Conference Program

2024 8th International Conference on

Big Data and Internet of Things

(BDIOT2024)

2024 6th International Conference on Virtual Reality and Image Processing

(VRIP2024)

University of Saint Joseph, Macao, SAR China | September 14-16, 2024

Address: 14-17 Estr. Marginal da Ilha Verde, Macao

地址: 中國澳門青洲河邊馬路14-17號

Co-sponsored by

Patrons

1

Table Of Contents

Organizing Committees

Conference Chairs

Simon Fong, University of Macau, Macao, SAR China
Alexandre Lobo, University of Saint Joseph, Macao, SAR China

Program Chairs

George Wencai Du, University of Saint Joseph, Macao, SAR China
Pedro Rodrigues, Universidade Catolica de Portugal UCP - Porto, Portugal
Hai-Ning Liang, Xi'an Jiaotong-Liverpool University, China
Xinghua Li, Wuhan University, China

Program Co-chairs

Ivan Arraut, University of Saint Joseph, Macao, SAR China
Lianggui Tang, Chongqing Technology and Business University, China
Fei Hao, Shaanxi Normal University, China
Richard C Millham, Durban University of Technology, South Africa

Publicity Chairs

R Kanesaraj Ramasamy, Multimedia University, Malaysia
Le Hong Trang, Ho Chi Minh City University of Technology, Vietnam
Wee Hong Ong, Universiti Brunei Darussalam, Brunei Darussalam

Technical Committee Members

Xing Tu, Shenzhen University, China
Qun Song, Chongqing Technology and Business University, China
Jacopo Soldani, University of Pisa, Italy
Luca Davoli, University of Parma, Italy
Euclides Chuma, Linkoping University, Sweden
Zhenhong (Brad) Lei, Rhode Island School of Design, USA
Masatoshi Arikawa, Akita University, Japan
Manolya Kavakli, Macquarie University, Australia
Rajan Shankaran, Macquarie University, Australia
Fangzhou Dong, Monash University, Australia
Dalvan Griebler, Pontifícia Universidade Pontifical Catholic University of Rio Grande do Sul (PUCRS), Brazil
Gabriel Gomes de Oliveira, University of Campinas, Brazil
Angela Maria Alves, Renato Archer Information Technology Center, Brazil
Yuzo Iano, University of Campinas, Brazil
Intan Salwani Mohamed, Universiti Teknologi MARA, Malaysia
Marshima Mohd Rosli, Universiti Teknologi MARA, Malaysia
Nabilah Rozzani, Management and Science University, Malaysia
Renjith V. Ravi, M.E.A Engineering College, India
Radhakrishna Bhat, Manipal Academy of Higher Education, India
Ali Saberi, Iranian Researchers Network, Iran
Mehdi Hassan, Air University Islamabad, Pakistan

3

Conference Venue

University of Saint Joseph (Macao)

Address: 14-17 Estr. Marginal da Ilha Verde, Macao

地址: 中國澳門青洲河邊馬路14-17號

■ Transportation

|	
From Macau International Airport	1) Metro: Board Macao Light Rapid Transit Taipa Line - take off at Barra Station- transfer to route MT4- take off at Qingmao / Av. Do. Comendador Ho Yin -walk to destination 2) Taxi: about 20 mins directly to the venue

■ Sign-in

Spot	USJ105, 1st floor, Academic Building
Time	13:30-17:00 | September 14th, 2024

■ Conference Room

Date	Activity	Meeting Room
Sept. 15th, 2024	Opening Ceremony & Signing Ceremony Keynote Speech & Invited Speech	Dom Bosco Auditorium, Ground floor, Academic Building
Sept. 15th, 2024	Session 1	USJ 105, **1st floor**, Academic Building
Sept. 15th, 2024	Session 2	USJ 107, **1st floor**, Academic Building

4
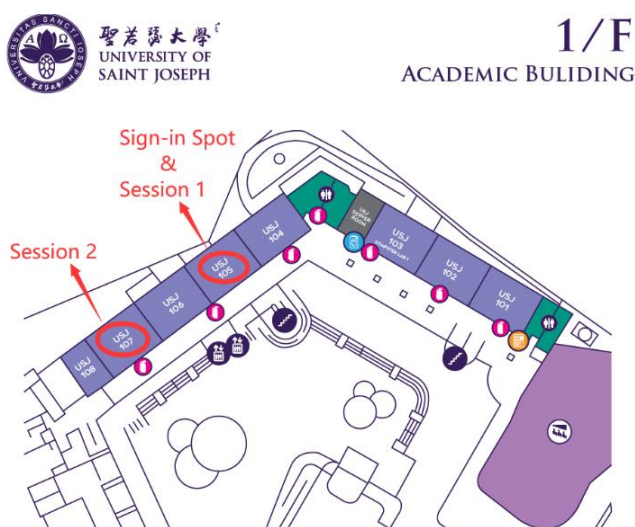
Onsite Instruction

■ Oral Presentation

1. Timing: a maximum of 15 minutes total, including speaking time and discussion. Please make sure your presentation is well timed. Please keep in mind that the program is full and that the speaker after you would like their allocated time available to them.

2. You can use USB flash drive (memory stick), make sure you scanned viruses in your own computer. Each speaker is required to meet her/his session chair in the corresponding session rooms 10 minutes before the session starts and copy the slide file (PPT or PDF) to the computer.

3. It is suggested that you email a copy of your presentation to your personal inbox as a backup. If for some reason the files can't be accessed from your flash drive, you will be able to download them to the computer from your email.

4. Please note that each session room will be equipped with a LCD projector, screen, point device, microphone, and a laptop with general presentation software such as Microsoft PowerPoint and Adobe Reader. Please make sure that your files are compatible and readable with our operation system by using commonly used fronts and symbols. If you plan to use your own computer, please try the connection and make sure it works before your presentation.

5. Videos: If your PowerPoint files contain video clips please make sure that they are well formatted and connected to the main files.

■ Dress Code

Please attend the conference in formal attire.

■ Safety Reminder: Secure Valuable Items at All Times

We remind you to secure your personal belongings at all times.

Please remember to:

* Wear your Conference Identification Badge at all times. Do not throw away Badge.

* If you are using a laptop computer, do not leave it unattended at any time.

* Keep your purse, wallet and other valuables with you at all times.

* The conference organizer will not be responsible for the loss or damage to any personal belongings.

Online Instruction

■ Time Zone

UTC+8

You're suggested to set up the time on your computer in advance.

■ Platform

We will be using **Zoom** for all our live stream sessions. So, if you haven't installed it, please download a Zoom client from:

https://zoom.us/download

The Zoom account is not mandatory to attend the conference. If you do not want to register the account, by entering meeting ID is also accessible to our conference.

Learn the Zoom skills at: https://support.zoom.us/hc/en-us/articles/206618765-Zoom-Video-Tutorials

■ Join the Test Session before the Formal Session

Date: September 14, 2024

Prior to the formal meeting, presenters shall join the test room to ensure everything is on the right track. Please check your test time on this program.

The Video presentation should be within 12 minutes, 3 minutes for Q&A, in total, one presentation is 15 minutes.

■ Equipment Needed

- A computer with internet connection and camera
- Headphones

■ Environment Needed

- Quiet Location
- Stable internet connection
- Proper lighting and background

■ Language

- Please make presentation in English. If necessary, it is allowed to restate some key points in Chinese.
- Please feel free to discuss in English or Chinese during Q&A.

■ Attention Please

The conference will be recorded. We will appreciate your proper behavior.

■ Presentation Recording and Broadcasting

The photograph(s) or video or audio recording(s) will be taken by conference organizer. It will be used in for conference program purpose. Each presentation will be recorded, if you don't want it, please inform our staff ahead of time.

