Conference Program

2026 4th Asia Conference on Advances in Image Processing (AAIP 2026)

June 12-14, 2026

Guilin, China

Sponsor

Organizer

Table of Contents

Welcome Message

Dear Colleagues, Researchers and Friends,

On behalf of the Conference Committee, we are delighted to welcome you to the 2026 4th Asia Conference on Advances in Image Processing (AAIP 2026) taking place during June 12-14, 2026, in the beautiful city of Guilin, China.

AAIP 2026 invites high-quality contributions from all frontier fields of advances of image processing and related technologies. Topics of interest include, but are not limited to: Imaging Acquisition and Storage, Image Processing, Image Analysis and Understanding, Computer Graphics, Image and Graphics Application and Protection of Images and Neural Networks.

The AAIP 2026 conference is designed to promote academic exchange, share cutting-edge achievements, strengthen international cooperation, and inspire future research directions for scholars, students, engineers, and practitioners around the world. Through high-level keynote speeches, invited speeches, paper presentations, academic discussions, and informal communications, we aim to build an open and professional platform for the dissemination of knowledge, the collision of ideas, and the establishment of long-term academic partnerships.

We are honored to have four distinguished keynote speakers to sharing their professional insights:
-Prof. Xudong Jiang (IEEE Fellow), Nanyang Technological University, Singapore
-Prof. Ruili Wang (Fellow of Engineering New Zealand, Stanford/Elsevier Top 2% Scientists List (2021-2025)), Massey University, New Zealand
-Prof. Leida Li (National Young Talent), Xidian University, China
-Prof. Jiantao Zhou, University of Macau, China

We extend our deepest gratitude to all committee members for their enthusiastic support and tireless efforts in making AAIP 2026 a successful academic event. We also express our sincere thanks to the organizing committee, the technical program committee, our invited speakers, and all the authors for their valuable contributions and active participation.

Welcome to AAIP 2026! We hope every participant an inspiring and memorable experience.

Respectfully yours,
Ching-Nung Yang, National Dong Hwa University
Conference Co-Chair
AAIP 2026

Conference Committee

Conference Chair
Xianxian Li, Guangxi Normal University, China

Conference Co-chair
Ching-Nung Yang, National Dong Hwa University

Program Chairs
Zhenjun Tang, Guangxi Normal University, China
Xinpeng Zhang, Fudan University, China
Huiyu Zhou, University of Leicester, UK
Yang Yue, Xi'an Jiaotong University, China

Program Co-chairs
Yew Kee WONG, Hong Kong Chu Hai College, China
Yanglong Lu, Hong Kong University of Science and Technology, China
Ke-Lin Du, Guangzhou Digital Technology Group Inc., China
Grigorios Beligiannis, University of Patras, Greece

Publication Chair
Chunqiang Yu, Guangxi Normal University, China

Publicity Chairs
Xiaoping Liang, Guangxi Normal University, China
Loc Nguyen, Sunflower Soft Company, Vietnam

Local Organizing Committee
Zhuoran Zhang, Guangxi Normal University, China
Jianhua Hu, Guangxi Normal University, China

Technical Program Committees
Dimitrios A. Karras, National and Kapodistrian University of Athens, Greece
Yi Chen, Guizhou University, China
Aleksandr Cariow, West Pomeranian University of Technology, Poland
Anwaar Ulhaq, Central Queensland University, Australia
Mohd Zamri Ibrahim, Universiti Malaysia Pahang Al-Sultan Abdullah (UMPSA), Malaysia
Chengde Lin, Guilin University of Electronic Technology, China
Junhao Wu, Shantou University, China
Tao Chen, Sun Yat-sen University, China
Ju Tang, Guangdong University of Technology, China
Syed Farooq Ali, University of Management & Technology, Pakistan

Akilan Thangarajah, Lakehead University, Canada
Xibin Jia, Beijing University of Technology, China
Zhennong Chen, Xi'an Jiaotong-Liverpool University, China
Dinesh Goyal, Poornima Institute of Engineering & Technology, India
Baha Ihnaini, Wenzhou-Kean University, China
Wattana Jindaluang, Chiang Mai University, Thailand
R. S. Hegadi, Central University of Karnataka, India
Mas Rina Mustaffa, Universiti Putra Malaysia, Malaysia
Zhihong Li, Wenzhou University, China
Sunil Kumar, Manipal University Jaipur, India
Lu Leng, Nanchang Hangkong University, China
Seokwon Yeom, Daegu University Gyeongsan, Korea
Qiang Cheng, University of Kentucky College of Engineering, USA
Suraiya Jabin, Jamia Millia Islamia, India
Boon Chin Yeo, Multimedia University, Malaysia
Olarik Surinta, Mahasarakham University, Thailand
Alicia Colmenero-Fernández, Universidad de Jaen, Spain
Yunli Lee, Sunway University, Malaysia
Wenbo Wan, Shandong Normal University, China
Stelvio Cimato, University of Milan, Italy
Cheonshik Kim, Sejong University, Korea
Lein Harn, University of Missouri Kansas City, USA
Hongxia Wang, Sichuan University, China
Xiaotian Wu, Jinan University, China
Lizhi Xiong, Nanjing University of Information Science and Technology, China
Wenbo Wan, Shandong Normal University, China
Rushit Dave, Minnesota State University, USA
Sukru Karaaslan, Firat University, Turkey
Philippe Durand, Conservatoire National des Arts et Métiers, France
Kanokrat Jirasatjankul, Phetchaburi Rajabhat University, Thailand
Thaweesak Trongtirakul, Rajamangala University of Technology Phra Nakhon, Thailand
João Manuel R. S. Tavares, Universidade do Porto (FEUP), Portugal
Kasturi Vasudevan, Indian Institute of Technology Kanpur, India
Kaushik Roy, West Bengal State University, India

Guidance

ONSITE ORAL PRESENTATIONS

- The duration of an oral presentation is **15 minutes**, including Q&A.
- Get your presentation PowerPoint or PDF files prepared and backed up with a flash disk or USB. Slide Background please visit:
- Laptops, projector & screen, laser sticks will be provided by the conference organizer.
- The regular presentation time arrangement is for reference only. In case any absence or some presentations are less than 15 minutes, please join your session before it starts. Your punctual arrival and active involvement in each session will be highly appreciated.
- Please keep all your belongings (bags and laptop etc.) at any time. The conference organizer does not assume any responsibility for the loss of personal belongings.
- For security purposes, delegates, speakers, exhibitors and staff are required to wear their name badge to all sessions and social functions. There will be NO access for people without a badge. Never discard your badge at will.

ONLINE ORAL PRESENTATIONS

- Platform: ZOOM Meeting. Download Link: https://zoom.us/download
- Virtual Background & Slide Background please visit: www.aaip.net/kits.zip
- ZOOM Link: https://us02web.zoom.us/j/89264147904 (Password: 202606)
- ZOOM Test: **June 12, 2026 (UTC+8)**
- The duration of an oral presentation is **15 minutes**, including Q&A, please join in the ZOOM room 10 minutes before the session.
- Equipment needed: A computer with stable internet connection and camera and headphones
- Please rename your screen name before entering the room.

Rename Screen Name Before Entering the Room	Examples
Author: Paper ID-Name	CG101-San Zhang
Delegate: Delegate-Name	Delegate-San Zhang
Keynote Speaker: Keynote-Name	Keynote-San Zhang
Invited Speaker: IS-Name	IS-San Zhang
Committee Member: Committee-Name	Committee-San Zhang

Conference Venue

Grand Link Hotel 桂山华星酒店
Address: No. 42 Chuanshan Road, Guilin, Guangxi Zhuang Autonomous Region, China | Phone: 0773-5639999
地址：中国广西壮族自治区桂林市穿山路 42 号 | 电话：0773-5639999
http://www.guishanhotel.com/index.html
ROOM: Conference Room 6&7 (2nd Floor) 会议室：二楼临桂&阳朔厅

Transportation Instruction

• Guilin Liangjiang International Airport 桂林两江国际机场

- Taxi: Approximately 30 km, 45 mins, about 40 CNY fare;
  Airport Shuttle: to "Minhang Building" first, then switch to a bus or taxi to reach Hotel, total about 1 hour 30 minutes, shuttle fare 20 CNY.
- 打车前往酒店约 30 公里 45 分钟，费用约 40 元；乘机场大巴到“民航大厦”，再换乘公交或打车前往酒店，全程约 1 小时 30 分钟，机场大巴票价 20 元。

• Guilin Railway Station 桂林站

- Taxi: Approximately 3.5 km, 10 mins, about 7 CNY fare;
  Bus: Take Bus 11 from "Guilin Railway Station" to "Seven-Star Park", then walk 900 meters to reach Hotel, total about 50 minutes, bus fare 2 CNY.
- 打车前往酒店约 3. 5 公里 10 分钟，费用约 7 元；乘坐 11 路公交，“桂林站”至“七星公园站”，再步行 900 米前往酒店，全程约 50 分钟，公交票价 2 元。

• Guilinbei Railway Station 桂林北站

- Taxi: Approximately 8 km, 25 mins, about 20 CNY fare;
  Bus: Take Bus 18 from "Guilinbei Railway Station Intersection" to "Seven-Star Park", then walk 900 meters to reach Hotel, total about 1 hour, bus fare 2 CNY.
- 打车前往酒店约 8 公里 25 分钟，费用约 20 元；乘坐 18 路公交，“桂林北站路口”至“七星公园站”，再步行 900 米前往酒店，全程约 1 小时，公交票价 2 元。

Emergency

- Emergency Call: 110
- First Aid Call: 120
- Fire Alarm Call: 119

Electricity

- 220V/50Hz sockets are mainly used in China.

