

# **2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW 2025)**

**Nashville, Tennessee, USA  
11-12 June 2025**

**Pages 1-603**



**IEEE Catalog Number: CFP2588A-POD  
ISBN: 979-8-3315-9995-9**

**Copyright © 2025 by the Institute of Electrical and Electronics Engineers, Inc.  
All Rights Reserved**

*Copyright and Reprint Permissions:* Abstracting is permitted with credit to the source. Libraries are permitted to photocopy beyond the limit of U.S. copyright law for private use of patrons those articles in this volume that carry a code at the bottom of the first page, provided the per-copy fee indicated in the code is paid through Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923.

For other copying, reprint or republication permission, write to IEEE Copyrights Manager, IEEE Service Center, 445 Hoes Lane, Piscataway, NJ 08854. All rights reserved.

***\*\*\* This is a print representation of what appears in the IEEE Digital Library. Some format issues inherent in the e-media version may also appear in this print version.***

IEEE Catalog Number:	CFP2588A-POD
ISBN (Print-On-Demand):	979-8-3315-9995-9
ISBN (Online):	979-8-3315-9994-2
ISSN:	2160-7508

**Additional Copies of This Publication Are Available From:**

Curran Associates, Inc  
57 Morehouse Lane  
Red Hook, NY 12571 USA  
Phone: (845) 758-0400  
Fax: (845) 758-2633  
E-mail: [curran@proceedings.com](mailto:curran@proceedings.com)  
Web: [www.proceedings.com](http://www.proceedings.com)

CURRAN ASSOCIATES INC.  
**proceedings**  
.com

# 2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW) **CVPRW 2025**

## Table of Contents

### LatinX in Computer Vision Research Workshop

NExNet Seg: Neuron Expansion Network for Medical Image Segmentation .....	1
<i>Abel Reyes-Angulo (Michigan Technological University) and Sidike Paheding (Fairfield University)</i>	
HeAL3D: Heuristical-enhanced Active Learning for 3D Object Detection .....	11
<i>Esteban Rivera (TUMünchen), Surya Prabhakaran (TUMünchen), and Markus Lienkamp (TUMünchen)</i>	
Scaling laws in zero-shot gender classification using CLIP .....	21
<i>Lucas M. Ceschini (Federal University of Rio Grande do Sul, Brazil), Gabriel O. Ramos (Universidade do Vale do Rio dos Sinos, Brazil), and Claudio R. Jung (Federal University of Rio Grande do Sul, Brazil)</i>	
Slot Attention-based Feature Filtering for Few-Shot Learning .....	30
<i>Javier Ródenas Cumplido (Universitat de Barcelona), Eduardo Aguilar (Universitat de Barcelona), and Petia Radeva (Universitat de Barcelona)</i>	
Emotions in LatAm: A new dataset and benchmark for emotion recognition in Latin America .....	41
<i>Pooja Kishore Kumar (University of Twente), Willams de Lima Costa (Voxar Labs, Centro de Informática, Universidade Federal de Pernambuco), Renato Nogueira Ferraz e Oliveira (Voxar Labs, Centro de Informática, Universidade Federal de Pernambuco), Veronica Teichrieb (Voxar Labs, Centro de Informática, Universidade Federal de Pernambuco), and Estefania Talavera Martinez (University of Twente)</i>	
Towards Faster and More Compact Foundation Models for Molecular Property Prediction .....	48
<i>Yasir Ghunaim (KAUST), Andrés Villa (KAUST), Gergo Ignacz (KAUST), Gyorgy Szekely (KAUST), Motasem Alfarra (KAUST), and Bernard Ghanem (KAUST)</i>	
Enhancing Vision Transformer Explainability Using Artificial Astrocytes .....	58
<i>Nicolas Echevarrieta-Catalan (University of Miami, USA), Ana Ribas-Rodriguez (Universidade da Coruna, Spain), Francisco Cedron (Universidade da Coruna, Spain), Odelia Schwartz (University of Miami, USA), and Vanessa Aguiar-Pulido (University of Miami, USA)</i>	

PineSORT: A Simple Online Real-time Tracking Framework for Drone Videos in Agriculture .....	65
<i>Danny Xie-Li (Instituto Tecnológico de Costa Rica) and Fabian Fallas-Moya (Universidad de Costa Rica)</i>	

## 8th Multimodal Learning and Applications Workshop

Exploring Missing Modality in Multimodal Egocentric Datasets .....	75
<i>Merey Ramazanova (KAUST), Alejandro Pardo (KAUST), Humam Alwassel (KAUST), and Bernard Ghanem (KAUST)</i>	
QID: Efficient Query-Informed ViTs in Data-Scarce Regimes for OCR-free Document Understanding .....	86
<i>Binh Le (Sungkyunkwan University), Shaoyuan Xu (Amazon), Jinmiao Fu (Amazon), Zhishen Huang (Amazon), Moyan Li (Amazon), Yanhui Guo (Amazon), Hongdong Li (Amazon, Australian National University), Sameera Ramasinghe (Pluralis Research), and Bryan Wang (Amazon)</i>	
Pose-Aware Weakly-Supervised Action Segmentation .....	97
<i>Zhihao Zhao (University of California, Berkeley; Honda Research Institute USA), Reza Ghoddoosian (Honda Research Institute USA), Isht Dwivedi (Honda Research Institute USA), Nakul Agarwal (Honda Research Institute USA), and Behzad Dariush (Honda Research Institute USA)</i>	
Location-Free Scene Graph Generation .....	108
<i>Ege Özsoy (TU Munich), Felix Holm (TU Munich), Chantal Pellegrini (TU Munich), Tobias Czempel (TU Munich), Mahdi Saleh (TU Munich), Nassir Navab (TU Munich), and Benjamin Busam (TU Munich)</i>	
SplatTouch: Explicit 3D Representation Binding Vision and Touch .....	118
<i>Antonio Luigi Stefani (University of Trento, Italy), Niccolò Bisagno (University of Trento, Italy), Nicola Conci (University of Trento, Italy), and Francesco De Natale (University of Trento, Italy)</i>	
Online Gaussian Test-Time Adaptation of Vision-Language Models. ....	128
<i>Clément Fuchs (UCLouvain), Maxime Zanella (UCLouvain, UMONS), and Christophe De Vleeschouwer (UCLouvain)</i>	
ICT-QA: Question Answering over Multi-modal Contexts including Image, Chart, and Text Modalities .....	138
<i>Youngrok Jang (LG AI Research), Hyesoo Kong (LG AI Research), Gyeonghun Kim (LG AI Research), Yejin Lee (LG AI Research), Jungkyu Choi (LG AI Research), and Kyunghoon Bae (LG AI Research)</i>	
Vocabulary-free few-shot learning for vision-language models .....	149
<i>Maxime Zanella (UCLouvain), Clément Fuchs (UCLouvain), Ismail Ben Ayed (ETS Montréal), and Christophe De Vleeschouwer (UCLouvain)</i>	
TLAC: Two-stage LMM Augmented CLIP for Zero-Shot Classification .....	159
<i>Ans Munir (Information Technology University, Pakistan), Faisal Z. Qureshi (University of Ontario Institute of Technology, Canada), Muhammad Haris Khan (Mohamed Bin Zayed University of Artificial Intelligence, UAE), and Mohsen Ali (Information Technology University, Pakistan)</i>	

TRISHUL: Towards Region Identification and Screen Hierarchy Understanding for Large VLM based GUI Agents .....	170
<i>Kunal Singh (Fractal AI, India ), Shreyas Singh (Fractal AI, India ), and Mukund Khanna (Fractal AI, India )</i>	
MVCM: Enhancing Multi-View and Cross-Modality Alignment for Medical Visual Question Answering and Medical Image-Text Retrieval .....	180
<i>Yuanhao Zou (UNIVERSITY OF MICHIGAN- ANN ARBOR) and Zhaozheng Yin (State University of New York, Stony Brook)</i>	
Multimodal Rationales for Explainable Visual Question Answering .....	191
<i>Kun Li (University of Twente, The Netherlands), George Vosselman (University of Twente, The Netherlands), and Michael Ying Yang (University of Bath, UK)</i>	
Improving multimodal hateful meme detection exploiting LMM-generated knowledge .....	202
<i>Maria Tzelepi (ITI-CERTH) and Vasileios Mezaris (ITI-CERTH)</i>	
Transformer-Based Lung Infection Severity Prediction with Cross Attention and Conditional TransMix Augmentation .....	212
<i>Bouthaina Slika (University of the Basque Country), Fadi Dornaika (IKERBASQUE, Basque Foundation for Science), Fares Bougourzi (Junia, UMR 8520, CNRS), and Karim Hammoudi (Université de Haute-Alsace, IRIMAS)</i>	
Skin Lesion Classification Using Dermoscopic Images and Clinical Metadata: Insights from Multimodal Models .....	222
<i>Sakib Ahammed (Manchester Metropolitan University), Xia Cui (Manchester Metropolitan University), Wenqi Lu (Manchester Metropolitan University), and Moi Hoon Yap (Manchester Metropolitan University)</i>	
LVP-CLIP: Revisiting CLIP for Continual Learning with Label Vector Pool .....	231
<i>Yue Ma (Syracuse University), Huantao Ren (Syracuse University), Boyu Wang (Syracuse University), Jingang Jin (Syracuse University), Senem Velipasalar (Syracuse University), and Qinru Qiu (Syracuse University)</i>	
Compositional Image-Text Matching and Retrieval by Grounding Entities .....	241
<i>Madhukar Reddy Vongala (George Mason University), Saurabh Srivastava (George Mason University), and Jana Kosecka (George Mason University)</i>	

## 4th edition of Computer Vision for Metaverse Workshop

Direction-Aware Hybrid Representation Learning for 3D Hand Pose and Shape Estimation .....	251
<i>Shiyong Liu (Huawei Noah's Ark Lab, China), Zhihao Li (Huawei Noah's Ark Lab, China), Xiao Tang (Huawei Noah's Ark Lab, China), and Jianzhuang Liu (Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, China)</i>	
What Makes for a Good Stereoscopic Image? .....	261
<i>Netanel Tamir (Apple), Shir Amir (Apple), Ranel Itzhaky (Apple), Noam Atia (Apple), Shobhita Sundaram (MIT), Stephanie Fu ( ), Ron Sokolovsky (Apple), Phillip Isola (MIT), Tali Dekel (weizmann institute), Richard Zhang (Adobe), and Miriam Farber (Apple)</i>	

ARC-NeRF: Area Ray Casting for Broader Unseen View Coverage in Few-shot Object Rendering ..	273
<i>Seunghyeon Seo (Seoul National University, South Korea), Yeonjin Chang (Seoul National University, South Korea), Jayeon Yoo (Seoul National University, South Korea), Seungwoo Lee (Seoul National University, South Korea), Hojun Lee (Seoul National University, South Korea), and Nojun Kwak (Seoul National University, South Korea)</i>	
OccludeNeRF: Geometry-aware 3D Scene Inpainting with Collaborative Score Distillation in NeRF .....	284
<i>Jingyu Shi (Purdue University), Achleshwar Luthra (Futurewei Technologies Inc.), Jiazhi Li (Futurewei Technologies Inc.), Xiang Gao (Futurewei Technologies Inc.), Xiyun Song (Futurewei Technologies Inc.), Zongfang Lin (Futurewei Technologies Inc.), David Gu (Stonybrook University), and Heather Yu (Futurewei Technologies Inc.)</i>	
PanoDreamer: Consistent Text to 360-Degree Scene Generation .....	295
<i>Zhexiao Xiong (Washington University in St. Louis), Zhang Chen (OPPO US Research Center), Zhong Li (OPPO US Research Center), Yi Xu (OPPO US Research Center), and Nathan Jacobs (Washington University in St. Louis)</i>	
SplatMesh: Interactive 3D Segmentation and Editing Using Mesh-Based Gaussian Splatting .....	305
<i>Kaichen Zhou (University of Oxford), Lanqing Hong (Huawei Noah's Ark Lab), Xinhai Chang (Peking University), Yingji Zhong (Hong Kong University of Science and Technology (HKUST)), Enze Xie (Huawei Noah's Ark Lab), Hao Dong (Peking University), Zhihao Li (Huawei Noah's Ark Lab), Yongxin Yang (Queen Mary University of London), Zhenguo Li (Huawei Noah's Ark Lab), and Wei Zhang (Huawei Noah's Ark Lab)</i>	
PluckerRF: A Line-based 3D Representation for Few-view Reconstruction .....	317
<i>Sam Bahrami (The Australian National University) and Dylan Campbell (The Australian National University)</i>	
ePBR: Extended PBR Materials in Image Synthesis .....	327
<i>Yu Guo (Futurewei Technologies, US), Zhiqiang Lao (Futurewei Technologies, US), Xiyun Song (Futurewei Technologies, US), Yubin Zhou (Futurewei Technologies, US), Zongfang Lin (Futurewei Technologies, US), and Heather Yu (Futurewei Technologies, US)</i>	
FaceGest: A Comprehensive Facial Gesture Dataset for Human-Computer Interaction .....	337
<i>Yaseen -- (Department of Electronics Engineering, Sejong University, Seoul) and Sonain Jamil (Department of Computer Science, University of Eastern Finland (UEF), Joensuu)</i>	
HumMorph: Generalized Dynamic Human Neural Fields from Few Views .....	348
<i>Jakub Zdrožný (University of Edinburgh, United Kingdom) and Hakan Bilen (University of Edinburgh, United Kingdom)</i>	
IL-NeRF: Incremental Learning for Neural Radiance Fields with Camera Pose Alignment .....	358
<i>Letian Zhang (Middle Tennessee State University), Ming Li (University of Central Florida), Chen Chen (University of Central Florida), and Jie Xu (University of Florida)</i>	

FreBIS: Frequency-Based Stratification for Neural Implicit Surface Representations .....	369
<i>Naoko Sawada (Mitsubishi Electric Research Laboratories, USA), Pedro Miraldo (Mitsubishi Electric Research Laboratories, USA), Suhas Lohit (Mitsubishi Electric Research Laboratories, USA), Tim K. Marks (Mitsubishi Electric Research Laboratories, USA), and Moitreya Chatterjee (Mitsubishi Electric Research Laboratories, USA)</i>	
DeclutterNeRF: Generative-Free 3D Scene Recovery for Occlusion Removal .....	380
<i>Wanzhou Liu (Washington University in St. Louis, USA), Zhexiao Xiong (Washington University in St. Louis, USA), Xinyu Li (Georgia Institute of Technology, USA), and Nathan Jacobs (Washington University in St. Louis, USA)</i>	