Do not record other presenters' presentation nor distribute it to or share with anyone unless the presenter gives written consent of agree. Failure to do so will be considered a serious academic violation subject to disciplinary/ lawful action.

Schedule on Sept. 14th

Onsite Sign-in

Time	Activity	Venue
13:30-17:00	**Sign-in** **Collecting Conference Materials for Onsite Participants** • Arrive at Registration desk • Inform the staff of your paper ID • Sign-in • Claim your conference kit including conference bag, conference identification badge, meal coupon, conference program	USJ105, **1**st floor, Academic Building

Online Meeting Test

Time	Sessions	Zoom ID
14:00-15:30	**Session 3** TP24-309, TP24-201, TP24-325, TP24-202, TP24-330, TP24-312, TP24-203, TP24-326 **Session 4** TP24-304, TP24-310, TP24-212, TP24-317 TP24-320, TP24-323, TP24-332, TP24-204, TP24-313	**815 0608 7777** (Zoom Link: https://us02web.zoom.us/j/81506087777)

7

Schedule on Sept. 15th

Time	Activity	Venue
Host: Alexandre Lobo, University of Saint Joseph, Macao, SAR China		
**Opening Ceremony	09:00-09:10**		
09:00-09:10	**Welcome Message from University of Saint Joseph, Macao, SAR China:** **Alexandre Lobo,** University of Saint Joseph, Macao, SAR China	
	Welcome Message from Conference Chair: **Simon Fong,** University of Macau, Macao, SAR China	
**Signing Ceremony	09:10-09:40**		
09:10-09:40	*"Consortium of BRICS E-health Research Taskforces Association (COBERTA)"*	
**Keynote Speech I	09:40-10:25**		
09:40-10:25	**Arumugam Nallanathan** Queen Mary University of London (QMUL), United Kingdom *"The Role Artificial Intelligence (AI) in Massive Internet of Things (mIoT)"*	**Dom Bosco Auditorium, Ground floor**
**Group Photo & Coffee Break	10:25-10:50**		
**Keynote Speech II	10:50-11:35**		
10:50-11:35	**Hai-Ning Liang** Xi'an Jiaotong-Liverpool University (XJTLU), China *"Re-conceptualizing Text Entry for Extended Reality Systems via Spatial Interfaces"*	
**Invited Speech	11:35-12:00**		
11:35-12:00	**Simon Fong** University of Macau, Macao, SAR China *"Federated Learning, Big Data, and IoT Synergy for e-Health Transformation: A Global Collaborative Framework for Scalable, Sustainable, and AI-Driven Healthcare Diagnosis"*	
**Lunch @ USJ Macau Canteen, ground floor	12:00-13:30**		
**Onsite Technical Sessions	13:30-15:45**		
13:30-15:45	Session 1 – *"Intelligent Image Analysis and Multimedia Application Technology"* TP24-206, TP24-315, TP24-324, TP24-207, TP24-331 TP24-314, TP24-328, TP24-210, TP24-311	**USJ 105, 1st floor**
13:30-15:45	Session 2 – *"Modern Information Theory and Data Analysis"* TP24-316, TP24-305, TP24-319, TP24-321, TP24-329 TP24-327-A, TP24-333, TP24-322, TP24-334	**USJ 107, 1st floor**
15:45-16:10	**Coffee Break**	
**Lab Visit	16:10-17:30**		
**Dinner Banquet @ USJ Macau Canteen, ground floor	17:45-19:00**		

Schedule on Sept. 16th

Time	Sessions
Online Technical Sessions	
Zoom ID: 815 0608 7777 (Link: https://us02web.zoom.us/j/81506087777)	

Time	Sessions
10:00-12:00	**Session 3 –** *"Digital Image Detection, Analysis and Application"* TP24-309, TP24-201 TP24-325, TP24-202 TP24-330, TP24-312 TP24-203, TP24-326
12:00-13:30	**Lunch**
13:30-15:45	**Session 4 –** *"Model Based Intelligent Information System and Interactive Network"* TP24-304, TP24-310 TP24-212, TP24-317 TP24-320, TP24-323 TP24-332, TP24-204 TP24-313

Keynote Speaker I

Prof. Arumugam Nallanathan

Queen Mary University of London (QMUL), United Kingdom

FIEEE, FIET, CEng, Web of Science Highly Cited Researcher

Time: 09:40-10:25, September 15th, 2024 (UTC+8)
Venue: Dom Bosco Auditorium

📚 Bio

Arumugam Nallanathan is Professor of Wireless Communications and the founding head of the Communication Systems Research (CSR) group in the School of Electronic Engineering and Computer Science at Queen Mary University of London since September 2017. He was with the Department of Informatics at King's College London from December 2007 to August 2017, where he was Professor of Wireless Communications from April 2013 to August 2017. He was an Assistant Professor in the Department of Electrical and Computer Engineering, National University of Singapore from August 2000 to December 2007. His research interests include 6G Wireless Networks, Internet of Things (IoT) and Molecular Communications. He published more than 500 technical papers in scientific journals and international conferences. He is a co-recipient of the Best Paper Awards presented at the IEEE International Conference on Communications 2016 (ICC'2016), IEEE Global Communications Conference 2017 (GLOBECOM'2017) and IEEE Vehicular Technology Conference 2017 (VTC'2017).

He is an Editor-at-Large for IEEE Transactions on Communications and a senior editor for IEEE Wireless Communications Letters. He was an Editor for IEEE Transactions on Wireless Communications (2006-2011), IEEE Transactions on Vehicular Technology (2006-2017), IEEE Signal Processing Letters and a Guest Editor for IEEE Journal on Selected Areas in Communications (JSAC). He served as the Chair for the Signal Processing and Computing for Communications (SPCC-TC) of IEEE Communications Society and Technical Program Chair and member of Technical Program Committees in numerous IEEE conferences. He received the IEEE Communications Society SPCE outstanding service award 2012 and IEEE Communications Society RCC outstanding service award 2014. He has been selected as a Web of Science (ISI) Highly Cited Researcher in 2016. He is an IEEE Fellow and IEEE Distinguished Lecturer.

🖼️ Speech Information

The Role Artificial Intelligence (AI) in Massive Internet of Things (mIoT)

Abstract: Massive Ultra-Reliable and Low-Latency Communications (mURLLC), which integrates URLLC with massive access, is emerging as a new and important service class in the next generation (6G) for time-sensitive traffics and has recently received tremendous research attention. However, realizing efficient, delay-bounded, and reliable communications for a massive number of user equipments (UEs) in mURLLC, is extremely challenging as it needs to simultaneously take into account the latency, reliability, and massive access requirements. To support these requirements, the third generation partnership project (3GPP) has introduced enhanced grant-free (GF) transmission in the uplink (UL), with multiple active configured-grants (CGs) for URLLC UEs. With multiple CGs (MCG) for UL, UE can choose any of these grants as soon as the data arrives, while with single CG (SCG), UE need to wait for the CG period to transmit the packet. In this talk, Machine Learning (ML) approaches in mURLLC systems will be presented. Promising research directions and possible ML solutions will also be discussed.

10

Keynote Speaker II

Prof. Hai-Ning Liang

Xi'an Jiaotong-Liverpool University (XJTLU), China

Time: 10:50-11:35, September 15th, 2024 (UTC+8)
Venue: Dom Bosco Auditorium

Bio

Hai-Ning Liang is a Professor in the Department of Computing at Xi'an Jiaotong-Liverpool University (XJTLU), China. He was the founding head of the department (2019-2023) and now serves as the Deputy Director of the Suzhou Key Laboratory of Intelligent Virtual Engineering and the XJTLU Virtual Engineering Center. He obtained his PhD from Western University in Canada and does research in human-computer interaction, focusing on virtual/augmented/mixed reality, gaming, and visualization technologies. He has co-authored over 200 peer-reviewed articles in these areas. He is an associate editor for The Visual Computer, Frontiers in Artificial Intelligence, and Frontiers in Virtual Reality.

