of Bologna, Italy
Pavlo Maruschak, Ternopil Ivan Puluj National Technical Universit, Ukraine
Pei-Chun Lin, Feng Chia University
Qiang (Shawn) Cheng, University of Kentucky College of Engineering, USA
Qiu Chen, Kogakuin University, Japan
Raenu A/L Kolandaisamy, UCSI University, Malaysia

Ranran Feng, University of Texas at Dallas, USA
Renjith V Ravi, Rashtriya Raksha University, India
Ria A. Sagum, Polytechnic University of the Philippines, Philippines
Robert S. Laramee, University of Nottingham, United Kingdom
Robin Ghosh, Arkansas Tech University, USA
Ruiheng Zhang, Beijing Institute of Technology, China
Rushit Dave, Minnesota State University, USA
Sai-Keung Wong, National Yang Ming Chiao Tung University
Shahina Begum, School of Innovation Design and Engineering, Mälardalen University, Västerås, Sweden
Shanjun Zhang, Kanagawa University, Japan
Shaoshuai Gao, University of Chinese Academy of Sciences, China
Shiqi Wang, City University of Hong Kong, China
Spiros Mancoridis, Drexel University, USA
Stephanie Chua, Universiti Malaysia Sarawak, Malaysia
Sugandima Mihirani Vidanagamachchi, University of Ruhuna, Sri Lanka
Sunny Joseph Kalayathankal, Jyothi Engineering College, India
Suraiya Jabin, Jamia Millia Islamia, India
T R Shivaram, Tata Consultancy Services, India
Takeshi Saitoh, Kyushu Institute of Technology, Japan
Tan Tse Guan, Universiti Malaysia Kelantan, Malaysia
Thumrongrat Amornraksa, King Mongkut's University of Technology Thonburi, Thailand
Vineet Dhanawat, Software Engineer at Meta
Vinod Kumar Verma, Sant Longowal Institute of Engineering and Technology, India
Wenbo Zhu, University of Chicago, USA
Wiesław Pawłowski, University of Gdańsk, Poland
Wuttinan Nunkaew, Thammasat University, Thailand
Xiaoqi Ma, Nottingham Trent University, UK
Xiwen Zhang, Beijing Language and Culture University, China
Xuechao Li, Auburn University, USA
Yaxin Bi, Ulster University, UK
Yeo Boon Chin, Multimedia University, Malaysia
Yoshito Mekada, Chukyo University, Japan
YuanKai Wang, Fu Jen Catholic University
Yuji Iwahori, Chubu University, Japan
Yuji Sekiya, The University of Tokyo, Japan
Yulia Kumar, Kean University, USA
Yung Gi WU, Chang Jung Christian University
Yunli Lee, Sunway University, Malaysia
Zahid Akhtar, State University of New York Polytechnic Institute, USA
Ze Jin, Tokyo Institute of Technology, Japan

Guideline

Important Notes

- Please enter the meeting room at least 15 minutes before your session. Your punctual arrival and active involvement will be highly appreciated.

- Please wear your name tag for all the conference activities. Lending it to others is not allowed. If you have any companying person, please do inform our staff in advance.

- Please keep all your belongings (laptop and camera etc.) at any time. The conference organizer does not assume any responsibility for the loss of personal belongings.

- Please show name tag and meal coupons when dining.

- Due to force majeure including but not limited to earthquake, natural disaster, war and country policy, the organizer reserves the rights to change the conference dates or venue with immediate effect and takes no responsibility.

- The conference venue does not provide WIFI. If you need to use mobile data, please prepare in advance.

Oral & Poster Presentation

- Regular oral presentation: 15 minutes (including Q&A).

- Get your presentation PPT or PDF files prepared. Presentations MUST be uploaded at the session room at least 15 minutes before the session starts.

- Laptop (with MS-Office & Adobe Reader), projector & screen, laser pointer will be provided in all oral session rooms.

- Poster Presenters should bring your poster to the conference venue and put it on designated place.
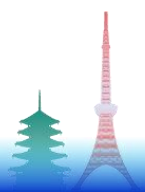
Guideline

FOR VIRTUAL ATTENDANCE

Platform: ZOOM

➢ Step 1: Download ZOOM from the link: https://zoom.us/download

How to use ZOOM

* A Zoom account is not required if you join a meeting as a participant, but you cannot change the virtual background or edit the profile picture.

➢ Rename: Before you enter the conference room, please change your name to Paper ID + Name

➢ Chat and raise your hand: During the session, if you have any questions, please let us know by clicking "raise your hands" and use "chat" to communicate with conference secretary.

➢ When you deliver your online speech, please open your camera.

➢ During the Question section, if you have any questions about keynote speakers or authors, you can also click "raise your hands" or "chat"

➢ Share Screen: Please open your power point first, and then click "share screen" when it's your turn to do the presentation.

How to join the conference online

➢ Find your paper ID and suitable meeting ID on the conference program.

➢ Open the ZOOM, click the join, paste the meeting ID, then you can join the conference.

➢ Click the stop share after you finish your presentation
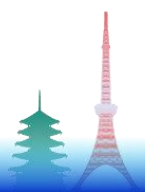
Time Zone

➢ Japan Standard Time, UTC/GMT+09:00

➢ Please make sure that both the clock and the time zone on your computer are set to the correct Japan Time

Device

➢ A computer with an internet connection (wired connection recommended)

➢ USB plug-in headset with a microphone (recommended for optimal audio quality)

➢ Webcam: built-in or USB plug-in

Online Room Information

Online Room Information

Zoom ID: 876 8000 7833

Zoom Link: https://us02web.zoom.us/j/87680007833

* Please rename your Zoom Screen Name in below format before entering meeting room

Role	Format	Example
Conference Committee	Position-Name	Conference Chair-Prof. XX
Keynote/ Invited Speaker	Position-Name	Invited Speaker-Prof. XX
Author	Session Number-Paper ID-Name	S1-SC0001-Name
Delegate	Delegate-Name	Delegate-Name
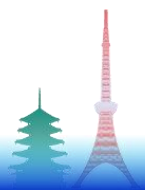
Conference Venue

International Conference Center of Waseda University

Address: 1 Chome-20-14 Nishiwaseda, Shinjuku-ku, Tokyo-to 169-0051 Japan, located in the Central Library of WASEDA University

Conference Rooms Information

| Conference Rooms	December 14	December 15
Ichijima Memorial Conference Room 市島記念会議室	10:00-12:00 Registration	9:10-12:00 Registration
Conference Room No.3 第三会議室	13:00-13:05 Opening Remarks 13:05-13:45 Keynote Speech 13:45-13:50 Group Photo 13:50-17:35 Session 1 & Session 3	9:20-12:00 Keynote & Invited Speeches 13:00-17:05 Session 5 & Session 7
Joint Research Room (7) 共同研究室(7)	13:50-17:35 Session 2 & Session 4	13:00-17:20 Session 6 & Session 8

11

Transportation to Waseda University

The conference will be held on Waseda University International Conference Center, which is located on the main campus of the University, near the Waseda station of Subway Tozai-line.
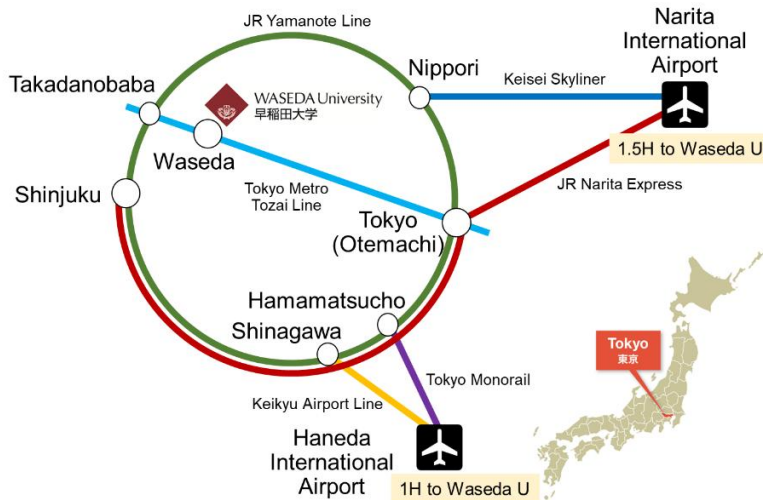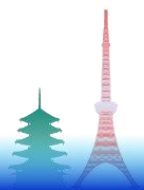

 * 20-minute walk from Takadanobaba Station on the JR Yamanote Line or Seibu-Shinjuku Line

 * 8-minute walk from Waseda Stataion on the Tozai Subway Line

 * 4-minute walk from Nishi-Waseda bus stop on the Tokyo City Bus route from Takadanobaba Station bound for Sodai Seimon

 * 5-minute walk from Waseda stop of Toden Arakawa Line

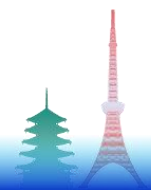You can choose prefer line to Waseda University as the map above.
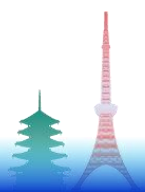
The Direction to International Conference Center

国際会議場・井深大記念ホール
Masaru Ibuka Auditorium -
International Conference Center

Conference Rooms Floor Layout

Group Photo Place

Entrance

Conference Room No.3

Ichijima Memorial Conference Room

3rd Floor

Joint Research Room 7

4th Floor

14

Simple Program

December 14, 2024

Onsite Registration

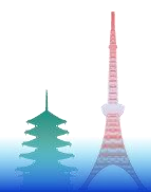| Registration Time: 10:00-12:00 |
|---|
| Venue: 3rd Floor, Ichijima Memorial Conference Room(市島記念会議室), International Conference Center-Waseda University |
| Address: 1 Chome-20-14 Nishiwaseda, Shinjuku-ku, Tokyo-to 169-0051 Japan, located in the Central Library of WASEDA University |
| 1. Arrive at the Ichijima Memorial Conference Room(市島記念会議室)of International Conference Center-Waseda University;
2. Inform the conference staff of your paper ID;
3. Sign your name on the Participants list;
4. Sign your name on Lunch & Dinner requirement list;
5. Check your conference kits;
6. Finish registration. |

Online Test
Time Zone: UTC+9

Zoom ID: 876 8000 7833	
Zoom Link: https://us02web.zoom.us/j/87680007833	
Time	**Speakers & Sessions**
10:00-10:10	Test for Invited Speaker- Prof. Chinthaka Premachandra
10:10-10:55	Test for Session Chair and papers: SC0040, SC0061, SC0067, SC0130, SC0131, SC0136
10:55-11:35	Test for Session Chair and papers: SC0018, SC0051, SC0071, SC1010, SC5015, SC0096
11:35-12:10	Test for papers: SC0002, SC0033, SC0052, SC0053, SC0087, SC0048, SC0035
15:00-15:10	Test for Invited Speaker- Prof. Pascal Lorenz

* If you want to do online zoom test after 15:10, please contact your conference secretary.
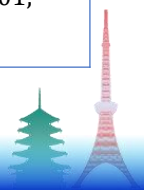
December 14, 2024

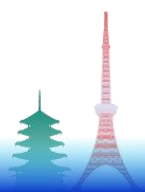| Afternoon Sessions		
Opening Remarks & Keynote Speech 1 **Venue: 3rd Floor,** Conference Room No.3 (第三会議室)		
13:00-13:05	**Opening Remarks**	**Prof. Seiichi OZAWA, Kobe University, Japan**
13:05-13:45	**Keynote Speech 1**	**Prof. Wenbing Zhao, Cleveland State University, USA** **Title:** Dos and Don'ts in Blockchain Research and Development
13:45-13:50	**Group Photo**	
Onsite Parallel Sessions		
	3rd Floor Conference Room No.3 (第三会議室)	**4th Floor** Joint Research Room 7 (共同研究室7)
13:50-15:35	**Session 1** **Topic:** Blockchain Theory and Information Security **Chair:** Prof. Ria A. Sagum, Polytechnic University of the Philippines, Philippines **Papers:** SC0028, SC0057, SC0050, SC0005, SC0065, SC0025-A, SC0030	**Session 2** **Topic:** Software and Information System Development **Chair:** Prof. Lu Sun, Rochester Institute of Technology, United States **Papers:** SC0027, SC0122, SC0026, SC0128, SC0017, SC0041, SC0032
15:35-15:50	**Coffee Break**	
15:50-17:35	**Session 3** **Topic:** Digital Image and Multimedia Processing Technology **Chair:** Prof. Keisuke Kameyama, University of Tsukuba, Japan **Papers:** SC0037, SC0056, SC0060, SC0115, SC1006, SC0021, SC5014	**Session 4** **Topic:** Machine Learning Theory and System Model **Chair:** Dr. Janelle Kyra A. Sagum, Polytechnic University of the Philippines, Philippines **Papers:** SC0031, SC0029, SC0042, SC0075, SC0014, SC1009-A

December 15, 2024

| Morning Sessions		
Keynote Speechs & Invites Speech **Host: Prof. Wenbing Zhao, Cleveland State University, USA** **Venue: 3rd Floor,** Conference Room No.3 (第三会議室)		
9:20-10:00	**Keynote Speech 2**	**Prof. Kenji Suzuki, Institute of Integrated Research at Institute of Science Tokyo, Japan** **Title:** Small-data Lightweight Deep Learning for AI-Aided Diagnosis
10:00-10:40	**Keynote Speech 3**	**Prof. Chia-Feng Juang (IEEE Fellow), National Chung Hsing University, Taiwan** **Title:** Towards Explainable AI Through Deep Fuzzy Neural Network for Image Classification
10:40-11:00	**Group Photo & Coffee Break**	
11:00-11:40	**Keynote Speech 4**	**Prof. Keiji Yanai, The University of Electro-Communications, Japan** **Title:** Food Multimedia Computing: Transforming Food with Technology
11:40-12:00	**Invited Speech 1**	**Prof. Manukid Parnichkun, Asian Institute of Technology, Thailand** **Title:** Self-Driving Car Technologies
12:00-13:00	**Lunch time**	
Afternoon Onsite Parallel Sessions		
	3rd Floor Conference Room No.3 (第三会議室)	4th Floor Joint Research Room 7 (共同研究室7)
13:00-15:00	**Session 5** **Topic:** Image Computing Model and Image Reconstruction **Chair:** Prof. Seokwon Yeom, Daegu University, Korea **Papers:** SC5005, SC5006, SC5003, SC5008, SC5010, SC5025-A, SC5009, SC5011	**Session 6** **Topic:** Next Generation Artificial Intelligence and Engineering Applications **Chair:** Prof. Manukid Parnichkun, Asian Institute of Technology, Thailand **Papers:** SC0012-A, SC0046, SC0047, SC0055-A, SC0116, SC0123-A, SC1001, SC0109

15:00-15:20	**Coffee Break & Poster Session** **Topic:** Intelligent image processing and artificial intelligence application technology (SC0019, SC0020, SC0054-A, SC0107-A, SC1014, SC5013, SC0072-A, SC5012)	
	3rd Floor Conference Room No.3 (第三会議室)	4th Floor Joint Research Room 7 (共同研究室7)
15:20-17:20	**Session 7** **Topic:** Semantic analysis and data processing **Chair:** Assoc. Prof. Yan Wang, Temple University, USA **Papers:** SC0003, SC0007, SC0077, SC0117, SC1015, SC1007, SC0125	**Session 8** **Topic:** Signal and Image Analysis, Data communication **Chair:** Prof. Harald Bolsinger, THWS Business School, Germany **Papers:** SC0104, SC0111, SC0126, SC0139, SC0045, SC0078, SC1004, SC1012-A
18:00-19:30	**Dinner Time** **Venue:** Dining ferio, RIHGA Royal Hotel Tokyo **Address:** 1-104-19 Totsuka-machi, Shinjuku-ku, Tokyo 169-8613, Japan	

December 16, 2024

| Morning Sessions	
Online Invited Speech & Sessions **Zoom ID:** 876 8000 7833 **Zoom Link:** https://us02web.zoom.us/j/87680007833	
Time	**Speech/Sessions**
10:15-10:35	**Invited Speech 2** **Prof. Chinthaka Premachandra, Shibaura Institute of Technology, Japan** **Title:** AI-Driven Hyperspectral Image Stitching and Applications
10:35-12:05	**Session 9** **Topic:** Machine Vision and Virtual Technology Applications **Chair:** Prof. Chien-Sing, LEE, Sunway University, Malaysia **Papers:** SC0051, SC0040, SC0067, SC0052, SC0048, SC0136
12:05-13:20	**Lunch Time**
Afternoon Sessions	
13:20-15:05	**Session 10** **Topic:** Next Generation Artificial Intelligence Theory and Application Technology **Chair:** Prof. Criselle J. Centeno, MIT, Philippines **Papers:** SC0130, SC0002, SC0033, SC0131, SC0053, SC0087, SC0035
15:05-15:25	**Invited Speech 3** **Prof. Pascal Lorenz, University of Haute-Alsace, France** **Title:** Advanced architectures of Next Generation Wireless Networks
15:25-17:25	**Session 11** **Topic:** Natural Language Processing, Machine Vision and Virtual Technology Applications **Chair:** Prof. Chinthaka Premachandra, Shibaura Institute of Technology, Japan **Papers:** SC1010, SC0071, SC0018, SC5015, SC0096, SC0061, SC0022, SC0143
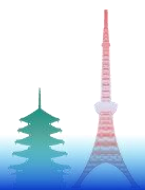
Detailed Program

December 14, 2024

13:00-13:05	**Opening Remarks**
Venue	**3rd Floor**, Conference Room No.3 (第三会議室)

Prof. Seiichi Ozawa
Conference Chair

Kobe University, Japan

Seiichi Ozawa received the B.E. and M.E. degrees in instrumentation engineering from Kobe University in 1987 and 1989, respectively. In 1998, he received his Dr. Eng. in computer science from Kobe University. He is currently the deputy director of Center for Mathematical and Data Sciences and a full professor with the Department of Electrical and Electronic Engineering, Graduate School of Engineering, Kobe University, Japan. He was a visiting researcher at Arizona State University in 2005. His current research interests are neural networks, machine learning, online learning, pattern recognition, big data analytics especially in cybersecurity, SNS and smart agriculture. He published 141 journal and refereed conference papers, and 9 book chapters/monographs. He is currently an associate editor of IEEE Trans. on Cybernetics, Evolving Systems Journal, Pattern Analysis and Applications Journal, and he was an associated IEEE Trans. on Neural Networks and Learning Systems for 6 years. Currently, he is a Pro Tempore Vice-President for Public Relations of International Neural Network Society (INNS), a vice-president for finance of Asia Pacific Neural Network Society (APNNS), and a special board of governor of Japan Neural Network Society (JNNS). He is a member of Neural Networks TC, Data Mining and Big Data Analytics TC, and Smart World TC of IEEE CI Society. He is serving as a general chair of INNS Conference on Big Data and Deep Learning 2018, Program Committee Chair of International Conference on Neural Information Processing 2018, Workshop Chair of 2018 IEEE Smart World Congress, and Program Committee Members of IJCNN 2018, INNS 2018, EAIS 2018, etc.
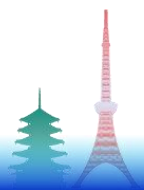
Keynote Speech 1

Host	Prof. Seiichi Ozawa	Time	13:05-13:45, December 14, 2024
	Kobe University, Japan	Venue	3rd Floor, Conference Room No.3 (第三会議室)

Prof. Wenbing Zhao
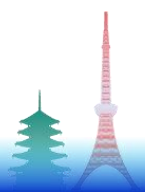
Cleveland State University, USA

Dr. Zhao is a Professor at the Department of Electrical Engineering and Computer Science, Cleveland State University. He got his BS and MS degrees from the Physics Department in Peking University. He earned his Ph.D. at University of California, Santa Barbara in 2002. He has over 250 peer-reviewed publications and the author of the research monograph titled "From Traditional Fault Tolerance to Blockchain." Dr. Zhao's research spans from dependable distributed systems, human centered smart systems, and engineering education. His research has been funded by the US NSF, the US Department of Energy, the US Department of Education, the US Department of Transportation, Ohio Bureau of Workers' Compensation, Ohio Department of Higher Education, the Ohio Development Services Agency, and Woodruff Foundation. He has delivered more than 10 keynotes, tutorials, public talks and demonstrations in various conferences, industry and academic venues. Dr. Zhao is an associate editor for IEEE Access, Human-centric Computing and Information Sciences, MDPI Computers, and PeerJ Computer Science, and a member of the editorial board of several international journals, including Sensors, Applied System Innovation, and Internal Journal of Parallel, Emergent and Distributed Systems. He is currently an IEEE Senior Member and serves on the executive committee of the IEEE Cleveland Section.