## 2nd MetaFood Workshop

Synthetic Data Augmentation using Pre-trained Diffusion Models for Long-tailed Food Image Classification .....	391
<i>GaYeon Koh (Korea University), Hyun-Jic Oh (Korea University), Jeonghyun Noh (Korea University), and Won-Ki Jeong (Korea University)</i>	
Extra-Lightweight AI-Based Privacy Preserving Framework for Egocentric Wearable Cameras .....	401
<i>Long Li (University of Alabama, U.S.A), Fengqing Zhu (Purdue University, U.S.A), Heather Eicher-Miller (Purdue University, U.S.A), J. Graham Thomas (Brown University, U.S.A), Yuning Huang (Purdue University, U.S.A), and Edward Sazonov (University of Alabama, U.S.A)</i>	
Privacy Preserving Ordinal-Meta Learning with VLMs for Fine-Grained Fruit Quality Prediction .....	411
<i>Riddhi Jain (TCS-Research), Manasi Patwardhan (TCS-Research), Aayush Mishra (TCS-Research), Parijat Deshpande (TCS-Research), and Beena Rai (TCS-Research)</i>	
Food Degradation Analysis Using Multimodal Fuzzy Clustering .....	420
<i>Julio Valdes (National Research Council Canada), Stephie Liu (National Research Council Canada), Shawn Yang (National Research Council Canada), Yuhao Chen (University of Waterloo), Alexander Wong (University of Waterloo), and Pengcheng Xi (National Research Council Canada)</i>	
Are Vision-Language Models Ready for Dietary Assessment? Exploring the Next Frontier in AI-Powered Food Image Recognition .....	430
<i>Sergio Romero-Tapiador (UNIVERSIDAD AUTONOMA DE MADRID), Ruben Tolosana (UNIVERSIDAD AUTONOMA DE MADRID), Blanca Lacruz-Pleguezuelos (IMDEA Food), Laura Judith Marcos-Zambrano (IMDEA Food), Guadalupe X. Bazán (IMDEA Food), Isabel Espinosa-Salinas (IMDEA Food), Julian Fierrez (UNIVERSIDAD AUTONOMA DE MADRID), Javier Ortega-Garcia (UNIVERSIDAD AUTONOMA DE MADRID), Enrique Carrillo de Santa Pau (IMDEA Food), and Aythami Morales (UNIVERSIDAD AUTONOMA DE MADRID)</i>	
Stochastic-based Patch Filtering for Few-Shot Learning .....	440
<i>Javier Ródenas Cumplido (Universitat de Barcelona), Eduardo Aguilar (Universitat de Barcelona), and Petia Radeva (Universitat de Barcelona)</i>	

VolTex: Food Volume Estimation using Text-Guided Segmentation and Neural Surface Reconstruction .....	450
<i>Ahmad AlMughrabi (University of Barcelona), Umair Haroon (University of Barcelona), Ricardo Marques (Pompeu Fabra University), and Petia Radeva (University of Barcelona)</i>	
FoodVideoQA: A Novel Baseline Framework for Dietary Monitoring .....	458
<i>Krish Shah (University of Waterloo), Siddharth Viswanath (University of Waterloo), Pengcheng Xi (National Research Council Canada), Alexander Wong (University of Waterloo), and Yuhao Chen (University of Waterloo)</i>	
SAMJAM: Zero-Shot Video Scene Graph Generation for Egocentric Kitchen Videos .....	467
<i>Joshua Li (University of Waterloo, Canada), Fernando Jose Pena Cantu (University of Waterloo, Canada), Emily Yu (University of Waterloo, Canada), Alexander Wong (University of Waterloo, Canada), Yuchen Cui (University of California, Los Angeles, USA), and Yuhao Chen (University of Waterloo, Canada)</i>	
Agro-Net: A Convolution-Attention Fusion based hyperspectral model for agro-food quality assessment .....	474
<i>Ocean Monjur (University of Illinois Urbana Champaign), Md. Toukir Ahmed (University of Illinois Urbana Champaign), Md Wadud Ahmed (University of Illinois Urbana Champaign), and Mohammed Kamruzzaman (University of Illinois Urbana Champaign)</i>	
Decomposing Food Images for Better Nutrition Analysis: A Nutritionist-Inspired Two-Step Multimodal LLM Approach .....	482
<i>Pitikorn Khlaisamniang (Artificial Intelligence Association of Thailand), Kun Kerdthaisong (Faculty of Engineering, Thammasat School of Engineering, Thammasat University), Supasate Vorathammathorn (Department of Computer Engineering, Faculty of Engineering, King Mongkut's University of Technology Thonburi), Nuchanon Yongsatianchot (Faculty of Engineering, Thammasat School of Engineering, Thammasat University), Hirunkul Phimsiri (Computer Engineering and Digital Technology Department of Computer Engineering, Faculty of Engineering, Chulalongkorn University, PreceptorAI team, CARIVA), Amrest Chinkamol (Vidyasirimedhi Institute of Science and Technology, PreceptorAI team, CARIVA), Teermade Thitseesaeng (National Health Security Office (NHSO)), Kanyakorn Veerakanjana (Faculty of Medicine Siriraj Hospital, Mahidol University, PreceptorAI team, CARIVA), Kaisorn Kachai (PreceptorAI team, CARIVA), Piyalitt Ittichaiwong (Faculty of Medicine Siriraj Hospital, Mahidol University, PreceptorAI team, CARIVA), and Tossaporn Saengja (Faculty of Medicine Siriraj Hospital, Mahidol University, PreceptorAI team, CARIVA)</i>	

# Expanding Horizons in AI Benchmarking: Multimodal Approaches

Behind the Magic, MERLIM: Multi-modal Evaluation Benchmark for Large Image-Language Models... 492

*Andrés Villa (King Abdullah University of Science and Technology (KAUST)), Juan León (King Abdullah University of Science and Technology (KAUST)), Alvaro Soto (Pontificia Universidad Católica de Chile), and Bernard Ghanem (King Abdullah University of Science and Technology (KAUST))*

Beyond Raw Videos: Understanding Edited Videos with Large Multimodal Model ..... 503

*Lu Xu (ByteDance), Sijie Zhu (ByteDance), Chunyuan Li (ByteDance), Chia-Wen Kuo (ByteDance), Fan Chen (ByteDance), Xinyao Wang (ByteDance), Guang Chen (ByteDance), Dawei Du (ByteDance), Ye Yuan (ByteDance), and Longyin Wen (ByteDance)*

Revisiting Referring Expression Comprehension Evaluation in the Era of Large Multimodal Models ..... 513

*Jierun Chen (The Hong Kong University of Science and Technology), Fangyun Wei (Microsoft Research Asia), Jinjing Zhao (The University of Sydney), Sizhe Song (The Hong Kong University of Science and Technology), Bohuai Wu (The Hong Kong University of Science and Technology), Zhuoxuan Peng (The Hong Kong University of Science and Technology), S.-H. Gary Chan (The Hong Kong University of Science and Technology), and Hongyang Zhang (University of Waterloo)*

TextInVision: Text and Prompt Complexity Driven Visual Text Generation Benchmark ..... 525

*Forouzan Fallah (Arizona State University), Maitreya Patel (Arizona State University), Agneet Chatterjee (Arizona State University), Vlad Morariu (Adobe Research), Chitta Baral (Arizona State University), and Yezhou Yang (Arizona State University)*

Choosing 'Right' from Wrong: A Closer Look at Selection Bias in Spatial Multiple-Choice Questions in Large Multimodal Models ..... 535

*Giselle Zeno (MIT Lincoln Laboratory), Nour Jedidi (MIT Lincoln Laboratory), and Steven Gomez (MIT Lincoln Laboratory)*

Quantum Federated Learning for Multimodal Data: A Modality-Agnostic Approach ..... 545

*Atit Pokharel (The University of Alabama in Huntsville, USA), Ratun Rahman (The University of Alabama in Huntsville, USA), Thomas Morris (The University of Alabama in Huntsville, USA), and Dinh C. Nguyen (The University of Alabama in Huntsville, USA)*

Revisiting Multi-Modal LLM Evaluation ..... 555

*Jian Lu (University of Rochester), Shikhar Srivastava (University of Rochester), Junyu Chen (University of Rochester), Robik Shrestha (University of Rochester), Manoj Acharya (SRI International), Kushal Kafle (Adobe), and Christopher Kanan (University of Rochester)*

MerCulture: A Comprehensive Benchmark to Evaluate Vision-Language Models on Cultural Understanding in Singapore ..... 565

*Pranav Tushar (SIT), Eshan Pandey (SIT), Lyka Diane Bala Austria (SIT), Yin Yin Loo (SIT), Jing Hao Lim (SIT), Indriyati Atmosukarto (SIT), and Donny Soh Cheng Lock (SIT)*

KOFFVQA: An Objectively Evaluated Free-form VQA Benchmark for Large Vision-Language Models in the Korean Language ..... 575

*Yoonshik Kim (MAUM.AI) and Jaeyoon Jung (MAUM.AI)*

## The 2nd Workshop on Computer Vision for Videogames

MoF-Image: Generating Mixture-of-Features Video Game Image Dataset via GPU Rendering Simulation .....	586
<i>Yu Wen (University of Houston), Xingke Yang (University of Houston), Aamir Bader Shah (University of Houston), Ruizhi Cao (Beihang University), Miao Pan (University of Houston), Chenhao Xie (University of Houston), and Xin Fu (University of Houston)</i>	
GRS: Generating Robotic Simulation Tasks from Real-World Images .....	594
<i>Alex Zook (NVIDIA), Fan-Yun Sun (Stanford University), Josef Spjut (NVIDIA), Valts Blukis (NVIDIA), Stan Birchfield (NVIDIA), and Jonathan Tremblay (NVIDIA)</i>	
G-Buffer Supported Neural Screen-space Refraction Baking for Real-Time Global Illumination.....	604
<i>Ziyang Zhang (Waseda University, Japan) and Edgar Simo-Serra (Waseda University, Japan)</i>	
A Generative AI Game Jam Case Study from October 2024 .....	612
<i>Josef Spjut (NVIDIA)</i>	

## The 6th Workshop on Fair, Data-Efficient, and Trusted Computer Vision

Towards Trustworthy Autonomous Vehicles with Vision-Language Models Under Targeted and Untargeted Adversarial Attacks .....	619
<i>Azwal Ahmed Fime (Southern Illinois University Carbondale, USA), Md Zarif Hossain (Southern Illinois University Carbondale, USA), Saika Zaman (Southern Illinois University Carbondale, USA), Abdur R Shahid (Southern Illinois University Carbondale, USA), and Ahmed Imteaj (Southern Illinois University Carbondale, USA)</i>	
Classification Drives Geographic Bias in Street Scene Segmentation .....	629
<i>Rahul Nair (Arizona State University), Bhanu Tokas (Arizona State University), Gabriel Tseng (Mila Quebec AI Institute), Esther Rolf (University of Colorado Boulder), and Hannah Kerner (Arizona State University)</i>	
Bridging Detection and Re-identification: Evaluating Trustworthiness and Error Propagation in Face Recognition Pipelines .....	639
<i>Kuan Yew Leong (A.I. System Research Co. Ltd) and Jaeseung Han (A.I. System Research Co. Ltd)</i>	
EigenLoRAx: Recycling Adapters to Find Principal Subspaces for Resource-Efficient Adaptation and Inference .....	649
<i>Prakhar Kaushik (Johns Hopkins University), Ankit Vaidya (Pune Institute of Computer Technology), Shravan Chaudhari (Johns Hopkins University), and Alan Yuille (Johns Hopkins University)</i>	
Intriguing Properties of Robust Classification .....	660
<i>Bernd Prach (Institute of Science and Technology Austria (ISTA)) and Christoph H. Lampert (Institute of Science and Technology Austria (ISTA))</i>	

A Large-Scale Analysis on Contextual Self-Supervised Video Representation Learning .....	670
<i>Akash Kumar (University of Central Florida), Ashlesha Kumar (Salesforce), Vibhav Vineet (Microsoft), and Yogesh S Rawat (UCF)</i>	
Is Temporal Prompting All We Need For Limited Labeled Action Recognition? .....	682
<i>Shreyank Gowda (University of Nottingham), Boyan Gao (University of Oxford), Xiao Gu (University of Oxford), and Xiabo Jin (Xi'an Jiaotong-Liverpool University)</i>	
Optimising Vision Transformer Performance on Limited Datasets: A Multi-Gradient Approach ....	693
<i>Mohsin Ali (School of Computer Science and Electronics Engineering, University of Essex), Haider Raza (School of Computer Science and Electronics Engineering, University of Essex), John Q. Gan (School of Computer Science and Electronics Engineering, University of Essex), and Muhammad Haris (Mohamed Bin Zayed University of Artificial Intelligence)</i>	
Defending Against Transfer-Based Adversarial Attacks Using SVD-Driven Feature Evolution .....	703
<i>Xinlei Liu (Information Engineering University), Tao Hu (Information Engineering University), Peng Yi (1. Information Engineering University; 2. Key Laboratory of Cyberspace Security, Ministry of Education of China), Qingtao Pan (College of Systems Engineering, National University of Defense Technology), Hailong Ma (1. Information Engineering University; 2. Key Laboratory of Cyberspace Security, Ministry of Education of China), Yiming Jiang (1. Information Engineering University; 2. Key Laboratory of Cyberspace Security, Ministry of Education of China), and Baolin Li (Information Engineering University)</i>	
ProtoPatchNet: An Interpretable Patch-Based Prototypical Network .....	712
<i>Mohana Singh (IISc), Vivek B S (TCS Research), Jayavardhana Gubbi (TCS Research), and R. Venkatesh Babu (Vision and AI Lab, IISc)</i>	
Balancing Privacy and Action Performance: A Penalty-Driven Approach to Image Anonymization	720
<i>Nazia Aslam (Aalborg University Denmark, Pioneer Center for AI Denmark) and Kamal Nasrollahi (Aalborg University Denmark, Milestone Systems Denmark)</i>	