Speech Information

Re-conceptualizing Text Entry for Extended Reality Systems via Spatial Interfaces

Abstract: Extended reality (XR) systems are becoming increasingly popular and more widely used, but they are still in their early stages of development and are considered maturing technologies. These wearable devices have the potential to revolutionize how people interact with digital information, offering an ideal platform to reimagine the next generation of interfaces that can support natural and intuitive spatial interactions. Just like touch-enabled mobile devices changed and enriched our understanding of how the keyboard should and could function as one of the primary interfaces between people and these devices, XR systems provide an opportunity for us to redesign and rethink how text entry can be achieved to support better the context and scenario in which these systems are used. In this talk, I will describe the challenges of text entry in XR systems and share several novel techniques and interfaces that our team has developed to improve user experience and performance. These techniques utilize input modalities available in XR systems and leverage people's spatial abilities to create more efficient and usable interfaces that move us away from our standard conception of text entry using constrained finger /hand movements.

Invited Speaker

Assoc. Prof. Simon James Fong

University of Macau, Macao, SAR China

Time: 11:35-12:00, September 15th, 2024 (UTC+8)
Venue: Dom Bosco Auditorium

📚 Bio

Simon James Fong graduated from La Trobe University, Australia, with a 1st Class Honours BEng. Computer Systems degree and a PhD. Computer Science degree in 1993 and 1998 respectively. Simon is now working as an Associate Professor at the Computer and Information Science Department of the University of Macau. Dr. Fong has published over 500 500 international conference and peer-reviewed journal papers, mostly in the areas of data mining and AI medical applications. He serves on the editorial boards of the Journal of Network and Computer Applications of Elsevier, IEEE IT Professional Magazine, and various special issues of SCIE-indexed journals.

🖥 Speech Information

Federated Learning, Big Data, and IoT Synergy for e-Health Transformation:
A Global Collaborative Framework for Scalable, Sustainable,
and AI-Driven Healthcare Diagnosis

Abstract: The integration of big data, AI, machine learning, and IoT health sensing applications is set to revolutionize healthcare by enabling scalable, real-time, and AI-powered diagnostic systems. In this invited speech, we would present an innovative, cloud-based framework that leverages federated learning to foster global collaboration among clinicians, researchers, and institutions. The framework facilitates decentralized AI model training across diverse healthcare centers while preserving data privacy, allowing the secure sharing of insights through a worldwide consortium. It empowers next-generation e-health diagnostics by combining data from distributed sources, neurocognitive experimentation, and IoT health sensors, enabling continuous model growth and improvements in real time. By connecting stakeholders across different geographical regions, the platform creates an ecosystem where global collaboration drives sustainable advancements in healthcare. Case studies such as glomerulonephritis diagnosis illustrate the model's capacity for personalized, adaptive, and accurate healthcare predictions, while IoT-driven health monitoring and neurocognitive experimentation broaden its application scope. As these collaborative networks evolve, the framework ensures diagnostic models remain accurate, scalable, and aligned with emerging innovations. This approach fosters a future where connected devices, shared knowledge, and AI-driven systems continually refine and optimize healthcare delivery worldwide. A real life case of Consortium of BRICS E-health Research Taskforces Association (COBERTA) will be presented.

Session 1

Sept. 15th | 13:30-15:45

Venue: USJ 105, 1st floor

- **Topic: Intelligent Image Analysis and Multimedia Application Technology**
- **Session Chair: Prof. Hai-Ning Liang, Xi'an Jiaotong-Liverpool University, China**

Note: The schedule of each presentation is for reference only. Authors are required to attend the whole session, in case there may be some changes on conference day. Please join in the room 5-10 minutes earlier.

13:30-13:45	(**TP24-206**) Directional Progress Indicator for Visualizing Off-Screen Point-of-Interest in Handheld Augmented Reality **Ren Kurosaki,** Akita University, Japan Abstract: This study introduces a novel visualization technique for off-screen Points of Interest (POIs) in handheld augmented reality (AR) environments, specifically within the context of the *ARTimeWalk* platform. This photo-sharing AR platform allows users to view historical photos at their original locations. Traditional methods have struggled to effectively signal the presence of off-screen POIs in handheld AR due to limited screen space and narrow fields of view. The proposed solution integrates a *progress indicator* with icons to intuitively display the direction and type of off-screen POIs, enhancing user experience by providing clearer directional cues and content information. The methodology includes a progress indicator composed of three elements: Arrowhead, Circular progress ring, and POI icon, indicating the necessary rotation angle and the nature of the content respectively. An evaluation involving seven participants demonstrated that the new technique significantly reduces the time required to discover POIs, providing an efficient and user-friendly navigation experience in AR environments.
13:45-14:00	(**TP24-315**) Facial Emotion Recognition Based on Optimized Xception Training **Jiajun Yang,** Chongqing Technology and Business University, China Abstract: This scholarly investigation is centered on the pursuit of high precision in the domain of facial emotion recognition (FER). The conventional approaches to feature extraction and classification within the realm of ER are often laborious and heavily reliant on the quality of manual feature extraction, which can introduce significant variability and subjectivity into the process. In an effort to address these limitations and enhance the efficiency and accuracy of ER, this study has made a deliberate selection of the FER2013 dataset, a benchmark for evaluating ER models. Leveraging the Xception model as a foundational framework, this research implements a transfer learning strategy with a pre-trained model to bolster the quality of the training model. The Xception architecture, known for its depth-wise separable convolutions, allows for a more efficient learning process and is adept at capturing intricate features from visual data. By utilizing a pre-trained model, the study capitalizes on the knowledge garnered from vast and diverse datasets, thus enabling a more robust and generalized feature representation for the task at hand. Furthermore, to augment the model's performance on the FER2013 dataset, this study employs data augmentation techniques. These methods introduce controlled variability into the training data by applying random transformations such as cropping, scaling, and flipping, which serve to increase the dataset's diversity and prevent overfitting. The data augmentation process aims to simulate a broader range of real-world conditions and expressions, thereby improving the model's ability to generalize and perform accurately across various emotional states. In the comprehensive comparative analysis conducted within this study, the optimized and fine-tuned Xception model is pitted against other prominent deep learning architectures that are mainstream in the current literature. The performance of the Xception model is evaluated on various metrics, including but not limited to accuracy, precision, recall, and F1 score. The results of this rigorous evaluation reveal that the Xception model, after undergoing the proposed enhancements,

13

surpasses the accuracy of the compared networks, demonstrating its superiority in the context of the FER2013 dataset.

The findings of this study contribute to the growing body of research that advocates for the adoption of deep learning methodologies in the field of computer vision, specifically in the nuanced task of facial emotion recognition. The outcomes underscore the potential of transfer learning and data augmentation in enhancing the performance of deep neural networks, offering a promising avenue for future research and development in the field. This study not only pushes the boundaries of what is achievable in automated facial emotion recognition but also presents a solid foundation for the integration of such models into practical applications, where the accurate interpretation of human emotions can significantly improve user experience and interaction in various sectors, including but not limited to healthcare, customer service, and human-computer interaction.