Speech Contents

Title: Dos and Don' ts in Blockchain Research and Development

Abstract: The goal of this talk is to clarify common misconceptions regarding the blockchain technology, pointing out some worrisome practices in the field of blockchain technology, and providing a guideline on appropriate use of the blockchain technology in practice. We first provide a brief introduction of the blockchain technology, highlighting its design principle, nuts and bolts, and the most fundamental innovation of the technology. Next, we elaborate on a key concept regarding data immutability. In particular, it is a mistake to equate blockchain to data immutability. Data immutability can only be achieved in large-scale public blockchains where the cost becomes an insurmountable barrier for any attempt to change the data recorded in the blockchain. The third topic we discuss is the differences between public, private, and consortium blockchains. Finally, we provide a guidance on the use of the blockchain technology, particularly, about when to use the private and consortium blockchains.

Keynote Speech 2

Host	Prof. Wenbing Zhao	Time	9:20-10:00, December 15, 2024
	Cleveland State University, USA	Venue	3rd Floor, Conference Room No.3 (第三会議室)

Prof. Kenji Suzuki

### Institute of Integrated Research at Institute of Science Tokyo, Japan

Kenji Suzuki, Ph.D. (Nagoya University) worked at Hitachi Medical Corp, Aichi Prefectural University, Japan, as a faculty member, in Department of Radiology, University of Chicago, as Assistant Professor, and Medical Imaging Research Center, Illinois Institute of Technology, as Associate Professor (Tenured). He is currently a Professor (Tenured) & Director of Biomedical Artificial Intelligence Research Unit in Institute of Innovative Research, Tokyo Institute of Technology, Japan. He published more than 395 papers (including 125 peer-reviewed journal papers). He has been actively researching on deep learning in medical imaging and AI-aided diagnosis in the past 25 years, especially his early deep-learning model was proposed in 1994. His papers were cited more than 16,000 times, and his h-index is 63. He is inventor on 38 patents (including ones of earliest deep-learning patents), which were licensed to several companies and commercialized. He published 15 books and edited 16 journal special issues. He has been awarded numerous grants including NIH, NEDO, and JST grants, totaling $17M. He serves as Editors of 34 leading international journals including Pattern Recognition. He chaired 114 international conferences. He received 27 awards, including 3 Best Paper Awards in leading journals. His research interests include Artificial intelligence, Deep learning, Machine learning, AI-aided system, Computer-aid diagnosis, Medical image analysis, Medical image processing, Data science.

Speech Contents

Title: Small-data Lightweight Deep Learning for AI-Aided Diagnosis

Abstract: Deep leaning became one of the most active areas of research in artificial intelligence (AI) and in medical diagnosis. The current limitations of deep learning in medical diagnosis are requirements of 1) "big data" for training and 2) heavy computational resources. To address those issues, my group has been investigating a lightweight deep learning model that can be trained with "small data". In this talk, our small-data deep learning model that can be trained with limited data is introduced. Our small-data lightweight model does not require heavy computational resources such as GPU servers, but ordinary consumer PC. We applied our models to develop AI-aided diagnostic systems with medical images, including AI systems for cancer detection, analysis, and diagnosis with medical images. Some of them have been commercialized via FDA approvals in the U.S., including the first FDA-approved deep-learning product, and in European and Asian countries.
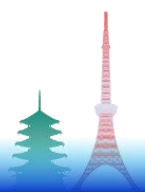
Keynote Speech 3

| Host	Prof. Wenbing Zhao	Time	10:00-10:40, December 15, 2024
	Cleveland State University, USA	Venue	3rd Floor, Conference Room No.3 (第三会議室)

Prof. Chia-Feng Juang

(IEEE Fellow)

National Chung Hsing University, Taiwan

Chia-Feng Juang received the B.S. and Ph.D. degrees in Control Engineering from the National Chiao-Tung University, Hsinchu, Taiwan, in 1993 and 1997, respectively. Since 2001, he has been with the Department of Electrical Engineering, National Chung Hsing University, Taichung, Taiwan, where he became a Full Professor in 2007 and has been a Distinguished Professor since 2009. He served as the Chapter Chair of IEEE Computational Intelligence, Taipei Chapter, in 2017-2018, during which the chapter won the Outstanding Chapter Award from IEEE Taipei Session. Dr. Juang has authored or coauthored over 110 journal papers (including over 65 IEEE journal papers), ten book chapters, and over 150 conference papers. His current research interests include computational intelligence, intelligent control, computer vision, intelligent robots, and AI-aided medical diagnosis.

Dr. Juang was the receipt of the Outstanding Automatic Control Engineering Award from Chinese Automatic Control Society (CACS), Taiwan, in 2014; the Outstanding Electrical Engineering Professor Award from Chinese Institute of Electrical Engineering, Taiwan, in 2019; and the Outstanding Research Award from Ministry of Science and Technology, Taiwan, in 2021. He was elevated to CACS Fellow in 2016, IEEE Fellow in 2019, and International Fuzzy Systems Association (IFSA) Fellow in 2023. He was an IEEE Computational Intelligence Society Distinguished Lecture during 2020-2023. He is currently an Associate Editor for IEEE TRANSACTIONS ON FUZZY SYSTEMS and Asian Journal of Control and an Area Editor for International Journal of Fuzzy Systems.
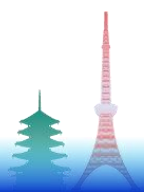
Speech Contents

Title: Towards Explainable AI Through Deep Fuzzy Neural Network for Image Classification

Abstract: AI has emerged as a popular research topic in recent years, demonstrating significant success across various applications. However, most AI models operate as black boxes, making it difficult to explain the reasoning behind their decisions. In response to this challenge, explainable AI (XAI) has garnered substantial attention from researchers. Approaches to achieving explainable AI generally fall into two categories: intrinsic model interpretability and post-hoc explanations. In this speech, I will introduce a visually interpretable fuzzy neural classification model that leverages deep convolutional feature maps alongside post-hoc analysis. This model, known as the Deep Feature Map-based Fuzzy Neural Classification Network (DFM-FNCN), intrinsically utilizes fuzzy if-then rules that are learned through both structural and parameter learning. To enable visual interpretation of these learned fuzzy rules, I will present a technique that employs a deep decoder to map the antecedent of a fuzzy rule to an image. Finally, I will discuss how Gradient-weighted Class Activation Mapping (Grad-CAM), a widely used method, is used for post-hoc explanations of image classification results within the DFM-FNCN.

Keynote Speech 4

Host	Prof. Wenbing Zhao	Time	11:00-11:40, December 15, 2024
	Cleveland State University, USA	Venue	3rd Floor, Conference Room No.3 (第三会議室)

Prof. Keiji Yanai

The University of Electro-Communications, Japan

Dr. Keiji Yanai is a Professor in the Department of Informatics at the University of Electro-Communications, Tokyo, Japan. His research interests include object recognition, web multimedia processing, social media mining, and food-related multimedia computing. Notably, he has been conducting research on food image recognition since 2008 and developed one of the most prominent food image datasets, UEC-FOOD 100/256. He has played key roles in organizing major conferences and workshops on the multimedia field. He served as a General Co-Chair for IEEE MIPR 2021, ACM Multimedia Asia 2022, and MMM 2025, as well as a Technical Program Co-Chair for ACM ICMR 2018. He has also co-organized the workshop series MADiMa (International Workshop on Multimedia Assisted Dietary Management) and has been an active Technical Program Committee member and reviewer for leading conferences and journals in the fields of multimedia, computer vision, and machine learning.

Speech Contents

### Title: Food Multimedia Computing: Transforming Food with Technology

Abstract: In the age of digital transformation, the intersection of food science and multimedia computing presents a unique opportunity to redefine how we perceive, interact with, and analyze food. This keynote, titled Food Multimedia Computing, explores cutting-edge research at the convergence of these domains, highlighting advancements and applications from the Multimedia Computing Laboratory (MMCL) at the University of Electro-Communications.

From dietary assessment to culinary creativity, food plays a pivotal role in our daily lives, influencing health, culture, and innovation. Leveraging multimedia technologies, such as image recognition, video analysis, and machine learning, researchers are now tackling critical challenges in understanding and interpreting food-related data. This keynote will showcase how food multimedia computing addresses these challenges through innovative solutions, driving breakthroughs in areas such as:

1) Automated Food Recognition and Segmentation: Harnessing advanced computer vision techniques to identify food items in images and videos with unprecedented accuracy. This technology underpins applications in dietary tracking and personalized nutrition management.

2) Calorie Estimation and Nutritional Analysis: Developing algorithms to estimate portion sizes and nutritional content from visual data, enabling users to make informed dietary choices.

3) Food Logging and Behavior Analysis: Exploring how continuous food logging through multimedia tools provides insights into eating habits, supporting research in health and wellness.

4) Cross-Cultural Culinary Exploration: Employing multimedia databases to analyze and bridge culinary practices across cultures, fostering innovation in food preparation and presentation.

Throughout the talk, we will explore the theoretical underpinnings, technological developments, and practical applications of these innovations. Examples from MMCL's pioneering work will illustrate the potential of multimedia computing to revolutionize food-related research and applications. These include projects on food recognition in complex environments, portion size estimation, and real-time food logging systems.

By combining multimedia technologies with a deep understanding of food science, Food Multimedia Computing aims to create a bridge between technology and our most fundamental human need: sustenance. This keynote will inspire researchers and practitioners alike to further advance the boundaries of this rapidly evolving field.
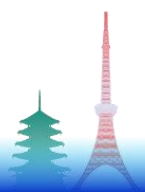
Invited Speech 1

Host	Prof. Wenbing Zhao	Time	11:40-12:00, December 15, 2024
	Cleveland State University, USA	Venue	3rd Floor, Conference Room No.3 (第三会議室)

Prof. Manukid Parnichkun

Asian Institute of Technology, Thailand

Manukid Parnichkun is currently a professor and program chair at Mechatronics and Machine Intelligence program, Asian Institute of Technology. He received B.Eng. from Mechanical Engineering, Chulalongkorn University in 1991, M.Eng. and Ph.D. from Precision Machinery Engineering, the University of Tokyo in 1993 and 1996 respectively. He joined Asian Institute of Technology as an assistant professor in 1996. He was promoted to Associate Professor in 2001, and professor in 2016. He supervised and graduated 23 doctoral students, and 208 master students. He was the founding committee of the Thai Robotics Society (TRS) and later became editor-in-chief of the society journal. He was elected to be the president of the Thai Robotics Society during 2003-2005. He organized and chaired several conferences including IEEE ICIT 2002, IEEE ROBIO 2008 and several robot competitions including Thailand Intelligent Vehicle Challenge 2007-2009, BicyRobo Thailand Championship 2010-2011, ABU Robocon Thailand Championship 2002-now. His research interests are Mechatronics, Robotics, Control, and Measurement.

Speech Contents

Title: Self-Driving Car Technologies

Abstract: Self-driving car was widely known for the first time from DARPA Grand Challenge in 2004 which was a competition of self-driving cars in dessert area. The competition was evolved to DARPA Urban Challenge in 2007 which was a competition in urban area. Since then, research and developments of self-driving cars have been conducted extensively for commercialization purposes for examples; self-driving car by Google, electrical car with driver assisting functions by Tesla. Today most of car manufacturers research and develop self-driving car actively.

This talk will present the key devices and control algorithms behind self-driving car technologies. Technologies used by Google car and Tesla will be firstly presented and compared. In the latter part of the talk, the control algorithms used for speed control, heading control, waypoints tracking, and obstacles avoidance of self-driving car, self-driving bicycle, and autonomous forklift developed at Asian Institute of Technology (AIT) will be presented.
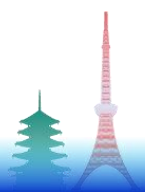
Invited Speech 2

Time	**10:15-10:35(UTC+9), December 16, 2024**
Room	**Zoom ID: 876 8000 7833** **Zoom Link: https://us02web.zoom.us/j/87680007833**

Prof. Chinthaka Premachandra

Shibaura Institute of Technology, Japan

Chinthaka Premachandra (Senior Member, IEEE) was born in Sri Lanka. He received the B.Sc. and M.Sc. degrees from Mie University, Tsu, Japan, in 2006 and 2008, respectively, and the Ph.D. degree from Nagoya University, Nagoya, Japan, in 2011. From 2012 to 2015, he was an Assistant Professor with the Department of Electrical Engineering, Faculty of Engineering, Tokyo University of Science, Tokyo, Japan. From 2016 to 2017, he was an Assistant Professor with the Department of Electronic Engineering, School of Engineering, Shibaura Institute of Technology, Tokyo, where he was an Associate Professor, from 2018 to 2022. In 2022, he was promoted to a Professor with the Department of Electronic Engineering, Graduate School of Engineering, Shibaura Institute of Technology, where he is currently the Manager of the Image Processing and Robotic Laboratory. His research interests include AI, UAV, image processing, audio processing, intelligent transport systems (ITS), and mobile robotics.

He is a member of IEEE, IEICE, Japan; SICE, Japan; RSJ, Japan; and SOFT, Japan. He received the IEEE SENSORS LETTERS Best Paper Award from the IEEE Sensors Council in 2022 and the IEEE Japan Medal from the IEEE Tokyo Section in 2022. He also received the FIT Best Paper Award and the FIT Young Researchers Award from IEICE and IPSJ, Japan, in 2009 and 2010, respectively. He has served as a steering committee member and an editor for many international conferences and journals. He is the Founding Chair of the International Conference on Image Processing and Robotics (ICIPRoB), which is technically co-sponsored by IEEE. He is currently serving as an Associate Editor for IEEE Robotics and Automation Letters (R-AL) and IEICE Transactions on Information and Systems.
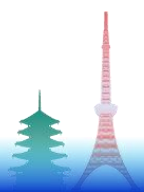
Speech Contents

Title: AI-Driven Hyperspectral Image Stitching and Applications

Abstract: Hyperspectral imaging has been widely applied in anomaly detection across various fields, including agriculture, medicine, and more. In agricultural applications, particularly those using UAVs equipped with hyperspectral cameras, image stitching is essential for generating wide-area ground coverage. Furthermore, for both UAV-based and lab-based hyperspectral images, dimensionality reduction plays a critical role in preparing the data for AI models. Since AI models are typically designed to process low-dimensional data, reducing the dimensionality of hyperspectral data improves training and testing efficiency in anomaly detection, object detection, and classification applications. This keynote speech explores advanced methods for image stitching and dimensionality reduction.
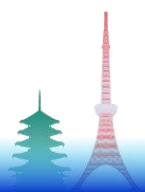
Invited Speech 3

Time	**15:05-15:25 (UTC+9), December 16, 2024**
Room	**Zoom ID: 876 8000 7833** **Zoom Link: https://us02web.zoom.us/j/87680007833**

Prof. Pascal Lorenz

University of Haute-Alsace, France

Pascal Lorenz (lorenz@ieee.org) received his M.Sc. (1990) and Ph.D. (1994) from the University of Nancy, France. Between 1990 and 1995 he was a research engineer at WorldFIP Europe and at Alcatel-Alsthom. He is a professor at the University of Haute-Alsace, France, since 1995. His research interests include QoS, wireless networks and high-speed networks. He is the author/co-author of 3 books, 3 patents and 200 international publications in refereed journals and conferences. He was Technical Editor of the IEEE Communications Magazine Editorial Board (2000-2006), IEEE Networks Magazine since 2015, IEEE Transactions on Vehicular Technology since 2017, Chair of IEEE ComSoc France (2014-2020), Financial chair of IEEE France (2017-2022), Chair of Vertical Issues in Communication Systems Technical Committee Cluster (2008-2009), Chair of the Communications Systems Integration and Modeling Technical Committee (2003-2009), Chair of the Communications Software Technical Committee (2008-2010) and Chair of the Technical Committee on Information Infrastructure and Networking (2016-2017), Chair of IEEE/ComSoc Satellite and Space Communications Technical (2022-2023), IEEE R8 Finance Committee (2022-2023), IEEE R8 Conference Coordination Committee (2023). He has served as Co-Program Chair of IEEE WCNC'2012 and ICC'2004, Executive Vice-Chair of ICC'2017, TPC Vice Chair of Globecom'2018, Panel sessions co-chair for Globecom'16, tutorial chair of VTC'2013 Spring and WCNC'2010, track chair of PIMRC'2012 and WCNC'2014, symposium Co-Chair at Globecom 2007-2011, Globecom'2019, ICC 2008-2010, ICC'2014 and '2016. He has served as Co-Guest Editor for special issues of IEEE Communications Magazine, Networks Magazine, Wireless Communications Magazine, Telecommunications Systems and LNCS. He is associate Editor
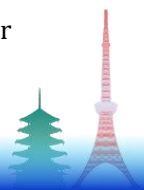
for International Journal of Communication Systems (IJCS-Wiley), Journal on Security and Communication Networks (SCN-Wiley) and International Journal of Business Data Communications and Networking, Journal of Network and Computer Applications (JNCA-Elsevier). He is senior member of the IEEE, IARIA fellow and member of many international program committees. He has organized many conferences, chaired several technical sessions and gave tutorials at major international conferences. He was IEEE ComSoc Distinguished Lecturer Tour during 2013-2014.