## 10th New Trends in Image Restoration and Enhancement Workshop and Challenges

Rethinking Compressive Sensing: A Compression Framework for Video Super-Resolution .....	730
<i>Ruthy Katz (Corephotonics, Israel), Adi Teitel (Corephotonics, Israel), Moran Mordechay (Corephotonics, Israel), Adi Falik (Corephotonics, Israel), Eli Bery (Corephotonics, Israel), and Maya Mayberg (Corephotonics, Israel)</i>	
MAD: Makeup All-in-One with Cross-Domain Diffusion Model .....	740
<i>Bo-Kai Ruan (National Yang Ming Chiao Tung University) and Hong-Han Shuai (National Yang Ming Chiao Tung University)</i>	
conSAMmé: Achieving Consistent Segmentations with SAM .....	750
<i>Josh Myers-Dean (University of Colorado Boulder), Kangning Liu (Adobe), Brian Price (Adobe), Yifei Fan (Adobe), Jason Kuen (Adobe), and Danna Gurari (University of Colorado Boulder)</i>	

STAPLE: Siamese Transformer Assisted Pseudo Label Ensembling for Unsupervised Domain Adaptation in No-Reference IQA .....	760
<i>Arshita Gupta (Samsung Research America), Zhe Zhu (Samsung Research America), and Tien Bau (Samsung Research America)</i>	
XYScanNet: A State Space Model for Single Image Deblurring .....	770
<i>Hanzhou Liu (Texas A&amp;M University), Chengkai Liu (Texas A&amp;M University), Jiacong Xu (Johns Hopkins University), Peng Jiang (Texas A&amp;M University), and Mi Lu (Texas A&amp;M University)</i>	
Text-Guided Patch Scoring and Local Distortion Guidance for Image Quality Assessment .....	781
<i>Juyong Park (Inha University), Jihun Song (Ulsan National Institute of Science and Technology (UNIST)), Gyewan Kim (Inha University), and Yoonsuk Hyun (Inha University)</i>	
MFSR-GAN: Multi-Frame Super-Resolution with Handheld Motion Modeling .....	791
<i>Fadeel Sher Khan (The University of Texas at Austin), Joshua Ebenezer (Samsung Research America), Hamid Sheikh (Samsung Research America), and Seok-Jun Lee (Samsung Research America)</i>	
Training Neural Networks on RAW and HDR Images for Restoration Tasks .....	801
<i>Andrew Yanzhe Ke (University of Cambridge), Lei Luo (Meta), Xiaoyu Xiang (Meta), Yuchen Fan (Meta), Rakesh Ranjan (Meta), Alexandre Chapiro (Meta), and Rafal Mantiuk (University of Cambridge)</i>	
DataFormer: Differential Additive Transformer for Lightweight Semantic Segmentation .....	811
<i>Mian Muhammad Naeem Abid (Computer Vision Lab, CAIDAS &amp; IFI, University of Würzburg, Germany), Nancy Mehta (Computer Vision Lab, CAIDAS &amp; IFI, University of Würzburg, Germany), Zongwei Wu (Computer Vision Lab, CAIDAS &amp; IFI, University of Würzburg, Germany), and Radu Timofte (Computer Vision Lab, CAIDAS &amp; IFI, University of Würzburg, Germany)</i>	
KernFusNet: Implicit Kernel Modulation and Fusion for Blind Super-resolution .....	823
<i>Nancy Mehta (Universität Würzburg), Akshay Dudhane (MBZUAI), Subrahmanyam Murala (Trinity College Dublin), and Radu Timofte (Universität Würzburg)</i>	
FUSION: Frequency-guided Underwater Spatial Image reconstruction .....	834
<i>Jaskaran Walia (Vellore Institute of Technology), Shravan Venkatraman (Vellore Institute of Technology), and Pavithra L K (Vellore Institute of Technology)</i>	
FALCON: Fast Image Haze Removal Leveraging Continuous Density Mask .....	844
<i>Donghyun Kim (Yonsei University), Seil Kang (Yonsei University), and Seong Jae Hwang (Yonsei University)</i>	
Efficient Burst Super-Resolution with One-step Diffusion .....	855
<i>Kento Kawai (Toyota Technological Institute), Takeru Oba (Toyota Technological Institute), Kyotaro Tokoro (Toyota Technological Institute), Kazutoshi Akita (Toyota Technological Institute), and Norimichi Ukita (Toyota Technological Institute)</i>	
SPIdepth: Strengthened Pose Information for Self-supervised Monocular Depth Estimation .....	865
<i>Mykola Lavreniuk (Space Research Institute, Ukraine) and Alla Lavreniuk (National Technical University of Ukraine «Igor Sikorsky Kyiv Polytechnic Institute», Ukraine)</i>	

FCTFANet: A Fused CNN-Transformer Feature Aggregator Network for Image Restoration .....	876
<i>Amit Monga (IIT Roorkee), Hemkant Nehete (IIT Roorkee), Partha Kaushik (IIT Roorkee), Tharun Kumar Reddy Bollu (IIT Roorkee), Balasubramanian Raman (IIT Roorkee), and Gaurav Sharma (University of Rochester)</i>	
A Simple Combination of Diffusion Models for Better Quality Trade-Offs in Image Denoising .....	886
<i>Jonas Dornbusch (Technical University of Munich), Emanuel Pfarr (University of Würzburg), Florin-Alexandru Vasluianu (University of Würzburg), Frank Werner (University of Würzburg), and Radu Timofte (University of Würzburg)</i>	
PromptNorm: Image Geometry Guides Ambient Light Normalization .....	896
<i>David Serrano-Lozano (Computer Vision Center, Spain), Francisco A. Molina-Bakhos (Computer Vision Center, Spain), Danna Xue (Computer Vision Center, Spain), Yixiong Yang (Computer Vision Center, Spain), Maria Pilligua (Computer Vision Center, Spain), Ramon Baldrich (Computer Vision Center), Maria Vanrell (Computer Vision Center), and Javier Vazquez-Corral (Computer Vision Center)</i>	
The Tenth NTIRE 2025 Efficient Super-Resolution Challenge Report .....	908
<i>Bin Ren (University of Trento), Hang Guo (Tsinghua University), Lei Sun (INSAIT, Sofia University “St. Kliment Ohridski”), Zongwei Wu (University of Würzburg), Radu Timofte (University of Würzburg), and Yawei Li (ETH Zürich)</i>	
Expanded SPAN for Efficient Super-Resolution .....	958
<i>Qing Wang (bytedance), Yang Wang (bytedance), Hongyu An (bytedance), Yi Liu (bytedance), Liou Zhang (bytedance), and Shijie Zhao (bytedance)</i>	
Leveraging Multimodal Large Language Models for Joint Discrete and Continuous Evaluation in Text-to-Image Alignment .....	968
<i>Zhichao Zhang (Shanghai Jiao Tong University), Xinyue Li (Shanghai Jiao Tong University), Wei Sun (Shanghai Jiao Tong University), Zicheng Zhang (Shanghai Jiao Tong University), Yunhao Li (Shanghai Jiao Tong University), Xiaohong Liu (Shanghai Jiao Tong University), and Guangtao Zhai (Shanghai Jiao Tong University)</i>	

NTIRE 2025 Challenge on HR Depth from Images of Specular and Transparent Surfaces .....	978
---	-----

*Pierluigi Zama Ramirez (University of Bologna), Fabio Tosi (University of Bologna), Luigi Di Stefano (University of Bologna), Radu Timofte (University of Wurzburg, Germany), Alex Costanzino (University of Bologna), Matteo Poggi (University of Bologna), Samuele Salti (University of Bologna), Stefano Mattoccia (University of Bologna), Zhe Zhang (Nanjing University of Science and Technology), Yang Yang (Nanjing University of Science and Technology), Wu Chen (Beijing University of Posts and Telecommunications), Anlong Ming (Beijing University of Posts and Telecommunications), Mingshuai Zhao (Beijing University of Posts and Telecommunications), Mengying Yu (Beijing University of Posts and Telecommunications), Shida Gao (Beijing University of Posts and Telecommunications), Xiangfeng Wang (Beijing University of Posts and Telecommunications), Feng Xue (University of Trento), Jun Shi (Samsung R&D Institute China-Beijing (SRC-B)), Yong Yang (Samsung R&D Institute China-Beijing (SRC-B)), Yong A (Samsung R&D Institute China-Beijing (SRC-B)), Yixiang Jin (Samsung R&D Institute China-Beijing (SRC-B)), Dingzhe Li (Samsung R&D Institute China-Beijing (SRC-B)), Aryan Shukla (École de technologie supérieure (ÉTS)), Liam Frija-Altarac (École de technologie supérieure (ETS)), Matthew Toews (École de technologie supérieure (ÉTS)), Hui Geng (National University of Defense Technology, Changsha, China), Tianjiao Wan (National University of Defense Technology, Changsha, China), Zijian Gao (National University of Defense Technology, Changsha, China), Qisheng Xu (National University of Defense Technology, Changsha, China), Kele Xu (National University of Defense Technology, Changsha, China), Zijian Zang (Fudan University), Jameer Babu (None), Kuldeep Purohit (None), Mykola Lavreniuk (Space Research Institute NASU-SSAU, Kyiv, Ukraine), Jing Cao (Harbin Institute of Technology), Shenyi Li (Harbin Institute of Technology), Kui Jiang (Harbin Institute of Technology), Junjun Jiang (Harbin Institute of Technology), and Yong Huang (Harbin Institute of Technology)*

NTIRE 2025 Challenge on Efficient Burst HDR and Restoration: Datasets, Methods, and Results .....	993
<p><i>Sangmin Lee (Samsung Electronics MX Business - Camera Innovation Group), Eunpil Park (Samsung Electronics MX Business - Camera Innovation Group), Angel Canelo (Samsung Electronics MX Business - Camera Innovation Group), Hyunhee Park (Samsung Electronics MX Business - Camera Innovation Group), Youngjo Kim (Samsung Electronics MX Business - Camera Innovation Group), Hyungju Chun (Samsung Electronics MX Business - Camera Innovation Group), Xin Jin (Nankai University), Chongyi Li (Nankai University), Chun-Le Guo (Nankai University), Radu Timofte (University of Wurzburg), Qi Wu (Imvision), Tianheng Qiu (Imvision), Yuchun Dong (Imvision), Shenglin Ding (Imvision), Guanghua Pan (Imvision), Weiyu Zhou (Northwestern Polytechnical University), Tao Hu (Northwestern Polytechnical University), Yixu Feng (Northwestern Polytechnical University), Duwei Dai (Second Affiliated Hospital Of Xi'an Jiaotong University), Yu Cao (Xi'an Institute of Optics and Precision Mechanics of CAS), Peng Wu (Northwestern Polytechnical University), Wei Dong (Xi'an University of Architecture and Technology), Yanning Zhang (Northwestern Polytechnical University), Qingsen Yan (Northwestern Polytechnical University), Simon Larsen (Esoft), Senyan Xu (University of Science and Technology of China), Xingbo Wang (University of Science and Technology of China), Ruixuan Jiang (University of Science and Technology of China), Xin Lu (University of Science and Technology of China), Marcos Conde (Cidaut AI), Javier Abad-Hernández (Cidaut AI), Álvaro García-Lara (Cidaut AI), Daniel Feijoo (Cidaut AI), Álvaro García (Cidaut AI), Zeyu Xiao (National University of Singapore), and Zhuoyuan Li (University of Science and Technology of China)</i></p>	
Flow-Guided Deformable Alignment with Channel-wise Self-Attention Reconstruct for Efficient Burst HDR Restoration .....	1009
<p><i>Weiyu Zhou (Northwestern Polytechnical University), Tao Hu (Northwestern Polytechnical University), Yixu Feng (Northwestern Polytechnical University), Duwei Dai (Second Affiliated Hospital Of Xi'an Jiaotong University), Yu Cao (Xi'an Institute of Optics and Precision Mechanics of CAS), Peng Wu (Northwestern Polytechnical University), Wei Dong (Xi'an University of Architecture and Technology), Yanning Zhang (Northwestern Polytechnical University), and Qingsen Yan (Northwestern Polytechnical University)</i></p>	
FusionNet: Multi-model Linear Fusion Framework for Low-light Image Enhancement .....	1019
<p><i>Kangbiao Shi (Northwestern Polytechnical University), Yixu Feng (Northwestern Polytechnical University), Tao Hu (Northwestern Polytechnical University), Yu Cao (Xi'an Institute of Optics and Precision Mechanics of CAS), Peng Wu (Northwestern Polytechnical University), Yijin Liang (Shanghai Institute of Satellite Engineering, Shanghai 201109, China), Yanning Zhang (Northwestern Polytechnical University), and Qingsen Yan (Northwestern Polytechnical University)</i></p>	