14:00-14:15

(**TP24-324**) Analysis of Deep Learning Algorithms for Emotion Classification Based on Facial Expression Recognition
Cheng Qian, University of Saint Joseph, China

Abstract: Facial expression recognition (FER) is essential for discerning human emotions and is applied extensively in big data analytics, healthcare, security, and user experience enhancement. This paper presents an empirical study that evaluates four existing deep learning models—VGG16, DenseNet, ResNet50, and GoogLeNet—utilizing the Facial Expression Recognition 2013 (FER2013) dataset. The dataset contains seven distinct emotional expressions: angry, disgust, fear, happy, neutral, sad, and surprise. Each model underwent rigorous assessment based on metrics including test accuracy, training duration, and weight file size to test their effectiveness in FER tasks. ResNet50 emerged as the top performer with a test accuracy of 69.46%, leveraging its residual learning architecture to effectively address challenges inherent in training deep neural networks. Conversely, GoogLeNet exhibited the lowest test accuracy among the models, suggesting potential architectural constraints in FER applications. VGG16, while competitive in accuracy, demonstrated lengthier training times and a larger weight file size (512MB), highlighting the inherent balance between model complexity and computational efficiency.

14:15-14:30

(**TP24-207**) Generative Live Commentaries Interacting with Geospatial Context for Promoting Local Festivals
Iori Sasaki, Akita University, Japan

Abstract: With the advancement of robust network connections and the rise of live streaming platforms, it has become possible to experience local events and walking tours remotely in real-time. This study introduces a city-wide audio augmented reality system designed to enhance engagement at local festivals by combining livestreamed video with character-driven commentaries on local history and event details. To ensure continuous storytelling, even in areas with few points of interest, this study proposes a novel geofencing architecture that considers multilayered city features, enabling seamless audio guide experiences. Additionally, this paper introduces prompt engineering for generating entertaining and listener-friendly guide scripts, tailored for large language models, named the geofence-to-conversation technique. A preliminary study of our prototype system, deployed at an actual local festival, demonstrates its effectiveness and potential for enhancing regional promotion.

14:30-14:45

(**TP24-331**) GlobalLocalSegNet: A Hybrid Model for Complex Medical Image Segmentation Combining Global and Local Features
Yong Zhou, Chongqing Technology and Business University, China

Abstract: Complex medical image segmentation holds critical importance in the medical field. However, renowned segmentation models such as U-Net [1] and TransUNet [2] often suffer from diminished performance when dealing with images that have complex backgrounds or high variability. These models also struggle to capture the dependencies between global and local features effectively, limiting their application in complex medical image segmentation tasks. To address these shortcomings, this paper introduces a novel U-shaped model, GlobalLocalSegNet (GLS-Net), specifically designed for precise segmentation of complex medical images. The GLS-Net model comprises two key sub-modules: a transformer model with positional encoding to extract global information, and a convolutional neural network tailored for local information extraction. Moreover, a U-shaped network structure based on the fusion of global and local features was designed to enhance feature extraction and detail capture.The model was tested using three publicly available complex medical image datasets and one non-medical complex image dataset to assess its scalability. A series of comparative experiments in this study confirm the robustness, scalability, and stability of the GLS-Net model.

14:45-15:00	(**TP24-314**) Impact of Convolutional Kernel Initialization based on Pixel Distribution Features on Table Recognition Performance **Zhixuan Chen,** Guangzhou City University of Technology, China Abstract: Due to the powerful feature learning and classification capabilities of convolutional neural networks, they have been widely applied in fields such as image recognition and classification, object detection, and semantic segmentation. Therefore, enhancing the efficiency and accuracy of convolutional neural networks in deep learning has become increasingly important. In convolutional neural networks, the initial design of the convolutional kernels significantly influences the efficiency of deep learning iterations. This paper focuses on the recognition and classification of image tables, proposing a method of customizing the initialization of convolutional kernel parameters based on several local elements of images and the similarity between their pixel distribution features and the initial parameters within the convolutional kernels. The deep learning of image tables through convolutional neural networks is then conducted on this basis, and compared experimentally with traditional initialization methods such as Normal and Xavier. The experimental results show that the parameter initialization method proposed in this paper achieves notably higher classification accuracy in the early stages of neural network training and overall higher classification accuracy for table recognition.
15:00-15:15	(**TP24-328**) Enhancing Dynamic Hand Gesture Recognition through Optimized Feature Selection using Double Machine Learning **Keyue Yan,** University of Macau, China Abstract: Causal machine learning combines causal inference and machine learning to understand and utilize causal relationships in data. While traditional machine learning focuses on missions of prediction and pattern recognition, causal machine learning goes a step further by revealing causal relationships between variables. In this research, we employ the double machine learning method to identify variables in the gesture recognition problem where independent variables have causal relationships with the final gesture. These variables are then selected for further classification and analysis. By comparing this approach with traditional feature selection methods, we find that the variables selected using double machine learning are more useful for classification and yield excellent results across different machine learning classification models. This new double machine learning based approach provides a valuable reference for researchers during the feature selection stage.
15:15-15:30	(**TP24-210**) Space-Division-Based Pseudo-Occlusion in 3D Trajectory Data Visualization for Indoor AR Navigations **Yuhan Jin,** Akita University, Japan Abstract: Augmented Reality (AR) for indoor navigation enhances users' ability to locate and understand their environment by overlaying digital information onto real-world views. A key aspect of effective AR is the correct alignment of virtual and real objects, known as occlusion. Traditional occlusion methods require extensive computational resources for 3D scene data reconstruction, making them unsuitable for real-time applications on mobile devices due to the need for low latency. Particularly, traditional occlusion methods have been more impractical for indoor AR navigation in large spaces, such as entire buildings, compared to confined areas. This paper introduces a novel approach called pseudo-occlusion, which leverages space-division and geofencing techniques to simplify real-time occlusion handling without the need for depth cameras. Experiments conducted at the Mineral Industry Museum, Akita University, demonstrate the feasibility and efficiency of this method. Pseudo-occlusion accurately maintains the relative positioning of objects, reducing computational load and improving user experience by providing a seamless AR trajectory visualization. The method's simplicity and real-time performance offer significant advantages over traditional occlusion techniques, making it a practical solution for indoor AR navigation applications in large spaces.
15:30-15:45	(**TP24-311**) Graph Energy Variety Network-based Out-of-distribution Detection **Dennis Wong,** Macao Polytechnic University, China Abstract: Graph relational structures are ubiquitous and prediction problems on graphs are very popular, such as node prediction vs. edge prediction. However, current models concentrating on improving test performance on intra-distributed data and largely ignoring the potential risks of out-of-distribution (OOD) test samples. In some cases, models may lead to negative results if they misclassify anomalous or out-of-distribution data. In this paper, we investigate the problem of OOD detection for graph-structured data and identify an effective OOD recogniser based on the loss of an energy function extracted directly from a graph neural network trained using standard classification losses, and constructed a way to use graph-based neural network learning in the context of energy theory. More importantly, it can further enhance the recognition of out-of-distribution data through unlearned hierarchical energy transfer mechanisms and energy attenuation schemes. For a comprehensive evaluation, we have conducted experiments using a recognised benchmark setup and our proposed method has achieved very good results in various experiments.
	Best Presentation Award & Session Group Photo

Session 2

Sept. 15th | 13:30-15:45

Venue: USJ 107, 1st floor

- **Topic: Modern Information Theory and Data Analysis**

- **Session Chair:** Assoc. Prof. Simon Fong, University of Macau, Macao, SAR China

*Note: The schedule of each presentation is for reference only. Authors are required to attend the whole session, in case there may be some changes on conference day. Please join in the room 5-10 minutes earlier.