Speech Contents

Title: Advanced architectures of Next Generation Wireless Networks

Abstract: Internet Quality of Service (QoS) mechanisms are expected to enable wide spread use of real time services. New standards and new communication architectures allowing guaranteed QoS services are now developed. We will cover the issues of QoS provisioning in heterogeneous networks, Internet access over 5G networks and discusses most emerging technologies in the area of networks and telecommunications such as IoT, SDN, Edge Computing and MEC networking. We will also present routing, security, baseline architectures of the inter-networking protocols and end-to-end traffic management issues.
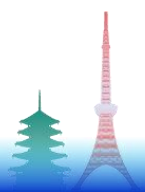
Session 1

December 14, 2024

Session 1: Blockchain Theory and Information Security

Room: 3rd Floor, Conference Room No.3 (第三会議室)

Time: 13:50-15:35

Session Chair: Prof. Ria A. Sagum, Polytechnic University of the Philippines, Philippines

Onsite

13:50-14:05	SC0028	**Title:** Enhancing Transparency and Traceability in Handicrafts Supply Chains Using Blockchain and NFTs **Author(s):** Duc Bui Tien, Khiem Huynh Gia and Trung Phan Hoang Tuan **Presenter:** *Trung Phan Hoang Tuan, FPT University, Viet Nam*
		Abstract- The handicraft industry, with its diverse and intricate supply chains, faces challenges in maintaining transparency and ensuring ethical sourcing due to incomplete traceability. This paper proposes a blockchain-based supply chain management system that incorporates smart contracts and Non-Fungible Tokens (NFTs) to enhance transparency and data integrity from the artisan to the end consumer. The system was evaluated on multiple EVM-compatible platforms, including Binance Smart Chain, Polygon, Fantom, and Celo. The evaluation focused on key functions such as recording supply chain information, minting NFTs for product traceability, and securely transferring NFTs. The findings provided insights into transaction speed, resource utilization, and economic feasibility, highlighting the potential for cost reduction and increased efficiency in supply chain operations within the handicraft industry. This study demonstrates how blockchain technology, combined with smart contracts and NFTs, can create a more transparent and trustworthy supply chain network for handicrafts.
14:05-14:20	SC0057	**Title:** Enhancing the AlphaGo Zero Algorithm Through Kolmogorov-Arnold Networks **Author(s):** Jianuo Lei, Hang Ouyang, Yang Tan, Qing Li, Haolan Wang and Kaixin Deng **Presenter:** *Kaixin Deng, Chengdu University of Technology, China*
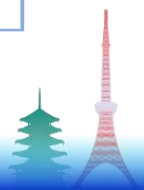
Abstract- With the rapid advancement of artificial intelligence, AlphaGo has revolutionized the game of Go, ultimately evolving into AlphaGo Zero. Despite its success, the neural network architecture of the AlphaGo Zero model, based on ResNet, is characterized by a high number of parameters and computational complexity, making it challenging to deploy on low-power devices. To address this issue, this paper presents an adaptation of the AlphaGo Zero algorithm by substituting its ResNet with a Kolmogorov-Arnold Network (KAN) structure. This modification is applied and evaluated within the game " Dots-and-boxes ", serving as a test case for potential real-world applications. The paper details the design and training process of the KAN-based model, alongside a rigorous comparison with the original AlphaGo Zero model. Experimental results indicate that the KAN-based model not only preserves competitive performance in Dots-and-boxes but also achieves a substantial reduction in model complexity and parameter count. These improvements demonstrate the model's enhanced suitability for resource-constrained environments, making it more efficient for low-power device deployment.

	SC0050	**Title:** Layer 2 Blockchains: An Introduction **Author(s):** Wenbing Zhao and Xiong Luo **Presenter:** *Wenbing Zhao, Cleveland State University, USA*

14:20-14:35

Abstract- In this paper, we provide an overview of a specific, but rather popular type of blockchain scaling solutions, which are usually referred to as Layer 2 blockchains. These solutions depend on the rollup scaling strategy. First, they all create batches of transactions known as rollups. Second, there are two primary approaches to rollup verifications. One approach is referred to as optimistic rollup, where it is assumed that fraudulent transactions rarely occur and should it happen, then some nodes could raise a challenge, which would lead to the resolution of the problem. The interaction between the asserter and the challenger is referred to as interactive fraud proof. The other approach is a conservative approach, where every rollup batch is associated with a validity proof. For all practical purposes, Ethereum is used as the Layer 1 blockchain because it is the first and only large-scale public blockchain that supports Turing-complete smart contracts. These Layer 2 blockchains typically deploy at least two smart contracts on the Layer 1 blockchain, one to receive rollups from the Layer 2 blockchain, and the other to verify the rollup batch. Furthermore, we provide technical details of two Layer 2 blockchains (i.e., Arbitrum and Optimism) that use the optimistic rollup mechanism, and two Layer 2 blockchains (i.e., Polygon and Starknet) that use validity rollup mechanism. Finally, we analyze the Layer 2 blockchain

December 14-16, 2024 | Tokyo, Japan

solutions in the framework of the blockchain trilemma theory. We show that all Layer 2 scaling solutions trade decentralization for better scalability.

14:35-14:50	SC0005	**Title:** Automated Translation between Bicolano and Tagalog using Transfer-Based Model **Author(s):** Ria Ambrocio Sagum **Presenter:** *Ria Ambrocio Sagum, Polytechnic University of the Philippines, Philippines*
	Abstract- This research proposes a transfer-based bi-directional machine translation system for Philippine Dialect, Bicolano and Tagalog. Language resources such as dictionaries and rules are used as well as statistical based approaches. Upon evaluation, the system obtained 74.75% for Tagalog to Bicolano translation and 71.7% for Bicolano to Tagalog translation.	

14:50-15:05	SC0065	**Title:** SuperSonic Glitch Physical Unclonable Function for Encryption and Device Authentication **Author(s):** Shu Takemoto, Yusuke Nozaki and Masaya Yoshikawa **Presenter:** *Shu Takemoto, Meijo University, Japan*
	Abstract- The fundamental technologies underlying digital social infrastructure and the corresponding cybersecurity countermeasures are important. To establish resilient systems, encryption and authentication at every stage of communication are indispensable. For industrial hardware devices, it is crucial to develop implementation techniques that not only integrate state-of-the-art cryptographic methods but also provide robust countermeasures against side-channel attacks and hardware Trojans. Accordingly, this study proposes a physically unclonable function (PUF) based on a low-latency block cipher. The proposed scheme incorporates SuperSonic as the cryptographic algorithm and utilizes glitch-based PUF as the security mechanism. This approach addresses critical issues related to computational security, tamper resistance, and counterfeiting of devices. Furthermore, by employing low-latency SuperSonic glitching for ID generation within the PUF, the proposed circuit achieves an authentication system with minimal overhead in terms of both implementation area and latency.	

15:05-15:20	SC0025-A	**Title:** Optimizing Trust Management cloud based networks **Author(s):** Hamid Al-Hamadi **Presenter:** *Hamid Al-Hamadi, Kuwait University, Kuwait*

37

Abstract- Despite the extensive research on Trust Management, there is a notable gap in practical applications that incorporate additional considerations such as energy efficiency and system longevity. For systems with strict energy constraints, trust management protocols that neglect lifetime and energy consumption are impractical. This study presents a generalized optimization model designed to maximize the lifetime of a typical LPWAN network. Our model determines the optimal parameters to maintain system security through trust management, balancing energy, lifetime, and reliability requirements to prevent excessive or insufficient trust management activities. This model can be broadly applied to various systems with similar constraints. We explore the trade-offs involved, emphasizing key decision parameters and their impacts.

	SC0030	**Title:** Cross-Chain Blockchain Voting System with NFT-Based Authentication **Author(s):** Duc Bui Tien, Nam Tran Ba and Khanh Vo Hong **Presenter:** *Khanh Vo Hong, FPT University, Vietnam*

15:20-15:35

Abstract- Traditional voting systems face numerous challenges, including security vulnerabilities, transparency issues, and operational inefficiencies, which undermine public confidence in electoral processes. Blockchain technology offers a promising solution with its immutable, decentralized, and cryptographically secure framework, addressing these critical issues. This paper presents a blockchainbased voting system implemented across multiple Ethereum Virtual Machine (EVM) platforms, including Binance Smart Chain, Fantom, Polygon, and Celo. The system leverages smart contracts for secure vote management and Non-Fungible Tokens (NFTs) for voter authentication, ensuring the uniqueness and authenticity of each vote. Our research includes a comprehensive evaluation of the system＇s performance, focusing on transaction costs, processing speed, and scalability. The findings demonstrate the potential of blockchain technology to efficiently handle large volumes of electoral data while maintaining security and integrity, thereby enhancing the reliability and transparency of voting systems.

Session 2

December 14, 2024

Session 2: Software and Information System Development

Onsite Room: 4th Floor, Joint Research Room 7 (共同研究室 7)

Time: 13:50-15:35

Session Chair: Prof. Lu Sun, Rochester Institute of Technology, United States

Onsite

13:50-14:05	SC0027	**Title:** A Multi-Criteria Decision Making Based on Fuzzy Inference System for Supplier Selection and Order Allocation Problems: A Case Study of Hard Disk Drives Manufacturer **Author(s):** Chailikhit Pukkham and Wuttinan Nunkaew **Presenter:** *Wuttinan Nunkaew, Thammasat University, Thailand*

Abstract- In the context of Hard Disk Drive (HDD) manufacturers, where supplier selection and order allocation are crucial, this research employed an application of Artificial Intelligence (AI), i.e., a Fuzzy Inference System (FIS) and Technique for Order Preference by Similarity to Ideal Solution (TOPSIS) for ranking potential suppliers before the order allocation. Beyond traditional criteria such as cost, quality, delivery, and service, this research included sustainability factors, specifically carbon dioxide emissions, electronic waste (e-waste) management, and product lifecycle extension based on Sustainable Development Goal 17 (SDG17) into the evaluation framework. Mamdani-type 1 was utilized in FIS to handle uncertainties using linguistic variables and trapezoidal membership functions. The fuzzy rules encapsulate the vague relationships between input criteria and supplier performance, generating crisp numerical outputs known as Center of Gravity (COG) scores. These scores are instrumental in applying the TOPSIS method, facilitating a clear ranking of suppliers based on their overall performance. The proposed method evaluated five suppliers of a case study HDD manufacturer. The results showed that Supplier SE emerges as the top performer with a score of 0.695, followed sequentially by SC, SA, SB, and SD. To emphasize the significance of SDG17, this study conducts a simulation excluding this criterion, revealing a substantial reshuffling of supplier rankings. Moreover, this research designs an order allocation plan to optimize supplier capacities against customer demand over 12 months. This

	operational strategy utilizes the ranked supplier list to allocate orders efficiently, the top-ranked supplier receives maximum capacity allocation, with subsequent suppliers allocated proportionally based on their rankings and capacities.	
14:05-14:20	SC0122	**Title:** A License Management System for Collaborative AI Engineering **Author(s):** Venkata Satya Sai Ajay Daliparthi, Kurt Tutschku, Nurul Momen, Miguel De Prado, Margaux Divernois, Nuria Pazos Escudero and Jean-Marc Bonnefous **Presenter:** *Venkata Satya Sai Ajay Daliparthi, Blekinge Institute of Technology, Sweden*

Abstract- The AI marketplace ecosystem accelerates multiple modules of the AI engineering pipeline by fostering collaboration between stakeholders. However, marketplace collaborators often face a dilemma in striking a balance between sharing artifacts and protecting intellectual property (IP) rights. Thus, there is a need for a license management system within the AI marketplace to facilitate the exchange of artifacts in a trusted and secure manner.

This work shares experiences while building such a license management system within the Bonseyes marketplace (BMP), a functional crowdsourcing AI marketplace that specializes in deploying real-time applications on edge devices. The BMP was developed, and its applicability is proven through the European H2020 project by a series of open calls and workshops, for gathering stakeholders and orchestrating the marketplace operations.

The main contributions of this work are (i) implementation of an end-to-end license management system that deals with selecting license templates, license agreement interaction between seller and buyer, and the generation and enforcement of human- and machinereadable license files, and (ii) introduction of "Synchronization licenses" concept from the music industry to the AI marketplace context where consumers acquire a license to integrate the artifact into another application, and a respective BMP use-case for collaborative AI engineering.

14:20-14:35	SC0026	**Title:** Malware Detection in Cloud Native Environments **Author(s):** Brian S Mitchell, Ansh Chandnani, John Carter, Danai Danai Roumelioti and Spiros Mancoridis **Presenter:** *Brian S Mitchell, Drexel University, USA*

Abstract- As cloud computing continues to grow, organizations are shifting to software architectures that depend on fully−managed cloud services. Unlike monolithic applications that run on virtual machines, modern cloud native systems are deployed on dynamic Function as a Service (FaaS) platforms, or managed container orchestration runtimes such as Kubernetes. These systems are difficult to monitor for correct operation because the many parts of the application can be terminated or (re)started at any time in the face of errors for resiliency, or for the economic benefits associated with autoscaling. When security vulnerabilities are discovered, the impacted runtime components need to be quickly monitored, identified and patched. This paper introduces a platform we created to observe and measure the health of cloud-native applications by applying machine learning techniques that benchmark normal behavior and can detect when the behavior drifts away from the benchmark due to security attacks. To achieve this goal, our platform must be able to accurately model the distributed and highly dynamic nature of cloud native systems. We present a case study to demonstrate the detection capabilities of our solution against two recent high−profile malware CVEs that target widely deployed open−source cloud software.

	SC0128	**Title:** Prototype for a Personalized Music Recommendation System Based on TL-CNN- GRU Model with 10-Fold Cross-Validation **Author(s):** Pei-Chun Lin, Chen-Yu Yu and Eric Odle **Presenter:** *Pei-Chun Lin, Feng Chia University*
14:35-14:50		**Abstract-** Music is a medium that connects cultures, expresses emotions, and can create memory connections that are deeply embedded in people's hearts. With the rapid expansion of digital music platforms, providing an effective personalized recommendation system that helps users browse huge song libraries can effectively promote platform usage. However, existing music recommendation systems usually fail to fully capture individual preferences. In this study, we aim to address these limitations by developing an advanced recommendation system focusing on two facets: 1) building a model for music genre classification, and 2) providing a platform with personalized recommendations based on user preferences. We herein integrate three models for music genre classification, transfer learning, convolutional neural networks, and gated recurrent units (TL + CNN + GRU), using the GTZAN dataset. Results show that the TL + CNN + GRU model can improve the accuracy (55%->71%) of personalized music recommendation systems by using 10-fold cross-validation. Finally, we introduce a prototype platform for

understanding user experience. In conclusion, our model not only improves the accuracy of recommendations, but also promotes user exposure to different music genres, redefining the user experience.

14:50-15:05	SC0017	**Title:** On the Automatic Identification of Misconfiguration Errors in Cloud Native Systems **Author(s):** Brian S Mitchell, Spiros Mancoridis and Jainam Kashyap **Presenter:** *Spiros Mancoridis, Drexel University, USA*

Abstract- The frequency of notable data breaches in cloud native systems has increased over the past several years causing many problems for both large and small organizations. These systems have a very large attack surface associated with the use of many publicly-accessible APIs that can make differentiating between normal and malicious behavior difficult to classify. The nature of managing every aspect of cloud native systems is prone to misconfiguration errors, requiring the setting and auditing of thousands of parameters in enterprise systems. Successful data breaches require bad actors to remain undetected for long periods of time so that data exfiltration does not trigger intrusion protection controls. In this paper we recreate a realistic cloud native environment to emulate a data breach attack using common API misconfiguration mistakes. We then introduce a tool we created to collect system call data from the kernels in a Kubernetes cluster that is used to train machine learning models capable of differentiating normal from suspicious activity during emulated stealthy attacks.

15:05-15:20	SC0041	**Title:** Modeling and evaluating the operation of smart medical wards during monitoring and treatment of patients **Author(s):** Dmitriy K. Levonevskiy and Anna I. Motienko **Presenter:** *Dmitriy K. Levonevskiy, Laboratory of big data technologies in socio-cyberphysical systems, St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences, St. Petersburg Federal Research Center of the Russian Academy of Sciences (SPC RAS), Russia*

Abstract- The application of smart technologies in medical wards offers significant advancements in patient care through enhancing the processes of monitoring and treatment. However, the complexity of these systems requires creation of frameworks for effective smart ward design and assessment. This paper presents a generalized architecture for smart medical wards, containing information and cyber-physical cycles, to

facilitate the component integration and implementation of medical tasks. We propose a hierarchical model that describes smart ward behavior and can be used to systematize technological solutions. It provides a formal approach to describing and assessing the smart ward operation. It makes possible the classification of existing solutions based on their functionality and simplifies the reuse of prototypes within new projects. The theoretical significance of this work consists in unification and formalization of smart ward behavior, thereby advancing our understanding of their operation in different medical scenarios, such as patient monitoring and intensive care. Practically, the proposed approach simplifies the design process, introduces uniform performance assessments, and facilitates analysis of smart ward efficiency under varying conditions.

	SC0032	**Title:** A Systematic Combination Approach Based on Fuzzy Inference System and Grey Relational Analysis for Evaluating and Prioritizing Online Food Delivery Platforms in Thailand **Author(s):** Marrisa Kimaporn and Wuttinan Nunkaew **Presenter:** *Wuttinan Nunkaew, Thammasat University, Thailand*

Abstract- The continuous growth of the online food delivery service, which has been significant since the COVID-19 era, has led to increased platforms providing these services. However, the rapid expansion of this business has resulted in heightened competition in terms of serving customer satisfaction. Consequently, this research aims to prioritize these platforms by evaluating user experience and preference based on specific criteria. The evaluation used the application of Artificial Intelligence (AI), i.e., a fuzzy inference system combined with the grey relational analysis method to study user behavior in selecting platforms. The study encompassed five platforms, divided into three scenarios: Bangkok Metropolitan, provincial areas, and all areas in Thailand. The findings indicate that Platform A consistently provided the best user experience and satisfaction across all three scenarios, with the highest grey relational grades of 0.9879, 0.9666, and 0.9829, respectively. The rankings for the other four platforms varied across the three scenarios. The results of this study can be used to analyze functional and operational elements of the platform that correspond to customer satisfaction in particular areas.

The time slot for this session is 15:20-15:35.

Session 3

December 14, 2024

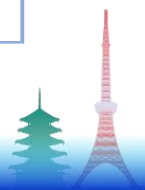**Session 3: Digital Image and Multimedia Processing Technology**

Room: 3rd Floor, Conference Room No.3 (第三会議室)

Time: 15:50-17:35

Session Chair: Prof. Keisuke Kameyama, University of Tsukuba, Japan

Onsite

15:50-16:05	SC0037	**Title:** Fourier Single Pixel Imaging Reconstruction using Deep Image Prior **Author(s):** Wei Lun Tey, Sing Yee Chua, Mau-Luen Tham and Yeong-Nan Phua **Presenter:** *Sing Yee Chua, Universiti Tunku Abdul Rahman, Malaysia*
		Abstract- Single pixel Imaging (SPI) is capable of handling imaging conditions in low light or unusual wavelengths. Fourier Single pixel Imaging (FSI) produces images through the acquisition of Fourier domain information by utilizing a single pixel detector, along with a sequence of light modulation patterns. To increase the efficiency of FSI, Compressed Sensing (CS) is implemented to reduce the number of samples needed. Traditionally, the image is subsequently reconstructed using the Total Variation (TV) method which is computationally expensive. Hence, this paper proposes Deep Image Prior (DIP) utilizing a deep untrained convolutional network as an alternative to reconstruct images in FSI. DIP has been proven to work well in the denoising field, treating the image as the noisy input, DIP can be implemented for FSI image reconstruction without any modification. The results indicate that replacing TV with the DIP method achieves overall better image quality and reduced processing time. Consequently, the reconstruction efficiency in FSI is improved.
16:05-16:20	SC0056	**Title:** ESPCA: An Efficient Spatial-Sensitive Partial Channel Attention Mechanism in CNNs **Author(s):** Tianze Huang, Yan Qian, Gaofei Sun, Jiahao Yu, Zhenjiang Qian, Feng Li and Xiaobing Xian **Presenter:** *Tianze Huang, Changshu Institute of Technology, China*

Abstract- Attention mechanisms, as a crucial component in neural networks, offer significant advantages in visual tasks. However, the pursuit of enhanced performance in existing methods often brings more feature redundancy. In this paper, we propose ESPCA: a novel Efficient Spatial-Sensitive Partial Channel Attention mechanism, to enhance spatial sensitivity while reducing redundant features. Inspired by previous works, we argue that the redundancy of spatial information hinges on the method of extraction, such as dimensionality reduction and partial channel convolution. Therefore, ESPCA adopts 1D convolution on partial channels without dimensionality reduction to aggregate features along the width and height directions of the feature maps. We utilize ResNet and MobileNetV2 as the backbone to evaluate our ESPCA. Experimental results demonstrate that ESPCA achieves superior computational efficiency while delivering better performance compared to other state-of-the-art attention mechanisms. Specifically, on the ImageNet-1k classification task, ESPCA achieves the highest Top-1 accuracy of 76.38% on ResNet18, surpassing the original ResNet18 by 1.08%. Additionally, on the semantic segmentation task, ESPCA improves mIoU by 1.10% and Pixel Accuracy by 1.85% compared to the original ResNet34.