Recursive Multi-Exposure Alignment with Spatiotemporal Decoupling for Efficient Burst HDR and Restoration .....	1029
<p><i>Tianheng Qiu (University of Science and Technology of China, Chinese Academy of Sciences), Qi Wu (Imvision), Yuchun Dong (Imvision), Shenglin Ding (Imvision), Xuan Huang (Chinese Academy of Sciences), Hu Wei (Chinese Academy of Sciences), and Guanghua Pan (Imvision)</i></p>	
NTIRE 2025 Challenge on Cross-Domain Few-Shot Object Detection: Methods and Results .....	1039
<p><i>Yuqian Fu (INSAT, Sofia University St. Kliment Ohridski, Bulgaria), Xingyu Qiu (Fudan University, China), Bin Ren (University of Pisa, Italy; University of Trento, Italy), Yanwei Fu (Fudan University, China), Radu Timofte (Computer Vision Lab, University of Wurzburg, Germany), Nicu Sebe (University of Trento, Italy), Ming-Hsuan Yang (University of California at Merced, United States), Luc Van Gool (INSAT, Sofia University St. Kliment Ohridski, Bulgaria), Kaijin Zhang, Qingpeng Nong, Xiugang Dong, Hong Gao, Xiangsheng Zhou, Jiancheng Pan, Yanxing Liu, Xiao He, Jiahao Li, Yuze Sun, Xiaomeng Huang, Zhenyu Zhang, Ran Ma, Yuhan Liu, Zijian Zhuang, Shuai Yi, Yixiong Zou, Lingyi Hong, Mingxi Chen, Runze Li, Xingdong Sheng, Wenqiang Zhang, Weisen Chen, Yongxin Yan, Xinguo Chen, Yuanjie Shao, Zhengrong Zuo, Nong Sang, Hao Wu, Haoran Sun, Shuming Hu, Yan Zhang, Zhiguang Shi, Yu Zhang, Chao Chen, Tao Wang, Da Feng, Linhai Zhuo, Ziming Lin, Yali Huang, Jie Me, Yiming Yang, Mi Guo, Mingyuan Jiu, Mingliang Xu, Maomao Xiong, Qunshu Zhang, Xinyu Cao, Yuqing Yang, Dianmo Sheng, Xuanpu Zhao, Zhiyu Li, Xuyang Ding, and Wenqian Li</i></p>	
Advancing Ambient Lighting Normalization via Diffusion Shadow Generation .....	1061
<p><i>Xin Lu (University of Science and Technology of China), Jiarong Yang (University of Science and Technology of China), Yuanfei Bao (University of Science and Technology of China), Zihao Fan (University of Science and Technology of China), Anya Hu (University of Science and Technology of China), Kunyu Wang (University of Science and Technology of China), Jie Xiao (University of Science and Technology of China), Xi Wang (University of Science and Technology of China), Hongjian Liu (University of Science and Technology of China), Xueyang Fu (University of Science and Technology of China), and Zheng-Jun Zha (University of Science and Technology of China)</i></p>	
EvenFormer: Dynamic Even Transformer for Real-World Image Restoration .....	1072
<p><i>Xin Lu (University of Science and Technology of China), Yuanfei Bao (University of Science and Technology of China), Jiarong Yang (University of Science and Technology of China), Anya Hu (University of Science and Technology of China), Jie Xiao (University of Science and Technology of China), Kunyu Wang (University of Science and Technology of China), Dong Li (University of Science and Technology of China), Senyan Xu (University of Science and Technology of China), Kean Liu (University of Science and Technology of China), Xueyang Fu (University of Science and Technology of China), and Zheng-Jun Zha (University of Science and Technology of China)</i></p>	

NTIRE 2025 Challenge on Short-form UGC Video Quality Assessment and Enhancement: Methods and Results ..... 1083

Xin Li (University of Science and Technology of China), Kun Yuan (Kuaishou Technology), Bingchen Li (University of Science and Technology of China), Fengbin Guan (University of Science and Technology of China), Yizhen Shao (Kuaishou Technology), Zihao Yu (University of Science and Technology of China), Xijun Wang (University of Science and Technology of China), Yiting Lu (University of Science and Technology of China), Wei Luo (University of Science and Technology of China), Suhang Yao (University of Science and Technology of China), Ming Sun (Kuaishou Technology), Chao Zhou (Beijing Kuaishou Technology Co., Ltd), Zhibo Chen (University of Science and Technology of China), Radu Timofte (University of Wurzburg & ETH Zurich), Yabin Zhang (Bytedance Inc.), Ao-Xiang Zhang (Bytedance Inc.), Tianwu Zhi (Bytedance Inc.), Jianzhao Liu (Bytedance Inc.), Yang Li (Bytedance Inc.), Jingwen Xu (Bytedance Inc.), Yiting Liao (Bytedance Inc.), Yushen Zuo (The Hong Kong Polytechnic University), Mingyang Wu (Texas A&M University), Renjie Li (Texas A&M University), Shengyun Zhong (Northeastern University), Zhengzhong Tu (Texas A&M University), Yufan Liu (Tencent Online Video), Xiangguang Chen (Tencent Online Video), Zuowei Cao (Tencent Online Video), Minhao Tang (Tencent Online Video), Shan Liu (Tencent Online Video), Kexin Zhang (Bytedance Inc.), Jingfen Xie (Bytedance Inc.), Yan Wang (Bytedance Inc.), Kai Chen (Bytedance Inc.), Shijie Zhao (Bytedance Inc.), Yunchen Zhang (ZTE Corporation), Xiangkai Xu (ZTE Corporation), Hong Gao (ZTE Corporation), Ji Shi (ZTE Corporation), Yiming Bao (ZTE Corporation), Xiugang Dong (ZTE Corporation), Xiangsheng Zhou (ZTE Corporation), Yaofeng Tu (ZTE Corporation), Ying Liang (Shanghai Jiao Tong University), Yiwen Wang (Shanghai Jiao Tong University), Xinning Chai (Shanghai Jiao Tong University), Yuxuan Zhang (Shanghai Jiao Tong University), Zhengxue Cheng (Shanghai Jiao Tong University), Yingsheng Qin (Transsion, China), Yucai Yang (Transsion, China), Rong Xie (Shanghai Jiao Tong University), Li Song (Shanghai Jiao Tong University), Wei Sun (East China Normal University), Kang Fu (Shanghai Jiao Tong University), Linhan Cao (Shanghai Jiao Tong University), Dandan Zhu (East China Normal University), Kaiwei Zhang (Shanghai Jiao Tong University), Yucheng Zhu (Shanghai Jiao Tong University), Zicheng Zhang (Shanghai Jiao Tong University), Menghan Hu (East China Normal University), Xionguo Min (Shanghai Jiao Tong University), Guangtao Zhai (Shanghai Jiao Tong University), Zhi Jin (Shenzhen Campus of Sun Yat-sen University), Jiawei Wu (Shenzhen Campus of Sun Yat-sen University), Wei Wang (Shenzhen Campus of Sun Yat-sen University), Wenjian Zhang (Shenzhen Campus of Sun Yat-sen University), Yuhai Lan (Tencent), Gaoxiong Yi (Tencent), Hengyuan Na (Sun Yat-sen University), Wang Luo (Sun Yat-sen University), Di Wu (Sun Yat-sen University), MingYin Bai (Nanjing University of Posts and Telecommunications), Jiawang Du (Nanjing University of Posts and Telecommunications), Zilong Lu (Nanjing University of Posts and Telecommunications), Zhenyu Jiang (Nanjing University of Posts and Telecommunications), Hui Zeng (Nanjing University of Posts and Telecommunications), Ziguan Cui (Nanjing University of Posts and Telecommunications), Zongliang Gan (Nanjing University of Posts and

Telecommunications), Guijin Tang (Nanjing University of Posts and Telecommunications), Xinglin Xie (XiDian University), Kehuan Song (XiDian University), Xiaoqiang Lu (XiDian University), Licheng Jiao (XiDian University), Fang Liu (XiDian University), Xu Liu (XiDian University), Puhua Chen (XiDian University), Ha Thu Nguyen (Norwegian University of Science and Technology), Katrien De Moor (Norwegian University of Science and Technology), Seyed Ali Amirshahi (Universit e de Poitiers, CNRS, XLIM, France), Mohamed-Chaker Larabi (Norwegian University of Science and Technology), Qi Tang (Beijing Jiaotong University), Linfeng He (Beijing Jiaotong University), Zhiyong Gao (Beijing Jiaotong University), Zixuan Gao (Beijing Jiaotong University), Guohua Zhang (Beijing Jiaotong University), Zhiye Huang (Beijing University of Posts and Telecommunications), Yi Deng (Beijing University of Posts and Telecommunications), Qingmiao Jiang (School of Information and Communication Engineering, Communication University of China), Lu Chen (School of Information and Communication Engineering, Communication University of China), Yi Yang (School of Information and Communication Engineering, Communication University of China), Xi Liao (School of Information and Communication Engineering, Communication University of China), Nourine Mohammed Nadir (None), Yuxuan Jiang (University of Bristol), Qiang Zhu (University of Electronic Science and Technology of China, China), Siyue Teng (University of Bristol, UK), Fan Zhang (University of Bristol, UK), Shuyuan Zhu (University of Electronic Science and Technology of China, China), Bing Zeng (University of Electronic Science and Technology of China, China), David Bull (University of Bristol, UK), Meiqin Liu (Beijing Jiaotong University), Chao Yao (Beijing Jiaotong University), and Yao Zhao (Beijing Jiaotong University)

NTIRE 2025 challenge on Text to Image Generation Model Quality Assessment .....	1095
---	------

*Shuhao Han (Nankai University), Haotian Fan (Bytedance), Fangyuan Kong (Bytedance), Wenjie Liao (Bytedance), Chunle Guo (Nankai University), Chongyi Li (Nankai University), Radu Timofte (University of Würzburg), Liang Li (Bytedance), Tao Li (Bytedance), Junhui Cui (Bytedance), Yunqiu Wang (Bytedance), Yang Tai (Bytedance), Jingwei Sun (Bytedance), Jianhui Sun (WeChat), Xinli Yue (Wuhan University), Tianyi Wang (WeChat), Huan Hou (WeChat), Junda Lu (WeChat), Xinyang Huang (WeChat), Zitang Zhou (WeChat), Zijian Zhang (Meituan), Xuhui Zheng (Meituan), Xuecheng Wu (Meituan), Chong Peng (Meituan), Xuezhi Cao (Meituan), Trong-Hieu Nguyen-Mau (University of Science, VNU-HCM), Minh-Hoang Le (University of Science, VNU-HCM), Minh-Khoa Le-Phan (University of Science, VNU-HCM), Duy-Nam Ly (University of Science, VNU-HCM), Hai-Dang Nguyen (University of Science, VNU-HCM), Minh-Triet Tran (University of Science, VNU-HCM), Yukang Lin (MICV), Yan Hong (MICV), Chuanbiao Song (MICV), Siyuan Li (MICV), Jun Lan (MICV), Zhichao Zhang (Shanghai Jiao Tong University), Xinyue Li (Shanghai Jiao Tong University), Wei Sun (Shanghai Jiao Tong University), Zicheng Zhang (Shanghai Jiao Tong University), Yunhao Li (Shanghai Jiao Tong University), Xiaohong Liu (Shanghai Jiao Tong University), Guangtao Zhai (Shanghai Jiao Tong University), Zitong Xu (Shanghai Jiao Tong University), Huiyu Duan (Shanghai Jiao Tong University), Jiarui Wang (Shanghai Jiao Tong University), Guangji Ma (Shanghai Jiao Tong University), Liu Yang (Shanghai Jiao Tong University), Lu Liu (Shanghai Jiao Tong University), Qiang Hu (Shanghai Jiao Tong University), Xiongkuo Min (Shanghai Jiao Tong University), Zichuan Wang (Institute of Automation, Chinese Academy of Sciences), Zhenchen Tang (Institute of Automation, Chinese Academy of Sciences), Bo Peng (Peng), Jing Dong (Institute of Automation, Chinese Academy of Sciences), Fengbin Guan (University of Science and Technology of China), Zihao Yu (University of Science and Technology of China), Yiting Lu (University of Science and Technology of China), Wei Luo (University of Science and Technology of China), Xin Li (University of Science and Technology of China), Minhao Lin (Hangzhou Dianzi University), Haofeng Chen (Hangzhou Dianzi University), Xuanxuan He (National University of Defense Technology), Kele Xu (National University of Defense Technology), Qisheng Xu (National University of Defense Technology), Zijian Gao (National University of Defense Technology), Tianjiao Wan (National University of Defense Technology), Bo-Cheng Qiu (National Cheng Kung University), Chih-Chung Hsu (Institute of Data Science, National Cheng Kung University, Taiwan), Chia-ming Lee (Institute of Data Science, National Cheng Kung University, Taiwan), Yu-Fan Lin (Institute of Data Science, National Cheng Kung University, Taiwan), Bo Yu (Beijing University of Posts and Telecommunications), Zehao Wang (Beijing University of Posts and Telecommunications), Da Mu (Beijing University of Posts and Telecommunications), Mingxiu Chen (Beijing University of Posts and Telecommunications), Junkang Fang (Beijing University of Posts and Telecommunications), Huamei Sun (Beijing University of Posts and Telecommunications), Wending Zhao (Beijing University of Posts and Telecommunications), Zhiyu Wang (Hunan University), Wang Liu (Hunan University), Weikang Yu (Technical University of Munich), Puhong*

Duan (Hunan University), Bin Sun (Hunan University), Xudong Kang (Hunan University), Shutao Li (Hunan University), Shuai He (OPDAI, Netease Game, Guangzhou China), Lingzhi Fu (OPDAI, Netease Game, Guangzhou, China), Heng Cong (OPDAI, Netease Game, Guangzhou, China), Rongyu Zhang (OPDAI, Netease Game, Guangzhou, China), Jiarong He (OPDAI, Netease Game, Guangzhou, China), Zhishan Qiao (missing), Yongqing Huang (missing), Zewen Chen (State Key Laboratory of Multimodal Artificial Intelligence System, Institute of Automation, Chinese Academy of Sciences), Zhe Pang (State Key Laboratory of Multimodal Artificial Intelligence Systems, Institute of Automation, Chinese Academy of Sciences), Juan Wang (State Key Laboratory of Multimodal Artificial Intelligence Systems, Institute of Automation, Chinese Academy of Sciences), Jian Guo (Beijing Union University), Zhizhuo Shao (School of Artificial Intelligence, University of Chinese Academy of Sciences), Ziyu Feng (Beijing Jiaotong University), Bing Li (State Key Laboratory of Multimodal Artificial Intelligence Systems, Institute of Automation, Chinese Academy of Sciences), Weiming Hu (State Key Laboratory of Multimodal Artificial Intelligence Systems, Institute of Automation, Chinese Academy of Sciences), Hesong Li (Transsion Inc.), Dehua Liu (Transsion Inc.), Zeming Liu (OPPO AI Center), Qingsong Xie (OPPO AI Center), Ruichen Wang (OPPO AI Center), Zhihao Li (OPPO AI Center), Yuqi Liang (Ant Group), Jianqi Bi (Ant Group), Jun Luo (Antgroup), Junfeng Yang (Xiangjiang Laboratory, Hunan University of Technology and Business), Can Li (Xiangjiang Laboratory, Hunan University of Technology and Business), Jing Fu (Xiangjiang Laboratory, Hunan University of Technology and Business), Hongwei Xu (Xiangjiang Laboratory, Hunan University of Technology and Business), Mingrui Long (Xiangjiang Laboratory, Hunan University of Technology and Business), and Lulin Tang (Xiangjiang Laboratory, Hunan University of Technology and Business)