Daily Schedule

* All times are China Standard Time (UTC+8).
* Online ZOOM Link: https://us02web.zoom.us/j/89264147904 (Password: 202606)

Day 1 – Friday, June 12, 2026		
Time	**Events**	**Venue**
10:00-17:00	Onsite Registration, Sign-in & Materials Collection	Lobby of Grand Link Hotel
10:00-17:00	Online ZOOM Test	ZOOM Link: https://us02web.zoom.us/j/89264147904 (Password: 202606)
Day 2 – Saturday, June 13, 2026		
Time	**Events**	**Venue**
08:30-08:40	**Opening Ceremony:** **Welcome Speech:** **Prof. Xianxian Li** Guangxi Normal University, China **Opening Remarks:** **Prof. Ching-Nung Yang** National Dong Hwa University **Host:** **Prof. Zhenjun Tang** Guangxi Normal University, China	Conference Room 6&7 (2nd Floor)
08:40-09:00	Group Photo	Main Entrance on 1st Floor
09:00-09:40	**Keynote Speech 1** **Prof. Ruili Wang** Fellow of Engineering New Zealand, Stanford/Elsevier Top 2% Scientists List (2021-2025) Massey University, New Zealand **Speech Title:** *Multimodal Data Processing* **Host:** **Prof. Ching-Nung Yang** National Dong Hwa University	Conference Room 6&7 (2nd Floor)

09:40-10:20	**Keynote Speech 2** **Prof. Xudong Jiang** IEEE Fellow Nanyang Technological University, Singapore **Speech Title:** *How Deep CNN and Transformer Solve Machine Learning Problems of Traditional ANN* **Host:** **Prof. Ching-Nung Yang** National Dong Hwa University	Conference Room 6&7 (2nd Floor) ZOOM Link: https://us02web.zoom.us/j/89264147904 (Password: 202606)
10:20-10:40	Coffee Break	Outside Corridor of Conference Room
10:40-11:20	**Keynote Speech 3** **Prof. Leida Li** National Young Talent Xidian University, China **Speech Title:** *Fine-grained Visual Quality Assessment* **Host:** **Prof. Jingli Wu** Guangxi Normal University, China	Conference Room 6&7 (2nd Floor)
11:20-12:00	**Keynote Speech 4** **Prof. Jiantao Zhou** University of Macau, China **Speech Title:** *Towards Robust Learning-Based Multimedia Forensics* **Host:** **Prof. Li'e Wang** Guangxi Normal University, China	
12:00-13:30	Lunch	Hexi Chinese Restaurant (first floor) 一楼和熙轩中餐厅
13:30-15:55	**Onsite Breakout Session 1:** *Multimodal Intelligent Image Fusion Analysis and Advanced Processing Methods*	Conference Room 6&7 (2nd Floor)
15:55-16:10	Coffee Break	Outside Corridor of Conference Room
16:10-18:05	**Onsite Breakout Session 2:** *High-Resolution Optical and Digital Imaging Methods*	Conference Room 6&7 (2nd Floor)

18:05-20:00	Dinner	Amber Room (2nd Floor) 2 楼琥珀厅
Day 3 – Sunday, June 14, 2026		
Time	**Events**	**Venue**
09:30-11:40	**Online Breakout Session 1:** *AI-Driven Digital Image Analysis and Processing Technologies*	ZOOM Link: https://us02web.zoom.us/j/89264147904 (Password: 202606)
14:00-16:25	**Online Breakout Session 2:** *Intelligent Image Analysis, Secure Computing, and Privacy Protection*	

Keynote Speaker

Prof. Ruili Wang

Fellow of Engineering New Zealand, Stanford/Elsevier Top 2% Scientists List (2021-2025)

Massey University, New Zealand

Speech Time: 09:00-09:40, June 13, 2026 (UTC+8)
Onsite Room: Conference Room 6&7 (2nd Floor)

Speech Title: Multimodal Data Processing

Abstract: In this presentation, we will introduce our recent progress in Multimodal Data Processing, especially in video captioning and infrared–visible image fusion.

Video captioning bridges computer vision and natural language processing and plays an essential role in various knowledge-driven systems within the streaming media era. Recent video captioning methods have achieved promising performance by leveraging external textual knowledge to better understand video content and generate more informative captions. Nevertheless, existing methods that rely excessively on knowledge graphs still suffer from several inherent limitations. To address these issues, we propose a novel knowledge enhancement and disentanglement learning framework for video captioning.

Image fusion is an important technique in computer vision and image processing. It integrates multiple images of the same scene — captured by different sensors, at different times or from different viewpoints — into a single high-quality composite image. Current text-driven infrared–visible image fusion methods mainly adopt sentence-level textual guidance. This paradigm easily introduces semantic noise caused by text redundancy and fails to fully exploit the deep semantic value of textual cues. To overcome these drawbacks, we propose a novel fusion approach dubbed Entity-Guided Multi-Task Learning for Infrared and Visible Image Fusion.

Biography: Professor Ruili Wang, Fellow of Engineering New Zealand, graduated from Huazhong University of Science and Technology, Northeastern University, and Dublin City University, where he obtained his B.E., M.E., and Ph.D., respectively. His research areas include AI, machine learning, computer vision, and speech and language processing. His research has been funded by multiple grants from the New Zealand government. Professor Wang is an associate editor/editorial board member of the following journals: IEEE

Transactions on Multimedia (TMM), IEEE Transactions on Circuits and Systems for Video Technology (TCSVT), IEEE Computational Intelligence Magazine, IEEE Transactions on Emerging Topics in Computational Intelligence (TETCI), ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM), etc.

Keynote Speaker

Prof. Xudong Jiang

IEEE Fellow

Nanyang Technological University, Singapore

Speech Time: 09:40-10:20, June 13, 2026 (UTC+8)
Onsite Room: Conference Room 6&7 (2nd Floor)
ZOOM Link: https://us02web.zoom.us/j/89264147904 (Password: 202606)

Speech Title: How Deep CNN and Transformer Solve Machine Learning Problems of Traditional ANN

Abstract: The powerfulness of machine learning was already proven more than 30 years ago in the boom of neural networks but its successful application to the real world is just in recent 10 years after the deep convolutional neural networks (CNN) have been developed. This is because the machine learning alone can only solve problems in the training data but the system is designed for the unknown data outside of the training set. This gap can be bridged by regularization: human knowledge guidance or interference to the machine learning. This speech will analyze these concepts and ideas from traditional neural networks such as MLP to the deep convolutional neural networks (CNN) and Transformer. It will answer the questions why the traditional neural networks fail to solve real world problems even after more than 30 years' intensive research and development and how the deep CNN and Transformer solve the problems of the traditional neural networks and now are very successful in solving various real world AI problems.

Biography: Xudong Jiang (Fellow of IEEE) received the B.E. and M.Eng degrees from the University of Electronic Science and Technology of China (UESTC), and the PhD degree from Helmut Schmidt University, Hamburg, Germany. From 1998 to 2004, he was with the Institute for Infocomm Research, A*STAR, Singapore, as a lead scientist, and the head of the Biometrics Laboratory. He joined Nanyang Technological University (NTU), Singapore, as a faculty member, in 2004, where he served as the director of the Centre for Information Security from 2005 to 2011. He is currently a professor with the School of EEE, NTU and serves as the director of the Centre for Information Sciences and Systems of School of EEE,

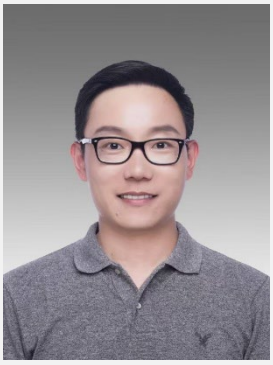
NTU. He has authored over 300 papers with over 80 papers in IEEE journals including 15 T-PAMI papers and over 20 T-IP papers. Dr Jiang has presented over 50 papers in top AI conferences CVPR/NeurIPS/ICML/ICCV/ECCV/ICLR/AAAI. His papers have been cited over 18 Thousand times with H-index 73. He served as IFS TC member of the IEEE Signal Processing Society from 2015 to 2017, associate editor for IEEE Signal Processing Letter from 2014 to 2018 and associate editor for IEEE Transactions on Image Processing from 2016 to 2020. Currently, he is an IEEE Fellow, serves as senior area editor for IEEE Transactions on Image Processing and editor-in-chief for IET Biometrics. He also served as Area chairs for AAAI, NeurIPS and IEEE ICIP. His current research interests include image processing, pattern recognition, computer vision, machine learning, and biometrics.