13:30-13:45	(**TP24-316**) Deep Learning Model Research and Performance Comparison Based on a Plant Classification Dataset **Zhuo Chen,** Chongqing Technology and Business University, China Abstract: Image recognition is one of the research directions in the field of computer vision. With the continuous development of computer technology and the escalating global trend towards intelligence, image recognition technology is being applied in various fields of daily life. The classification and recognition of plant species are crucial aspects of the comprehensive promotion of ecological civilization. The traditional methods of plant species classification are not only time-consuming and labor-intensive but also require a high level of expertise in botany, making it difficult to meet the demand for rapid and comprehensive investigation of plant resources in the domestic efforts to advance ecological civilization. This study establishes a dataset consisting of images of 30 plant species. Training is conducted using the ResNet34 model and the Swin Transformer model under the PyTorch framework. A comparative analysis is performed with traditional basic Convolutional Neural Network (CNN), MobileNetV2 model, and VGG11 model to analyze the performance differences among the models. Suggestions for improvement are proposed for certain models.
13:45-14:00	(**TP24-305**) Bibliometric Research on Business Data Analysis Based on Knowledge Graph **Jiajun Hou,** Guangzhou City University of Technology, China Abstract: With the development of big data technology, data plays a significant role in assisting business decisions and maximizing business value and benefits. In recent years, there have been numerous academic discussions on business data analysis, with scholars researching from different perspectives on how to analyze and mine data in the business field and how to make business data serve businesses. This paper collects literature on business data analysis from the CNKI database, including publications from the PKU core, CSSCI, and CSCD journals from 2013 to 2023. The CiteSpace literature visualization software is employed to conduct descriptive statistical analysis, research hotspot analysis, and research frontier analysis of the research achievements in this field. The results indicate that the research hotspots in this field can be categorized into three main themes: business data profitability and performance analysis, business data mining and intelligent transformation of business models, and business data risk analysis and security protection. Future research should focus on exploring the theme of business data sharing under the backdrop of the digital economy.
14:00-14:15	(**TP24-319**) Study on the Differences in the Impact of the American Welfare System on the Health of Different Races Based on Data Mining **Chenxia Li,** Chongqing Technology and Business University, China Abstract: In the context of increasing social diversity and racial differences, the impact of welfare systems on various social groups has gained widespread attention. Focusing on the welfare system in the United States, this study tries to explore the optimal combination of welfare plans to enhance the overall health level of society. Utilizing data mining methods and based on the Apriori algorithm, a new SAP-Apriori algorithm is proposed. It incorporates dynamic parameter adjustment to enhance model accuracy and robustness. Meanwhile, multidimensional evaluation criteria are employed to

16

	comprehensively assess the quality and credibility of association rules. The analysis reveals that there are health disparities between different racial groups under the same welfare conditions. Furthermore, within the same racial group, there is a correlation between the number of welfare plans received and the health status of that race. Based on these findings, some suggestions for relevant welfare plan combinations are given to the government. This study offers a new perspective on understanding racial health disparities within the U.S. welfare system and provides a basis for formulating more effective public policies.
14:15-14:30	(**TP24-321**) Credit Card Fraud Detection Based on MiniKM-SVMSMOTE-XGBoost Model **Yanzhao Gu,** Macao Polytechnic University, China Abstract: In recent years, the problem of credit card fraud has become more acute with the digitisation of credit cards. For the high data volume, high dimensionality and extreme imbalance of credit card transaction data. This paper explores the application in the field of credit card fraud detection based on MiniBatchKMeans-SVMSMOTE-XGBoost model. Through combining clustering, oversampling and classification algorithms, an improved fraud detection method is proposed. The experimental results show that the model performs well in handling unbalanced data with high accuracy and generalisation ability.
14:30-14:45	(**TP24-329**) TMemAE: Handwriting Abnormality Analysis for Parkinson Disease **Jing Chen,** Chongqing Technology and Business University, China Abstract: Causal machine learning combines causal inference and machine learning to understand and utilize causal relationships in data. While traditional machine learning focuses on missions of prediction and pattern recognition, causal machine learning goes a step further by revealing causal relationships between variables. In this research, we employ the double machine learning method to identify variables in the gesture recognition problem where independent variables have causal relationships with the final gesture. These variables are then selected for further classification and analysis. By comparing this approach with traditional feature selection methods, we find that the variables selected using double machine learning are more useful for classification and yield excellent results across different machine learning classification models. This new double machine learning based approach provides a valuable reference for researchers during the feature selection stage.
14:45-15:00	(**TP24-327-A**) CD63 As a Potential Diagnostic Marker of Depression and Its Correlation with Immune Infiltrates **Mengyu Liu,** Southeast University, China Abstract: Major depressive disorder (MDD) is the most common severe mental illness and the leading cause of suicide and disability. The study attempted to screen the diagnostic markers of peripheral blood of MDD and their relationship with immune infiltration. From the differential gene intersection set of the two data sets (GSE98793 and GSE76826), we screened to 72 up-regulated genes and 139 down-regulated genes. Gene Ontology (GO) enrichment analysis revealed that identified genes were primarily related to mononuclear cell differentiation, lymphocyte differentiation and T cell selection. Protein-Protein interaction (PPI) network analysis, random forest algorithm and Lasso regression algorithm obtained the hub genes: CD63. ROC analysis showed that CD63 had a good diagnostic and predictive effect on MDD patients. The CIBERSORT algorithm was used to assess immune cell infiltration and identify common immune characteristics in MDD. We analyzed the proteins that interact with CD63 using GeneMANIA databases, the miRNAs that interact with CD63 using RNAInter, miRWalk and starBase databases, and the traditional Chinese medicine ingredients that regulate CD63 using the COREMINE databases. Finally, we further verified the expression of CD63 using qPCR and WB on MDD patients' plasm, depressive model mice, and in vitro cells. These findings indicate that CD63 associated the infiltration of immune cells may be potential diagnostic indicators for MDD.
15:00-15:15	(**TP24-333**) PAT-Net: A Method for Alzheimer's Detection Using Long-Duration Speech **Lei Chen,** Chongqing Technology and Business University, China Abstract: This study presents an automated method based on temporal algorithms for detecting Alzheimer's Disease (AD) from long speech samples. Addressing the limitations of traditional speech detection methods in feature processing, we utilized Mel-Frequency Cepstral Coefficients (MFCCs) as key features and fed them into our innovative PAM-Net model. The PAM-Net model integrates TimesNet with Pyramid Attention Mechanism (PAM) and effectively extracts and processes temporal features through a Multi-Scale Module (CSCM). Additionally, based on an analysis of the

17

characteristics of MFCC-based handcrafted features, we updated the activation functions to better suit the model. Experimental results on the NCMMSC2021 AD and ADRESS-M datasets demonstrate that the PAM-Net model achieved prediction accuracies of 89.5% and 76.1%, respectively, surpassing the baseline models by 9.7% and 2.2%. Furthermore, the model also exhibited a 73.1% accuracy on the Ivanova Spanish-language dataset, demonstrating robust cross-lingual performance. These findings not only highlight the potential of the PAM-Net model in speech detection but also provide a new technological avenue for the early diagnosis of Alzheimer's Disease.

15:15-15:30

(**TP24-322**) Emotional Analysis in Animated Films Using Big Data and IoT: An In-Depth Study of 'Krek

Xing Tu, ShenZhen University, China

Abstract: This study explores the intersection of narrative structure, color variation, and audience emotional response in the classic animated film "Krek" from the Zagreb School of Animation. Utilizing a combination of Big Data and IoT technologies, we captured over 500 frames at one-second intervals using the Color Summarizer platform to extract key color properties such as hue, chroma, and lightness. This granular dataset enabled a detailed analysis of color dynamics throughout the film.
Audience emotional responses were quantitatively assessed using manual annotation and plotted to generate emotional curves. Pearson correlation analysis was conducted using SPSS to examine the relationships between emotional intensity, specific emotional states (pleasure, sadness, and disappointment), and color attributes. The results revealed a strong correlation between audience disappointment and color attributes, particularly chroma, highlighting the film's effective use of color to evoke specific emotional responses.
Our findings demonstrate the significant impact of color changes on emotional modulation in the film 'Krek' and underscore the unique narrative and visual techniques employed by the Zagreb School of Animation. This research contributes to a deeper understanding of how visual and narrative elements interact to shape audience emotions, offering valuable insights for the design of emotionally engaging animated films.