16:20-16:35	SC0060	**Title:** Detecting Diseases in Corn through Convolutional Neural Network Architectures and Ensemble **Author(s):** James Vincent Verdeflor Bacus, Caitlin Mariel Leyson Lindsay, Christian V. Maderazo and Gerard Ompad **Presenter:** *James Vincent Verdeflor Bacus, University of San Carlos, Philippines*

Abstract- Crop diseases reduce yield, affecting a nation's agricultural sector. This is more pronounced in nations such as the Philippines, where agriculture is the nation's foundation. To decrease the cost and time needed for disease detection, this study utilized Transfer Learning on three Convolutional Neural Networks (DenseNet201, EfficientNetV2M, and InceptionResNetV2) to identify three corn diseases: Common Rust, Blight, and Gray Leaf Spot. Data Augmentation was used to diversify the image dataset, resulting in 4,000 images per category. This research presents a comparison between the different architectures. It also illustrates the effect of the Soft Voting Ensemble Method.

|		
16:35-16:50	SC0115	**Title:** Combination of Global Maximum Pooling and Local Average Pooling for Unsupervised Fine-Grained Image Retrieval **Author(s):** Chang-Hsing Lee, Jau-Ling Shih, Wen-Li Su, Cheng-Chang Lien and Chin-Chuan Han

		Presenter: *Chang-Hsing Lee, Ming Chi University of Technology*

Abstract- In this paper, a novel pooling approach which combines global maximum pooling (GMP) and local average pooling (LAP), called GMP-LAP, is designed to extract global and local CNN features for unsupervised fine-grained image retrieval (FGIR), without any fine-tuning or re-training of the CNN backbone model. First, each image is inputted into a pre-trained CNN model to get a number of feature maps in the last convolutional layer. GMP is first used to extract the global feature from each feature map. Meanwhile, LAP is proposed to obtain some local features from selected salient regions. GMP and LAP are then combined for unsupervised FGIR. As a result, unsupervised FGIR is realized since no manual labeling or re-training/fine-tuning of the CNN model is involved. To evaluate the performance of the proposed GMP-LAP approach, we have conducted experiments on six datasets, including the Stanford Cars, Stanford Aircrafts, CUB-200-2011, Oxford Flowers 102, Stanford Dogs, and Oxford-IIIT Pets, for unsupervised FGIR. Compared with the most widely used global pooling approaches, such as global average pooling (GAP) and GMP, the experimental results have shown that GMP-LAP can improve the retrieval performance in term of mean average precision (MAP).

	SC1006	**Title:** Development of Enumerator on Android Platform for Counting Elispot **Author(s):** Somkiat Tangjitsitcharoen, Jettanong Klaewsongkram, Supranee Buranapraditkun, Naris Lawankowit, Ravee Bunduwongse, Suphakit Anonglekha and Siwat Assavakijphanich **Presenter:** *Somkiat Tangjitsitcharoen, Chulalongkorn University, Thailand*

16:50-17:05

Abstract- The objective of this research is to create an object-detecting Android application for ELISpot immunoassay reading. The ability to accurately, quickly, and cheaply conduct immunoassay tests will guarantee that public healthcare industries are operating and helping patients in rural areas effectively. The past reliance on purchased laboratory equipment, and the rampant cost of health services has led to the privatization of healthcare industries and the negligence of rural area healthcare facilities. Therefore, this research is intended to help spread ELISpot test capability to smaller less-funded programs by creating an accessible ELISpot reading system in the form of an application and the hardware required. Object detection solutions present 2 issues; computational intensity and image quality. An Android application containing the enumeration algorithm was also designed. In order to

implement object detection capability on mobile devices, the parameters used were extracted from a Faster R-CNN neural network model trained from a set of immunoassay images provided by the laboratory. The annotation of the images was done using LabelImage, whose output of .xml was converted to .csv and tfrecord files respectively. The computational accuracy tested over the samples averages out at 98% accuracy. While other design parameters such as image processing speed, system weight, and process throughput time are all satisfactory. With the project completed successfully, it is believed that with the upcoming technologies of deep learning super sampling (DLSS) With this kept in mind, true deep learning capable object detection applications operated on a mobile device will soon become more widespread and accessible.

17:05-17:20	SC0021	**Title:** Predicting Coding Parameters for Video Coding in Cloud-Based Transcoding Systems **Author(s):** Jaeil Kim, Jeong-Mee Moon, Taeseung Hwang, Sangmin Yoo and Seongsoo Park **Presenter:** *Jaeil Kim, Dept. of Media R&D, SK Telecom Co.*

Abstract- This paper proposes a machine learning-based algorithm for predicting initial compression parameters to optimize video content encoding within cloud-based computing environments. Utilizing extensive video data from the commercial system, we generated training data that accounts for diverse characteristics and applied an ML algorithm using Amazon SageMaker Autopilot. The implementation of this prediction algorithm significantly enhanced the encoding process, achieving an average time saving of 12.85% and a prediction accuracy of 95.80%. These improvements have substantially reduced costs and increased encoding efficiency in cloud settings, demonstrating the potential of integrating advanced machine learning techniques into video encoding workflows.

17:20-17:35	SC5014	**Title:** HRSTA: A Pipeline for Reference-Based High-Resolution Skin Tone Adjustment **Author(s):** Boqi Wu, Ulrich Jung and Hasan Tercan **Presenter:** *Boqi Wu, University of Wuppertal, Germany*

Abstract- Different skin tones reflect the diversity of the world's population. In the dynamic global retail environment, especially in the fashion industry, there is a growing need to adapt marketing strategies to authentically represent different ethnicities, thereby increasing consumer engagement and likelihood of purchase. However, creating photo shoots that include a wide range of skin tones is expensive and time-consuming.
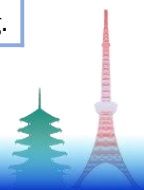
Moreover, human visual perception is particularly adept at noticing the fine details and natural appearance of skin tones. Our study presents an Artificial Intelligence (AI)-based solution for cost-effectively modifying skin tones in images. We propose a High-Resolution Skin Tone Adjustment (HRSTA) pipeline that employs segmentation models, Generative Adversarial Networks (GANs), and digital image processing techniques to adjust skin tones to match references under varying lighting conditions, without compromising image resolution or authenticity. Our experiments demonstrate that the HRSTA is a robust baseline for skin tone adjustment, offering a promising solution for the fashion, beauty, etc. industries.
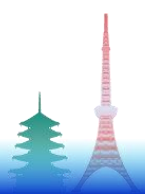
Session 4

December 14, 2024

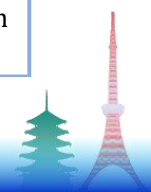**Session 4: Machine Learning Theory and System Model**

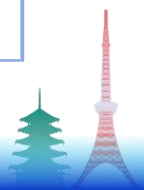**Room: 4th Floor, Joint Research Room 7 (共同研究室 7)**

Time: 15:50-17:20

Session Chair: Dr. Janelle Kyra A. Sagum, Polytechnic University of the Philippines, Philippines
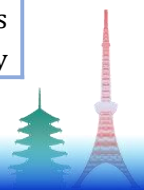
Onsite

15:50-16:05	SC0031	**Title:** A Fuzzy Inference System-Based Multi-Criteria Inventory Classification with Three-Sigma Control Limits **Author(s):** Pagamalueang Nakfon and Wuttinan Nunkaew **Presenter:** *Pagamalueang Nakfon, Thammasat University, Thailand*
		Abstract- Inventory item classification is one of the crucial topics in inventory management, directly affecting the benefits to a business. Effective inventory management can lead to better cost control and more efficient budget allocation. However, inventory classification is subject to several criteria. Therefore, this paper proposes a new approach considering three criteria of Annual Dollar Usage (ADU), Average Unit Cost (AUC), and Lead Time (LT) for inventory item classification. The Fuzzy Inference System (FIS), an application of Artificial Intelligence (AI), was applied as a powerful tool to assess inventory item evaluation scores for each item based on multi-criteria and fuzzy rules. The systematic item categorization referring to the three-sigma control limits concept was also introduced in this paper. The proposed method identifies and assigns items into classes "A+", "A", "B", "C", and "C-". Unlike other inventory classification methods, items with extremely high evaluation scores are classified into Class A+, while items with extremely low scores are categorized into Class C-. This method gives more information to the decision-makers to rigorously monitor the sensitive items. Moreover, the comparative study firmly showed that the proposed method obtained effective solutions, making inventory item classification more reasonable and efficient.
16:05-16:20	SC0029	**Title:** A Blockchain-Based Framework for Managing Electronic Health Records of Farm Animals **Author(s):** Duc Bui Tien, Bang Khanh Le and Triet M. Nguyen **Presenter:** *Triet M. Nguyen, FPT University, Vietnam*

Abstract- The transition from paper-based to digital records has significantly enhanced healthcare practices, including the management of farm animal health. Electronic Health Records (EHRs), referred to as Animal Medical Records (AMRs), are essential for improving clinical workflows, supporting evidence-based decisions, and advancing veterinary research. However, the adoption of digital records introduces challenges related to data security, quality, and interoperability. This paper explores the use of blockchain technology to address these challenges. We propose a blockchain-based system for managing AMRs that leverages Non-Fungible Tokens (NFTs) and smart contracts to ensure data integrity, transparency, and accessibility. The system is evaluated through a proof-of-concept deployment on four EVM-supported platforms: BNB Smart Chain, Celo, Fantom, and Polygon. Our findings indicate that blockchain technology can enhance the security and efficiency of managing electronic health records for farm animals.

	SC0042	**Title:** Enhancing Process Yield Though Quality Prediction Using Machine Learning Techniques **Author(s):** Nattapon Wongngern and Busaba Phruksaphanrat **Presenter:** *Busaba Phruksaphanrat, Thammasat University, Thailand*

16:20-16:35

Abstract- This research investigates the application of Machine learning (ML) techniques to enhance yield of optical modules in manufacturing process of fiber optics. The study focuses on the quality process, which consists of four stages: FCAL, OPM, OPMT, and EXS. The objective is to improve process yield by identifying and removing parts likely to fail testing at an early stage to save time and resources. Initially, data collection and preprocessing were performed. Then, statistical hypothesis testing of mean and variance were utilized for feature selection. After that, various ML techniques were employed to classify parts as either "fail" or "pass," including Random Forest, XGBoost, Gradient Boosting, Neural Networks, Logistic Regression, Decision Tree, Naïve Bayes, and Nearest Neighbors. The research found that Random Forest, XGBoost, and Gradient Boosting models outperformed other models in predicting the quality of part across different stages. Among these, Random Forest and XGBoost were the most effective in improving the cumulative yield of the overall processes. Implementing these models allowed for the early removal of defective parts, resulting in a significant increased in cumulative yield to approximately 97%.
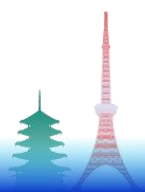
December 14-16, 2024 | Tokyo, Japan

16:35-16:50	SC0075	**Title:** Cross-Disciplinary Integration of Machine Learning and Heuristic Techniques for Enhanced Customer Insights **Author(s):** Tzu-Chien Wang **Presenter:** *Tzu-Chien Wang, Department of Artificial Intelligence and Medical Applications, MacKay Junior College of Medicine, Nursing, and Management*

Abstract- E-commerce and enterprise systems are increasingly pivotal for sustainable business growth, especially in the aftermath of the COVID-19 pandemic, which has permanently reshaped the business landscape. The abundance of rich, unstructured data and customer online activity within enterprise systems holds tremendous potential and value. Strategically harnessing this information can efficiently manage customer journeys, thereby enriching customer experience and fostering service innovation. This study proposes a comprehensive machine learning and heuristic algorithm optimization technique for customer insight analysis in areas such as customer journey management, three-stage customer purchase cycle planning, multi-channel customer value optimization, and product specification recommendation. The system integrates techniques such as cluster analysis, Latent Dirichlet Allocation (LDA) topic analysis, machine learning, integer programming, and heuristic algorithms. Results demonstrate that the framework optimizes the customer purchase cycle and product development, significantly enhances the enterprise's ability to systematically process both numerical and unstructured data, and facilitates dynamic and automated data analysis while efficiently allocating enterprise resources.

16:50-17:05	SC0014	**Title:** Parallel Ensemble Approach for Named Entity Recognition in Filipino Text **Author(s):** Ria Ambrocio Sagum and Janelle Kyra A. Sagum **Presenter:** *Janelle Kyra A. Sagum, Polytechnic University of the Philippines, Philippines*

Abstract- Named Entities play a crucial role in various natural language applications such as automatic text summarization, information retrieval, information extraction, question answering, and machine translation. Systems designed for Named Entity Recognition (NER) employ diverse methodologies to label words with specific named entities like persons, organizations, and locations. Research efforts focusing on NER for the Filipino language have utilized handcrafted rules and various classifiers such as SVM, CRF, HMM, and TBL. These studies span different domains for testing the recognition capabilities of their systems. While achieving promising accuracy rates within their respective domains, these systems exhibit varied performance when applied to the same texts. This study

aims to develop an NER system capable of recognizing diverse entities across different text domains. Employing a parallel ensemble method, the researchers experimented with optimizing the weight combinations of five algorithms. The model underwent testing across 100 files encompassing domains such as News, History, Novel, Political, and Speech, achieving computed accuracy metrics of Precision: 89.56%, Recall: 88.37%, and F-Measure: 88.62%.The results indicate that the developed system performs acceptably well, demonstrating its capability to recognize entities effectively across diverse domains using a unified Named Entity Recognizer.

	SC1009-A	**Title:** A Predictive Model for Turbocharged Engine Calibration Through Deep Neural Network and Optimization Algorithm **Author(s):** Christopher Chan Yew Fai, Chiong Meng Soon and Srithar Rajoo **Presenter:** *Srithar Rajoo, UTM-LoCARtic, Institute for Sustainable Transport, Universiti Teknologi Malaysia, Malaysia*
17:05-17:20		**Abstract-** Vehicular emissions are one of the main contributors for global warming and its mitigation is a key to achieving net zero targets. Internal combustion engines have been criticized for its emissions, nevertheless its significant role in our soceity cannot be denied. Hence, internal combustion engines, especially with turbochargers, will still be on the road for many years to come, however it has been to continously improved. Optimizing the turbocharged engine parameters, such as lambda, can play an important role in reducing vehicle emissions and contributes to the mitigation of global warming. However, experiments involving the study and calibration of internal combustion engines are not only resource-intensive but also time-consuming and complex, especially when accounting for turbocharger designs. To address these challenges, a machine learning approach has been developed and tested to streamline and expedite the engine calibration process, particularly in relation to turbocharger matching. Using a dataset of 626 engine experimental samples involving various turbochargers, the data was divided into training and validation sets using an 8:2 ratio. Multiple deep learning neural network models were constructed to predict different engine output parameters, with turbocharger turbine features serving as categorical inputs in the model architecture. Sensitivity analysis and feature selection was carried out to determine the importance of each turbine feature for every engine output parameter. Data preprocessing, Monte Carlo Cross Validation, hyperparameter tuning and feature selection were executed to minimize the Mean Absolute Error (MAE) for

each output parameter. The finalized models demonstrated good predictive capability, yielding an average validation MAE of 4.53%. The range of validation MAE varied from 2.29% to 8.47%. The main engine parameters DNN achieved low validation MAE of 2.8% and 2.42% for torque and fuel flow rate respectively. Single objective optimization was utilized to optimize the engine operation by maximizing torque or minimizing fuel consumptions in different loads and speeds. The comparison of Metaheuristic optimization (Genetic Algorithm (GA)) and local optimization algorithms (SLSQP and trust region) paired with global optimization method (Basin Hopping and SHGO) revealed that the SHGO with SLSQP approach had great potential in optimizing the turbocharged engine performance. This approach is significantly faster compared to GA, achieving speeds approximately ten times faster. The optimization outcomes for maximizing torque showcased an average 41.36% torque with a simultaneous 19.49% decrease in fuel flow rate. On the other hand, the optimization to minimize fuel flow rate resulted in an average reduction of 33.9% with a simultaneous 27.85% increase in torque.

Session 5

December 15, 2024

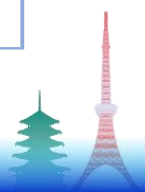**Session 5: Image Computing Model and Image Reconstruction**
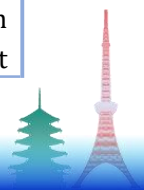**Room: 3rd Floor, Conference Room No.3 (第三会議室)**
Time: 13:00-15:00
Session Chair: Prof. Seokwon Yeom, Daegu University, Korea

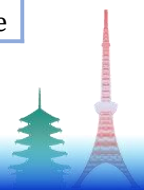**Onsite**

13:00-13:15	SC5005	**Title:** LiteAttentionNet: Lightweight Attention Model for Underwater Image Enhancement **Author(s):** May Thet Tun, Yosuke Sugiura, Tetsuya Shimamura **Presenter:** *Tetsuya Shimamura, Saitama University, Saitama, Japan*

Abstract- In this paper, we propose a novel approach to underwater image enhancement (UIE) by incorporating the impulse response of a low-pass filter (LPF) image with the raw underwater image through skip connections to enhance underwater image features into baseline Shallow-UWnet. Furthermore, a simple, parameter-free attention module (SimAM) is added as a layer to enhance the quality of the underwater images, which enables the baseline Shallow-UWnet to focus on important features. This integration strengthens the model's robustness and enhances noise reduction, allowing it to perform effectively across diverse underwater conditions, and leads to a more efficient processing pipeline. Additionally, this integration not only improves feature extraction but also reduces the model's computational load, making it more suitable for deployment in resource-constrained devices due to lightweight and efficient architecture. This approach is referred to as LiteAttentionNet in this study. Quantitative and qualitative experimental results on three benchmark underwater image datasets such as UFO-120, UIEB, and EUVP_Dark show that the proposed LiteAttentionNet outperforms Shallow-UWnet and state-of-the-art (SOTA) methods. Additionally, an ablation study on non-reference underwater image datasets such as U45, ColorCheck7 and OceanDark datasets confirm the effectiveness of LiteAttentionNet for UIE. Moreover, LiteAttentionNet has fewer trainable parameters and faster training times, which make a suitable choice for underwater robots.
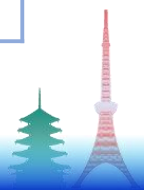
	SC5006	**Title:** Token selection using Grad-CAM in Vision Transformer to reduce computational costs **Author(s):** Rinka Kiriyama, Akio Sashima, Ikuko Shimizu **Presenter:** *Rinka Kiriyama, Tokyo University of Agriculture and Technology, Japan*
13:15-13:30		**Abstract-** While the vision transformer is a useful deep learning model used in various situations, its computational cost is increasing. Since the computational time is on the order of the square of the number of tokens, it is effective to reduce the number of tokens for reducing the computational time. In this paper, we propose a method for sampling tokens according to the score calculated by Grad-CAM. Specifically, in the learning phase, the activation map which means the importance of the token is learned using the Grad-CAM scores by using the weight and tokenizer which are pre-trained and fixed during this step. Then, in the inference phase, the activation map is calculated for each image, and tokens are selected from those with larger values. In caltech101 class recognition, the proposed method recorded higher accuracy than randomly selecting tokens when the number of tokens was small, especially when the number of tokens was reduced by a factor of five, an 11% improvement in accuracy was observed.
	SC5003	**Title:** Inheritance of Attention Region in Model Compression by Distillation in Image Classification **Author(s):** Takumi Shirahama, Keisuke Kameyama **Presenter:** *Takumi Shirahama, University of Tsukuba, Japan*
13:30-13:45		**Abstract-** Large convolutional neural networks (CNNs) have achieved ex-cellent results in image recognition and classification, but their high number of parameters makes implementation and operation on small devices difficult. Therefore, there is a demand for high-performance, compact CNNs. One solution to this problem is knowl-edge distillation (KD), which transfers the output of a large CNN (teacher model) to a smaller CNN (student model) to create a light-weight model. Since the introduction of this knowledge distillation technique, various methods have been proposed to transfer different types of knowledge from the teacher model to the student model. The purpose of this research is to improve the efficiency and the quality of model compression by proposing a distillation method that aims to inherit not only the output of the teacher model, but also its focus areas for CNNs aimed at image classification. Class Activation Mapping (CAM) is known for visualizing the areas fo-cused by CNNs in image classification. In the proposed method, in addition to the KD that only inherits the teacher model's out-put, we aim to inherit the teacher's focus areas extracted through CAM in the student

model's learning process. This approach is expected to allow the student model to learn which features in the input image to focus on during inference, potentially improving the classification performance efficiently. Experiments demonstrated that the proposed KD method using the teacher model's focus area successfully allowed the student model to inherit a similar nature of focusing. This method was particularly effective in the early stages of the student model's learning. By applying the proposed method in the initial learning phase of the student model, it showed superior classification performance when compared to the existing methods on image classification task using the CIFAR-10 dataset. There, it was demonstrated that the Wide-ResNet teacher could be compressed to a small student having only 1/13 parameters.