NTIRE 2025 Challenge on Short-form UGC Video Quality Assessment and Enhancement: KwaiSR Dataset and Study ..... 1117

Xin Li (University of Science and Technology of China), Xijun Wang (University of Science and Technology of China), Bingchen Li (University of Science and Technology of China), Kun Yuan (Kuaishou Technology), Yizhen Shao (Kuaishou Technology), Suhang Yao (University of Science and Technology of China), Ming Sun (Kuaishou Technology), Chao Zhou (Beijing Kuaishou Technology Co., Ltd), Radu Timofte (University of Wurzburg & ETH Zurich), and Zhibo Chen (University of Science and Technology of China)

LFTramba: Comprehensive Information Learning for Light Field Image Super-Resolution via A Hybrid Transformer-Mamba Framework ..... 1128

Haosong Liu (Huaqiao University), Xiancheng Zhu (Huaqiao University), Huanqiang Zeng (Huaqiao University, Xiamen University of Technology), Jianqing Zhu (Huaqiao University), Yifan Shi (Huaqiao University), Jing Chen (Huaqiao University), and Junhui Hou (City University of Hong Kong)

NTIRE 2025 Challenge on RAW Image Restoration and Super-Resolution .....	1139
--	------

*Marcos Conde (University of Würzburg & Sony PlayStation), Radu Timofte (University of Würzburg), Zihao Lu (University of Würzburg), Xiangyu Kong, Xiaoxia Xing, Fan Wang, Suejin Han, MinKyu Park, Tianyu Hao, Yuhong He, Ruoqi Li, Yueqi Yang, Jianyang Yu, Kele Xu, Zisheng Xu, Yong Dou, Watchara Ruangsang, Ruixuan Jiang, Senyan Xu, Siyuan Jiang, Xueyang Fu, Zheng-Jun Zha, Jiajie Lu, Xiang Yu (Nanjing University), Minmin Yi, Yuanjia Chen, Liwen Zhang, Zijie Jin, Tianyu Zhang, Xin Luo, Yeda Chen, Dong Liu, Li Pang, Yuhang Yang, Hongzhong Wang, Xiangyong Cao, Cheng Li, Lian Liu, Wei Song, Heng Sun, Tianyu Hao, Yuhong He, Ruoqi Li, Yueqi Yang, Yubo Wang, Jinghua Wang, and Guanlan Hong (Nanjing University)*

NTIRE 2025 Challenge on Day and Night Raindrop Removal for Dual-Focused Images: Methods  
and Results ..... 1163

Xin Li (University of Science and Technology of China), Yeying Jin (National University of Singapore), Xin Jin (Eastern Institute of Technology, Ningbo), Zongwei Wu (University of Würzburg), Bingchen Li (University of Science and Technology of China), Yufei Wang (Snap Research), Wenhan Yang (Pengcheng Laboratory), Yu Li (IDEA), Zhibo Chen (University of Science and Technology of China), Bihan Wen (Nanyang Technological University), Robby Tan (National University of Singapore), Radu Timofte (University of Würzburg), Qiyu Rong (Beijing Union University), Hongyuan Jing (Beijing Union University), Mengmeng Zhang (Beijing Union University), Jinglong Li (Beijing Union University), Xiangyu Lu (Beijing Union University), Yi Ren (Beijing Union University), Yuting Liu (Beijing Union University), Meng Zhang (Beijing Union University), Xiang Chen (Nanjing University of Science and Technology), Qiyuan Guan (Dalian Polytechnic University), Jiangxin Dong (Nanjing University of Science and Technology), Jinshan Pan (Nanjing University of Science and Technology), Conglin Gou (Tianjin University), Qirui Yang (Tianjin University), Fangpu Zhang (Tianjin University), Yunlong Lin (Xiamen University), Sixiang Chen (The Hong Kong University of Science and Technology, Guangzhou), Guoxi Huang (University of Bristol, UK), Ruirui Lin (University of Bristol, UK), Yan Zhang (National University of Singapore, Singapore), Jingyu Yang (Tianjin University), Huanjing Yue (Tianjin University), Jiyuan Chen (The Hong Kong Polytechnic University), Qiaosi Yi (The Hong Kong Polytechnic University), Hongjun Wang (The University of Tokyo), Chenxi Xie (The Hong Kong Polytechnic University), Shuai Li (The Hong Kong Polytechnic University), Yuhui Wu (The Hong Kong Polytechnic University), Kaiyi Ma (Xidian University), Jiakui Hu (Xidian University), Juncheng Li (Shanghai University), Liwen Pan (Shanghai University), Guangwei Gao (Nanjing University of Posts and Telecommunications), Wenjie Li (Beijing University of Posts and Telecommunications), Zhenyu Jin (Beijing University of Posts and Telecommunications), Heng Guo (Beijing University of Posts and Telecommunications), Zhanyu Ma (Beijing University of Posts and Telecommunications), Yubo Wang (Harbin Institute of Technology (Shenzhen)), Jinghua Wang (Harbin Institute of Technology (Shenzhen)), Wangzhi Xing (Griffith University), Anjusree Karnavar (Griffith University), Diqi Chen (Griffith University), Mohammad Aminul Islam (Griffith University), Hao Yang (Beijing Institute of Technology), Ruikun Zhang (Beijing Institute of Technology), Liyuan Pan (Beijing Institute of Technology), Qianhao Luo (CISDI Information Technology CO., LTD), Xin Cao (CISDI Information Technology CO., LTD), Han Zhou (Department of Electrical and Computer Engineering, McMaster University), Yan Min (Department of Electrical and Computer Engineering, McMaster University), Wei Dong (Department of Electrical and Computer Engineering, McMaster University), Jun Chen (Department of Electrical and Computer Engineering, McMaster University), Taoyi Wu (Tongji University), Weijia Dou (Tongji University), Yu Wang (Tongji University), Shengjie Zhao (Tongji University), Yongcheng Huang (Delft University of Technology), Xingyu Han (Delft University of Technology), Anyan Huang (Delft University of Technology), Hongtao Wu

(Medical Artificial Intelligence Laboratory, West Lake University),  
 Hong Wang (School of Life Science and Technology, Xi'an Jiaotong  
 University), Yefeng Zheng (Medical Artificial Intelligence Laboratory,  
 West Lake University), Abhijeet Kumar (Indian Institute of Technology  
 Madras), Aman Kumar (Indian Institute of Technology Madras), Marcos V.  
 Conde (Cidaut AI), Paula Garrido (Cidaut AI), Daniel Feijoo (Cidaut  
 AI), Juan C. Benito (Cidaut AI), Guanglu Dong (Sichuan University),  
 Xin Lin (Sichuan University), Siyuan Liu (Sichuan University),  
 Tianheng Zheng (Sichuan University), Jiayu Zhong (Sichuan University),  
 Shouyi Wang (Sichuan University), Xiangtai Li (Sichuan University),  
 Lanqing Guo (Sichuan University), Lu Qi (Sichuan University), Chao Ren  
 (Sichuan University), Shuaibo Wang (Xidian University), Shilong Zhang  
 (Xidian University), Wanyu Zhou (Xidian University), Yunze Wu (Xidian  
 University), Qinzong Tan (Xidian University), Jieyuan Pei (Zhejiang  
 University of Technology), Zhuoxuan Li (Tongji University), Jiayu Wang  
 (University of Electronic Science and Technology of China), Haoyu  
 Bian (University of Electronic Science and Technology of China),  
 Haoran Sun (University of Electronic Science and Technology of China),  
 Subhajit Paul (Space Applications Centre), Ni Tang (School of  
 Informatics, Xiamen University), Junhao Huang (School of Informatics,  
 Xiamen University), Zihan Cheng (School of Informatics, Xiamen  
 University), Hongyun Zhu (School of Informatics, Xiamen University),  
 Yuehan Wu (School of Informatics, Xiamen University), Kaixin Deng  
 (Hokkaido University), Huang Ouyang (Chengdu University of  
 Technology), Tianxin Xiao (Chengdu University of Technology), Fan Yang  
 (Chengdu University of Technology), Zhizun Luo (Chengdu University of  
 Technology), Zeyu Xiao (National University of Singapore), Zhuoyuan Li  
 (University of Science and Technology of China), Pham Hoang Le Nguyen  
 (University of Information Technology, Ho Chi Minh City, Vietnam),  
 Dinh Thien An (University of Information Technology, Ho Chi Minh City,  
 Vietnam), Luu Thanh Son (University of Information Technology, Ho Chi  
 Minh City, Vietnam), Kiet Van Nguyen (Ho Chi Minh City, Vietnam),  
 Ronghua Xu (School of Big Data and Software Engineering, Chongqing  
 University), Xianmin Tian (School of Big Data and Software  
 Engineering, Chongqing University), Weijian Zhou (School of Big Data  
 and Software Engineering, Chongqing University), Jiacheng Zhang  
 (School of Big Data and Software Engineering, Chongqing University),  
 Yuqian Chen (Southeast University), Yihang Duan (Xidian University),  
 Yujie Wu (Xidian University), Suresh Raikwar (NA), Arsh Garg (NA),  
 Kritika Kritika (NA), Jianhua Zheng (Zhongkai University of  
 Agriculture and Engineering), Xiaoshan Ma (Zhongkai University of  
 Agriculture and Engineering), Ruolin Zhao (Zhongkai University of  
 Agriculture and Engineering), Yongyu Yang (Zhongkai University of  
 Agriculture and Engineering), Yongsheng Liang (Zhongkai University of  
 Agriculture and Engineering), Guiming Huang (Zhongkai University of  
 Agriculture and Engineering), Qiang Li (Shanghai University), Hongbin  
 Zhang (Shanghai University), Xiangyu Zheng (Shanghai University), and  
 A.N. Rajagopalan (Indian Institute of Technology Madras)

NTIRE 2025 Challenge on Video Quality Enhancement for Video Conferencing: Datasets, Methods and Results .....	1175
<i>Varun Jain (Microsoft), Zongwei Wu (University of Wurzburg), Quan Zou (Microsoft), Louis Florentin (Microsoft), Henrik Turbell (Microsoft), Sandeep Siddhartha (Microsoft), Radu Timofte (University of Wurzburg), Qifan Gao (Tencent, China), Linyan Jiang (Tencent, China), Qing Luo (Tencent, China), Jie Song (Xidian University, China), Yaqing Li (Tencent, China), Summer Luo (Independent researcher, China), Mae Chen (Nankai University, China), Stefan Liu (Shanghai Jiao Tong University, China), Danie Song (Shenzhen University, China), Huimin Zeng (Northeastern University, USA), Qi Chen (Johns Hopkins University, USA), Ajeet Kumar Verma (IIT Jammu, India), Shweta Tripathi (IIT Jammu, India), Vinit Jakhetiya (IIT Jammu, India), Badri N Subhdhi (IIT Jammu, India), and Sunil Jaiswal (K\ Lens, GmbH, Germany)</i>	
LFTransMamba: A Hybrid Mamba-Transformer Model for Light Field Image Super-Resolution ..	1186
<i>Kai Jin (Bigo Technology Pte. Ltd.), Zeqiang Wei (School of Intelligent Engineering and Automation, Beijing University of Posts and Telecommunications; Global Explorer Ltd., Suzhou China), Angulia Yang (Bigo Technology Pte. Ltd.), Di Wu (School of Artificial Intelligence, Beijing University of Posts and Telecommunications), Mingzhi Gao (Bigo Technology Pte. Ltd.), and Xiuzhuang Zhou (School of Intelligent Engineering and Automation, Beijing University of Posts and Telecommunication; Beijing Ketai Industrial Intelligence (BKII), China)</i>	

NTIRE 2025 Challenge on Low Light Image Enhancement: Methods and Results .....	1196
--	------

Xiaoning Liu (University of Electronic Science and Technology of China), Zongwei Wu (University of Wurzburg), Florin-Alexandru Vasluianu (University of Wurzburg), Hailong Yan (University of Electronic Science and Technology of China), Bin Ren (University of Trento), Yulun Zhang (Shanghai Jiao Tong University), Shuhang Gu (University of Electronic Science and Technology of China), Le Zhang (University of Electronic Science and Technology of China), Ce Zhu (University of Electronic Science and Technology of China), Radu Timofte (University of Wurzburg), Kangbiao Shi (School of Computer Science, Northwestern Polytechnical University, China), Yixu Feng (School of Computer Science, Northwestern Polytechnical University, China), Tao Hu (School of Computer Science, Northwestern Polytechnical University, China), Yu Cao (Xi'an Institute of Optics and Precision Mechanics of CAS, China), Peng Wu (School of Computer Science, Northwestern Polytechnical University, China), Yijin Liang (Shanghai Institute of Satellite Engineering, Shanghai 201109, China), Yanning Zhang (School of Computer Science, Northwestern Polytechnical University, China), Qingsen Yan (School of Computer Science, Northwestern Polytechnical University, China), Han Zhou (McMaster University), Wei Dong (McMaster University), Yan Min (McMaster University), Mohab Kishawy (McMaster University), Jun Chen (McMaster University), Pengpeng Yu (School of Electronics and Communication Engineering, Sun Yat-sen University, China; Pengcheng Laboratory, China), Anjin Park (Intelligent Technology Research Center, Korea Photonics Technology Institute, Korea), Seung-Soo Lee (Doowon Electronics and Telecom Co., LTD., Korea), Young-Joon Park (Doowon Electronics and Telecom Co., LTD., Korea), Zixiao Hu (University of Electronic Science and Technology of China, China), Junyu Liu (University of Electronic Science and Technology of China, China), Huilin Zhang (University of Electronic Science and Technology of China, China), Jun Zhang (University of Electronic Science and Technology of China, China), Fei Wan (Beijing Union University), Bingxin Xu (Beijing Union University), Hongzhe Liu (Beijing Union University), Cheng Xu (Beijing Union University), Weiguo Pan (Beijing Union University), Songyin Dai (Beijing Union University), Xunpeng Yi (Electronic Information School, Wuhan University, China), Qinglong Yan (Electronic Information School, Wuhan University, China), Yibing Zhang (Electronic Information School, Wuhan University, China), Jiayi Ma (Electronic Information School, Wuhan University, China), Changhui Hu (Nanjing University of Posts and Telecommunications), Kerui Hu (Nanjing University of Posts and Telecommunications), Donghang Jing (Nanjing University of Posts and Telecommunications), Tiesheng Chen (Nanjing University of Posts and Telecommunications), Zhi Jin (Shenzhen Campus of Sun Yat-sen University), Hongjun Wu (Shenzhen Campus of Sun Yat-sen University), Biao Huang (Shenzhen Campus of Sun Yat-sen University), Haitao Ling (Shenzhen Campus of Sun Yat-sen University), Jiahao Wu (Shenzhen Campus of Sun Yat-sen University), Dandan Zhan (Shenzhen Campus of Sun Yat-sen University), G Gyaneshwar Rao (KLE Technological University), Vijayalaxmi Ashok Aralikatti (KLE Technological University), Nikhil Akalwadi (KLE Technological University), Ramesh