Keynote Speaker

Prof. Leida Li

National Young Talent

Xidian University, China



Speech Time: 10:40-11:20, June 13, 2026 (UTC+8)
Onsite Room: Conference Room 6&7 (2nd Floor)

Speech Title: Fine-grained Visual Quality Assessment

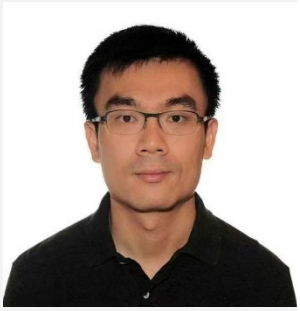
Abstract: Visual quality assessment measures the perceptual quality of images by simulating the characteristics of the Human Visual System (HVS). As a common technology, it has important applications in many fields such as low-level vision, imaging optimization, smart photography, and AIGC. After more than 20 years of rapid development, a large number of algorithms have been proposed. However, the existing methods typically suffer from insufficient discrimination ability when used in real-world environments. This talk focuses on the key differences between coarse-grained and fine-grained visual quality assessment, the main research progress in fine-grained visual quality assessment, as well as applications in camera tuning and aesthetic recommendation, etc.

Biography: Leida Li is a Full Professor at Xidian University, recognized as a National Young Talent. His research interests include computer vision, visual quality assessment, and computational aesthetics. He has published over 100 papers in top-tier journals and conferences like IEEE TPAMI, IEEE TIP, CVPR, ICCV, and AAAI, with about 10,000 citations. He has led five projects supported by the National Natural Science Foundation of China and has actively engaged in industry-academia collaborations with top companies such as Huawei, OPPO, and Tencent. He was awarded the "Outstanding Industry-Academia Collaboration Partner" by OPPO, and his research outcomes have been applied in smart phone and live-streaming cameras. He is an Associate Editor of IEEE Transactions on Image Processing (TIP) and Journal of Visual Communication and Image Representation (Best Associate Editor Award 2021/2023), and serves as Area Chair/Senior Program Committee member for top international conferences such as AAAI, IJCAI, and ACM MM. He is a Senior Member of IEEE/CCF/CSIG.

Keynote Speaker

Prof. Jiantao Zhou

University of Macau, China



Speech Time: 11:20-12:00, June 13, 2026 (UTC+8)
Onsite Room: Conference Room 6&7 (2nd Floor)

Speech Title: Towards Robust Learning-Based Multimedia Forensics

Abstract: In recent years, the proliferation of sophisticated multimedia generation and manipulation technologies, such as deepfakes and advanced image/video editing tools, has significantly blurred the line between authentic and fabricated content. As multimedia plays an increasingly crucial role in information dissemination, legal evidence, and social interactions, ensuring its integrity has become a pressing concern. How can we effectively distinguish genuine media from skillfully crafted forgeries, especially when traditional forensic techniques struggle to keep pace with rapidly evolving tampering methods? Moreover, the challenges are compounded by the degradation of forensic features during media transmission and the vulnerability of detection models to adversarial attacks. In the realm of multimedia forensics, learning-based approaches offer a promising avenue for tackling these complex issues. However, there is a pressing need to enhance their robustness against various distortions, obfuscation strategies, and dynamic threats. This talk explores the latest advancements in robust learning-based multimedia forensics, delving into novel methodologies designed to fortify detection capabilities. From developing innovative feature extraction techniques that can withstand transmission-induced degradation to creating resilient models that can counter adversarial manipulations, the presentation aims to outline a comprehensive research direction for achieving reliable and trustworthy multimedia forensics in an increasingly digital and deceptive world.

Biography: Dr. Jiantao Zhou is a Full Professor at the Department of Computer and Information Science, and the State Key Laboratory of Internet of Things for Smart City, University of Macau, where he also serves as the Director for Research Services and Knowledge Transfer Office. He graduated from the Hong Kong University of Science and Technology in 2009 with a PhD in Electrical and Computer Engineering. He was a Fulbright

Junior Scholar at the University of Illinois at Urbana-Champaign (UIUC). Professor Zhou's research focuses on AI security, multimedia information privacy protection and forensics, and intelligent multimedia information processing. He has published more than 200 papers in top journals such as IEEE T-PAMI, IEEE T-IP, IEEE T-SP, IEEE T-IFS, IEEE T-AC and other top conferences such as CVPR, ICCV, ICML, and AAAI. He currently serves as the Associate Editor for IEEE Trans. Multimedia and IEEE Trans. Dependable and Secure Computing, the top journals in the field of multimedia information processing and security and was the Editor-in-Chief of APSIPA Newsletters. He is the Chair for the Multimedia Systems and Applications Technical Committee in IEEE Circuits and Systems Society and was the TPC Co-Chair of ICME 2023 and the General-Chair of APSIPA ASC 2024. He received the 2022 Macau Science and Technology Award (Third Prize, Natural Science Award) and the 2023 Alibaba Outstanding Academic Cooperation Project Award.

Invited Speaker

Prof. Xibin Jia

Beijing University of Technology, China

Speech Time: 13:30-13:50, June 13, 2026 (UTC+8)
Onsite Room: Conference Room 6&7 (2nd Floor)

Speech Title: Clinically Oriented Medical Image Analysis

Abstract: Medical imaging plays a fundamental role in modern clinical practice, supporting disease screening, diagnosis, and treatment planning across diverse modalities such as X-ray, MRIs, and CT. Despite significant advances in medical image analysis, two major challenges remain: poor generalization across different clinical environments and a heavy reliance on large-scale annotated datasets. In this talk, I will present our recent work on clinically oriented medical image analysis, focusing on improving model robustness and data efficiency. Specifically, we explore domain generalization, aiming to develop models that can generalize effectively to unseen hospitals or imaging conditions without accessing target-domain data. In addition, we investigate learning with limited annotations, including semi-supervised and few-shot learning approaches, as well as vision–language pre-training, to reduce dependence on costly expert annotations while maintaining high performance. We further introduce our efforts in leveraging emerging architectures, such as Mamba, to capture long-range dependencies while enabling effective local modeling. Finally, we briefly extend our research to causal modeling for complex tasks such as long-term action recognition. Overall, this work aims to bridge the gap between methodological advances and real-world clinical applicability, enabling more effective and efficient medical image analysis.

Biography: Xibin Jia a Professor of College of Computer science at the Beijing University of Technology and Beijing Institute of Artificial Intelligence. She received her Ph.D. in Computer Application Technology from Beijing University of Technology in 2007. She was a one-year visiting scholar at University of California Riverside U.S. and a half-year visiting scholar at Flinders University Australia respectively. She has nearly 30 years of experience in

computer science teaching and research. She is a Ph.D. and M.S. supervisor. Her research interests include computer vision, deep representation learning, and multi-modality deep learning. Her current work focuses on Intelligent Medical Image Analysis and Diagnosis, Affective Computing, and behavior recognition. She has served as Principal Investigator (PI) or co-PI on dozens of projects supported by grants from the National Natural Science Foundation of China, the Beijing Natural Science Foundation, and others. She has published dozens of papers in refereed journals and conference proceedings, including ACM MM 2025, MICCAI 2025, IEEE JBHI, ESWA, EAAI, and IEEE/ACM-TCBB. She is a Distinguished Member of CCF and a member of several technical committees, including CCF-CV, CSIG-BVD, CSIG-MV, and CAAI-IM. She has served as an Area Editor and Editor for the SCI-indexed journal KSII-TIIS. Her papers were awarded the IEEE MedAI 2024 Best Paper Award.

Invited Speaker

Prof. Seokwon Yeom

Daegu University, Korea

Speech Time: 16:50-17:10, June 13, 2026 (UTC+8)
Onsite Room: Conference Room 6&7 (2nd Floor)

Speech Title: Thermal Image Detection and Tracking with a Multi-rotor Drone

Abstract: This talk presents thermal image-based object detection and tracking using a multi-rotor drone. Thermal objects are detected by a YOLO detection model trained on a custom dataset, and the detected objects are tracked using Kalman or IMM filters. Through track association and fusion, the most reliable tracks are selected and fused to generate continuous trajectories. In addition, the track segment association connects track segments that are temporally disconnected. In the experiment, three hikers on a mountain were captured using a drone-mounted thermal imaging camera. The proposed approach achieved excellent tracking performance in terms of total track lifetime, mean track lifetime, and track purity.

Biography: Seokwon Yeom has been a faculty member of Daegu University since 2007. He has a Ph.D. in Electrical and Computer Engineering from the University of Connecticut in 2006.