	SC5008	**Title:** Braille Block Recognition with Deep Learning Utilizing Autocorrelation Map **Author(s):** Hiroyuki Arai **Presenter:** *Hiroyuki Arai, Nippon Institute of Technology, Japan*
13:45-14:00		**Abstract-** This study proposes a method for accurately recognizing Braille Blocks from images using deep learning techniques. We combine a CNN trained on raw images with a CNN trained on autocorrelation maps to enhance the detection accuracy of dot-pattern blocks, which are crucial for indicating warnings or stops. We implemented and evaluated two methods for integrating the CNNs. The first method involves training and predicting with the two CNNs independently, followed by making a decision based on their output values. The second method involves combining the two CNNs through a fully connected layer. In both integration methods, we confirmed that the accuracy improved over that of the conventional CNN using only raw images. This suggests that the information obtained from autocorrelation maps contains data that the conventional CNN alone cannot capture.
14:00-14:15	SC5010	**Title:** Enhancing Daytime Train Image Detection with YOLOv8 through Data Augmentation Techniques **Author(s):** Yuta Inoue, Rebeka Sultana, Yosuke Nishino, Ikuko Shimizu **Presenter:** *Yuta Inoue, Tokyo University of Agriculture and Technology, Tokyo, Japan*
		Abstract- Public transportation delays in Japan, stemming from various causes, interrupt business operations and necessitate prompt information sharing with commuters. Conventional systems, such as the

	"Camera-Based Train Delay Management and Information System" and Cascade Classifier, face challenges in accuracy due to fluctuating conditions. Deep learning models, particularly YOLO, have proven highly effective in object detection tasks. One-stage models are favored for real-time detection, and data augmentation techniques improve model performance, especially with limited datasets. This study, therefore, seeks to enhance train detection accuracy in daytime images by employing YOLOv8 and straightforward data augmentation methods, resulting in notable accuracy improvements.	
14:15-14:30	SC5025-A	**Title:** Thermal Target Tracking with a Multi-rotor Drone **Author(s):** Seokwon Yeom **Presenter:** *Seokwon Yeom, Daegu University, Korea*
	Abstract- This work discusses people tracking with drone-based thermal imaging for search and rescue (SAR) missions. The YOLO detection model detects people in thermal videos. Multiple-target tracking is performed via track initialization, maintenance, and termination. Tracks are maintained by a state estimator, Kalman filter. A bounding box gating rule is adopted for the measurement-to-track association. The track-to-track association selects the fittest track for redundant tracks and fuses them. In the experiments, several videos of three hikers simulating being lost in the mountains were captured using a thermal imaging camera on a drone. People walk closely or occluded, and the drone flies in arbitrary directions. Robust tracking results were obtained in terms of average total track life, averamge mean track life, and average track purity.	
14:30-14:45	SC5009	**Title:** Image Memorability Manipulation with Intermediate Guided Diffusion **Author(s):** Jingyi Huang, Edgar Simo-Serra **Presenter:** *Jingyi Huang, Ridge-i Inc., Tokyo, Japan*
	Abstract- In this work we target at manipulating a novel property of images − memorability, which can be interpreted as the degree people can remember an image after viewing it. Image memorability is consid-ered an elusive yet intrinsic attribute, thus is able to be objectively quantified. We consider implementing this on top of the diffusion model framework, where we don't perform forward inference all the way to a completely noisy state but start the reverse process from an intermediate step to preserve the rough features of the original image. In the meanwhile the generated sample is guided by the memorability signal, where we try tuning the intensity and the direction of the guidance to investigate the possibly best strategy. In this way we can automate the content design process for more efficient memorable product creation.	

14:45-15:00	SC5011	**Title:** Empirical Analysis of Image-Based 3D Reconstruction Techniques for Museum Artifacts Incorporating Artificial Features **Author(s):** Tomáš Fiala, Muto Meilan, Rebeka Sultana, Radim Šára, Ikuko Shimizu **Presenter:** *Muto Meilan, Tokyo University of Agriculture and Technology, Tokyo, Japan*

Abstract- This study explores the low-cost enhancement of image-based 3D reconstruction techniques for museum artifacts by incorporating artificial features into the scene. Traditional methods often struggle with reconstructing objects that have reflective surfaces, poor textures, or complex geometries, leading to incomplete or inaccurate digital models. To address these challenges, we introduce artificial features into the reconstruction process, specifically applying them to a case study involving a museum artifact—a vintage sewing machine. By comparing the reconstruction results with and without these artificial features, we demonstrate improvements in the quality of the 3D models. The inclusion of artificial features facilitated better camera parameter estimation and reduced issues such as part duplication. This approach provides a cost-effective solution that can be widely adopted, particularly in the preservation and digitization of cultural heritage artifacts, where traditional surface treatment methods may not be feasible. Our findings suggest that integrating artificial features can be a valuable enhancement for a broader range of 3D reconstruction applications, contributing to more reliable and detailed digital preservation efforts.

Session 6

December 15, 2024

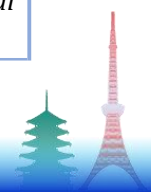**Session 6: Next Generation Artificial Intelligence and Engineering Applications**

Room: 4th Floor, Joint Research Room 7 (共同研究室 7)

Time: 13:00-15:00

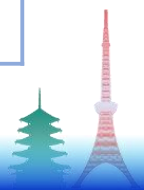**Session Chair: Prof. Manukid Parnichkun, Asian Institute of Technology, Thailand**

Onsite

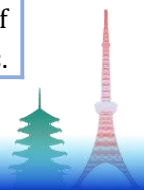|		
13:00-13:15	SC0012-A	**Title:** Use of Artificial Intelligence in Prosocial Micro Lending Platforms **Author(s):** Sanjog Ray **Presenter:** *Sanjog Ray, Indian Institute of Management Indore, India*
		Abstract- Artificial intelligence features are increasingly being implemented in digital platforms. The need for decision systems and abundance of high-quality data on these platforms provide multiple opportunities for the application of machine learning and artificial intelligence algorithms. This paper investigates the application of AI in Prosocial Micro Lending Platforms. Prosocial Microlending Platforms provide a technology platform through which financially weak entrepreneurs can borrow money from individual lenders and lending teams. This paper reviews the extant literature to provide a summary of different aspects of decision support systems problems in the platforms and the use of AI in resolving the issues. The paper primarily focuses on literature regarding recommendation problems, fairness, and algorithm aversion. The main recommender system problem in this domain is matching the borrower to a lender. But because of the complexity of the domain there exists various other problems that recommender systems can help resolve. Factors impacting the adoption of decision support systems in these platforms are also examined.
13:15-13:30	SC0046	**Title:** Adoption of AI and Sensor Technologies in Thai Dairy Farming: Analyzing Factors Influencing Farmer Intentions Using TAM and IDT Frameworks **Author(s):** Davids Makararpong **Presenter:** *Davids Makararpong, University of the Thai Chamber of Commerce, Thailand*

Abstract- The adoption of AI and sensor technologies in dairy farming is becoming increasingly vital for enhancing productivity, improving animal welfare, and reducing operational costs, particularly in Thailand, where smallholder farms dominate the agricultural sector. This study examines the factors influencing the intention of 218 Thai dairy farmers to adopt AI-assisted support technologies, utilizing the Technology Acceptance Model (TAM) and Innovation Diffusion Theory (IDT) as theoretical frameworks. By analyzing compatibility, perceived usefulness, attitudes, and trialability, the research highlights how these technologies can address knowledge gaps and operational challenges, ultimately leading to more efficient and sustainable farming practices. The findings aim to inform policy and support mechanisms to enhance the integration of AI in Thai dairy farming, contributing to the sector's overall growth and modernization.

	SC0047	**Title:** Network Traffic Prediction Using Gradient Boosting Ensemble Method **Author(s):** Haesik Kim **Presenter:** *Haesik Kim, VTT Technical Research Centre of Finland, Finland*

13:30-13:45

Abstract- 6G networks will be complex heterogeneous networks. One of key challenges is how 6G mobile operators deploy and manage their networks efficiently and satisfy different requirements of mobile users. Thus, it is important to monitor the network traffics and evaluate the network performances in real time. The network traffic prediction is one of key tasks when managing the mobile networks. There are many developed network traffic models to capture the statistical characteristics of the actual network traffics. However, it is still difficult to predict network traffic accurately as the networks are getting more complex and the mobile operators should consider many different aspects such as user behaviours, traffic congestion, different network types and so on. There are many attempts to adopt AI algorithms for 6G systems because AI allows us to find network traffic patterns and adopt to varying network conditions. In order to improve the network performance and deploy the network efficiently, network traffic classification, network traffic prediction, anomaly detection and fault detection using AI are widely investigated now. In this paper, we investigate a low complexity network traffic prediction method using gradient boosting ensemble methods and evaluate its performance in a short-term network traffic period. In order to reduce the complexity and convergence time, we find the optimal number of base leaners while not degrading the performance. Under the

given dataset and simulation configuration, we obtain the accurate network traffic prediction method with a low RMSE value and predict the future network traffics.

13:45-14:00	SC0055-A	**Title:** Humanity and AI: Grounding Values for a bright Future !? **Author(s):** Harald Bolsinger **Presenter:** *Harald Bolsinger, Technical University of Applied Sciences Würzburg-Schweinfurt, THWS Business School, Germany*
	Abstract- The presentation shows various key ethical foundations of digitalization and ideological influences behind common statements on AI. For this the basics of worldview analyses and the main functions of human values is introduced. A liberal plead for the retention of ideological diversity in the approaches to the design and use of AI follows. Based on existing global norms like the UN Human Rights Charter a minimal portfolio of basic values is proposed to keep artificial intelligence human as well as digitization. This prepares the ground for a discourse on how to deal with new challenges through AI and how to set a a professional code of conduct for the AI industry. The author would like to see the proposals discussed within the scientific AI community in order to refine, supplement or modify the proposed portfolio of values.		
14:00-14:15	SC0116	**Title:** Adversarial Examples for Field Programmable Gate Array based AI Module **Author(s):** Yusuke Nozaki, Shu Takemoto and Masaya Yoshikawa **Presenter:** *Yusuke Nozaki, Meijo University, Japan*
	Abstract- Applications of edge AI technology are expected in a variety of fields. On the other hand, various threats of attacks have been reported in the application of AI technologies. It is important to evaluate the security of these attacks. Since adversarial examples (AEs) can lead AI to make incorrect inferences, it is required to implement countermeasures against them and to evaluate security. However, there have been insufficient studies on security threats of AE for LUT-Network, which is a field programmable gate array (FPGA) oriented AI implementation. Therefore, this study proposes an AE generation method for FPGA oriented AI module and evaluates its security.		

14:15-14:30	SC0123-A	**Title:** Trust_Gen_Z: Trustworthy Generative AI for Advanced Industrial DigitaliZation **Author(s):** Shahina Begum, Shaibal Barua, Mobyen Uddin Ahmed, Emmanuel Weiten, Rafia Inam and Mikael Nordenstjerna **Presenter: *Shahina Begum**, Mälardalen University, Sweden*

Abstract- The convergence of advanced analytics and AI has paved the way for transformative applications in industrial digitalization. Among these, "prescriptive analytics" stands out as a powerful tool for optimizing decision-making processes by providing actionable insights and recommendations. However, the efficacy of prescriptive analytics in industrial settings relies heavily on the underlying AI models. In recent years, generative AI (gAI) has been one of the most promising advancements in AI, which holds immense potential for revolutionizing industrial digitalization. gAI refers to algorithms capable of creating new content, such as images, text, or even entire virtual environments, based on patterns learned from data. Nevertheless, the black-box nature of gAI poses challenges in understanding the rationale behind prescriptive recommendations, raising concerns about safety, and bias issues. Here, the primary objective is to develop a multi-modal framework for advancing "prescriptive analytics" in industrial digitalization through gAI.

It develops XAI methods based on probabilistic modelling to explain gAI's decisions and investigates multi-modal learning tools for image synthesis and automated analysis of unstructured free text. The proposed prescriptive analytics framework utilizing gAI will drive sustainable industrial digital transformation. The demonstration covers two use cases: Antenna optimization and network maintenance and Undercarriage wear inspection, monitoring, and maintenance.

14:30-14:45	SC1001	**Title:** Exploring the Effectiveness of AI Virtual Humans in Digital Marketing **Author(s):** Ting-Sheng Weng, Ting-Hsuan Lin, Jin-Suei Zheng and Yu-Chieh Chen **Presenter: *Ting-Sheng Weng**, National Chiayi University*

Abstract- Virtual humans have found a broader scope of applications in recent years. Whether in entertainment, education, or business, they have produced ideal results, demonstrating their increasingly prominent role in society. Digital marketing has become a vital domain in today's world. To showcase their strengths and attract more students, universities have transitioned to more effective promotional strategies. As a new means of communication, virtual humans can enhance possibilities for universities.

This study employed artificial intelligence (AI) technology to produce a situation video featuring virtual teachers introducing departments while engaging with virtual students in virtual campuses and classrooms. The video was uploaded to YouTube as part of the university's digital marketing efforts to promote its departments. A total of 250 viewers watched the video and completed a survey afterward. The analysis and corroboration results indicate that the use of virtual humans is practical and impactful for university admissions. The study reveals that after watching the video, most students expressed approval and enjoyment of its content and presentation. They found that virtual humans significantly attracted their attention and had a notably positive effect on promoting admissions.

14:45-15:00	SC0109	**Title:** Leveraging Innovative Technologies and Risk Management Strategies for Smart, Sustainable Warehouses: A Cyber-Physical Systems Case Study in SMEs. **Author(s):** Mariam M. Moufaddal, Asmaa A. Benghabrit, Abderahman A. Kriouile and Oumaima O. Dali **Presenter:** *Asmaa Benghabrit, LMAID Laboratory, Higher National School of Mines Rabat (ENSMR), Morocco*

Abstract- This paper presents a comprehensive approach that includes technological innovations, regulatory frameworks, and a transition towards sustainable practices in warehouse operations. It investigates the incorporation of emerging technologies — such as the Internet of Things, artificial intelligence, and cyber-physical systems — within warehouses to enhance risk management and foster sustainable development. Through a case study of a Moroccan SME utilizing a cyber-physical system as a collaborative robot, we illustrate how these technologies can optimize operations, reduce costs, and lessen environmental impact. The findings reveal notable improvements in safety, customer satisfaction, and resource management, highlighting the transformative potential of these innovations in logistics within an industrial setting.

Session 7

December 15, 2024

Session 7: Semantic Analysis and Data Processing

Room: 3rd Floor, Conference Room No.3 (第三会議室)

Time: 15:20-17:05

Session Chair: Assoc. Prof. Yan Wang, Temple University, USA

Onsite

15:20-15:35	SC0003	**Title:** Unlocking Insights into Mathematics Learning Proficiency: A Data-driven Approach with ANFIS and Data Mining **Author(s):** Mohamad Ariffin Abu Bakar, Ahmad Termimi Ab Ghani and Mohd Lazim Abdullah **Presenter:** *Ahmad Termimi Ab Ghani, Universiti Malaysia Terengganu, Terengganu, Malaysia*

Abstract- The utilization of an Adaptive Neuro-Fuzzy Inference System (ANFIS) and data mining for educational data management represents a remarkable advancement in artificial intelligence applications within the educational evaluation domain. This approach offers significant implications and enhances the efficiency of data handling and analysis. However, the complexity of data orientation and operational intricacies pose challenges for educators in understanding and implementing these techniques. This complexity underscores the necessity for a comprehensive exploration of ANFIS and data mining applications to enhance students' mathematics learning proficiency, addressing both methodological hurdles and practical implications for educational practice. Therefore, this study aims to contribute to the advancement of data-driven approaches for personalised learning interventions and instructional decision-making in mathematics education by introducing a simplified and user-friendly framework, enabling educators to harness the benefits of ANFIS and data mining, particularly in developing predictive models for educational assessment. The paper presents an experimental study based on the introduced framework, highlighting the potential and advantages of ANFIS and data mining in predicting students' proficiency. Through the framework's implementation, a model for predicting student problem-solving proficiency in mathematics learning is described. The obtained values for R-square, Root Mean Square Error (RMSE), and Mean Bias Error (MBE) demonstrate the model's effectiveness in handling real

64

data. The study's implications extend to the enhancement of knowledge and practices in educational assessment, aligning with the pursuit of Sustainable Development Goals (SDG) to improve the quality of education.

15:35-15:50	SC0007	**Title:** Data Pricing for Data Exchange: Technology and AI **Author(s):** Minnu Malieckal, Anjula Gurtoo and Rupsa Majumdar **Presenter:** *Minnu Malieckal, Indian Institute of Science, India*

Abstract- The data economy has grown significantly in the last several years, and governments are now marketing data as a commodity that can be traded, opening up the arena for "data pricing." Despite the rise of data markets, there is lack of research to establish pricing models and specifications for maximising profits when pricing datasets. Furthermore, the magnitude of data providers and massive volume of data impedes data pricing specifications as well. The paper fills both the gaps by first, systematically examining the major data pricing approaches using content analysis methodology, and secondly, applying technological interventions using the thematic analysis methodology. The analysis categorizes the pricing strategies into five main themes: general pricing, quality-based pricing, query-based pricing, privacy-based pricing, and special cases. The technology models applied indicate the forecasting techniques and technology application assessment that are key to pricing frameworks and strategies. The paper lists out different technologies used in various types of data pricing. In doing so, the article establishes an outline for developing technically sound data pricing strategies.

15:50-16:05	SC0077	**Title:** Study on Edge Server Placement using Data Clustering Approach **Author(s):** Hoa Nguyen Thanh Doan, Phat Truong Quang, Duy Hoang Khuong and Tri Nguyen Duc **Presenter:** *Hoa Nguyen Thanh Doan, FPT University, Vietnam*

Abstract- The surge in mobile device usage and the growing demand for rapid data processing have spurred the adoption of edge computing solutions. This paper introduces an innovative data clustering method aimed at optimizing the deployment of edge servers to improve processing efficiency, using Shanghai taxi booking data as a practical illustration. Our goal is to identify the best locations for edge servers, treating each taxi as an individual edge device. These servers will handle incoming requests and route them to the appropriate edge devices within their assigned groups. The primary objective is to minimize server count

65
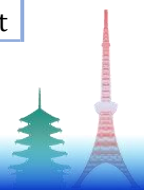
while ensuring the clustering algorithm's quality and accuracy. To achieve this, we explore various clustering algorithms, including Balanced Iterative Reducing and Clustering using Hierarchies, KMEANS, Ordering Points to Identify the Clustering Structure, and Density-Based Spatial Clustering of Applications. We utilize the Silhouette method to determine the optimal number of clusters. Our results reveal that the Balanced Iterative Reducing and Clustering using Hierarchies algorithm produces the most favorable outcomes with the optimal cluster count. This study presents a promising avenue for implementing an edge computing framework to enhance processing efficiency in urban environments.