Ashok Tabib (KLE Technological University), Uma Mudanagudi (KLE Technological University), Ruirui Lin (University of Bristol), Guoxi Huang (University of Bristol), Nantheera Anantrasirichai (University of Bristol), Qirui Yang (Tianjin University), Alexandru Brateanu (University of Manchester), Ciprian Orhei (Polytechnic University Timisoara), Cosmin Ancuti (Polytechnic University Timisoara), Daniel Feijoo (Cidaut AI), Juan C. Benito (Cidaut AI), Alvaro Garcia (Cidaut AI), Marcos V. Conde (University of Wurzburg), Yang Qin (Nanjing University of Information Science and Technology), Raul Balmez (University of Manchester), Anas M. Ali (Prince Sultan University), Bilel Benjdira (Prince Sultan University), Wadii Boulila (Prince Sultan University), Tianyi Mao (Hefei University of Technology), Huan Zheng (University of Macau), Yanyan Wei (Hefei University of Technology), Shengeng Tang (Hefei University of Technology), Dan Guo (Hefei University of Technology), Zhao Zhang (Hefei University of Technology), Sabari Nathan (Couger In ,Japan), K Uma (Sasi Institute of Technology & Engineering, India), A Sasithradevi (Vellore Institute of Technology, India), B Sathya Bama (Thiagarajar college of engineering, India), S. Mohamed Mansoor Roomi (Thiagarajar college of engineering, India), Ao Li (University of Electronic Science and Technology of China), Xiangtao Zhang (University of Electronic Science and Technology of China), Zhe Liu (University of Electronic Science and Technology of China), Yijie Tang (Chongqing University), Jialong Tang (Zhejiang University), Zhicheng Fu (Lenovo Research), Gong Chen (Lenovo Research), Joe Nasti (Lenovo Research), John Nicholson (Lenovo Research), Zeyu Xiao (National University of Singapore), Zhuoyuan Li (University of Science and Technology of China), Ashutosh Kulkarni (Indian Institute of Technology Ropar), Prashant W. Patil (Indian Institute of Technology Guwahati), Santosh Kumar Vipparthi (Indian Institute of Technology Ropar), Subrahmanyam Murala (Trinity College Dublin), Duan Liu (Jinhua University of Vocational Technology), Weile Li (Jinhua University of Vocational Technology), Hangyuan Lu (Jinhua University of Vocational Technology), Rixian Liu (Jinhua University of Vocational Technology), Tengfeng Wang (Chongqing University of Posts and Telecommunications), Jinxing Liang (Wuhan Textile University), and Chenxin Yu (Wuhan Textile University)

#### Frequency-Prior Enhanced Ambient Lighting Normalization via Visual Perceptual Refinement ... 1207

Yuanfei Bao (University of Science and Technology of China, China), Xin Lu (University of Science and Technology of China, China), Xingbo Wang (University of Science and Technology of China, China), Jiarong Yang (University of Science and Technology of China, China), Anya Hu (University of Science and Technology of China, China), Kunyu Wang (University of Science and Technology of China, China), Jie Xiao (University of Science and Technology of China, China), Dong Li (University of Science and Technology of China, China), Xueyang Fu (University of Science and Technology of China, China), and Zheng-Jun Zha (University of Science and Technology of China, China)

NTIRE 2025 Challenge on Light Field Image Super-Resolution: Methods and Results .....	1218
<p>Yingqian Wang (National University of Defense Technology), Zhengyu Liang (National University of Defense Technology), Fengyuan Zhang (National University of Defense Technology), Loli Tian (National University of Defense Technology), Longguang Wang (Aviation University of Air Force), Juncheng Li (Shanghai University), Jungang Yang (National University of Defense Technology), Radu Timofte (University of Würzburg), Yulan Guo (Sun Yat-sen University), Kai Jin, Zeqiang Wei, Angulia Yang, Di Wu, Mingzhi Gao, Xiuzhuang Zhou, Yue Yan, Yuaho Wang, Shuang Chen, Zeping Tian, Yizhi Hu, Yao Lu, Haosong Liu, Xiancheng Zhu, Huanqiang Zeng, Jianqing Zhu, Yifan Shi, Junhui Hou, Mingyang Yu, Zhijian Wu, Dingjiang Huang, Wenli Zheng, Zekai Xu, Huiyuan Fu, Heng Zhang, Zhijuan Huang, Hongyuan Yu, Zeke Zexi Hu, Haodong Chen, Vera Yuk Ying Chung, Xiaoming Chen, Zean Chen, Yeyao Chen, Gangyi Jiang, Haiyong Xu, Ting Luo, Guanglong Liao, Danhao Zhang, Siyu Zhang, Wendong Mao, Zhongfeng Wang, Sunita Arya, Abhishek Kumar Sinha, S Manthira Moorthi, Hao Zhang, Hao Sheng, Da Yang, Zhenglong Cui, Shuai Wang, Haotian Zhang, Xingzheng Wang, Yuanbo Huang, Jiahao Lin, Yuhang Lin, Ahmed Salem, Ebrahim Elkady, Hatem Ibrahim, Jae-Won Suh, Hyun-Soo Kang, Changguang Wu, Hao Hou, Pengpeng Li, Peng Huang, Jiangxin Dong, and Jinhui Tang</p>	
Q-CIDNet: Perceptual Quality aware Color and Intensity Decoupling Network for Video Quality Enhancement .....	1238
<p>Ajeet Kumar Verma (Indian Institute of Technology Jammu), Shweta Tripathi (Indian Institute of Technology Jammu), Vinit Jakhetiya (Indian Institute of Technology Jammu), Badri N Subudhi (Indian Institute of Technology Jammu), and Sunil Jaiswal (K Lens GmbH)</p>	
RAW Image Reconstruction from RGB on Smartphones. NTIRE 2025 Challenge Report .....	1245
<p>Marcos Conde (University of Würzburg), Radu Timofte (University of Würzburg), Radu Berdan (Sony AI), Beril Besbinar (Sony AI), and Daisuke Iso (Sony AI)</p>	
NTIRE 2025 the 2nd Restore Any Image Model (RAIM) in the Wild Challenge .....	1260
<p>Jie Liang (OPPO Research Institute), Radu Timofte (University of Würzburg &amp; ETH Zurich), Qiaosi Yi (The Hong Kong Polytechnic University), Zhengqiang Zhang (The Hong Kong Polytechnic University), Shuaizheng Liu (The Hong Kong Polytechnic University), Lingchen Sun (The Hong Kong Polytechnic University), Rongyuan Wu (The Hong Kong Polytechnic University), Xindong Zhang (OPPO Research Institute), Hui Zeng (OPPO Research Institute), Lei Zhang (Hong Kong Polytechnic University, Hong Kong, China), Tianyu Hao, Lin Wang, Zhe Xiao, Pengzhou Ji, Shu-Peng Zhong, Xiangming Wang, Jiaqi Yan, Sishun Pan, Ce Wang, Yibin Huang, Zhang Sheng Wang, Haobo Liang, Zhenghao Pan, Jinjian Wu, Yushen Zuo, and Yuanbo Zhou</p>	
TokenFocus-VQA: Enhancing Text-to-Image Alignment with Position-Aware Focus and Multi-Perspective Aggregations on LVLMS .....	1270
<p>Zijian Zhang (Meituan Inc), Xuhui Zheng (Nanjing University), Xuecheng Wu (Xi'an Jiaotong University), Chong Peng (Meituan Inc), and Xuezhi Cao (Meituan Inc)</p>	

NTIRE 2025 Ambient Lighting Normalization Challenge Report .....	1280
--	------

Florin-Alexandru Vasluianu (Computer Vision Lab, University of Wurzburg), Tim Seizinger (Computer Vision Lab, University of Wurzburg), Zhuyun Zhou (CNRS - Université de Bourgogne), Zongwei Wu (Computer Vision Lab, University of Wurzburg), Radu Timofte (Computer Vision Lab, University of Wurzburg), Yuanfei Bao (University of Science and Technology of China (USTC)), Xingbo Wang (University of Science and Technology of China (USTC)), Xin Lu (University of Science and Technology of China (USTC)), Jiarong Yang (University of Science and Technology of China (USTC)), Anya Hu (University of Science and Technology of China (USTC)), Kunyu Wang (University of Science and Technology of China (USTC)), Jie Xiao (University of Science and Technology of China (USTC)), Dong Li (University of Science and Technology of China (USTC)), Xueyang Fu (University of Science and Technology of China (USTC)), Zheng-Jun Zha (University of Science and Technology of China (USTC)), Zihao Fan (University of Science and Technology of China (USTC)), Xi Wang (University of Science and Technology of China (USTC)), Yurui Zhu (University of Science and Technology of China (USTC)), Kean Liu (University of Science and Technology of China (USTC)), Senyan Xu (University of Science and Technology of China (USTC)), Hongjian Liu (University of Science and Technology of China (USTC)), Yupeng Xiao (University of Science and Technology of China (USTC)), David Serrano-Lozano (Computer Vision Center and Universitat Autònoma de Barcelona), Francisco A. Molina-Bakhos (Computer Vision Center and Universitat Autònoma de Barcelona), Danna Xue (Computer Vision Center and Universitat Autònoma de Barcelona), Yixiong Yang (Computer Vision Center and Universitat Autònoma de Barcelona), Maria Pilligua (Computer Vision Center and Universitat Autònoma de Barcelona), Ramon Baldrich (Computer Vision Center and Universitat Autònoma de Barcelona), Maria Vanrell (Computer Vision Center and Universitat Autònoma de Barcelona), Javier Vazquez-Corral (Computer Vision Center and Universitat Autònoma de Barcelona), Xuan Sun (MT Lab, Meitu Inc.), Zijie Lou (MT Lab, Meitu Inc.), Ting Liu (MT Lab, Meitu Inc.), Kuldeep Purohit (Independent Researchers), Jameer Babu Pinjari (Independent Researchers), Yilin Zhang (Hefei University of Technology), Huan Zheng (University of Macau), Yanyan Wei (Hefei University of Technology), Suiyi Zhao (Hefei University of Technology), Shengeng Tang (Hefei University of Technology), Zhao Zhang (Hefei University of Technology), Yushen Zuo (The Hong Kong Polytechnic University), Zongqi He (The Hong Kong Polytechnic University), Zhe Xiao (The Hong Kong Polytechnic University), Cuixin Yang (The Hong Kong Polytechnic University), Rongkang Dong (The Hong Kong Polytechnic University), Jun Xiao (The Hong Kong Polytechnic University), Kin-Man Lam (The Hong Kong Polytechnic University), Nikhil Akalwadi (School of Computer Science and Engineering, KLE Technological University), Vijayalaxmi Ashok Aralikatti (School of Computer Science and Engineering, KLE Technological University), Dheeraj Damodhar Hegde (School of Electronics and Communication Engineering, KLE Technological University), Ramesh Ashok Tabib (School of Electronics and Communication Engineering, KLE Technological University), Uma Mudenagudi (School of Electronics and Communication Engineering, KLE

*Technological University), Anas M. Ali (Robotics and Internet-of-Things Laboratory, Prince Sultan University), Bilel Benjdira (Robotics and Internet-of-Things Laboratory, Prince Sultan University), and Wadii Boulila (Robotics and Internet-of-Things Laboratory, Prince Sultan University)*

NTIRE 2025 Challenge on Single Image Reflection Removal in the Wild: Datasets, Methods and Results ..... 1292

*Kangning Yang (OPPO AI Center), Jie Cai (OPPO AI Center), Ling Ouyang (OPPO AI Center), Florin-Alexandru Vasluianu (University of Würzburg), Radu Timofte (University of Würzburg), Jiaming Ding (OPPO AI Center), Huiming Sun (OPPO AI Center), Lan Fu (OPPO AI Center), Jinlong Li (OPPO AI Center), Chiu Man Ho (OPPO AI Center), Zibo Meng (OPPO AI Center), Mingjia Li (Tianjin University), Hainuo Wang (Tianjin University), Qiming Hu (Tianjin University), Jiarui Wang (Tianjin University), Hao Zhao (Tianjin University), Jin Hu (Tianjin University), Xiaojie Guo (Tianjin University), Mengru Yang (Harbin Institute of Technolog), Jing He (Xidian University), Yiqing Wang (Xidian University), Zhiyang Chen (Shandong University), Hao Fang (Shandong University), Wei Zhang (Shandong University), Runmin Cong (Shandong University), Dheeraj Damodhar Hegde (KLE Technological University), Jatin Kalal (KLE Technological University), Nikhil Akalwadi (KLE Technological University), Ramesh Ashok Tabib (KLE Technological University), Uma Mudenagudi (KLE Technological University), Yu-Fan Lin (National Cheng Kung University), Chia-Ming Lee (National Cheng Kung University), Chih-Chung Hsu (National Cheng Kung University), Mengxin Zhang (Beijing meitu company), Sabari Nathan (Couger Inc), K Uma (Sasi Institute of Technology & Engineering), A Sasithradevi (Vellore Institute of Technology), B Sathya Bama (Thiagarajar College of engineering), S. Mohamed Mansoor Roomi (Thiagarajar College of engineering), Bilel Benjdira (Prince Sultan University), Anas M. Ali (Prince Sultan University), Wadii Boulila (Prince Sultan University), Wei Dong (McMaster University), Yunzhe Li (McMaster University), Ali Hussein (McMaster University), Han Zhou (McMaster University), Jun Chen (McMaster University), Zeyu Xiao (National University of Singapore), and Zhuoyuan Li (University of Science and Technology of China)*