15:30-15:45

(**TP24-334**) Integrating Convolutional Neural Network and Broad Learning System for Moisture Detection in Oil via Near-Infrared Spectroscopy

Qing Li, Chongqing Technology and Business University, China

Abstract: Accurate determination of moisture content in transformer oil is critical for assessing transformer insulation, planning maintenance, extending life expectancy, and reducing failure rates. However, existing detection methods are often limited by high costs, time inefficiencies, susceptibility to human error, and low accuracy. To overcome these challenges, this study proposes an oil moisture content detection method that integrates a convolutional neural network (CNN) with a broad learning system (BLS) using near-infrared spectroscopy (NIR). This approach aims to achieve efficient, precise, rapid, and environmentally friendly quantitative analysis. The research began with the creation of a NIR spectral database for trace water content in transformer oil, followed by the application of the Savitzky-Golay convolution smoothing method for data preprocessing. The CNN was employed to extract features from the preprocessed spectral data, with the output from the penultimate fully connected layer serving as input for the BLS. The model's performance was further enhanced through training and parameter optimization. Validation and testing were conducted using a real NIR spectral dataset of transformer oil with trace water content. Experimental results show that the proposed method outperforms CNN, BLS, and PLSR models, achieving a Root Mean Square Error of Prediction (RMSEP=0.064) and Prediction Coefficient of Determination (R^2=0.926). These findings confirm the method's effectiveness and superiority in detecting moisture content in oil.

Best Presentation Award & Session Group Photo

Sept. 16th | 10:00-12:00
Zoom ID: 815 0608 7777

- ■ **Topic: Digital Image Detection, Analysis and Application**
- ■ **Session Chair: Dr. Gabriel Gomes de Oliveira, CTI - Renato Archer Information Technology Center, Brazil**

*Note: The schedule of each presentation is for reference only. Authors are required to attend the whole session, in case there may be some changes on conference day. Please join in the room 5-10 minutes earlier.

10:00-10:15	(**TP24-309**) Research on Multi-Image Panorama Stitching Method Based on Improved ORB+GMS Algorithm **Guangling Yu,** Chongqing Technology and Business University, China Abstract: Panoramic image stitching is one of the pivotal tasks involved in merging multiple overlapping images into a panoramic view. The ORB (Oriented FAST and Rotated BRIEF) algorithm and the GMS (Grid-based Motion Statistics) algorithm are crucial techniques for image feature extraction and matching. This paper proposes an improved ORB+GMS method, employing the AGAST (Adaptive and Generic Accelerated Segment Test) corner detector to replace the FAST corner detector in the original ORB algorithm to enhance corner detection performance and robustness. As the GMS algorithm utilizes a 3x3 grid-based feature statistic and grid partitioning, this study employs the Descending Grid Scale Method for feature statistic computation to enhance matching speed. The improved ORB algorithm is integrated into the enhanced GMS algorithm, wherein a homography matrix is computed based on high-quality matched point pairs, and mismatched feature points are eliminated to improve matching accuracy. The proposed improved ORB+GMS method is applied to various sets of different image samples to validate its matching effectiveness. Experimental results demonstrate that the matching performance, image panorama stitching effect, matching time, and matching rate of the proposed fusion of GMS and ORB feature extraction and matching algorithm are significantly superior to those of the ORB algorithm and other similar algorithms. The matching rate can reach over 85%, with an increase in correctly matched points relative to the ORB and SIFT algorithms by 8.11% and 23.3%, respectively. The matching time is reduced to the level of 1.995 milliseconds, which is decreased by 2.468 seconds and 0.35 seconds compared to the ORB algorithm and the SIFT algorithm, respectively.
10:15-10:30	(**TP24-201**) Evaluation Modeling of The Main Flight Display Interface in The Cockpit of The Boeing 747 Airplane Based on Fluency Theory **Jiayu Guo,** Southeast University, China Abstract: Based on the continuous development of artificial intelligence (AI) in interface recognition technology, the purpose of this article is to build a PFD interface fluency classification model, providing fluency level evaluation for future interface optimization design. The article introduces the influencing factors of fluency theory on the interface of the Boeing 747 PFD (Main Flight Display), and explores the effects of low, medium, and high fluency levels on participants browsing PFD interface information. By optimizing three interface factors, including font size, color contrast, and graphic outline, different levels of smoothness were constructed. Experiments were designed to collect participants' reaction time, accuracy, blink frequency, and subjective evaluations for tasks with different fluency levels. Finally, an interface fluency recognition and classification model was established using the Fisher discriminant method, and its effectiveness was verified. This model replaces the traditional single subjective emotional evaluation and forms an intuitive and quantitative interface design evaluation method. This method provides designers with a new evaluation approach, breaking free from the limitations of subjective evaluation.

10:30-10:45	**(TP24-325)** A Method for Detecting Deepfake Faces by Integrating Multi-Scale Deep-Shallow Spatial and Frequency Features **Chunyue Zhang,** Chongqing Technology and Business University, China Abstract: In recent years, deep learning technology has been excessively used in the creation of fake videos. Deepfake technology alters or replaces the facial information of the original video, synthesizes false voices, and is used to create pornographic films, fake news, political rumors, and so on, posing significant challenges to the authenticity of information and personal privacy. Currently, most face deepfake detection networks are based on single-scale and single-source detection research, and the proposed networks and methods often lack the ability to recognize unknown attack methods and generalization ability. To address these issues, this paper proposes a face deepfake detection method that integrates deep and shallow multi-scale spatial and frequency-domain features to achieve high-accuracy recognition of deepfake faces. The network is based on feature extraction from shallow and deep layers, where the shallow branch captures detailed information from low-level features to detect small-scale targets such as minor noise, while the deep branch focuses on detecting larger and higher-level fake features with a larger receptive field. Each branch adopts a method to fuse multi-scale spatial and frequency-domain features. Multi-scale features can capture information at different levels and granularities, while frequency-domain features can supplement fake traces that are difficult to detect in the spatial domain. Extensive experiments show that tthis framework outperforms most face spoofing detection networks and achieves good performance.
10:45-11:00	**(TP24-202)** Building Metaverse: Propose a design framework for the metaverse based on the requirements **Yunyan Wu,** Southeast University, China Abstract: In 2021, the term "metaverse" burst onto the scene, quickly capturing the public's imagination. This digital and virtual realm offers a space where people can interact through digital avatars, reflecting society's desire for a virtual existence. The allure of the metaverse lies in its promise of an immersive escape: a fantastical domain where individuals can break free from the physical world's limitations, create their own norms, and find temporary respite from real-world pressures. Despite its rapid rise in popularity, the metaverse's definition remains elusive, and the current state of technology infrastructure poses significant challenges to its widespread adoption. Research into the metaverse is still evolving and lacks a unified theoretical framework to guide the design of these virtual spaces. Moreover, existing technology faces considerable obstacles in supporting millions of users simultaneously in these environments. To bring the vision of the metaverse to fruition, it's crucial to develop and adopt fundamental theoretical principles. Establishing these standards will accelerate the metaverse's development, making it more accessible and enjoyable for the general public. This paper charts the progression of human needs through different epochs, emphasizing the crucial development of virtual environments that align with these dynamic requirements. It highlights the metaverse's rapid ascent as a reflection of current human needs within the digital age. We introduce a pyramid model featuring six key design criteria crucial for virtual world construction. This model presents a structured approach, detailing the essential factors in developing virtual spaces, ranging from content creation to addressing technical complexities. Additionally, the paper offers a detailed examination of the core technical foundations needed to build virtual environments, identifying existing gaps in academic research. It outlines essential technological breakthroughs and strategic directions that could accelerate the metaverse's implementation. In its conclusion, the document forecasts the metaverse's potential applications across various sectors, anticipating its broad impact on numerous disciplines.
11:00-11:15	**(TP24-330)** Self-Training Based Instance Calibration for Unsupervised Domain Adaptation Semantic Segmentation of Remote Sensing Images **Mei Ai,** Shaanxi Normal University, China Abstract: Self-training based unsupervised domain adaptation approaches play a pivotal role in mitigating the domain shift and improving the segmentation performance of the target domain, where a trained model generates pseudo-labels for the target domain. However, how to improve the quality of the pseudo-labels of the target domain containing noise has received considerable critical attention. Moreover, remote sensing images have an imbalanced number of samples across classes. In this paper, we propose a self-training based instance calibration for unsupervised domain adaptation semantic segmentation of remote sensing images. Firstly, we encourage the different