	SC0117	**Title:** Application of PCA-Whitening Enhanced BERT for Chinese Semantic Understanding in Short Text Product Name Classification **Author(s):** Yifan Zhu, Hanpin Hou, Wei Zhou, Ruiyang Qi, Qiuxia Zhang and Shiyu Jiang **Presenter:** *Yifan Zhu, Beijing Jiaotong University, Beijing, China*
16:05-16:20		

Abstract- This study examines the utility of PCA-Whitening enhanced BERT for Chinese short text product name classification within the context of extreme multi-class mapping. Although traditional machine learning models offer strong interpretability and rapid execution, they are ill-equipped to handle the complexities of semantic relationships inherent in extensive product catalogues. In contrast, BERT exhibits superior semantic comprehension, particularly in the context of polysemy and intricate phrasing. By employing keyword extraction techniques, the input data is effectively reduced in noise, thereby enhancing mapping accuracy. The PCA-Whitening technique optimizes the original embedding space, resulting in a more uniform vector distribution and improved computational efficiency. The experimental results demonstrate that the PCA-Whitening optimized BERT model attains a top-5 accuracy of 95% on brick level (the smallest classifications). Moreover, when the model was applied to a high-quality dataset, the top-1 accuracy reached 99%, which serves to illustrate the considerable impact that data quality has on model performance. The modular design of the model permits prospective enhancements, including multi-standard mapping and multilingual adaptability. In conclusion, this research contributes to the efficiency and effectiveness of product classification systems, with substantial implications for e-commerce, supply chain management, and digital standardization.

	SC1015	**Title:** Online Action Detection via Temporal Dependency Modeling with Semantics Embedding **Author(s):** Sensen Wang and Yuehu Liu **Presenter:** *Sensen Wang, Xi'an Jiaotong University, China*
16:20-16:35		**Abstract-** Online Action Detection aims to identify ongoing incomplete actions from online video streams. The key is to model the inter-frame relationships related to the current action, i.e., temporal dependency. However, frames with similar appearances may contain different action semantics. Therefore, modeling the attention interaction between frame features without considering the action semantics may result in incorrect dependency between frames with similar appearances but different semantics. To address this problem, we propose to utilize action semantics and frame features to jointly model temporal dependency. Specifically, to mine frame-level action semantics, our proposed Semantic Correlation Transformer (SecTR) first extracts representative features from redundant and diverse action frames features by performing K-means clustering for each action category. Considering that a single frame only contains local action information, SecTR reconstructs the frame feature based on clustering features from the perspective of various categories, which is used to mine the action semantics of frames. Then, according to consistency with the semantics of the current frame, SecTR further embeds the frame semantics into the frame features to jointly model temporal dependency. Moreover, to further optimize the modeling quality, we propose an auxiliary task called current action anticipation, which predicts the current action based on historical frames. The effectiveness and superiority of SecTR are validated on THUMOS'14, TVSeries and HDD.
	SC1007	**Title:** Distributed Incremental Ontology Reasoning over Dynamic T-boxes **Author(s):** Bruno Rucy Carneiro Alves De Lima, Merlin Kramer and Victor Henrique Cabral Pinheiro **Presenter:** *Bruno Rucy Carneiro Alves De Lima, University of Tartu, Estonia*
16:35-16:50		**Abstract-** With the advent of Retrieval Augmented Generation (RAG), Knowledge Graphs (KGs) have yet again had a surge in interest in both Academia and Industry, as their use allows for extending the context of Large Language Models (LLMs) by combining traditional vector search with reasoning over Ontologies or Property Graphs encoded as KGs. RAG is a highly dynamic scenario, where the LLM agent might not only retrieve information from a KG or vector store but mutate it as well. This implies eventually there being a greater demand for equally-dynamic KG reasoning systems. We provide a solution to this for the popular ontology

language RDF-Schema (RDFS) by showing that computing entailment as a bottom-up query over RDFS graphs with dynamic Terminological Boxes (TBox) and Assertional Boxes (ABox), those where edges and nodes belonging to both boxes can be freely added and removed, can be expressed as an incremental DBSP computation. This computation is then implemented with the distributed computation framework Differential Dataflow (DD), that subsumes DBSP, and compared with a state-of-the-art commercial ontology reasoner. We find that our approach provides more even performance across additions and deletions and a higher potential for scalability across benchmarks with up to 250 GBs of data.

	SC0125	**Title:** Role of Multi-modal Machine Learning, Explainable AI and Human-AI Teaming in Trusted Intelligent Systems for Remote Digital Towers **Author(s):** Mobyen Uddin Ahmed, Shaibal Barua, Shahina Begum, Waleed Jmoona, Ricky Stanley D Cruze, Alexandre Veyrie, SARA KEBIR and CHRISTOPHE HURTER **Presenter:** *Mobyen Uddin Ahmed, Mälardalen University, Sweden*

Abstract-Remote digital towers (RDTs) represent a transformative advancement in air traffic management (ATM), leveraging cutting-edge technology to enable remote operation by air traffic controllers (ATCOs) while improving efficiency and safety. In the context of RDTs, artificial intelligence (AI), Multimodal Machine Learning (MML) and eXplainable AI (XAI) are playing an increasingly pivotal role in enhancing operational efficiency and safety. However, several challenges need to be addressed, including the development of AI, MML and XAI, research into functional requirements, and the identification of inputs for user and machine interfaces, as well as customization options. This study explores the use of XAI in addressing specific air traffic control challenges and by offering transparent, comprehensible, and actionable insights, XAI fosters resilience, efficient, and closer collaboration between human operators and AI systems. Here, the study defines the specifications for taxiway and runway monitoring and decision support within the RDT domain. It outlines the functional requirements for customized solutions, including XAI, human-centred XAI, human-machine interfaces (HMI), and human-AI teaming (HAIT). A systematic literature review is conducted to assess transparency in AI, with a focus on explainability, HMI, and graphical user interfaces (GUI) within human-centred XAI for RDTs. Additionally, the research identifies state-of-the-art techniques for interactive data visualization, human-centric AI model development, hAIi interfaces, and HAIT, providing a multi-modal agent framework for future development in the RDT domain.

The time slot for this entry is 16:50-17:05.

Session 8

December 15, 2024

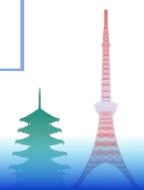**Session 8: Signal and Image Analysis, Data communication**

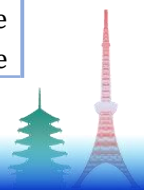**Room: 4th Floor, Joint Research Room 7 (共同研究室 7)**

Time: 15:20-17:20

Session Chair: Prof. Harald Bolsinger, THWS Business School, Germany
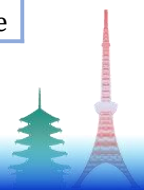
Onsite

15:20-15:35	SC0104	**Title:** Evaluating the Correlation Between Intertrigonal Distance and Anterior Mitral Leaflet Height for Predicting Mitral Annulus Shape Using 3D Modeling **Author(s):** Chirattikan Srisook **Presenter:** *Chirattikan Srisook, Thammasat University, Thailand*

Abstract- Mitral annuloplasty is a widely adopted surgical technique for correcting mitral valve insufficiency. However, concerns have been raised about the selection process for repair devices, which is often guided by the surgeon's experience rather than by geometrical considerations tailored to the patient's specific pathology. Intraoperative sizing techniques for these devices are inconsistent, as they rely on the height of the anterior mitral leaflet — a parameter affected by dilation. The anterior leaflet height frequently increases due to dilation, rendering it unreliable for determining the valve's healthy size. Conversely, the intertrigonal distance remains relatively unaffected by dilation, offering a more stable measurement. This study seeks to analyze the relationship between the mitral valve's key dimensions and to propose a more accurate sizing methodology that could enhance surgical outcomes in mitral valve repair. Using 34 human mitral valves from cadavers, we measured the intertrigonal distance and the anterior mitral leaflet height digitally from using 3D scanner to create 3D models of the mitral valve. The data showed a significant linear relationship with an R^2 value of 0.987. However, individual mitral valve shapes displayed considerable variation, with some cases diverging from the average trend. The findings suggest that a 4:3 ratio between the intertrigonal distance and the anterior mitral leaflet height represents a healthy mitral valve. Incorporating these measurements into device selection may lead to more accurate sizing and better postoperative results.

	SC0111	**Title:** Design and Development of Energy-efficient Edge AI Models for 12-lead ECG Diagnosis System in Rural Areas **Author(s):** Yu-Chi Lin, Azmatullah Khan, Sirapop Nuannimnoi, Milzam Wafi Azhar, Gnangze Elie Gildas Traore, Zheng-Ru Wen and Ching-Yao Huang **Presenter:** *Yu-Chi Lin, National Yang Ming Chiao Tung University*
15:35-15:50		**Abstract-** Artificial intelligence (AI), particularly deep learning, has recently made significant strides in healthcare, offering promising tools for the early detection and management of heart disease, a leading cause of mortality globally. However, the accessibility and quality of such care are often limited in rural areas. Convolutional Neural Networks (CNNs) have been particularly successful in analyzing 12-lead electrocardiograms (ECGs) to detect cardiac anomalies crucial for accurate diagnosis. In this study, we present a comparative analysis of various deep learning models tailored for deployment in resource-constrained rural healthcare environments, emphasizing energy efficiency, low computational demands, and user-friendliness. We introduce a novel Residual Convolutional Gated Recurrent Network (Residual ConvGRU) and evaluate its performance against established benchmarks using 10-fold cross-validation on three widely recognized datasets for classifying different types of cardiac arrhythmias (CAs). We also conduct further model compression experiments on the top 3 models and our proposed model to explore potentials for future deployment in environments with limited computational resources such as in rural areas.
	SC0126	**Title:** Generating Swarm Satellite Data using LSTM and GAN-based models for Detecting Seismic Precursors **Author(s):** Arzaan Kankudti and Yaxin Bi **Presenter:** *Yaxin Bi, School of Computing, Ulster University, UK*
15:50-16:05		**Abstract-** Precise understanding of the evolution mechanisms of earthquakes remains challenging in geophysics, seismology and artificial intelligence (AI) related areas. This work leverages the advance of deep learning predictive (Long Short-Term Memory (LSTM)) and generative (Generative Adversarial Networks (GANs)) approaches to develop two models for generating synthetic time series data based on Swarm satellite data, which is then used for detecting anomalies over seismic evolution. Our results demonstrated the LSTM-based model is capable of making use of the dependence learnt from time-series data to predict synthetic data. By contrast, although the GAN-based model is able to capture the data distribution of the time series, it is not able to discriminate non-informative values produced by the generator very well. These

findings highlight both the promise and challenges associated with applying deep learning techniques to generate synthetic electromagnetic satellite data, underscoring the potential of deep learning approaches for seismic anomaly detection.

	SC0139	**Title:** Analysis of Connectivity Data in Shanghai Telecom Network for Service Optimization **Author(s):** Tuan. Dang Thanh, Duc. Mai Xuan, Hieu. Ha Quang, Trung. Phan Quoc and Khang, Nguyen Van **Presenter:** *Tuan. Dang Thanh*, *FPT University, Vietnam*

16:05-16:20

Abstract- In this paper, we conduct a comprehensive analysis of connectivity data collected from Shanghai Telecom's network, encompassing over 7.2 million connection records from 3,233 base stations and 9,481 mobile devices over six months. Our study aims to evaluate the current network performance, uncover insights into user connectivity patterns, and identify areas for service enhancement. To achieve these goals, we introduce a structured set of 20 analytical questions ranging from foundational metrics to advanced exploratory analyses, designed to assess connectivity quality, geographic usage distributions, peak demand times, and potential for edge system deployment. By addressing these questions, we offer an in-depth understanding of network load distribution, user mobility, service quality, and demand growth trends, ultimately guiding optimization efforts and suggesting strategies for improved service quality and resource allocation across Shanghai Telecom's network. The results and methodolo- gies outlined here provide a foundation for further research in telecom network optimization and edge computing deployment, contributing valuable insights for enhancing connectivity services and planning future infrastructure expansions.

	SC0045	**Title:** Ergonomic Risk Assessment Using Human Pose Estimation with MediaPipe Pose **Author(s):** Chansiri Singhtaun, Suriya Natsupakpong and Pollakrit Lorprasertkul **Presenter:** *Chansiri Singhtaun*, *Kasetsart University, Thailand*

16:20-16:35

Abstract- This research utilizes Human Pose Estimation (HPE) to detect improper working postures that may lead to Work-Related Musculoskeletal Disorders. These postures are evaluated according to ergonomic principles using the OWAS method, which assesses the back, arms, legs, and the weight of the load during the task with codes indicating risk categories. In the experiment, both Two-dimensional (2D) and three-dimensional (3D) keypoints of body joints obtained from MediaPipe

71

	Pose and the angles of the joints relative to 72 different postures, encompassing all combinations of the code for the three body parts, were used as features for setting the criteria in posture prediction. The best classifier obtained from the experiment provided accuracy rates of 90.74% for the back, 92.71% for the arms, and 85.53% for the legs, respectively. When the program was applied to evaluate working postures in a simulated packing station, the accuracy rates for the back, arms, and legs were 70.42%, 82.78%, and 71.25%, respectively. The accuracy in identifying the risk category, which considers all body parts simultaneously, is 77.64%.	
16:35-16:50	SC0078	**Title:** Multi-Objective Optimization for Resource Allocation in Edge Computing: A MOGA-Based Approach **Author(s):** Hoa. Nguyen Thanh Doan, Nghia. Phan Duc, Phu. Nguyen Ngoc Thien, Lap. Hoang Van and Tri. Nguyen Thanh **Presenter:** *Nghia. Phan Duc, FPT University, Vietnam*
	Abstract- Edge computing has emerged as a promising paradigm for addressing the computational demands of IoT applications. In this paper, we present a novel approach to optimize resource allocation in edge computing environments. Our proposed model leverages multi-objective optimization techniques to allocate task processing resources to edge devices effectively. Specifically, we formulate the resource allocation problem as a multi-objective optimization problem aiming to minimize both the distance between request sources and edge devices and the utilization of edge device capacities. To solve this complex problem, we employ the MOGA algorithm (Multi-Objective Genetic Algorithm), which is capable of efficiently exploring the solution space and identifying Pareto-optimal solutions. Through extensive experimentation and comparative analysis with existing models, we demonstrate the superiority of our proposed approach in terms of processing time and resource utilization. Our findings underscore the effectiveness of multi-objective optimization in enhancing resource allocation strategies for edge computing environments, offering significant improvements in both performance and efficiency.	
16:50-17:05	SC1004	**Title:** Enhancing Anomaly Detection in Cycling Paths Using a Hybrid LSTM-VQ-VAE Deep Learning Model **Author(s):** Kewei Ren and Chun-Hsing Ho **Presenter:** *Chun-Hsing Ho, University of nebraska-Lincoln, USA*
	Abstract- Detecting anomalies in bike lanes using acceleration data presents significant challenges, particularly with unlabeled datasets. To address this, we introduce a novel hybrid model that integrates a Vector	

	Quantized Variational Autoencoder (VQ-VAE) with a Long Short-Term Memory (LSTM) network. This architecture leverages the VQ-VAE's capability to map input samples to discrete vectors, effectively filtering out noise and extraneous details, while the LSTM component captures the essential temporal correlations for accurate anomaly detection in time-series data. By synergizing these technologies, our model not only identifies crucial features but also compresses them into a latent space representation, which is refined through a multi-stage training process. Our experiments demonstrate that the pretrained LSTM-VQ-VAE model can effectively identify and localize pavement anomalies, achieving up to 90% accuracy across various distress types such as potholes and cracking. The integration of LSTM with VQ-VAE not only conserves human and material resources but also marks a substantial advancement in the automated detection of transportation infrastructure anomalies.	
17:05-17:20	SC1012-A	**Title:** Collision Avoidance Algorithm to Optimize the Path of the Drone System against Dynamic Obstacles **Author(s):** C. H. Kim and Dowan Cha **Presenter:** *Dowan Cha, Pai Chai University, Korea National Defense University*
	Abstract- In this study, we present a new Collision Avoidance Algorithm(CAA) to optimize the path of the drone system against dynamic obstacles. We utilized a camera and a LiDar to detect dynamic obstacles. The drone system calculates the relative velocity of dynamic obstacles, and the flight radius and obstacle progress path are compared to determine the dangerous state of the drone system based on proposed CAA. CAA determines one of among the three cases (deceleration-stop, evasive maneuver, normal flight), and the flight of the drone system proceeds according to the pre-set flight purpose (fast flight, stable flight). We evaluated CAA using Trajectory Planner in Multi-Agent and Dynamic Environments(MADER). As a result, CAA guaranteed safety of the drone system against dynamic obstacles and showed up to 50.3% reduction in the flight time compared to basic CAA provided by MADER.	

73

Poster Session

December 15, 2024

Poster Session: Intelligent Image Processing and Artificial Intelligence Application Technology
Time: 15:00-15:20

Onsite

SC0019	**Title:** An Improved Infectious Disease Risk Prediction Model Based on Attention Mechanism **Author(s):** Yingshuai Wang, Yanli Wan, Qingkun Chen, Xingyun Lei, Yan Wang, Guoqiang Sun, Xiaoze Li and Hongpu Hu **Affiliation:** *Chinese Academy of Medical Sciences & Peking Union Medical College, China*

Abstract- Most patients infected with COVID-19, caused by the coronavirus, exhibit mild to moderate respiratory symptoms and have the ability to recover on their own. However, for the elderly and those with underlying conditions such as cardiovascular disease, diabetes, chronic respiratory diseases, and cancer, infection with COVID-19 can lead to severe illness and even pose a life-threatening risk. In this global pandemic, the scarcity and rational allocation of medical resources have become particularly critical. Therefore, accurately predicting the mortality risk of patients and accordingly allocating medical resources under limited conditions is crucial for effectively reducing the burden on the healthcare system and ensuring the health and safety of the general public. This paper aims to develop an improved machine learning model capable of predicting infectious disease risk. The model will assess whether COVID-19 patients are at risk of death or in a high-risk state based on their real-time symptoms, current health status, and medical history.

SC0020	**Title:** A Study on Top-level Design and Implementation Pathway of Intelligent Primary Medical and Health Services **Author(s):** Hongpu Hu, Yingshuai Wang, Yanli Wan, Xingyun Lei, Liqin Xie, Xiaoze Li, Guoqiang Sun, Jian Li and Jianhong Yao **Affiliation:** Chinese Academy of Medical Sciences & Peking Union Medical College, China

Abstract- With a focus on the needs of service collaboration, the study aimed to address the challenge of building, sharing, and applying quality primary health knowledge and data resources. It presented a top-level design and implementation path for the first-ever establishment of an intelligent primary health service system in China, which integrates management, systems, data, standards, and knowledge application, among other aspects. A
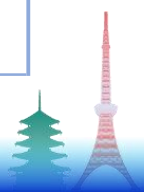
multi-source heterogeneous data fusion model had been created using the complex multi-source heterogeneous data fusion technology; By utilizing crucial techniques such as knowledge mining, representation, and knowledge base construction, and leveraging expert advantages and quality resources, a standardized and professional intelligent primary health knowledge base had been built, forming quality primary health knowledge resources; Based on the advanced technology for constructing an intelligent primary health data center, an integrated and unified intelligent primary health data resource center, with high process integration, stable model tools, and interpretable knowledge, centered on professional knowledge and model bases, has been established for the first time in China, promoting the sharing and application of quality resources through multi-scenario intelligent decision-making applications for primary health services. Based on a data and knowledge dual-driven approach, this study can help address the challenges posed by digitization and intellectualization of primary health, promote the healthy development of national primary health informatization, improve the quality and efficiency of primary health services, strengthen disease prevention and treatment capabilities, and ensure the quality of health and medical big data from the source.