NTIRE 2025 Image Shadow Removal Challenge Report .....	1303
--	------

Florin-Alexandru Vasluianu (University of Würzburg), Tim Seizinger (University of Würzburg), Zhuyun Zhou (CNRS - Université de Bourgogne), Cailan Chen (Shanghai Jiao Tong University), Zongwei Wu (University of Würzburg), Radu Timofte (University of Würzburg), Mingjia Li (Tianjin University), Jin Hu (Tianjin University), Hainuo Wang (Tianjin University), Hengxing Liu (Tianjin University), Jiarui Wang (Tianjin University), Qiming Hu (Tianjin University), Xiaojie Guo (Tianjin University), Xin Lu (University of Science and Technology of China, Hefei), Jiarong Yang (University of Science and Technology of China, Hefei), Yuanfei Bao (University of Science and Technology of China, Hefei), Anya Hu (University of Science and Technology of China, Hefei), Zihao Fan (University of Science and Technology of China, Hefei), Kunyu Wang (University of Science and Technology of China, Hefei), Jie Xiao (University of Science and Technology of China, Hefei), Xi Wang (University of Science and Technology of China, Hefei), Xueyang Fu (University of Science and Technology of China, Hefei), Zheng-Jun Zha (University of Science and Technology of China, Hefei), Yu-Fan Lin (National Cheng Kung University), Chia-Ming Lee (National Cheng Kung University), Chih-Chung Hsu (National Yang Ming Chiao Tung University), Xingbo Wang (University of Science and Technology of China, Hefei), Dong Li (University of Science and Technology of China, Hefei), Yuxu Chen (Fuzhou University), Bin Chen (Fuzhou University), Yuanbo Zhou (Fuzhou University), Yuanbin Chen (Fuzhou University), Hongwei Wang (Fuzhou University), Jiannan Lin (Fuzhou University), Qinquan Gao (Fuzhou University), Tong Tong (Fuzhou University), Zhao Zhang (Hefei University of Technology), Yanyan Wei (Hefei University of Technology), Wei Dong (McMaster University), Han Zhou (McMaster University), Seyed Amirreza Mousavi (McMaster University), Jun Chen (McMaster University), Haobo Liang (Lanzhou Jiaotong University), Jiajie Jing (Lanzhou Jiaotong University), Junyu Li (Lanzhou Jiaotong University), Yan Yang (Lanzhou Jiaotong University), Seoyeon Lee (Kookmin University), Chaewon Kim (Kookmin University), Ziyu Feng (Beijing Jiaotong University), Shidi Chen (Beijing Jiaotong University), Bowen Luan (Beijing Jiaotong University), Zewen Chen (Beijing Jiaotong University), Vijayalaxmi Ashok Aralikatti (KLE Technological University), G Gyaneshwar Rao (KLE Technological University), Nikhil Akalwadi (KLE Technological University), Chaitra Desai (KLE Technological University), Ramesh Ashok Tabib (KLE Technological University), Uma Mudenagudi (KLE Technological University), Anas M. Ali (Prince Sultan University), Bilel Benjdira (Prince Sultan University), Wadii Boulila (Prince Sultan University), Alexandru Brateanu (University of Manchester), Cosmin Ancuti (Politehnica University Timisoara), Tanmay Chaturvedi (Sardar Vallabhbhai National Institute of Technology, Surat), Manish Kumar (Sardar Vallabhbhai National Institute of Technology, Surat), Anmol Srivastav (Sardar Vallabhbhai National Institute of Technology, Surat), Daksh Trivedi (Sardar Vallabhbhai National Institute of Technology, Surat), Shashwat Thakur (Sardar Vallabhbhai National Institute of Technology, Surat), Kishor Upla (Sardar Vallabhbhai National Institute of Technology, Surat), Zeyu Xiao (National University of Singapore), Zhuoyuan Li (University of Science and

Technology of China, Hefei), Boda Zhou (Independent), Shashank Shekhar (Jawaharlal Nehru University), Kele Xu (National University of Defense Technology), Qisheng Xu (National University of Defense Technology), Zijian Gao (National University of Defense Technology), Tianjiao Wan (National University of Defense Technology), Suiyi Zhao (Hefei University of Technology), Bo Wang (Hefei University of Technology), Yan Luo (Hefei University of Technology), Mingshen Wang (Hefei University of Technology), and Yilin Zhang (Hefei University of Technology)

NTIRE 2025 Challenge on Event-Based Image Deblurring: Methods and Results ..... 1315

Lei Sun (INSAIT, Sofia University "St. Kliment Ohridski"), Andrea Alfarano (INSAIT, Sofia University "St. Kliment Ohridski"), Peiqi Duan (Peking University), Shaolin Su (Computer Vision Center), Kaiwei Wang (Zhejiang University), Boxin Shi (Peiking University), Radu Timofte (University of Wurzburg), Danda Pani Paudel (INSAIT, Sofia University "St. Kliment Ohridski"), Luc Van Gool (INSAIT, Sofia University "St. Kliment Ohridski"), Qinglin Liu, Wei Yu, Xiaoquian Lv, Lu Yang, Shuigen Wang, Shengping Zhang, Xiangyang Ji, Long Bao, Yuqiang Yang, Jinao Song, Ziyi Wang, Shuang Wen, Heng Sun, Kean Liu, Mingchen Zhong, Senyan Xu, Zhijing Sun, Jiaying Zhu, Chengjie Ge, Xingbo Wang, Yidi Liu, Xin Lu, Xueyang Fu, Zheng-Jun Zha, Dawei Fan, Dafeng Zhang, Yong Yang, Siru Zhang, Qinghua Yang, Hao Kang, Huiyuan Fu, Heng Zhang, Hongyuan Yu, Zhijuan Huang, Shouyan Wei, Feng Li, Runming Cong, Weiqi Luo, Mingyun Lin, Chenxu Jiang, Hongyi Liu, Lei Yu, Weilun Li, Jiajun Zhai, Tingting Lin, Shuang Ma, Sai Zhou, Zhanwen Liu, Yang Wang, Eiffel Chong, Nuwan Bandara, Thivya Kandappu, Archan Misra, Yihang Chen, Zhan Li, Weijun Yuan, Wenzhuo Wang, Boyang Yao, Zhanglu Chen, Yijing Sun, Tianjiao Wan, Zijian Gao, Qisheng Xu, Kele Xu, Yukun Zhang, Yu He, Xiaoyan Xie, Tao Fu, Yashu Guatamkumar Patel, Vihar Ramesh Jain, Divesh Basina, Rishik Ashili, Manish Kumar Manjhi, Sourav Kumar, Prinon Benny, Himanshu Ghunawat, B Sri Sairam Gautam, Anett Varghese, and Abhishek Yadav

The Tenth NTIRE 2025 Image Denoising Challenge Report ..... 1333

Lei Sun (INSAIT, Sofia University "St. Kliment Ohridski"), Hang Guo (Tsinghua University), Bin Ren (University of Trento), Luc Van Gool (INSAIT, Sofia University "St. Kliment Ohridski"), Radu Timofte (University of Wurzburg), and Yawei Li (ETH Zurich)

STRNet: Semantics-guided Two-stage Raindrop Removal Network ..... 1361

Qiyu Rong (Beijing Union University), Hongyuan Jing (Beijing Union University), Mengmeng Zhang (Beijing Union University), Jinlong Li (Beijing Union University), and Mengfei Han (Beijing Union University)

Instruction-augmented Multimodal Alignment for Image-Text and Element Matching ..... 1370

Xinli Yue (Wuhan University), JianHui Sun (WeChat), Junda Lu (WeChat), Liangchao Yao (WeChat), Fan Xia (WeChat), Tianyi Wang (WeChat), Fengyun Rao (WeChat), Jing Lyu (WeChat), and Yuetang Deng (WeChat)

NTIRE 2025 XGC Quality Assessment Challenge: Methods and Results .....	1380
--	------

Xiaohong Liu (Shanghai Jiao Tong University), Xiongkuo Min (Shanghai Jiao Tong University), Qiang Hu (Shanghai Jiao Tong University), Xiaoyun Zhang (Shanghai Jiao Tong University), Jie Guo (Peng Cheng Laboratory), Guangtao Zhai (Shanghai Jiao Tong University), Shushi Wang (Shanghai Jiao Tong University), Yingjie Zhou (Shanghai Jiao Tong University), Lu Liu (Shanghai Jiao Tong University), Jingxin Li (China National Information Technology Standardization Committee Multimedia Subcommittee), Liu Yang (Shanghai Jiao Tong University), Farong Wen (Shanghai Jiao Tong University), Li Xu (Peng Cheng Laboratory), Yanwei Jiang (Shanghai Jiao Tong University), Xilei Zhu (Shanghai Jiao Tong University), Chunyi Li (Shanghai Jiao Tong University), Zicheng Zhang (Shanghai Jiao Tong University), Huiyu Duan (Shanghai Jiao Tong University), Xiele Wu (Shanghai Jiao Tong University), Yixuan Gao (Shanghai Jiao Tong University), Yubin Cao (Shanghai Jiao Tong University), Jun Jia (Shanghai Jiao Tong University), Wei Sun (Shanghai Jiao Tong University), Jiezhong Cao (Harvard University), Radu Timofte (University of Würzburg), Baojun Li (shopline), Jiamian Huang (shopline), Dan Luo (shopline), Tao Liu (shopline), Weixia Zhang (Shanghai Jiao Tong University), Bingkun Zheng (Shanghai Jiao Tong University), Junlin Chen (Shanghai Jiao Tong University), Ruikai Zhou (Xiaomi Technology Co., Ltd), Meiya Chen (Xiaomi Technology Co., Ltd), Yu Wang (Xiaomi Technology Co., Ltd), Hao Jiang (Xiaomi Technology Co., Ltd), Xiantao Li (JD), Yuxiang Jiang (JD), Jun Tang (JD), Yimeng Zhao (Chongqing University of Posts and Telecommunications), Bo Hu (Chongqing University of Posts and Telecommunications), Zelu Qi (Communication University of China), Chaoyang Zhang (Communication University of China), Fei Zhao (Communication University of China), Ping Shi (Communication University of China), Lingzhi Fu (Netease), Heng Cong (Netease), Shuai He (Netease), Rongyu Zhang (Netease), Jiarong He (Netease), Zongyao Hu (Beijing Institute of Technology), Wei Luo (University of Science and Technology of China), Zihao Yu (University of Science and Technology of China), Fengbin Guan (University of Science and Technology of China), Yiting Lu (University of Science and Technology of China), Xin Li (University of Science and Technology of China), Zhibo Chen (University of Science and Technology of China), Mengjing Su (University of Mining and Technology China), Yi Wang (University of Mining and Technology China), Tuo Chen (University of Mining and Technology China), Chunxiao Li (University of Mining and Technology China), Shuaiyu Zhao (University of Mining and Technology China), Jiaxin Wen (University of Mining and Technology China), Chuyi Lin (University of Mining and Technology China), Sitong Liu (University of Mining and Technology China), Ningxin Chu (University of Mining and Technology China), Jing Wan (University of Mining and Technology China), Yu Zhou (University of Mining and Technology China), Baoying Chen (Alibaba Group), Jishen Zeng (Alibaba Group), Jiarui Liu (Alibaba Group), Xianjin Liu (Alibaba Group), Xin Chen (AutoHome), Lanzhi Zhou (Beihang University China), Hangyu Li (AutoHome), You Han (AutoHome), Bibo Xiang (AutoHome), Zhenjie Liu (University of Science and Technology of China), Jianzhang Lu (University of Science and Technology of China), Jialin Gui (University of Science and

Technology of China), Renjie Lu (University of Science and Technology of China), Shangfei Wang (University of Science and Technology of China), Donghao Zhou (The Chinese University of Hong Kong), Jingyu Lin (Monash University), Quanjian Song (Monash University), Jiancheng Huang (Shenzhen Institute of Advanced Technology), Yufeng Yang (Southern University of Science and Technology), Changwei Wang (Shandong Computer Science Center), Shupeng Zhong (Nanjing University of Science and Technology), Yang Yang (Nanjing University of Science and Technology), Lihuo He (School of Electronic Engineering, Xidian University), Jia Liu (School of Electronic Engineering, Xidian University), Yuting Xing (School of Electronic Engineering, Xidian University), Tida Fang (School of Electronic Engineering, Xidian University), and Yuchun Jin (School of Electronic Engineering, Xidian University)	
An Empirical Study for Efficient Video Quality Assessment .....	1394
Wei Sun (East China Normal University), Kang Fu (Shanghai Jiao Tong University), Linhan Cao (Shanghai Jiao Tong University), Dandan Zhu (East China Normal University), Kaiwei Zhang (Shanghai Jiao Tong University), Yucheng Zhu (Shanghai Jiao Tong University), Zicheng Zhang (Shanghai Jiao Tong University), Menghan Hu (East China Normal University), Xiongkuo Min (Shanghai Jiao Tong University), and Guangtao Zhai (Shanghai Jiao Tong University)	
Quality Assessment for Talking Head Videos via Multi-modal Feature Representation .....	1405
Mengjing Su (China University of Mining and Technology, Jiangsu), Yi Wang (China University of Mining and Technology, Jiangsu), Tuo Chen (China University of Mining and Technology, Jiangsu), Chunxiao Li (China University of Mining and Technology, Jiangsu), Shuaiyu Zhao (China University of Mining and Technology, Jiangsu), Jiaxin Wen (China University of Mining and Technology, Jiangsu), Chuyi Lin (China University of Mining and Technology, Jiangsu), Sitong Liu (China University of Mining and Technology, Jiangsu), Ningxin Chu (China University of Mining and Technology, Jiangsu), and Yu Zhou (China University of Mining and Technology, Jiangsu)	
Enhanced Semantic Extraction and Guidance for UGC Image Super Resolution .....	1412
Yiwen Wang (Shanghai Jiao Tong University, China), Ying Liang (Shanghai Jiao Tong University, China), Yuxuan Zhang (Shanghai Jiao Tong University, China), Xinming Chai (Shanghai Jiao Tong University, China), Zhengxue Cheng (Shanghai Jiao Tong University, China), Yingsheng Qin (Transsion, China), Yucai Yang (Transsion, China), Rong Xie (Shanghai Jiao Tong University, China), and Li Song (Shanghai Jiao Tong University, China)	
Distillation-Supervised Convolutional Low-Rank Adaptation for Efficient Image Super-Resolution .....	1422
Xinming Chai (Shanghai Jiao Tong University, China), Yao Zhang (Shanghai Jiao Tong University, China), Yuxuan Zhang (Shanghai Jiao Tong University, China), Zhengxue Cheng (Shanghai Jiao Tong University, China), Yingsheng Qin (Transsion, China), Yucai Yang (Transsion, China), and Li Song (Shanghai Jiao Tong University, China)	