He has been a guest editor of Applied Sciences and Drones in MDPI since 2019. He has served as a board member of the Korean Institute of Intelligent Systems since 2016, and a member of the board of directors of the Korean Institute of Convergence Signal Processing since 2014. He has been program chair of several international conferences. He was a vice director of the AI homecare center and a head of the department of IT convergence engineering at Daegu University in 2020-2023, a visiting scholar at the University of Maryland in 2014, and a director of the Gyeongbuk techno-park specialization center in 2013. He has been a keynote or invited speaker at several international conferences. His research interests are intelligent image and optical information processing, deep and machine learning, target tracking, and state estimation of drones.

Invited Speaker

Prof. Suraiya Jabin

Jamia Millia Islamia, India

Speech Time: 14:00-14:20, June 14, 2026 (UTC+8)
Online Room: https://us02web.zoom.us/j/89264147904 (Password: 202606)

Speech Title: AI-based Adult Age Threshold Modeling using Orthopantomographs for Forensic and Judicial Applications

Abstract: Dental development visible in OPG is considered one of the most reliable biological indicators for age estimation for persons. This is the most legally sensitive age group in India: 17 or 18 years can change court jurisdiction, determines juvenile vs adult trial, impacts sentencing severity, etc. Therefore, OPG-based AI-assisted adult/non-adult classification constitutes a scientifically valid, ethically sound, and legally relevant tool to aid forensic age determination in India.

EfficientNet emphasizes efficiency through compound scaling, jointly optimizing network depth, width, and input resolution, resulting in exceptional accuracy; efficiency trade-offs using MBConv blocks and squeeze-and-excitation attention. We fine-tuned these variants of deep convolution networks viz. EfficientNet_b0, ResNext50_32X4d, ConvNext_base using 6860 OPG collected from Faculty of Dentistry with meta data towards computational modelling of age threshold. We followed standard practices of Train/Test splits (85:15), judging performance of all models on the same independent Test data set of size 1029 OPGs. After Train/Test split, data augmentation with random horizontal flip, random rotation, and color jitters, etc were added on the fly to the Train

dataset. As only less than 20% of samples were belonging to Non-adult class, special set of hyper-parameters such as weighted loss, OPG reliant image shape (384X768), performance-based scheduler, etc. were incorporated while training of models. An ensemble of EfficientNet_b0 outperformed all other models with F1 score of 0.890 for Non-adult class, F1 score of 0.985 for Adult class, and F1 macro of 0.937 on an independent Test dataset of size 1029. We deployed the proposed model on a dynamic web server publicly available at: http://115.241.23.53:8000/ under tab “OPG Forensics” for radiologists to upload their OPG

data and get results and Grad-CAM visualization within a few seconds. With our highly accurate model (F1 macro of 0.937), we aim to reduce the workload of forensic experts and support more efficient and objective judicial decision-making. This contribution marks an early and significant step toward AI-driven mathematical modeling of adult age threshold classification.

Biography: Dr. Suraiya Jabin is a Professor of Computer Science in Jamia Millia Islamia, central university in New Delhi, India. She has more than 23 years of experience in computer science teaching and research. Her research interests include Artificial Intelligence in Healthcare, Behavioural Biometrics, Smartphone signature biometrics, social media analysis, and computational biology. Her current work focuses on providing AI and deep learning-based solutions to problems in postgenomic biology. She has taught courses on machine learning, AI, digital image processing, deep learning, advanced DBMS, Compiler Design, Discrete Mathematics, etc. to Masters and UG, and Pre-PhD courses during her teaching career of 23 years.

She has contributed over 70 research articles in journals, conference proceedings, and book chapters, books, including 15 papers in SCIE/Scopus indexed journals, and 2 books on Machine Learning with Wiley India publisher. She owns an Indian patent titled “Mobile-Biometric Signature based Authentication System” dated Feb 2017, and 2 patents are under progress. She was PI for the departmental infrastructural grant Bioinformatics Infrastructure Facility Center funded by DBT, GoI for 8 years from 2012 to 2020. She successfully supervised 8 Ph.D. scholars, and currently, supervising 5 Ph.D. scholars working on various problems such as disease outbreak prediction using social media, crowd monitoring using deep learning, and Mental Task classification using EEG data, etc. She is a member of several professional bodies including ACM Professional Member since Feb 2024 (in recognition of her reviewer assignments for ACM journals), and a life member of Indian Society for Technical Education (ISTE) since 2005. She is an editorial board member of prestigious journals Nature Scientific Reports, Frontiers in Computer Science, IGI Global, etc. She is an active reviewer for various journals of IEEE, ACM, Springer, iScience, InderScience, Taylor & Francis, Sage, etc. Along with teaching & research, she has been serving several administrative responsibilities such as Teacher Placement Coordinator, UG/PG Curriculum Design Coordinator in the present, and non-resident warden, assistant proctor, etc. in the past in JMI.

Invited Speaker

Prof. Xiwen Zhang

Beijing Language and Culture University, China

Speech Time: 09:30-09:50, June 14, 2026 (UTC+8)
Online Room: https://us02web.zoom.us/j/89264147904 (Password: 202606)

Speech Title: Three views on intelligently extracting and generating information from image

Abstract: Due to pattern recognition and deep learning, various information can be extracted and generated from image. Our work has focused on using the proposed hierarchy models, local homogeneity, and adversarial generation.

Various digital images are processed, such as ones scanned from mechanical paper drawings and paper text, face images, portrait ones with line drawings, and microscopic bone marrow images.

Various information is extracted using the proposed hierarchy models. Graphics and their multi-levels compounded objects are extracted and recognized from images scanned from mechanical paper drawings using a hierarchy model of engineering drawings. Faces and their components are extracted from photos using a facial model.

Various information is extracted using the proposed local homogeneity. Karyocytes and their components from microscopic bone marrow images based on regional color features.

Various information is extracted and generated from image using cycle-Consistent adversarial networks. Text is separated from grid background using cycle-Consistent adversarial networks. Digital images of Chinese classical upper-class lady paintings are generated from images with line drawings using conditional generative adversarial networks.

Biography: Professor, Doctoral Supervisor, Beijing Language and Culture University Biography: XiWen Zhang is currently a full professor of Digital Media Department, School of Information Science, Beijing Language and Culture University.

Prof. Zhang worked as an associated professor from 2002 to 2007 at the Human-computer interaction Laboratory, Institute of Software, Chinese Academy of Sciences. From 2005 to

2006 he was a Post doctor advised by Prof. Michael R. Lyu in the Department of Computer Science and Engineering, the Chinese University of Hong Kong. From 2000 to 2002 he was a Post doctor advised by Prof. ShiJie Cai in the Computer Science and Technology department, Nanjing University. Prof. Zhang's research interests include pattern recognition, computer vision, and human-computer interaction, as well as their applications in digital image, video, and ink. Prof. Zhang has published over 60 refereed journal and conference papers. His SCI papers are published in Pattern Recognition, IEEE Transactions on Systems Man and Cybernetics B, Computer-Aided Design. He has published more than twenty EI papers.

Prof. Zhang received his B.E. in Chemical equipment and machinery from Fushun Petroleum Institute (became Liaoning Shihua University since 2002) in 1995, and his Ph.D. advised by Prof. ZongYing Ou in Mechanical manufacturing and automation from Dalian University of Technology in 2000.

Invited Speaker

Prof. Dinesh Goyal

Poornima Institute of Engineering & Technology, India

Speech Time: 14:20-14:40, June 14, 2026 (UTC+8)
Online Room: https://us02web.zoom.us/j/89264147904 (Password: 202606)

Speech Title: AI-Enabled Image Analysis Framework for Healthcare Applications

Abstract: Artificial intelligence (AI) has become one of the most effective medical image analysis tools, that allows it to interpret complicated visual patterns automatically to aid in clinical diagnosis and decision-making. The inconsistency of medical imaging data, preprocessing strategies, and fragmented evaluation practices, however, remain a limiting factor to the reliability and reproducibility of available methods. The paper describes a systematic AI-based image analysis system in healthcare, aimed at offering a framework of unified and repeatable medical image processing and analysis pipeline. The proposed architecture incorporates standardized image preprocessing, feature learning through deep learning, systematic model evaluation, and performance analysis in a modular architecture that can be tailored to be used in various imaging modalities. With the help of the available convolutional neural networks' architecture, the framework is tested on representative healthcare imaging tasks, such as disease classification and medical image segmentation. The outcomes of the experiment indicate that deep learning models can be effectively utilized as diagnostic tools with regards to the accuracy, sensitivity, specificity, and AUC- ROC when integrated into a clear analytical framework. The results suggest that the suggested strategy enhances robustness and generalization on a wide range of imaging tasks and is still clinically relevant. The article highlights the significance of organized AI pipelines in the development of trusted and highly scalable medical image analysis to actual healthcare settings.