SC0054-A	**Title:** Development of Ensemble Learning Approaches for the Automatic Detection of Sleep-Disordered Breathing based on Single-lead ECG signal **Author(s):** Tae Gwan Jang[1], So Yeon Park[1], Ha Young Park[2], Ji Ae Lee[2], E-Loo Kim[3], Gahyeon Yun[3], Doljinsuren Enkhbayar[3] and Erdenebayar Urtnasan[2,4] **Affiliation:** 1: Yonsei Wonju Health System, Yonsei University, Republic of Korea; 2: Yonsei Insititue of AI Data Convergence Science, Yonsei University, Republic of Korea; 3: Department of Biomedical Engineering, Yonsei University, Republic of Korea; 4: Artificial Intelligence Big Data Medical Center, Yonsei Wonju Health System, Yonsei University, Republic of Korea

Abstract- Sleep-disordered breathing (SDB) is a common sleep disorder characterized by partial or complete pauses in breathing during sleep due to upper airway obstruction. According to the 2024 American Thoracic Society (ATS), over 85% of individuals with SDB remain undiagnosed, emphasizing the importance of early detection and diagnosis.

This study aims to propose ensemble learning methods for the automatic detection of SDB using single-lead electrocardiography (ECG) signals. The approach involved four ensemble learning algorithms: Random Forest, XGBoost, LightGBM, and CatBoost, which were fine-tuned for optimal performance. ECG signals from 100 participants were segmented into 30-second intervals and classified into three categories: 82,036 normal instances, 87,141 hypopnea instances, and 57,945 apnea instances. Data were split 70% for training and 30% for testing. The Random Forest model achieved an overall accuracy of 69.6%. Class-specific accuracy showed 66.85% for normal cases, 78.35% for hypopnea, and 60.38% for apnea cases. Single-lead ECG data can effectively detect SDB including hypopnea and apnea events, it can potentially contribute to the development of early and easy detection of SDB in sleep medicine.

SC0107-A	**Title:** Real-Time Gesture-Controlled eTheremin with AI-Driven Audio Generation **Author(s):** Tsen-Fang Lin[1] and Liang-Bi Chen[2] **Affiliation:** 1: Southern Taiwan University of Science and Technology 2: National Penghu University of Science and Technology

Abstract- This paper presents the development of a real-time eTheremin performance system, enhanced with AI-driven audio generation. The system utilizes hand gestures to control pitch and volume, employing MediaPipe for hand tracking and TensorFlow-based AI models, such as Magenta's NSynth, for dynamic sound synthesis. Left-hand movements adjust volume, while right-hand movements control pitch, offering an intuitive and immersive musical experience. By leveraging the computational power of the Mac Mini M2 Pro, the system generates complex, AI-driven soundscapes in real-time, providing a seamless interaction between user gestures and audio output. User testing revealed that the system's innovative combination of gesture control and AI-generated sound offers an engaging and expressive platform for both professional musicians and enthusiasts. Potential applications include music education, interactive performances, and digital sound design. This work demonstrates the integration of AI in musical instrument interfaces, bridging the gap between traditional instruments and modern, AI-enhanced music technologies.

SC1014	**Title:** LHE-PSN: A Photometric Stereo Network with Lightweight High-frequency Enhancement Module **Author(s):** Chenning Kang, Yonghong Song and Yuehu Liu **Affiliation:** Xi'an Jiaotong University, China

Abstract- Photometric stereo is a well-established technique in computer vision for recovering surface normals from a set of images captured under varying lighting conditions. Traditional approaches often rely on simplified reflectance models to approximate the characteristics of non-Lambertian surfaces. On the contrary, modern methodologies have endeavored to employ deep learning in photometric stereo, a technique that has received widespread acclaim for effectively addressing the complexities associated with non-Lambertian surfaces. Nevertheless, challenges persist in the high-frequency regions, prompting the development of various techniques to enhance recovery outcomes. However, these improvements often involve more intricate network structures, leading to increased demands on parameters and training process. To address these challenges, we present an innovative approach to calibrated photometric stereo, employing a network architecture coupled with dilated convolutions. In comparison to existing methods, this integration significantly enhances the accuracy and robustness of surface normal vector estimation, all without the addition of more parameters. In this paper, we (1) incorporated a high-frequency enhancement module consisting of dilated convolutions to augment high-frequency information, (2) proposed an innovative feature aggregation module designed to capture both local and global features simultaneously, and (3) validated the performance of our method on real-world scenarios, affirming the efficacy of our approach.
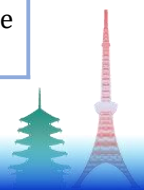
SC5013	**Title:** Ablation Study of Attention Modules integrated with Deep Learning Networks for Ship Detection **Author(s):** Hyunjin Jo and Byung Wook Kim **Affiliation:** Changwon National University, South Korea

Abstract- Despite the ship detection using Synthetic Aperture Radar (SAR) images is crucial for various marine applications, small-sized ship detection is still difficult due to factors such as viewpoint, backgrounds, illumination occlusion, and multi-scale. In this paper, we presented the ablation study of various attention modules in improving the performance of deep learning-based ship detection models on SAR images. To identify the most effective attention modules based on object size, we conducted a series of experiments using ResNet34 and ResNet50 with the backbone network. In addition, we investigated image preprocessing techniques that are effective for ship detection. The experimental results show that convolutional block attention module (CBAM) performs better for small object detection in simple networks such as ResNet34, and spatial attention modules perform better for complex networks such as ResNet50. Furthermore, adjusting image scales according to object size markedly enhanced detection performance, particularly for small ships. These findings not only highlight the importance of tailoring attention mechanisms to the network architecture and object size but also provide valuable insights into designing more effective SAR ship detection models. These results can contribute to designing effective networks in the field of SAR ship detection.

SC0072-A	**Title:** A Design of Inference Engine for Recurrent Neural Networks with Block-Circulant Weight Matrices Using an FFT-based SoC-FPGA Approach **Author(s):** Song-Nien Tang and Ting Chang **Affiliation:** Chung Yuan Christian University

Abstract- Inference engines for recurrent neural networks (RNNs) are increasingly important in edge artificial intelligence applications. To improve the operation capabilities in an RNN inference engine handling large weight matrices, recent research suggests using a block circular (BC) weight matrix method that reduces the ordoer of each n-dimensional weight matrix block from $O(n2)$ to $O(n)$, thereby improving the RNN weight storage efficiecny in a resource-limited edge computing environment. Based on the BC weight matrix scheme, we present an inference engine design for RNNs using a software-hardware (SW-HW) co-design strategy on an SoC-FPGA platform. In the proposed design, the SW side is responsible for the RNN inference operation control, while the HW site conducts the gate operation that mainly consists of BC-matrix vector multiplication using the Fast Fourier Transform (FFT) computation. Based on the proposed SW-HW co-design approach, flexible and HW-efficient RNN inference operations can be performed using an SoC-FPGA platform. Our RNN inference engine design was successfully validated through the experiment of phoneme classification using the TIMIT dataset.
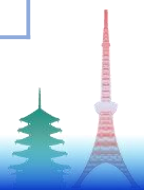
SC5012	**Title:** A Fast Statistical Image Reconstruction Method for Spectral Photon-Counting CT **Author(s):** Heejeong Kim, Tomohiro Reo, Hiroyuki Kudo **Affiliation:** University of Tsukuba, Japan

Abstract- In this study, we propose a statistical image reconstruction method for spectral photon-counting CT (SPCCT). The proposed method consists of two approaches: Method 1, which formulates image reconstruction using energy images as variables, and Method2 which formulates image reconstruction using material images as variables. Both methods incorporate the following key innovations: (1) Utilizing a weighted least squares (WLS) data term that accounts for statistical noise following a Poisson distribution. (2) Incorporating a regularization term in the energy domain that exploits the sparsity of materials constituting the object, in addition to the spatial total variation (TV) regularization term. (3) Reducing computation time through a fast-convergence iterative method based on Dykstra-like splitting, as detailed in [1]. Significantly, the development of an iterative method that exceptionally fast convergence using Dykstra-like splitting to mitigate computational challenges is considered a major achievement.
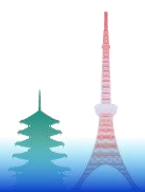
Session 9

December 16, 2024
Time Zone: UTC+9

Session 9: Machine Vision and Virtual Technology Applications
Zoom ID: 876 8000 7833
Zoom Link: https://us02web.zoom.us/j/87680007833
Time: 10:35-12:05
Cahir: Prof. Chien-Sing, LEE, Sunway University, Malaysia

Online

10:35-10:50	SC0051	**Title:** Advancements and Challenges in Cloud Computing: Multi-Cloud Management, Security, and AI-Driven Threat Mitigation **Author(s):** Robert Dilworth, Charan Gudla and Md Shohel Rana **Presenter:** *Robert Dilworth, Mississippi State University, USA*

Abstract- The rapid adoption of cloud computing has introduced a new set of challenges for organizations, particularly with the rise of multicloud architectures. The complexity of managing multiple interconnected cloud instances exponentially increases the risk of security vulnerabilities, such as misconfigurations and virtualization mismanagement. As the cloud surface area expands, the probability of human error also rises, leading to a heightened susceptibility to DDoS attacks. In this paper, we examine the interplay between multi-cloud environments, virtualization, system misconfigurations, and the threat of Distributed Denial-of-Service (DDoS) attacks. This review analyzes recent research to understand the current state of cloud computing security, focusing on the challenges posed by multi-cloud topologies. The findings suggest that the increased attack surface and the inherent difficulties in maintaining consistent configurations across multiple cloud providers create an environment ripe for exploitation by malicious actors. By synthesizing existing scholarship, this paper aims to shed light on novel solutions proposed in the literature and the efficacy of contemporary practices in mitigating the risks associated with multi-cloud environments, virtualization, misconfiguration, and DDoS attacks. The insights gained from this study can inform cloud service providers and organizations in developing more robust and secure cloud computing strategies.

10:50-11:05	SC0040	**Title:** Towards Resilient and Efficient LLMs: A Comparative Study of Efficiency, Performance, and Adversarial Robustness **Author(s):** Xiaojing Fan and Chunliang Tao **Presenter:** *Xiaojing Fan, New York University, USA*

Abstract- With the increasing demand for practical applications of Large Language Models (LLMs), many attention-efficient models have been developed to balance performance and computational cost. However, the adversarial robustness of these models remains under-explored. In this work, we design a framework to investigate the trade-off between efficiency, performance, and adversarial robustness of LLMs and conduct extensive experiments on three prominent models with varying levels of complexity and efficiency – Transformer++, Gated Linear Attention (GLA) Transformer, and MatMul-Free LM – utilizing the GLUE and AdvGLUE datasets. The AdvGLUE dataset extends the GLUE dataset with adversarial samples designed to challenge model robustness. Our results show that while the GLA Transformer and MatMul-Free LM achieve slightly lower accuracy on GLUE tasks, they demonstrate higher efficiency and either superior or comparative robustness on AdvGLUE tasks compared to Transformer++ across different attack levels. These findings highlight the potential of simplified architectures to achieve a compelling balance between efficiency, performance, and adversarial robustness, offering valuable insights for applications where resource constraints and resilience to adversarial attacks are critical.

11:05-11:20	SC0067	**Title:** Confidential Trends: Assessing the Feasibility of Transformers in Private Time Series Forecasting. **Author(s):** Bhagiradh Kantheti and Carlos A. Paz De Araujo **Presenter:** *Bhagiradh Kantheti, University of Colorado Colorado Springs, USA*

Abstract- Transformers have become the state-of-the-art architectures for Natural Language Processing (NLP) by capturing the long-range dependencies with the help of attention mechanisms. As the use of Transformers is increasing in the field of time series analysis, we conducted our research on assessing the feasibility of applying secure inference to time series Transformers using CrypTen, a privacy-preserving framework for machine learning applications through Multi Party Computation (MPC). Our approach uses only the encoder part of Transformers for training on plaintext data and will conduct inference on the CrypTen. This methodology combines the predictive power of Transformers and the security power of CrypTen, and we propose a novel foundational framework for privacy-preserving time series analysis using Transformers. Despite adding private inference, we got performance

	results similar to those of Informer model. This work represents a significant step towards secure predictive analytics showcasing the potential of Transformers in privacy-preserving time series forecasting.	
11:20-11:35	SC0052	**Title:** Cloud Computing and Security: An Overview of Vulnerabilities, Cyber Attacks, and AI-Driven Solutions **Author(s):** Robert Dilworth and Charan Gudla **Presenter:** *Robert Dilworth, Mississippi State University, USA*
	Abstract- Cloud computing has emerged as a cornerstone of modern IT infrastructure, offering scalability, flexibility, and cost-efficiency. However, despite its widespread adoption and apparent advantages, cloud environments remain susceptible to various vulnerabilities and cyber attacks. This paper delves into the intricacies of cloud computing, highlighting its significance while shedding light on the persistent challenges posed by security threats. By intertwining the realms of Artificial Intelligence (AI), Machine Learning (ML), Language Learning Models (LLMs), Generative Adversarial Networks (GANs), and Data Science, we explore the potential of these technologies in fortifying cloud security. Our discourse extends beyond mere observation, presenting a comprehensive analysis of vulnerabilities, cyber attacks, and AI-driven solutions. We propose innovative approaches such as automation misconfiguration detection, threat detection, and log analysis, underlining the pivotal role of AI in mitigating security risks in cloud environments.		
11:35-11:50	SC0048	**Title:** Multi-modal clothing recommendation model based on large model and VAE enhancement **Author(s):** Bingjie Huang, Qingyi Lu, Shuaishuai Huang, Xue-She Wang and Haowei Yang **Presenter:** *Bingjie Huang, Independent researcher, Sunnyvale, CA, USA*
	Abstract- Accurately recommending products has long been a subject requiring in-depth research. This study proposes a multimodal paradigm for clothing recommendations. Specifically, it designs a multimodal analysis method that integrates clothing description texts and images, utilizing a pre-trained large language model to deeply explore the hidden meanings of users and products. Additionally, a variational encoder is employed to learn the relationship between user information and products to address the cold start problem in recommendation systems. This study also validates the significant performance advantages of this method over various recommendation system methods through extensive ablation experiments, providing crucial practical guidance for the comprehensive optimization of recommendation systems.		

	SC0136	**Title:** ChatGPT Does Not Understand Moore **Author(s):** Aminata Sabané and Tegawendé F. Bissyandé **Presenter:** *Aminata Sabané, Université Joseph Ki-Zerbo, Burkina Faso&CITADEL, Burkina Faso*
11:50-12:05		**Abstract-** The advent of Large Language Models (LLMs) likely constitutes a turning point in the history of mankind. LLMs have the potential to revolutionize the way we interact with computers and each other. While current state of the art LLMs mostly support western languages, such as English, it is paramount that we consider the opportunity of integrating African languages into LLMs. First, LLMs can be used to generate text, translate languages, and answer questions in African languages. This can help to make information and resources more accessible to people who speak African languages. Second, LLMs can be used to develop African language keyboards, chatbots, and voice assistants. This can help to make African languages more accessible and user-friendly. Third, LLMs can be used to generate new content in African languages, translate African language texts into other languages, and create educational resources in African languages. This can help to ensure that African languages are used and spoken in the future. Despite the potential benefits of LLMs for African languages, there are still a number of challenges that need to be addressed before LLMs can be widely deployed for indigenous communities in Africa. In this paper, we will discuss the challenges for training a large language model with African languages in more detail. We will also discuss some of the recent efforts to address these challenges.

Session 10

Session 10: Next Generation Artificial Intelligence Theory and Application Technology
Zoom ID: 876 8000 7833
Zoom Link: https://us02web.zoom.us/j/87680007833
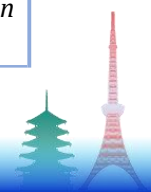Time: 13:20-15:05

Chair: Prof. Criselle J. Centeno, MIT, Philippines

Online

	SC0130	**Title:** MHSLEARN: E-Learning System with Literacy Test and Automatic Speech Recognition Technology for Manila High School **Author(s):** Mariah Carmella Panlilio Santander, Catherine Nicole Perez, Falve James Bucasas, Criselle Centeno, Ariel Antwaun Rolando Sison and Jethro San Diego **Presenter:** *Mariah Carmella Panlilio Santander, Pamantasan ng Lungsod ng Maynila, Philippines*
13:20-13:35		**Abstract-** MHSLEARN is a web-based e-learning system developed for Manila High School to support both regular and Alternative Learning System (ALS) students. It provides essential tools for remote learning, including access to educational materials, user dashboards, and progress tracking. A key feature is the enhanced Functional Literacy Test (FLT), designed to assess reading and comprehension levels to gauge ALS students' readiness for various curriculum stages, with new classification and real-time feedback capabilities. Additionally, Automatic Speech Recognition (ASR) technology enables reading assessments, offering immediate feedback to support literacy development. Built using the Agile Software Development Life Cycle (SDLC), MHSLEARN received high satisfaction ratings from students, teachers, and IT experts in functional suitability, portability, and usability. These findings underscore MHSLEARN's effectiveness in addressing remote learning and literacy assessment needs, providing a comprehensive, user-friendly solution for diverse learners at Manila High School.
13:35-13:50	SC0002	**Title:** Optimization of AI large models for scientific research **Author(s):** TianYuan **Presenter:** *TianYuan, Qingdao Enrolment and Examination Institute, China*

Abstract- This paper analyzes the limitations and challenges of general large models in scientific research, proposes enhanced technology paths including data enhancement, computation enhancement, model enhancement, security enhancement, and domain enhancement. It designs a circular optimization process that includes enhanced algorithms, professional fine-tuning, and human-computer collaboration, and outputs optimized scenarios for the application of large scientific research models.

13:50-14:05	SC0033	**Title:** Challenges and Limitations of Using Artificial Intelligence Generated Content (AIGC) with ChatGPT in Programming Curriculum: A Systematic Literature Review **Author(s):** Chi In Chang, Wan Chong Choi and Iek Chong Choi **Presenter:** *Wan Chong Choi, Illinois Institute of Technology, USA*

Abstract- This systematic literature review examined the challenges and limitations of integrating Artificial Intelligence Generated Content (AIGC) tools into programming curricula. Following Kitchenham's framework, a comprehensive search was conducted across ACM Digital Library and Web of Science databases to select 22 relevant peer-reviewed papers. Key issues identified included increased plagiarism risk due to AI-generated unique code, accuracy and reliability concerns of AI outputs, privacy and data security risks, curriculum changes to incorporate prompt engineering and critical thinking, and challenges in traditional assessment methods. Additionally, the review highlighted the potential over-reliance on AI tools, which could hinder the development of fundamental programming skills. Implementation challenges and inadequate training for educators further complicate the integration process. Addressing these challenges is crucial for leveraging the benefits of AIGC tools while maintaining academic integrity and enhancing learning outcomes in programming education. Future research should focus on developing strategies to mitigate these issues and ensure the responsible use of AI in education.