augmented results of the same pixel of the target domain images to not only have the same class prediction but also have the similar relationship with the pixel instances of the source domain. Secondly, variations related to the number of samples of different classes are added to the logits of the pixels to narrow the gap between the feature areas of different classes. Experiments on International Society for Photogrammetry and Remote Sensing (ISPRS) and Tibet Plateau remote sensing datasets demonstrate that the proposed model can effectively improve the quality of the pseudo-labels of the target domain and mitigate low segmentation performance of the target domain caused by the imbalanced number of samples across classes.

11:15-11:30

(**TP24-312**) Self-Attention Facial Attribute Editing
Hang Zhang, Chongqing Technology and Business University, China

Abstract: With the advancement of facial image processing technologies, facial attribute editing has become a key research area, addressing challenges like semantic direction discovery, unnatural edits, and poor attribute disentanglement. This paper introduces a novel supervised method, the Multi-layer Self-Attention Network (MSAN), which uses self-attention to learn semantic information for controllable facial attribute editing. A labeled datasets was constructed for training, and a perceptual image patch method was introduced for image quality control. Extensive experiments show that the proposed method offers superior control over facial attribute editing while preserving image quality and identity, outperforming three state-of-the-art methods by 5% to 10% in all metrics.

11:30-11:45

(**TP24-203**) A Study of Icon Design Styles for Recognition in HUD Environments
Dihui Chu, Southeast University, China

Abstract: In field of AR research and icon design, a problem worthy of further discussion is whether the design style change will affect the user's cognitive efficiency. With the development of intelligent automobile, the display of icons is no longer limited to the mobile phone and computer, but extends to the intelligent cockpits. The icon design style inside the human-computer interaction interface in automobile HMI and HUD directly relates to the driver's cognitive performance and driving safety. Especially in the HUD, due to the vibration of the car itself, the display effect of the icon is significantly different from the traditional display interface. Therefore, the design style of the icon may have different effects on the driver's cognition in different driving environments. In order to investigate this issue in depth, this paper designed a visual search task through simulation experiments to compare the effects of three different design styles of icons (linear icons, filled icons, and combined linear-filled Icon) on drivers' recognition speeds in different vibration directions (horizontal and vertical) and at three vibration frequency levels (5 Hz, 10 Hz, and 15 Hz). The experimental results show that the linear design style icons display particularly well in the vibration environment, while in the vibration direction, the icons with transverse vibration are more recognizable compared to those with longitudinal vibration. These findings provide a reference for optimizing the design of icons in in-vehicle HUDs, which can help to improve drivers' cognitive efficiency and driving safety.

11:45-12:00

(**TP24-326**) Feature Fusion for Multi-Condition Controllable Image Generation
Pengfei Zhao, Tianjin University of Technology, China

Abstract: The rapid advancement of AI-generated content (AIGC) has led to methods that produce highly diverse images from text descriptions. However, relying solely on text descriptions often fails to meet the creator's precise needs. To address this issue, researchers have proposed various methods to enhance the controllability of generated images, which have achieved notable success. Nonetheless, a single condition often fails to fully express creative intentions. Combining multiple conditions with textual information can significantly enhance result controllability. This paper proposes a method for multi-condition controllable image generation based on feature fusion. It categorizes the given conditions into two types: details (e.g., pose maps, edge maps) and macro conditions (e.g., semantic segmentation maps, depth maps). These categories describe the image's detailed structure and background style, respectively. Leveraging their distinct characteristics and requirements, the method uses a combination of structural and global attention mechanisms for feature fusion. This approach extracts high-level and low-level conditional features, which are then fused into an image incorporating all given conditions using a feature pyramid. The composite image is fed into the control module of the generation model, guiding the generation process. Through experimental visualization and data metrics analysis, our method demonstrates improved controllability of generated results compared to existing approaches. It effectively satisfies multiple descriptive conditions while ensuring a certain level of quality and diversity in the generated images.

Best Presentation Award & Session Group Photo

Session 4

Sept. 16th | 13:30-15:45
Zoom ID: 815 0608 7777

- **Topic: Model Based Intelligent Information System and Interactive Network**
- **Session Chair: Assoc. Prof. Wei Song, Beijing Institute of Graphic Communication, China**

*Note: The schedule of each presentation is for reference only. Authors are required to attend the whole session, in case there may be some changes on conference day. Please join in the room 5-10 minutes earlier.

13:30-13:45	(**TP24-304**) Theory of Blockchain and Smart Healthcare: Highlights **Yuzo Iano,** State University of Campinas, Brazil Abstract: Smart healthcare incorporates cutting-edge technologies and intelligent systems into the healthcare sector to enhance productivity and accuracy. Blockchain technology is popular in the smart healthcare field because of its advantages such as reducing transaction cost, improving transaction time, and reducing fraudulent activities by removing the intermediate agents. This emphasizes technological advancements from Artificial Intelligence (AI) and the Internet of Things (IoT) to blockchain. This fulfills the needs of the smart healthcare industry. This paper starts with an introduction to blockchain and IoT and discusses the application of blockchain in smart healthcare.
13:45-14:00	(**TP24-310**) SAMFENet: A Lightweight Scale-Aware Modulation single-modal feature extraction network for Face Anti-Spoofing **Zihan Shang,** Chongqing Technology and Business University, China Abstract: In the field of Face Anti-Spoofing (FAS), lightweight models primarily rely on Convolutional Neural Networks (CNN) and em- ploy deep separable convolution to optimize computational requirements and model parameters.Generally, CNNs are primarily sensitive to local details and limited in their ability to capture global perception.However, Transformers can capture global perception, their substantial parameter count poses challenges for deployment on resource-constrained devices such as mobile or low-computing-power devices.In FAS, both detailed features and global features of face images hold equal significance. Acquiring local detailed features and global semantic features of face images enhances the recognition of genuine and deceptive faces. In order to solve the problems, this paper proposes SAMFENet, a lightweight feature extraction model for Face Anti-Spoofing, which integrates the strengths of CNN and Transformers. The model introduces an optimized Feature Extraction Module (FEM), a Scale-Aware Modulation (SAM) module, and a transformer Multi-Head Self-Attention (MSA) module. By stacking the SAM and MSA modules, it effectively simulates the transition from capturing local dependencies to global dependencies as the network's depth increases. These three modules serve as the backbone for feature extraction in this paper's model, enabling effective feature extraction. The model exhibits a lightweight design and can be deployed in single-mode embedded device scenarios. Experimental results show that the model performs well on multiple datasets, with an accuracy of 99.896% and an average error rate of 0.112% on the Depth mode (CASIA-SURF) dataset. The accuracy of the SIW dataset was 99.74%, with an average error rate of 0.16%. Furthermore, the model achieves zero error rates on two additional datasets, Replay-Attack and CASIA-FASD. The performance of this model is significantly better than most models, and the number of model parameters is only 0.5221M, which is much smaller than the Transformer series model and less than most hybrid models of CNN and Transformer.