Biography: Dr. Dinesh Goyal, is an Professor in Computer Science & Engineering and is currently designated as Principal and Director at Poornima Institute of Engineering & Technology (PIET), Jaipur, with over 25 years of experience in teaching, research, and administration. He holds B.E, M.Tech, and Ph.D. degrees in Computer Science & Engineering. Dr. Goyal's research interests include Cloud Security, Image Processing, Data Analytics, and

Information Security, he has been pivotal in many turnkey projects & research and development activities. He has been pivotal in establishing advanced research labs—such as AICTE-sponsored IDEA Lab and Deep Learning Lab. He has been invited speaker and conference chair in various conferences organized globally in countries like China, Japan etc. His career includes successful organization of conferences, workshops, and faculty development programs, and he is an empanelled assessor for NAAC since 2021, contributing nationally to academic quality improvement. He has also received Grants-in-Aid for Research, Development, Conference & workshops, amounting more than Rs. 81 Lakh, from agencies like AICTE, TEQIP, ISTE etc, that include Combined Research Project, MODROBS, AICTE-IDEA Lab Dr. Goyal has published extensively and is known for driving innovation and outcome-based education in higher technical education. He has also completed his CMI level 5 Award in "Management and Leadership", under AICTE-UKIERI Technical Leadership Development Program in association with Dudley University, United Kingdom. He has 36 Full patents published 2 Granted & 1 Copyright under his name. He has successfully published 16 edited books with big publishing giants like Springer, Wiley, IGI Global, Apple Academic Press, Taylor & Francis and Eureka. He has published 5 SCI & 116 Scopus and 52 Web of Science indexed papers & is editor of 2 SCI & 5 Scopus Indexed Journals, special issues. He has also attended more than 25 International Conferences & has been invited speaker for more than 15 Conferences & Seminars. He is Senior Member of IEEE, life member of ISC, CSI, IETE & ISTE and fellow member of ACM.

Invited Speaker

Assoc. Prof. Xiaolong Hu

Zhejiang University, China

Speech Time: 16:10-16:30, June 13, 2026 (UTC+8)
Onsite Room: Conference Room 6&7 (2nd Floor)

Speech Title: LiDAR Imaging with High-Performance Fractal Superconducting Nanowire Single-Photon Detectors

Abstract: Traditional superconducting nanowire single-photon detectors (SNSPDs) with meander-nanowire structures can efficiently detect single photons in specific states of polarization while offering excellent timing performance and ultra-low dark count rates, making them widely used in quantum and classical faint-light detection. However, enabling SNSPDs to efficiently detect photons in arbitrary states of polarization while maintaining their other superior performance remains a challenge. To address this issue, we proposed the fractal SNSPDs capable of efficiently detecting incident photons in arbitrary states of polarization and developed: (1) a fractal SNSPD system coupled with single-mode fiber, featuring plug-and-play operation, achieving a system detection efficiency (SDE) of 91% for arbitrary polarization states at a wavelength of 1540 nm; (2) an 8-channel fractal SNSPD system operating in the 930–940 nm band, with an average SDE of 90% for arbitrary polarization states across the eight channels and a maximum SDE of 96% in the channel with the highest detection efficiency; (3) a fractal SNSPD with a large photosensitive area, coupled with a 50-μm core-diameter multimode fiber, achieving an SDE of 78% for arbitrary polarization states at a wavelength of 1530 nm. Using fractal SNSPDs, we demonstrated the following applications: (1) photon time-of-flight (ToF) LiDAR imaging, (2) full-Stokes LiDAR imaging, (3) non-line-of-sight imaging, (4) high-precision dual-comb ranging, and (5) three-dimensional single-pixel imaging.

Biography: Dr. Xiaolong Hu is a tenured associate professor at the College of Information and Electronic Engineering, Zhejiang University. His research focuses on micro- and nano-

optoelectronic devices and quantum photonic devices, and he has achieved a series of research results in the field of SNSPDs: (1) He proposed waveguide-integrated SNSPDs, which has now become the mainstream device structure for SNSPDs on integrated quantum optical chips; (2) He revealed two mechanisms of device timing jitter of SNSPDs and developed the device physics of SNSPDs; (3) He proposed and developed fractal SNSPDs with high system detection efficiency and low timing jitter for incident photons in all states of polarization, and developed practical systems that are fiber-coupled and plug-and-play; (4) He applied the fractal SNSPDs to dual-comb precision ranging, full-Stokes lidar imaging, infrared non-line-of-sight imaging, and three-dimensional single-pixel imaging. The fractal SNSPD systems have been used in many institutions such as Peking University, Tsinghua University, Zhejiang University, Sun Yat-sen University, Wuhan University of Technology, and Songshan Lake Materials Laboratory. The research results have been published in journals such as Nature, Nature Photonics, and Nature Nanotechnology. The fractal SNSPDs have received medium coverage from Nature Photonics (News & Views), the Optica, IEEE Photonics Society, IEEE Spectrum, and Science and Technology Daily; the fractal SNSPDs have won the golden medal at the 28th National Invention Exhibition, the first-grade prize of the Invention and Entrepreneurship Award of the China Association of Inventions, the gold medal at the 50th Geneva International Invention Exhibition, and a nomination for the Chip 10 Science Award in 2024. Dr. Hu is now serving as an associate editor of Optics Express.

Invited Speaker

Assoc. Prof. Chen Li

North China University of Technology, China

Speech Time: 10:30-10:50, June 14, 2026 (UTC+8)
Online Room: https://us02web.zoom.us/j/89264147904 (Password: 202606)

Speech Title: Computational Analysis of Histopathological Whole Slide Images for Multifaceted Tumor Assessment

Abstract: Whole slide pathological images (WSIs) contain abundant and distinctive tumor histomorphological information. Computational pathological analysis based on WSIs has emerged as a powerful and effective tool for auxiliary tumor diagnosis and precise clinical management. Our group has conducted solid research on WSI-based prognostic survival prediction for tumor patients.

In medical pathological analysis, prediction accuracy alone cannot fully meet clinical demands. Benefiting from the particularity of medical application scenarios, model interpretability is regarded as an essential guarantee for clinical transformation. Therefore, we are currently conducting targeted explorations on interpretable pathological prognostic modeling.

In addition, histological inference of gene mutation status and automatic identification of primary tumor origins based on WSI morphological features are recognized as promising and clinically valuable research directions in computational pathology. This series of studies aims to facilitate the development of accurate, credible and clinically translatable digital pathological diagnosis techniques.

Biography: Li Chen is an Associate Professor at North China University of Technology (NCUT). She was a Visiting Scholar at the Intelligent Systems and Vision Laboratory, University of California, Riverside (UCR).

She has been selected into the training programs of Top-notch Young Talents and Key Young Teachers sponsored by the Beijing Municipal Education Commission. Her main research

interests include Artificial Intelligence, Medicine-Engineering Integration, Image and Video Content Security, as well as multi-sensor information fusion and perception. She has presided over and participated in more than ten national and provincial/ministerial-level research projects. She has published dozens of academic achievements in authoritative journals and conferences including ASCO, and has won 4 provincial and ministerial-level research awards with 7 authorized national invention patents. She has been invited to deliver oral presentations at international academic conferences, and to serve as Session Chair and Local Chair of international conferences.

Invited Speaker

Senior Lecturer Anwaar Ulhaq

Central Queensland University, Australia

Speech Time: 09:50-10:10, June 14, 2026 (UTC+8)
Online Room: https://us02web.zoom.us/j/89264147904 (Password: 202606)

Speech Title: From Artificial to Real Intelligence

Abstract: Artificial intelligence can recognise patterns and learn from data, but it still lacks the flexibility and efficiency seen in biological systems. This talk explores how real intelligence can be understood by studying living neural structures. I will present recent work on brain organoids, neural imaging, and computational modelling, including our research on organoid mitosis analysis using computer vision. These studies show how biological systems organise, adapt, and learn in ways that current algorithms do not. By combining ideas from neuroscience, imaging, and machine learning, we can design models that are more robust and closer to real intelligence. The talk also outlines future research directions connecting computer vision, brain science, and neuromorphic computing.

Biography: Dr Anwaar Ulhaq is a Senior Lecturer in Artificial Intelligence at Central Queensland University's Sydney Campus, Australia. He holds a PhD in Artificial Intelligence from Monash University, Australia; a Graduate Certificate in Machine Learning and Artificial Intelligence from Massachusetts Institute of Technology, USA; and a professional certificate in Business Analytics from Harvard Business School. Additionally, he completed the Oxford Executive Leadership Program at the University of Oxford's Said Business School. He is the current President of the Australian Pattern Recognition Society and is a member of the Australian Academy of Sciences, the Australian Computer Society, and IEEE Signal Processing. With a diverse academic background, Dr Ulhaq has contributed significantly to the field of computer vision, serving as the General Chair of DICTA 2023, a major Australian computer vision conference. He is an associate editor of IEEE Transactions on Image Processing (TIP). He has published over 100 peer-reviewed journal and conference papers, received teaching and research excellence awards, and secured approximately $3.15 million

in research funding as a principal investigator and co-investigator. Dr Ulhaq's research interests span computer vision, image and signal processing, responsible AI, and remote sensing, and his work has been featured in national and international media more than 30 times, underscoring the relevance and significance of his contributions.