14:05-14:20	SC0131	**Title:** DDYSLEXIAGAME: An Educational Mobile Application Game for Children with Double Deficit Dyslexia Using Multisensory Orton-Gillingham Activity **Author(s):** Mary Chaylze Rivo Ignacio, Lionel Ray C. Purugganan, Rhiyen Lee Soriano, Criselle J. Centeno, Jethro L. San Diego and Mark Anthony S. Mercado **Presenter:** *Mary Chaylze Rivo Ignacio, Pamantasan ng Lungsod ng Maynila, Philippines*
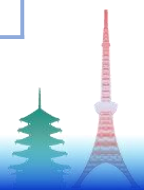
Abstract- Children with Double Deficit Dyslexia (DDD) face significant challenges with phonological processing and rapid naming, which hinder their reading and writing skills. To address these difficulties, the study titled "DDDYSLEXIAGAME: An Educational Mobile Application Game for Children with Double Deficit Dyslexia Using Multisensory Orton-Gillingham Activity" developed an interactive mobile app based on the Orton-Gillingham method. This approach emphasizes multisensory learning, integrating text-to-speech (TTS) and speech-to-text (STT) technologies to create a richer learning experience for children.

The study evaluated the app's effectiveness using surveys and questionnaires from parents and educators of children with DDD, following ISO-25010 standards to assess various quality factors. Results indicated that the TTS and STT features significantly improved children's phonological awareness and rapid naming abilities, with a functional completeness score of 3.11. Usability tests showed a satisfaction rate of 3.05, demonstrating that the app was both easy to use and engaging for children and their parents. Overall, DDDYSLEXIAGAME effectively supports children with Double Deficit Dyslexia by providing fun, multisensory activities that enhance their literacy skills. The study underscores the potential of mobile applications as valuable educational tools and suggests that further development could improve accessibility and customization, making these resources even more effective for diverse learners.
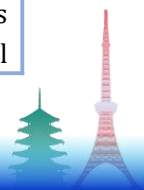
	SC0053	**Title:** A New Cross-Cultural Exploration of English Loose and Tight Vowel Listening Ability of Han and Uyghur College Students by AI Fusion of Deep Learning and Big Data under the E-Prime Platform **Author(s):** YingXin Gao, ZhiWen Zhang and Lei Guo **Presenter:** *YingXin Gao, Northwest Minzu University, China*
14:20-14:35		**Abstract-** Utilizing the E-Prime platform, this study pioneeringly integrates advanced AI algorithms with deep learning and big data technologies to undertake an innovative cross-cultural examination of the distinct abilities of Han and Uyghur college students in discriminating between loose and tight vowels in English. Rooted in the theories of language petrification and native language transfer, the research systematically probes into and compares the unique performance of these two ethnic groups in the realm of speech perception through a rigorously designed experimental framework. The findings reveal that Uyghur students exhibit a notable advantage in discerning English loose and tight vowels, likely stemming from the specific phonological similarities between their mother tongue and English vowels, as well as potential

cumulative effects of English language exposure. Additionally, the study underscores the significant impact of specialized English education, as English majors consistently demonstrate higher accuracy in vowel discrimination compared to their non-major peers.This research contributes a valuable cross-cultural lens to the field of English language teaching, elucidating the intricate mechanisms by which diverse ethnic backgrounds influence students' speech perception capabilities. Furthermore, it establishes a solid theoretical foundation and offers practical guidance for comprehensively enhancing the speech perception abilities of multi-ethnic student populations. By harnessing the power of AI, deep learning, and big data technologies, this study pioneers new avenues for future speech instruction and research, fostering innovation and progress in the realm of educational technology.

	SC0087	**Title:** Eco-Friendly AI-Based Classification System for Payment Card Support Tickets **Author(s):** Metouole Mwinbe Yves Ghislain Somda, Othmane Tahar, Thierry Kondengar, Rodrique Kafando, Mamadou Ba, Samuel Ouya and Gervais Mendy **Presenter:** *Metouole Mwinbe Yves Ghislain Somda*, *Université Cheikh Anta Diop, Dakar, Sénégal*
14:35-14:50		**Abstract-** The payment card industry has revolutionized the way we use and interact with money. The Card Payment Support Department is responsible for customer management, ensuring that bank cards, ATMs, POS terminals, and online stores function properly, allowing users to make transactions with their bank cards anytime and anywhere. The Payment Service Center serves as the first point of contact for customers, and its main challenges include the ability to quickly categorize customer requests, consistently resolve tickets within the SLA (Service Level Agreement), and enhance the user experience. In this paper, we propose a fine-tuned Large Language Model (LLM) to automatically classify incoming customer requests. Our model achieved an average accuracy of 76.74% across 51 categories, using only 3 epochs. The complete code and model are available online. We also developed a web application with Streamlit that leverages our model to predict the category of incoming tickets and automatically assign the request to the correct department when the model's confidence is higher than 40%. Through this study, we demonstrate that it is possible to achieve good results in AI model training with minimal computational resources, which is a key consideration for the financial sector in West Africa and also beneficial for the environment.

14:50-15:05	SC0035	**Title:** Self-Supervised Feature Transmission for improving Supervised Representations **Author(s):** Meiyi Yang, Zilin Ding and Xiaomin Wang **Presenter:** *Meiyi Yang, University of Electronic Science and Technology of China, China*

Abstract- Pretext tasks are effective technological means to improve the performance of target tasks by constructing semantically meaningful image representations. Traditional pretext tasks are designed by transforming the input image, which causes multiple expansions of training data accompanied followed by immense training costs. For the semantic information of high-level features are similar to images, we argue that feature-based pretext task is conducive to representation learning with less training costs. This work shows that feature transformations can also be regarded as supervisory signals to construct the pretext task. Specifically, we propose a general feature-based pretext task to transform the feature maps by discarding different regions. And then, original labels are expanded to joint labels to identify the discarded part during training. We apply this method to supervised classification tasks and find it is a simple yet effective augmentation strategy for improving supervised representation. In experiments, we demonstrate that feature transformations consistently enhance the quality of supervised representations with a low computation cost. Furthermore, we confirm our approach can be compatible with existing methods with negligible computational overhead via extensive contrastive studies across model and dataset sizes.

Session 11

December 16, 2024

Time Zone: UTC+9

Session 11: Natural Language Processing, Machine Vision and Virtual Technology Applications
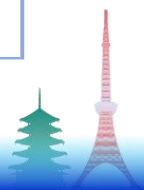**Zoom ID: 876 8000 7833**
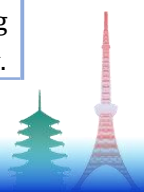**Zoom Link: https://us02web.zoom.us/j/87680007833**
Time: 15:25-17:25
Session Chair: Prof. Chinthaka Premachandra, Shibaura Institute of Technology, Japan

Online

|		
	SC1010	**Title:** VGuard: Development of Mobile Real-Time Vehicle Damage Detection Application with Image Recognition using YOLOv8 **Author(s):** Criselle J. Centeno, Jethro R. San Diego and Mark Christoper L. Blanco **Presenter:** *Criselle J. Centeno, Pamantasan ng Lungsod ng Maynila, Philippines*
15:25-15:40		**Abstract-** Innovations in the modern automotive industry have transformed vehicle maintenance, but inspections remain to be labor-intensive and lack visual representation. This study proposes VGuard, a mobile application that integrates On-Board Diagnostics (OBD-II) data retrieval and You Only Look Once-based (YOLOv8) image recognition for external damage detection, and a dual visualization of 2D and 3D to streamline vehicle health inspections. The app utilizes an ELM 327 scanner for engine data, while a dataset of 4,609 images of vehicle damage, specifically of rust, scratches, dents, and cracks, were used to train the image recognition model. This achieved an mAP of 72.1%, 82.7% precision, and 63.7% recall. Developed with Flutter, VGuard was evaluated by sixty (60) respondents using a 4-point Likert scale survey based on ISO 25010 standards. The app received an overall rating of 3.36, corresponding to "Excellent" and demonstrating its effectiveness in data retrieval, damage detection, and display. Recommendations include improvements in data dependency, detection accuracy, usability, user interface, and additional features for further development.

15:40-15:55	SC0071	**Title:** Advanced Liver Tumor Detection in Abdominal CT Scans Using Modified YOLOv8 **Author(s):** Zhenjiang Li and Jiajun Zhou **Presenter:** *Zhenjiang Li, University of Electronic Science and Technology of China, Chengdu, China*
	Abstract- Liver tumors are one of the leading causes of cancer-related mortality, and early, precise detection is crucial for effective treatment. Abdominal CT scans, particularly in the arterial phase (AP) and portal venous phase (PVP), provide critical information for tumor detection. In this paper, we propose an advanced liver tumor detection method based on a modified YOLOv8 model, incorporating several key innovations. First, the traditional Intersection over Union (IoU) loss is replaced with Normalized Wasserstein Distance (NWD) loss, improving bounding box accuracy. Second, we substitute standard convolution layers with deformable convolutional networks (DCNv4), enhancing the model's ability to capture complex deformations in tumor shapes. Finally, we integrate the BiFormer attention mechanism into the final layer of YOLOv8 to better capture intricate features. Extensive experiments on abdominal CT data demonstrate the superiority of our approach, yielding high detection accuracy and robustness, particularly in detecting small tumors.		
15:55-16:10	SC0018	**Title:** Enhancing Autonomous Driving Safety through PRDQN and Zero-Shot segmentation in Real-Time Traffic Scenarios **Author(s):** Aoran Li and Hong Liu **Presenter:** *Aoran Li, East China Normal University, China*
	Abstract- Safety is a key prerequisite for autonomous driving systems, yet the many unpredictable corner cases on public transportation remain a huge hazard. By definition, a corner case is the presence of unpredictable and relevant objects/categories at the location in question, including sudden traffic accidents, unmarked roadblocks, and so on. To this end, we introduce a perceive everything autonomous approach that can still perceive shapes and categories in real-time traffic scenes with zero-shot learning. In addition, considering the scarcity of corner cases, we implement the DQN algorithm with prioritized experience replay (PER) to effectively balance the empirical equilibrium between corner cases and generic cases. Finally, we designed four different trajectories on CARLA simulator, a real-time simulator for autonomous driving, and compared them with other autonomous driving algorithms to achieve very excellent results. In addition, we perform ablation experimental analyses of our own models to validate the effectiveness of the segmentation everything algorithm module and the DQN module with prioritized experience replay.		

16:10-16:25	SC5015	**Title:** Emotion-Guided Image to Music Generation **Author(s):** Souraja Kundu, Saket Singh, Yuji Iwahori **Presenter:** *Souraja Kundu, Indian Institute of Technology Guwahati, India*

Abstract- Generating music from images can enhance various applications, including background music for photo slideshows, social media experiences, and video creation. This paper presents an emotion-guided image-to-music generation framework that leverages the Valence-Arousal (VA) emotional space to produce music that aligns with the emotional tone of a given image. Unlike previous models that rely on contrastive learning for emotional consistency, the proposed approach directly integrates a VA loss function to enable accurate emotional alignment. The model employs a CNN-Transformer architecture, featuring pre-trained CNN image feature extractors and three Transformer encoders to capture complex, high-level emotional features from MIDI music. Three Transformer decoders refine these features to generate musically and emotionally consistent MIDI sequences. Experimental results on a newly curated emotionally paired image-MIDI dataset demonstrate the proposed model's superior performance across metrics such as Polyphony Rate, Pitch Entropy, Groove Consistency, and loss convergence.

16:25-16:40	SC0096	**Title:** Cognigun: IOT - Integrated Firearm Safeguarding with YOLOv9 Algorithm for Forensic Analysis **Author(s):** Charles Tirones Catacutan, Keith Mendoza Pama, Zilloh Ramos Punzalan, Criselle J. Centeno, Jethro San Diego and Raymund M. Dioses **Presenter:** *Zilloh Ramos Punzalan, Pamantasan ng Lungsod ng Maynila, Philippines*

Abstract- Negligent use of firearms has been a rampant problem in the Philippines, may it be the New Year or the unregulated use of loose firearms. Gun control policies do exist but weak enforcement tends to make these policies inefficient. This leads to the extensive use of illegal firearms, with many being unregistered or lacking proper licensing, and improper firearm handling by both civilians and law enforcement. Due to these factors, the researchers designed CogniGun, an IoT-integrated firearm safeguarding application that utilizes the YOLOv9 algorithm for enhanced, real-time forensic analysis. This system aims to address the challenges faced when it comes to firearm safety and investigation by providing real-time object identification and location through a mobile application. CogniGun utilizes the Arduino 33 IoT as the main microcontroller and the NEO M8N Flight Controller GPS Module as its

90

tracking module. These integrated circuits ensure tracking precision and data processing efficiency. CogniGun also utilizes the use of deep learning architectures for firearm recognition. A biometric authentication system using fingerprint technology was successfully implemented in handguns, addressing concerns related to unauthorized access and firearm misuse. These integrations allow law enforcement officers to conduct forensic ballistics analysis on-site with greater speed and precision. CogniGun was able to encompass the real-time object detection, location, and identification using YOLOv9. CogniGun also works smoothly on mobile phones as stated in a survey for user satisfaction and encompasses the different flaws in the current problems with regards to gun policies.

	SC0061	**Title:** Revolutionizing IoT Security: A Blockchain and Federated Learning-Based Anomaly Detection System **Author(s):** Haya Saleh Alharthi, Suhair Zafir Alshehri and Manal Matoq Kalkatawi **Presenter:** *Haya Saleh Alharthi, King Abdulaziz University, Jeddah, Saudi Arabia*
16:40-16:55		**Abstract-** The rapid development of the Internet and smart devices has led to a surge in network traffic, increasing the complexity and vulnerability of infrastructure to malicious activities. Intrusion Detection Systems (IDS) are crucial in safeguarding the security and privacy of Internet of Things (IoT) devices. In IDS, Machine Learning techniques are often employed to identify deviations from normal behavior, enabling effective threat detection. However, traditional approaches face challenges such as centralized architectures, which raise data privacy concerns and limited resources on large-scale datasets. Federated Learning (FL) aligns well with a privacy-preserving decentralized learning method that does not transfer data but trains models locally. Only model parameters are transferred to the centralized server. The inherent risk of relying on a central server for model aggregation is mitigated by incorporating Blockchain technology, which provides a verifiable ledger for recording model updates and ensures secure aggregation without needing a central authority. This research introduces a novel conceptual framework for IDS integrating FL with Blockchain to enhance privacy, scalability, and resilience in IoT networks.
16:55-17:10	SC0022	**Title:** Towards Classification of Covariance Matrices via Bures-Wasserstein-Based Machine Learning **Author(s):** Michael Zirpoli, Yuyan Yi, Shu-Chin Lin, Linqiang Ge and Jingyi Zheng **Presenter:** *Michael Zirpoli, Auburn University, USA*
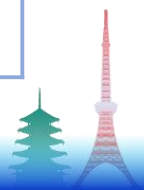
Abstract- Spatial-temporal data is a prevalent data type in biomedical domains, encompassing instances like multi-channel EEG and fMRI. In the analysis of such data, the connectivity matrix (e.g., functional connectivity derived from fMRI, covariance matrix derived from EEG) is widely extracted and analyzed. Rather than analyzing these matrices within the Euclidean space, this paper considers each matrix as a point situated on the manifold of positive semi-definite (PSD) matrices coupled with Bures-Wasserstein (BW) metric. Within this framework, two machine learning models based on the BW metric are proposed for the classification of PSD matrices on the manifold. Specifically, projection map techniques, based on the BW metric, have been introduced and integrated into machine learning models such as support vector machines and random forest. In comparison with Euclidean methods, our approach considers the geometry of the Riemannian manifold where PSD matrices reside. Moreover, compared with prevalent Affine-Invariant (AI) metrics, our framework does not require matrix regularization and is computationally efficient. To comprehensively evaluate the proposed methods, four fMRI datasets and three brain-computer interface datasets with varying dimensions and quantities have been utilized. The results demonstrate comparable and even superior performance of the proposed methods compared with Euclidean and AI-based approaches.

|		
	SC0143	**Title:** A NEW M-DDPG ALGORITHM FOR COMPUTING TASK SCHEDULING AND RESOURCE ALLOCATION IN EDGE ENVIRONMENTS **Author(s):** Dongfang Shang, Chen Zhang and Jiang Wang **Presenter:** ***Dongfang Shang***, *Tianjin Research Institute for Water Transport Engineering, Ministry of Transport, China*

Abstract- The rapid expansion of wireless communication and the Internet of Things (IoT) has led to a growing need for real-time data processing. Traditional cloud computing architectures often struggle to meet the low latency, high bandwidth, and low energy consumption requirements of certain applications. In response, Mobile Edge Computing (MEC) has emerged as a promising solution, offloading data processing closer to the user's device at the network's edge, significantly reducing system overhead. This enables more efficient handling of latency-sensitive and energy-efficient applications. One critical challenge in MEC research is optimizing the offloading of computational tasks from user devices to edge nodes, while fully utilizing idle resources to minimize operational overhead and enhance Quality of Service (QoS). This paper focuses on offloading algorithms in a multi-user, multi-edge-node environment, where each user device generates tasks with dependency constraints. The goal is to minimize total system overhead. We propose a parametric

The time slot **17:10-17:25** appears in the left column aligned with the second abstract.

modeling approach for task offloading and introduce a Multi-agent Deep Deterministic Policy Gradient (M-DDPG) algorithm to solve the offloading strategy. The M-DDPG algorithm, deployed at the edge layer, is designed to optimize offloading decisions based on state space, observations, action space, and reward mechanisms. Experimental results show that the M-DDPG algorithm outperforms the traditional S-DDPG algorithm in terms of convergence and achieves lower total system overhead, improving overall performance in the edge network environment.
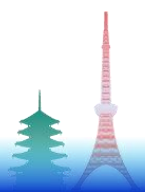
Call for Papers

Kawasaki, Japan, March 4-6, 2025	Submission deadline: 2025-1-9

2025 7th International Conference on Image, Video and Signal Processing (IVSP 2025)

Conference Website: **https://ivsp.net**

IVSP 2025 papers will be published into *ACM international conference proceedings, which will be indexed by Ei Compendex and Scopus.*

Submission methods: e-mail: **ivsp@acm-sg.org**

Zmeeting Submission System: **https://www.zmeeting.org/submission/IVSP2025**

Keynote Speaker: **Prof. Xudong Jiang (IEEE Fellow) from Nanyang Technological University, Singapore, Prof. Kenneth K. M. Lam from The Hong Kong Polytechnic University, China, Prof. Kiyoshi Hoshino from Meiji University, (Professor Emeritus) University of Tsukuba, Japan and Prof. Tae-Kyun Kim from Korea Advanced Institute of Science and Technology, Korea.**

Conference will be held in **Meiji University Ikuta Campus, Kawasaki, Japan**

Tsukuba, Japan, June 25-27, 2025	Submission deadline: 2025-1-10

2025 7th Blockchain and Internet of Things Conference (BIOTC 2025)

Conference Website: **https://biotc.net/**

BIOTC 2025 papers will be published into *International Conference Proceedings, which will be indexed by Ei Compendex and Scopus.*

Submission methods: e-mail: **biotc.contact@gmail.com**

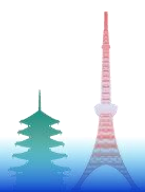Zmeeting Submission System: **https://www.zmeeting.org/submission/biotc2025**

Keynote Speaker: **Prof. Qun Jin from Waseda University, Japan and Prof. Chin-Chen Chang (IEEE Fellow/ IEE Fellow) from Feng Chia University, Taiwan.**

Conference will be held in **University of Tsukuba, Tsukuba, Japan**

Notes

Notes