22

14:00-14:15	**(TP24-212)** Tracked Inspection Robot：Teleoperation based on ROS and Unity 3D **Ruochen Du,** Southeast University, China Abstract: Deficiencies in depth and perspective perception often hinder teleoperation based on video observation, especially in monocular settings. The integration of virtual reality (VR) allows the operator to be more immersed in the working environment, which can optimize the operation process and reduce the difficulty of operation. Taking cable trench inspection as an application scenario, this paper proposed a teleoperation system based on robot operating system (ROS) and Unity. The proposed method achieved the synchronization of both virtual and real robot actions, enabling the transmission and visualization of multimodal sensor data, and eventually allowing operators to utilize the virtual interface for immersive robot operation. The results of the user survey verified the feasibility of using the tracked robot based on ROS to achieve teleoperation combined with VR, and the introduction of VR reduces the cognitive load.
14:15-14:30	**(TP24-317)** I-WOA: An Optimization of K-Means Clustering **Pinyao He,** Chongqing Technology and Business University, China Abstract: As the big data technology is widely used in various research fields, cluster analysis has been generally regarded as one of the most effective methods for processing different types of data. In response to the shortcomings of the K-means algorithm, such as sensitivity to the initial clustering points and a propensity to fall into local optima, an Improved Whale Optimization Algorithm (I-WOA) is proposed to optimize the K-means clustering algorithm. In the proposed I-WOA algorithm, the Sobol sequence is introduced to initialize the whale population, thus enriching the diversity of the whale population. An elite reverse learning strategy is also employed to improve the position updating method of the optimal whale in the whale population, hence enhancing the quality of the solution. Additionally, the convergence factor （a) and contraction probability (p) in the whale algorithm are adjusted to balance global and local searches, therefore improving convergence speed and optimization precision. Finally, the I-WOA algorithm is evaluated by six famous benchmark functions. The experimental results show that the proposed algorithm has better cluster quality and global search capability compared to traditional clustering algorithms, and outperforms some popular algorithms.
14:30-14:45	**(TP24-320)** Analysis of Education Model Based on Student Big Data Portrait **Xinxin Wang,** Shandong Institute of Commerce and Technology, China Abstract: This article mainly studies student profiling behavior in marketing education in universities. Portraiting marketing students can help solve the problem of identifying and judging their marketing professional abilities. The process of constructing a student portrait system in marketing education mainly includes four steps: data collection and processing, label extraction, forming a label system, and portrait formation. After the marketing ability of students is specifically represented through portraits, we can provide personalized and accurate marketing education resources, teaching methods, and evaluation systems based on the label characteristics and needs of different students.
14:45-15:00	**(TP24-323)** Case Study on Urban Mobility in the USA. Applying Multivariate Linear Regression **Gabriel Gomes de Oliveira,** CTI - Renato Archer Information Technology Center, Brazil Abstract: We have been experiencing technological modernization in all segments of cities, which enables the advancement of the smart cities concept through the use of technology to solve current and common problems in cities. Currently, numerous innovative technologies, such as artificial intelligence, big data, the Internet of Things, and sensors, are used to collect and analyze data regarding aspects of cities. Notably, more precise investments in data science make it possible to solve countless problems more safely and efficiently. More than 800 billion dollars are invested annually in the USA to make the urban mobility scenario safer and more efficient. In this perspective, the present work focuses on the application of a mathematical model, Multivariate Linear Regression, to a dataset that contains data about traffic accidents in the USA to understand the aspects, natural or not, that can characterize a higher risk of accidents in the USA.

September 14-16, 2024

University of Saint Joseph, Macao, SAR China

15:00-15:15	(**TP24-332**) Network Optimization for Deterministic Services in Industrial Internet **Meng Song,** China Unicom Smart City Research Institute, China Abstract: This article proposes a new solution that combines network awareness and business characteristics to optimize network service quality. Firstly, in-band network telemetry is used to enhance the network awareness. Secondly, end-to-end deterministic network service capabilities are achieved through dynamic resource scheduling and routing optimization. Finally, the applicability of this method in the industrial Internet business scenario is verified by conducting test in real business scenarios.
15:15-15:30	(**TP24-204**) Development and Research on Virtual Antarctic Experience System **Jierui Xing,** Beijing Institute of Graphic Communication, China Abstract: Experiencing the world based on virtual reality technology is increasingly valued. The aim of this study is to develop an Antarctic experience system that integrates multiple key technologies using the Unreal Engine platform, in order to promote a deeper understanding and recognition of the natural environment in Antarctica. This system integrates advanced technologies such as Antarctic scene reconstruction, physics-based aurora simulation, penguin behavior logic programming, realistic day and night alternation simulation, interactive operation design of snowmobiles, and automatic generation of complex terrain. Through high-precision modeling and rendering techniques, this system accurately reproduces the natural environment of Antarctica and provides users with an immersive exploration experience by simulating natural phenomena such as auroras, penguin behavior, and day-night alternation. In addition, the interactive design and diverse terrain generation of snowmobiles further enhance the interactivity and exploration of the system. This study will also provide powerful tools for polar scientific research and education, and also offer some ideas for the application of virtual reality technology in natural environment simulation and experiences.
15:30-15:45	(**TP24-313**) A GPT-PERNIE Model for Short Text Sentiment Analysis **Jingyue Huang,** Chongqing Technology and Business University, China Abstract: Recently, deep learning techniques have been widely used for text sentiment classification in the domain of natural language processing. Effective text representation plays a critical role in improving the classification performance of deep learning models in this field. Due to the limited emotional contents in short texts and the susceptibility to noise during training, we propose a GPT-PERNIE model for short text sentiment analysis. This model incorporates GPT with adversarial training (P). Initially, GPT is used to enrich the expressiveness of the text. Subsequently, a single-tower model is used to evaluate text similarity, followed by vectorization of input text through the ERNIE pre-training model, which enables preliminary extraction of emotional features from the text. Later, noise interference is introduced into the output vector of the ERNIE pre-training model, leading to the generation of adversarial samples by attacking the origin. Then, these adversarial samples are used for the adversarial training of the classification model, enhancing the model's robustness against noise attacks. Experimental results indicate that the GPT-PERNIE model demonstrates superior performance and generalization capability in text classification tasks. This achievement is not only significant within the realm of short text sentiment analysis but also paves the way for novel approaches to utilizing deep learning in natural language processing tasks.

Best Presentation Award & Session Group Photo

USJ Macao Introduction

University of Saint Joseph (Macao)

The **University of Saint Joseph (Macao)** is a Catholic university of China, rooted in a long tradition of intercultural exchange that has been the hallmark of Catholic higher education in Macao since the College of Saint Paul, founded in 1594.

The project of the Jesuit fathers Valignano, Rugieri and Ricci inspires our view of higher education with a universal character capable of promoting creative dialogue between the cultural, scientific, ethical and philosophical traditions of China and Europe.

The combined local and Portuguese roots of the University make it particularly suited to embody at the academic level Macao's strategic objective to act as a platform from which to boost relations between China and Portuguese-speaking countries.

"Scientia et Virtus" (Knowledge and Virtue), the motto of the Diocese of Macao, points our academic community to a harmonious pursuit of truth and the common good for the benefit of humanity, the Chinese people and Macao's multicultural society.

NOTE