Invited Speaker

Assoc. Prof. Ts. Dr. Mas Rina Mustaffa
Universiti Putra Malaysia, Malaysia

Speech Time: 10:10-10:30, June 14, 2026 (UTC+8)
Online Room: https://us02web.zoom.us/j/89264147904 (Password: 202606)

Speech Title: From Fabric Patterns to Garment Visualisation Using Image Processing and Generative Modelling

Abstract: A considerable number of individuals rely on visual representation when making design-related decisions, creating a need for tools that can effectively preview how selected fabrics will appear when formed into garments. This study proposes a computational framework grounded in image processing and generative modelling to simulate the transformation of fabric patterns into traditional attire, with a focus on structured garment visualisation. The approach utilises binary silhouette mapping combined with morphological image processing techniques, including erosion and dilation, to generate refined two-dimensional garment representations. To enhance realism and allow flexible customisation, a diffusion-based inpainting model is incorporated to produce visually coherent and context-aware fabric patterns. The framework is implemented within a mobile-based prototype, enabling users to interactively apply and adjust fabric designs in near real time. Evaluation results indicate that the system performs effectively in terms of usability, functional capability, and user satisfaction. Overall, the proposed framework demonstrates the potential of integrating image processing techniques with generative modelling for practical garment visualisation, supporting more informed decision-making and improving the overall fabric selection experience.

Biography: Assoc. Prof. Ts. Dr. Mas Rina Mustaffa is an Associate Professor at the Multimedia Department, Faculty of Computer Science and Information Technology, Universiti Putra Malaysia (UPM), and a registered Professional Technologist (Ts.). She received her PhD in Multimedia Systems from Universiti Putra Malaysia in 2012. A Senior Member of IEEE, she specializes in computer vision, multimedia analytics, pattern

recognition, and multimedia information retrieval, with a focus on AI-driven solutions for education, agriculture, and intelligent systems.

She has led and collaborated on major national and international projects, including Malaysia's FRGS grants and the European Union's Horizon 2020 ULTRACEPT initiative, and recently completed a research secondment at the University of Leicester, UK (2024). Dr. Mas Rina has published extensively in high-impact journals and international conferences (IEEE, ACM, Springer) and is a recipient of multiple Best Paper and Presentation Awards. She is Editor-in-Chief of the Journal of Intelligent Media Computing, Vice President of PECAMP Malaysia (2025–2027), and currently serves as Publication Chair and Track Chair for several international conferences, advancing global research in intelligent multimedia representation and retrieval.

Invited Speaker

Assoc. Prof. Syed Farooq Ali

University of Management & Technology, Pakistan

Speech Time: 14:40-15:00, June 14, 2026 (UTC+8)
Online Room: https://us02web.zoom.us/j/89264147904 (Password: 202606)

Speech Title: A boosting framework for human posture recognition using spatio-temporal features along with radon transform

Abstract: Automatic human posture recognition in surveillance videos has real world applications in monitoring old-homes, restoration centers, hospitals, disability, and child-care centers. It also has applications in other areas such as security and surveillance, sports, and abnormal activity recognition. Human posture recognition is a challenging problem due to occlusion, background clutter, illumination variations, camouflage, and noise in the captured video signal. In the current study, which is an extension of our previous work (Ali et al. Sensors, 18(6):1918, 2018), we propose a novel combination of a number of spatio-temporal features computed over human blobs in a temporal window. These features include aspect ratios, shape descriptors, geometric centroids, ellipse axes ratio, silhouette angles, and silhouette speed. In addition to these features, we also exploit the radon transform to get better shape based analysis. In order to obtain improved posture classification accuracy, we used J48 classifier under a boosting framework by employing the AdaBoost algorithm. The proposed algorithm is compared with eighteen existing state-of-the-art approaches on four publicly available datasets including MCF, UR Fall detection, KARD, and NUCLA. Our results demonstrate the excellent performance of the proposed algorithm compared to these existing methods.

Biography: Dr. Syed Farooq Ali has around 18+ years of teaching and research experience. His areas of specialization include Computer Vision, Digital Image Processing, Medical Imaging, and Image & Video Coding. He did his BS from NUCES- FAST, Lahore, and later earned an MS degree from LUMS with Dean's Honor List and 4th position out of a batch of around 130 graduate students. During his stay in MS, he was on a LUMS fellowship. He also

completed his MS from Ohio State University, (Col.), USA. Moreover, he passed the Ph.D. Comprehensive Exam (Qualifier Exam) from Ohio State University, Col. USA. Later, he transferred his Ph.D. from Ohio State University USA to UMT and completed it.

During his graduate studies at LUMS and Ohio State University, he studied a total of 34 courses. Out of these, he studied ten courses related to his depth area (area of interest) that included Computer Vision, Medical Imaging, Image and Video Coding, Digital Image Processing, Artificial Intelligence, Pattern Recognition, Machine Learning, Computer Graphics, Signals and Systems, and Data Mining.

Currently, he is Associate Professor and Chair Vision and Image Processing Research Group at UMT. He also served two times as a Director Projects. Dr. Ali has around 39 publications including 23 journals, a book chapter, and 15 international conference papers. Dr. Ali is also a reviewer of many international conferences and journals. He also won 8 Fundings from National Grassroots ICT Research Initiative (NGIRI) for Final Year Projects.

Invited Speaker

Senior Lecturer Philippe Durand

Conservatoire National des Arts et Métiers, France

Speech Time: 13:50-14:10, June 13, 2026 (UTC+8)
Onsite Room: Conference Room 6&7 (2nd Floor)

Speech Title: Morphological and Topological Methods for Urban Extraction from Noisy Radar Imagery: Application to the City of Le Luc (France)

Abstract: This invited lecture presents a complete mathematical framework for extracting urban structures—built-up areas, residential blocks, and transport networks— from highly noisy radar imagery acquired over the city of Le Luc in South-Eastern France. Radar images of semi-urban environments are complex to interpret due to speckle noise, non-Gaussian backscatter distributions, and the multiple scattering mechanisms generated by buildings and roads. Our goal is to derive a coherent representation of the urban fabric comparable to the structures visible in the optical aerial photograph, while relying exclusively on the raw radar acquisition.

We develop an integrated approach combining two mathematical paradigms: (1) Mathematical Morphology, including Alternating Sequential Filters, granulometry, directional openings, and morphological skeletonization; and (2) Topological Data Analysis (TDA), based on persistent homology, H0–H1 generators, and Wasserstein distances. Morphology provides a multi-scale geometric description, capable of reducing speckle and isolating coherent radar responses associated with buildings and road segments. TDA offers complementary information on the global structure of urban textures, revealing stable topological patterns (loops, cavities, block organisation) which remain robust under noise and local fluctuations.

Applied to the Le Luc dataset, the combined methodology enables the recovery of major urban zones (historical nuclei VN1, VN2, VN3), pavillonnaire housing areas, dense HLM blocks, and the main transport axes (highway, major roads, railway). The morphological segmentation is strongly supported by the TDA signatures: urban regions exhibit a rich distribution of

persistent H1 generators, while background areas show simple connectivity profiles dominated by H0 components.

This talk highlights the synergy between geometry and topology for radar image analysis, demonstrating that the fusion of morphological filtering with TDA produces a robust, interpretable, and noise-tolerant pipeline for urban structure extraction, even under severe speckle conditions. Perspectives include extensions to multi-temporal radar sequences, polarimetric acquisitions, and integration into near-real-time environmental monitoring systems.

Biography: Philippe Durand is Senior Lecturer in the Mathematics and Statistics Department of the National Conservatory of Arts and Crafts in the Mathematical and Numerical Modeling Department (M2N), he works on the interaction between mathematical engineering and the theoretical tools of mathematics including usage has been increasing since the introduction of modern mathematics in the early sixties. He is interested in the mathematization of gauge theories in physics and string theory, he also works on tensor analysis applied to networks as well as the application of topological and statistical methods to image processing. In image processing, he used remote sensing images and especially radar images, he invested different methods of pattern recognition, and in particular the tools of mathematical morphology for the extraction of texture information. Currently I am focusing on the use of topological data analysis and different approaches to applying classical or quantum neural networks to image processing. He published his results in various journals of math-ematical engineering, and various proceedings of image processing conferences. Philippe Durand is assistant professor in Department of Mathematics (mod´elisation math´ematique et num´erique), Conservatoire National des Arts et M´etiers, 292 rue Saint Martin, 75141 Paris, FRANCE, (e-mail: philippe.durand@lecnam.net).

Invited Speaker

Asst. Prof. Zhennong Chen

Xi'an Jiaotong-Liverpool University, China

Speech Time: 15:00-15:20, June 14, 2026 (UTC+8)
Online Room: https://us02web.zoom.us/j/89264147904 (Password: 202606)

Speech Title: AI-empowered brain CT motion correction

Abstract: Motion artifacts remain one of the primary sources of image degradation in computed tomography (CT). Patient movement violates the static-object assumption underlying conventional reconstruction algorithms, leading to image blurring, distortion, and reduced diagnostic accuracy. To address this challenge, we have developed two complementary AI-based approaches that operate in the sinogram and image domains, respectively. In the sinogram domain, we designed a convolutional neural network (CNN) that takes partial-angle reconstructions (PARs) as input to estimate patient motion and perform motion compensation prior to final image reconstruction. In the image domain, we proposed a diffusion generative model incorporating histogram equalization and an elucidated diffusion framework for motion artifact correction. Furthermore, we conducted a human observer study on a real-world portable brain CT scan dataset to validate that the diffusion-based correction not only enhances image quality but also preserves lesion detectability.

Biography: Dr. Zhennong Chen is an Assistant Professor at the School of Artificial Intelligence and Advanced Computing (AIAC), Xi'an Jiaotong-Liverpool University. He received both his Ph.D. and Bachelor's degree in Bioengineering from the University of California, San Diego, and subsequently worked as a Postdoctoral Research Fellow at Harvard Medical School and Massachusetts General Hospital. Dr. Chen's research focuses on developing innovative AI methods for medical imaging, with particular emphasis on computed tomography (CT) and cardiovascular imaging. He has received the Trainee Research Prize at RSNA and the Best All-Conference Paper Finalist at SPIE Conference.

Invited Speaker

Asst. Prof. Zhi Lu

Tsinghua University, China

Speech Time: 15:20-15:40, June 14, 2026 (UTC+8)
Online Room: https://us02web.zoom.us/j/89264147904 (Password: 202606)

Speech Title: Leveraging spatial-angular redundancy for self-supervised denoising of 3D fluorescence imaging without temporal dependency

Abstract: Photon noise is a major bottleneck for extracting reliable biological information from fluorescence microscopy, especially when imaging fast and volumetric biological dynamics. However, most existing self-supervised denoising strategies depend on repeated measurements in time or strong spatial assumptions, which inevitably reduce temporal or spatial fidelity. Here we introduce LF-denoising, a transformer-based framework that exploits the rich spatial–angular redundancy encoded in light field measurements, enabling accurate denoising without requiring temporal repetition. Across simulations and diverse intravital experiments, LF-denoising robustly improves 3D imaging quality under extremely low excitation power, and generalizes across species including zebrafish, Drosophila and mice. By preserving the true temporal evolution of biological processes while substantially reducing noise, LF-denoising opens a path toward more reliable and accessible high-speed 3D imaging for quantitative biology, with broad applications.

Biography: Dr. Zhi Lu is currently an Assistant Professor at Tsinghua University. He received the B.S. and Ph.D. degrees in Control Science and Engineering from Tsinghua University, Beijing, China, in 2018 and 2023, followed by postdoctoral research from 2023 to 2025. His research interests include computational imaging and intelligent microscopy. In recent years, He has published papers in journals including Cell, Nature Methods, Nature Biotechnology, Nature Protocols and Nature Communications, with over twenty granted patents. He is honorably on Forbes China 30 Under 30 List (2023), and supported the National Postdoctoral Program for Innovative Talents (2023) and by the Young Top-notch Talent of National High-Level Talent Special Support Program (2024). Additionally, he received SAIL awards at the

World Artificial Intelligence Conference (2022, 2024), the Dimitris N. Chorafas Prize (2024), Best Paper Award in Computational Optics at IBCS (2023), and the Award in Ten Advances for Optics in China (2022).

Invited Speaker

Dr. Loc Nguyen
Sunflower Soft Company, Vietnam

Speech Time: 10:50-11:10, June 14, 2026 (UTC+8)
Online Room: https://us02web.zoom.us/j/89264147904 (Password: 202606)

Speech Title: Is matrix neural network the alternative of convolutional neural network?

Abstract: Currently, deep learning is the most important and popular methodology in artificial intelligence (AI) and artificial neural network (ANN) is the foundation of deep learning. The main drawback of ANN is the boom problem of a huge number of parametric weights when ANN in deep learning establishes a large number of hidden layers. The excellent solution for image processing within context of deep learning is convolutional neural network (CNN) equipped filtering kernel. Another solution of the boom problem is that large parametric weight vector is organized as matrix, which leads to a so-called matrix neural network (MNN). Computation cost of MNN is decreased significantly in comparison with ANN but it is necessary to test the main hypothesis "whether MNN is the alternative of CNN". Moreover, transformer which is the new trend in AI and deep learning, which aims to improve/replace traditional ANN by self-supervised learning, in which attention is the significant mechanism of self-supervised learning. Therefore, the implicit deep meanings of attention and filtering kernel are similar, which represents feature of data, which does not go beyond parametric weights too. In general, the research has two goals: 1) explaining and implementing ANN, CNN, and transformer (attention) and 2) applying analysis of variance (ANOVA) into evaluating the effectiveness of ANN, CNN, and transformer (attention) within context of image classification. The ultimate result is that it is not asserted that MNN is the alternative of CNN but MNN can be an optional choice for implementing ANN instead of focusing on the unique CNN solution. Moreover, the incorporation of MNN and attention in implementing transformer produces a compromising solution of high performance and computational cost.

Biography: Loc Nguyen is an independent scholar from 2017. He holds Master degree in Computer Science from University of Science, Vietnam in 2005. He holds PhD degree in Computer Science and Education at Ho Chi Minh University of Science in 2009. His PhD

dissertation was honored by World Engineering Education Forum (WEEF) and awarded by Standard Scientific Research and Essays as excellent PhD dissertation in 2014. He holds Postdoctoral degree in Computer Science from 2013, certified by Institute for Systems and Technologies of Information, Control and Communication (INSTICC) by 2015. Now he is interested in poetry, computer science, statistics, mathematics, education, and medicine. He serves as reviewer, editor, speaker, and lecturer in a wide range of international journals and conferences from 2014. He is volunteer of Statistics Without Borders from 2015. He was granted as Mathematician by London Mathematical Society for Postdoctoral research in Mathematics from 2016. He is awarded as Professor by Scientific Advances and Science Publishing Group from 2016. He was awarded Doctorate of Statistical Medicine by Ho Chi Minh City Society for Reproductive Medicine (HOSREM) from 2016. He was awarded and glorified as contributive scientist by International Cross-cultural Exchange and Professional Development-Thailand (ICEPD-Thailand) from 2021 and by Eudoxia Research University USA (ERU) and Eudoxia Research Centre India (ERC) from 2022. He has published 101 papers and preprints in journals, books, conference proceedings, and preprint services. He is author of 5 scientific books. He is author and creator of 10 scientific and technological products.

Onsite Breakout Sessions

Onsite Breakout Session 1

Multimodal Intelligent Image Fusion Analysis and Advanced Processing Methods

Session Chair: Prof. Xibin Jia, Beijing University of Technology, China

Time: 13:30-15:55, June 13, 2026 (UTC+8) **Venue: Conference Room 6&7 (2nd Floor)**	
*Presenters are recommended to enter the meeting room 10 mins in advance. **Presenters are recommended to stay for the whole session in case of any absence. ***After the session, there will be a group photo for all presenters in this session.	
Invited Speech 13:30-13:50	**Speech Title:** Clinically Oriented Medical Image Analysis **Invited Speaker:** Prof. Xibin Jia, Beijing University of Technology, China
Invited Speech 13:50-14:10	**Speech Title:** Morphological and Topological Methods for Urban Extraction from Noisy Radar Imagery: Application to the City of Le Luc (France) **Invited Speaker:** Senior Lecturer Philippe Durand, Conservatoire National des Arts et Métiers, France
CG103 14:10-14:25	Reversible Data Hiding in Shared JPEG Images for Distributed Framework **Author(s):** Shengyi Cao, Anqi Qiu, Yingchao Yang, Lingchen Gu and Wenbo Wan **Presenter:** Anqi Qiu, Shandong Normal University, China
CG109 14:25-14:40	Structural and Appearance Regularization for Robust Retinal Vessel Segmentation **Author(s):** Jun Fang and Hui Ma **Presenter:** Hui Ma, Heilongjiang University, China
CG113 14:40-14:55	No-Reference Image Quality Assessment via Multi-Head Cross-Scale Attention Fusion and Contrastive Ranking Loss **Author(s):** Shengyu Pei, Yule An, Sisi Fan and Huan Lao **Presenter:** Shengyu Pei, Guangxi Minzu University, China
CG114 14:55-15:10	Encrypted Deep Robust Reversible Watermarking With Preset Private Key **Author(s):** Zixuan Zhang, Chunqiang Yu, Xianquan Zhang and Zhenjun Tang **Presenter:** Zixuan Zhang, Guangxi Normal University, China
CG119 15:10-15:25	MSHA-Net: Learning with Multi-Scale Hierarchical Discriminative Attention for Few-Shot Medical Segmentation **Author(s):** Yandong Fang, Xibin Jia, Hui Xu, Zeyu Xu, Chao Fan, Senhui Jia and Luo Wang **Presenter:** Luo Wang, Beijing University of Technology, China
CG120 15:25-15:40	Aesthetic Evaluation Network with Refined Style Attributes for Art Images **Author(s):** Dayu Yang, Yihua Chen, Kejing Wu, Feijian Huang, Yujian Gan and Zhenjun Tang **Presenter:** Dayu Yang, Guangxi Normal University, China

CG123 15:40-15:55	Diagnosis of Alzheimer's Disease Based on 3D MRI Features and Large Language Model Embeddings **Author(s):** Huan Lao, Sisi Fan, Yule An, Shengyu Pei and Di Zhao **Presenter:** Huan Lao, Guangxi Minzu University, China

Onsite Breakout Sessions

Onsite Breakout Session 2

High-Resolution Optical and Digital Imaging Methods

Session Chair: Assoc. Prof. Xiaolong Hu, Zhejiang University, China

Time: 16:10-18:05, June 13, 2026 (UTC+8) **Venue: Conference Room 6&7 (2nd Floor)**	
*Presenters are recommended to enter the meeting room 10 mins in advance. **Presenters are recommended to stay for the whole session in case of any absence. ***After the session, there will be a group photo for all presenters in this session.	
Invited Speech 16:10-16:30	**Speech Title:** LiDAR Imaging with High-Performance Fractal Superconducting Nanowire Single-Photon Detectors **Invited Speaker:** Assoc. Prof. Xiaolong Hu, Zhejiang University, China
Invited Speech 16:30-16:50	**Speech Title:** Thermal Image Detection and Tracking with a Multi-rotor Drone **Invited Speaker:** Prof. Seokwon Yeom, Daegu University, Korea
CG107 16:50-17:05	Self-supervised denoising of light field microscopy for high-dimensional neural imaging **Author(s):** Wentao Chen and Zhi Lu **Presenter:** Wentao Chen, Fudan University, China
CG115 17:05-17:20	Analysis of Poisson processes in quantum random number generator chip based on CMOS image sensor **Author(s):** Xiaoyuan Yu, Zaiguang Li, Xiao Shi, Yuyang Pei and Congjun Xia **Presenter:** Xiaoyuan Yu, Xianhai (Shanghai) Quantum Tech. Co., Ltd., China
CG117 17:20-17:35	A Structured Light Simulation Platform for 3D Measurement with Realistic Optics **Author(s):** Rui Ma, Qihe Huang, Shenglian Lu, Shijie Li, Ming Chen and Fengli Lu **Presenters:** Rui Ma, Guangxi Normal University, China Qihe Huang, Guangxi Normal University, China
CG118 17:35-17:50	Phase Compensation Method for Multi-Look Coherent Processing in Wide-Area Squint SAR **Author(s):** Shize Shang, Yang Zhou, Kesai Ouyang, Qiang Cheng, Yuhao Yang and Pin Li **Presenter:** Shize Shang, Nanjing Research Institute of Electronics Technology, China
CG121 17:50-18:05	LBF-Net: Lightweight Bidirectional Fusion for Hybrid Voxel-Point Place Recognition **Author(s):** Lineng Chen, Haiying Xia and Shuxiang Song **Presenter:** Lineng Chen, Guangxi Normal University, China

Online Breakout Sessions

Online Breakout Session 1

AI-Driven Digital Image Analysis and Processing Technologies

Session Chair: Prof. Xiwen Zhang, Beijing Language and Culture University, China

Time: 09:30-11:40, June 14, 2026 (UTC+8) **ZOOM ID: 89264147904** **Online Room: https://us02web.zoom.us/j/89264147904 (Password: 202606)**	
*Presenters are recommended to enter the meeting room 10 minutes in advance. **Presenters are recommended to stay for the whole session in case of any absence. ***After the session, there will be a group photo for all presenters in this session.	
Invited Speech 09:30-09:50	**Speech Title:** Three views on intelligently extracting and generating information from image **Invited Speaker:** Prof. Xiwen Zhang, Beijing Language and Culture University, China
Invited Speech 09:50-10:10	**Speech Title:** From Artificial to Real Intelligence **Invited Speaker:** Senior Lecturer Anwaar Ulhaq, Central Queensland University, Australia
Invited Speech 10:10-10:30	**Speech Title:** From Fabric Patterns to Garment Visualisation Using Image Processing and Generative Modelling **Invited Speaker:** Assoc. Prof. Mas Rina Mustaffa, Universiti Putra Malaysia, Malaysia
Invited Speech 10:30-10:50	**Speech Title:** Computational Analysis of Histopathological Whole Slide Images for Multifaceted Tumor Assessment **Invited Speaker:** Assoc. Prof. Chen Li, North China University of Technology, China
Invited Speech 10:50-11:10	**Speech Title:** Is matrix neural network the alternative of convolutional neural network? **Invited Speaker:** Dr. Loc Nguyen, Sunflower Soft Company, Vietnam
CG105 11:10-11:25	A Neighborhood Prediction-Based Quality Enhancement Method for CRT Progressive Image Reconstruction **Author(s):** Zhengtian Zhu, Dan Tang and Hongliang Cai **Presenter:** Zhengtian Zhu, Chengdu University of Information Technology, China
CG112 11:25-11:40	Full-Reference Image Quality Assessment Based on LS Convolution **Author(s):** Huan Lao, Sisi Fan, Yule An and Shengyu Pei **Presenter:** Sisi Fan, Guangxi Minzu University, China

Online Breakout Sessions

Online Breakout Session 2

Intelligent Image Analysis, Secure Computing, and Privacy Protection

Session Chair: Prof. Suraiya Jabin, Jamia Millia Islamia, India

Time: 14:00-16:25, June 14, 2026 (UTC+8) **ZOOM ID: 89264147904** **Online Room: https://us02web.zoom.us/j/89264147904 (Password: 202606)**	
*Presenters are recommended to enter the meeting room 10 minutes in advance. **Presenters are recommended to stay for the whole session in case of any absence. ***After the session, there will be a group photo for all presenters in this session.	
Invited Speech 14:00-14:20	**Speech Title:** AI-based Adult Age Threshold Modeling using Orthopantomographs for Forensic and Judicial Applications **Invited Speaker:** Prof. Suraiya Jabin, Jamia Millia Islamia, India
Invited Speech 14:20-14:40	**Speech Title:** AI-Enabled Image Analysis Framework for Healthcare Applications **Invited Speaker:** Prof. Dinesh Goyal, Poornima Institute of Engineering & Technology, India
Invited Speech 14:40-15:00	**Speech Title:** A boosting framework for human posture recognition using spatio-temporal features along with radon transform **Invited Speaker:** Assoc. Prof. Syed Farooq Ali, University of Management & Technology, Pakistan
Invited Speech 15:00-15:20	**Speech Title:** AI-empowered brain CT motion correction **Invited Speaker:** Asst. Prof. Zhennong Chen, Xi'an Jiaotong-Liverpool University, China
Invited Speech 15:20-15:40	**Speech Title:** Leveraging spatial-angular redundancy for self-supervised denoising of 3D fluorescence imaging without temporal dependency **Invited Speaker:** Asst. Prof. Zhi Lu, Tsinghua University, China
CG104 15:40-15:55	Multi-Party Reversible Data Hiding in Encrypted Image Based on Secret Sharing over Dual Galois Fields **Author(s):** Bing Chen, Yulong Zhou, Kai Cai, Jingqi Luo and Xiaotian Wu **Presenter:** Yulong Zhou, Guangdong Polytechnic Normal University, China
CG201-A 15:55-16:10	Processing Perfect Images with Obsolete Standards: The Temporal Crisis in Vision AI **Author(s):** Himanshu Goel **Presenter:** Himanshu Goel, Independent AI Researcher, USA
CG110 16:10-16:25	Lens Defect Detection Based on Feature Enhancement and Improved RT-DETR **Author(s):** Yifu Che, Zengzhen Mi and Yuchengzhi Lei **Presenter:** Yifu Chen, Chongqing University of Technology, China

Guilin Attractions

Li River

Originating from the "First Peak of South China" — the Yuecheng Ridge in northern Guangxi — the Li River is a national 5A-rated scenic area and a key national tourist attraction. It is renowned for its "verdant mountains, crystal-clear waters, fantastic caves, and stunning rocks", as well as its "green islets, perilous shoals, deep pools, and cascading waterfalls". It is often said, "A hundred miles of the Li River unfolds a hundred miles of gallery".

.

Seven-Star Park

Located on the east bank of the Li River in Guilin, the park is named after the Seven-Star Hill and its Seven-Star Cave. As the largest comprehensive ecological scenic area in Guilin, it covers approximately 134.7 hectares and features verdant mountains, crystal-clear waters, fantastic caves, and stunning rocks. It has been a renowned tourist destination since the Sui and Tang dynasties

Elephant Trunk Hill

Originating from the "First Peak of South China" — the Yuecheng Ridge in northern Guangxi — the Li River is a national 5A-rated scenic area and a key national tourist attraction. It is renowned for its "verdant mountains, crystal-clear waters, fantastic caves, and stunning rocks", as well as its "green islets, perilous shoals, deep pools, and cascading waterfalls". It is often said, "A hundred miles of the Li River unfolds a hundred miles of gallery".

Memo