Pureformer: Transformer-Based Image Denoising .....	1432
<i>Arnim Gautam (Government Engineering College Ajmer), Aditi Pawar (Government Engineering College Ajmer), Aishwarya Joshi (Government Engineering College Ajmer), Satya Narayan Tazi (Government Engineering College Ajmer), Sachin Chaudhary (UPES Dehradun), Praful Hambarde (IIT Mandi), Akshay Dudhane (Mohamed bin Zayed University of Artificial Intelligence), Santosh Vipparthi (IIT Ropar), and Subrahamanyam Murala (IIT Ropar)</i>	
LFMix: A Lightweight Hybrid Architecture for Light Field Super-Resolution .....	1441
<i>Mingyang Yu (East China Normal University), Zhijian Wu (East China Normal University), and Dingjiang Huang (East China Normal University)</i>	
Towards Scale-Aware Low-Light Enhancement via Structure-Guided Transformer Design .....	1451
<i>Wei Dong (McMaster University), Yan Min (McMaster University), Han Zhou (McMaster University), and Jun Chen (McMaster University)</i>	
Retinex-Guided Histogram Transformer for Mask-Free Shadow Removal .....	1462
<i>Wei Dong (McMaster University), Han Zhou (McMaster University), Seyed Amirreza Mousavi (McMaster University), and Jun Chen (McMaster University)</i>	
Event-Conditioned Dual-Modal Fusion for Motion Deblurring .....	1473
<i>Kean Liu (University of Science and Technology of China), Mingchen Zhong (University of Science and Technology of China), Senyan Xu (University of Science and Technology of China), Zhijing Sun (University of Science and Technology of China), Jiaying Zhu (University of Science and Technology of China), Chengjie Ge (University of Science and Technology of China), Xingbo Wang (University of Science and Technology of China), Xin Lu (University of Science and Technology of China), Xueyang Fu (University of Science and Technology of China), and Zheng-Jun Zha (University of Science and Technology of China)</i>	
Towards Holistic Visual Quality Assessment of AI-Generated Videos: A LLM-Based Multi-Dimensional Evaluation Model .....	1484
<i>Zelu Qi (Communication University of China), Ping Shi (Communication University of China), Chaoyang Zhang (Communication University of China), Shuqi Wang (Communication University of China), Fei Zhao (Communication University of China), Da Pan (Communication University of China), and Zefeng Ying (Communication University of China)</i>	

NTIRE 2025 Challenge on UGC Video Enhancement: Methods and Results .....	1494
--	------

*Nickolay Safonov (MSU), Alexey Bryntsev (Lomonosov Moscow State University), Andrey Moskalenko (Lomonosov Moscow State University), Dmitry Kulikov (Lomonosov Moscow State University), Dmitriy Vatolin (Lomonosov Moscow State University), Radu Timofte (University of Würzburg), Haibo Lei (Tencent), Qifan Gao (Tencent), Qing Luo (Tencent), Yaqing Li (Tencent), Jie Song (DeepView), Shaozhe Hao (DeepView), Meisong Zheng (Alibaba), Jingyi Xu (Department of Tao Technology, Alibaba Group, China), Chengbin Wu (Department of Tao Technology, Alibaba Group, China), Jiahui Liu (Department of Tao Technology, Alibaba Group, China), Ying Chen (Department of Tao Technology, Alibaba Group, China), Xin Deng (Beihang University, Beijing, China), Mai Xu (Beihang University, Beijing, China), Peipei Liang (Longyuan (Beijing), Jie Ma (China Telecom Digital Intelligence Technology Co., Ltd., China), Junjie Jin (Key Laboratory of Optical Astronomy, National Astronomical Observatories, Chinese Academy of Sciences, China), Yingxue Pang (University of Science and Technology of China), Fangzhou Luo (Bytedance), Kai Chen, Shijie Zhao (MMLab, ByteDance Inc), Mingyang Wu (Texas A&M University), Renjie Li (Texas A&M University, USA), Yushen Zuo (Texas A&M University, USA), Zhengzhong Tu (Texas A&M University, USA), and Shengyun Zhong (Northeastern University, USA)*

NTIRE 2025 Challenge on Night Photography Rendering .....	1505
---	------

*Egor Ershov (AIRI), Sergey Korchagin (IITP RAS), Alexei Khalin (IITP RAS), Artyom Panshin (IITP RAS), Arseniy Terekhin (IITP RAS), Ekaterina Zaychenkova (IITP RAS), Georgiy Lobarev (IITP RAS), Vsevolod Plokhotnyuk (IITP RAS), Denis Abramov (IITP RAS), Elisey Zhdanov (IITP RAS), Sofia Dorogova (IITP RAS), Yasin Mamedov (IITP RAS), Nikola Banic (Gideon Brothers), Georgii Perevozchikov (University of Wurzburg), Radu Timofte (University of Wurzburg), Lize Zhang (Xiaomi Inc), Yuqian Zhang (Xiaomi Inc), Shuai Liu (Xiaomi Inc), Chaoyu Feng (Xiaomi Inc), Luyang Wang (Xiaomi Inc), Yibin Huang (Xiaomi Inc), Guangqi Shao (Xiaomi Inc), Xiaotao Wang (Xiaomi Inc), Lei Lei (Xiaomi Inc), Sishun Pan (Nanjing University of Science and Technology), Zhiqiang Zhong (Nanjing University of Science and Technology), Yang Yang (Nanjing University of Science and Technology), Anas M. Ali (Prince Sultan University), Hamad Aloqayli (Prince Sultan University), Bilel Benjdira (Prince Sultan University), Wadii Boulila (Prince Sultan University), Ma Xiaoyang (China Agriculture University), Gao Zijun (China Agriculture University), Xing Leyi (Beihang University), Zongqi He (The Hong Kong Polytechnic University), Yushen Zuo (The Hong Kong Polytechnic University), Zhe Xiao (The Hong Kong Polytechnic University), Kin-Chung Chan (The Hong Kong Polytechnic University), Hanmin Li (Sun Yat-sen University), Jun Xiao (The Hong Kong Polytechnic University), Kin-Man Lam (The Hong Kong Polytechnic University), Yunpeng Wu (Nanchang University), Dmitriy Manzura (ITMO University), Daniil Storonkin (ITMO University), Weixin Guo (National University of Defense Technology), Kele Xu (National University of Defense Technology), Qisheng Xu (National University of Defense Technology), Zijian Gao (National University of Defense Technology), Tianjiao Wan (National University of Defense Technology), Buda Vampilov (ITMO University), Furkan Kinli (Ozyegin University), and Furkan Kirac (T-Fashion Inc)*

NTIRE 2025 Challenge on Image Super-Resolution (×4): Methods and Results .....	1516
--	------

Zheng Chen (Shanghai Jiao Tong University), Kai Liu (Shanghai Jiao Tong University), Jue Gong (Shanghai Jiao Tong University), Jingkai Wang (Shanghai Jiao Tong University), Lei Sun (ETH Zurich), Zongwei Wu (University of Wurzburg), Radu Timofte (University of Wurzburg & ETH Zurich), Yulun Zhang (Shanghai Jiao Tong University), Xiangyu Kong (Samsung R&D Institute China - Beijing, China), Xiaoxuan Yu (Samsung R&D Institute China - Beijing, China), Hyunhee Park (The Department of Camera Innovation Group, Samsung Electronics, Republic of Korea), Suejin Han (The Department of Camera Innovation Group, Samsung Electronics, Republic of Korea), Hakjae Jeon (The Department of Camera Innovation Group, Samsung Electronics, Republic of Korea), Dafeng Zhang (Samsung R&D Institute China - Beijing, China), Hyung-Ju Chun (The Department of Camera Innovation Group, Samsung Electronics, Republic of Korea), Donghun Ryou (Seoul National University, Republic of Korea), Inju Ha (Seoul National University, Republic of Korea), Bohyung Han (Seoul National University, Republic of Korea), Lu Zhao (Intsig Information Co., Ltd., China), Yuyi Zhang (Intsig Information Co., Ltd., China), Pengyu Yan (Intsig Information Co., Ltd., China), Jiawei Hu (Intsig Information Co., Ltd., China), Pengwei Liu (Intsig Information Co., Ltd., China), Fengjun Guo (Intsig Information Co., Ltd., China), Hongyuan Yu (Multimedia Department, Xiaomi Inc., China), Pufan Xu (School of Integrated Circuits, Tsinghua University, China), Zhijuan Huang (Multimedia Department, Xiaomi Inc., China), Shuyuan Cui (Huatai Insurance Group Co., Ltd., China), Peng Guo (Hanhai Information Technology (Shanghai) Co., Ltd., China), Jiahui Liu (Multimedia Department, Xiaomi Inc., China), Dongkai Zhang (Multimedia Department, Xiaomi Inc., China), Heng Zhang (Multimedia Department, Xiaomi Inc., China), Huiyuan Fu (Beijing University of Posts and Telecommunications, China), Huadong Ma (Beijing University of Posts and Telecommunications, China), Yanhui Guo (McMaster University, Hamilton, Ontario, Canada), Sisi Tian (Northeastern University, Seattle, Washington, United States), Xin Liu (Nanjing University, China), Jinwen Liang (Nanjing University, China), Jie Liu (Nanjing University, China), Jie Tang (Nanjing University, China), Gangshan Wu (Nanjing University, China), Zeyu Xiao (National University of Singapore, Singapore), Zhuoyuan Li (University of Science and Technology of China, China), Yinxiang Zhang (Northwestern Polytechnical University, Xi'an, China), Wenxuan Cai (Northwestern Polytechnical University, Xi'an, China), Vijayalaxmi Ashok Aralikatti (KLE Technological University, INDIA), Nikhil Akalwadi (KLE Technological University, INDIA), G Gyaneshwar Rao (KLE Technological University, INDIA), Chaitra Desai (KLE Technological University, INDIA), Ramesh Ashok Tabib (KLE Technological University, INDIA), Uma Mudenagudi (KLE Technological University, INDIA), Marcos V. Conde (Fundación CIdaut, Spain), Alejandro Merino (Fundación CIdaut, Spain), Bruno Longarela (Fundación CIdaut, Spain), Javier Abad (Fundación CIdaut, Spain), WeiJun Yuan (Jinan University, Guangzhou, China), Zhan Li (Jinan University, Guangzhou, China), Zhanglu Chen (Jinan University, Guangzhou, China), Boyang Yao (Jinan University, Guangzhou, China), Aagam Jain (Sardar Vallabhbhai National Institute of Technology, India), Milan Kumar Singh (Sardar Vallabhbhai National

Institute of Technology, India), Ankit Kumar (Sardar Vallabhbhai National Institute of Technology, India), Shubh Kawa (Sardar Vallabhbhai National Institute of Technology, India), Divyavardhan Singh (Sardar Vallabhbhai National Institute of Technology, India), Anjali Sarvaiya (Sardar Vallabhbhai National Institute of Technology, India), Kishor Upla (Sardar Vallabhbhai National Institute of Technology, India), Raghavendra Ramachandra (Norwegian University of Science and Technology, Norway), Chia Ming Lee (Institute of Data Science, National Cheng Kung University, Taiwan), Yu-Fan Lin (Institute of Data Science, National Cheng Kung University, Taiwan), Chih-Chung Hsu (Institute of Data Science, National Cheng Kung University, Taiwan), Risheek V Hiremath (KLE Technological University, India), Yashaswini Palani (KLE Technological University, India), Yuxuan Jiang (University of Bristol, United Kingdom), Qiang Zhu (University of Electronic Science and Technology of China, China), Siyue Teng (University of Bristol, United Kingdom), Fan Zhang (University of Bristol, United Kingdom), Shuyuan Zhu (University of Electronic Science and Technology of China, China), Bing Zeng (University of Electronic Science and Technology of China, China), David Bull (University of Bristol, United Kingdom), Jingwei Liao (George Mason University, USA), Yuqing Yang (George Mason University, USA), Wenda Shao (George Mason University, USA), Junyi Zhao (Key Laboratory for Parallel and Distributed Processing, Changsha, China), Qisheng Xu (Key Laboratory for Parallel and Distributed Processing, Changsha, China), Kele Xu (Key Laboratory for Parallel and Distributed Processing, Changsha, China), Sunder Ali Khawaja (Dublin City University, Dublin, Ireland), Ik Hyun Lee (IKLab and Tu Korea, u Korea, Siheung-Si, Republic of Korea), Snehal Singh Tomar (Stony Brook University, USA), Rajarshi Ray (Stony Brook University, USA), Klaus Mueller (Stony Brook University, USA), Sachin Chaudhary (UPES Dehradun, India), Surya Vashisth (Amity University, Punjab, India), Akshay Dudhane (Mohamed bin Zayed University of Artificial Intelligence, Abu Dhabi, United Arab Emirates), Praful Hambarde (Indian Institute of Technology Mandi, India), Satya Naryan Tazi (Government Engineering College Ajmer, India), Prashant Patil (Indian Institute of Technology Guwahati, India), Santosh Kumar Vipparthi (Indian Institute of Technology Ropar, India), Subrahmanyam Murala (Trinity College Dublin, Ireland), Bilel Benjdira (Robotics and Internet-of-Things Laboratory, Prince Sultan University, Riyadh 12435, Saudi Arabia), Anas M. Ali (Robotics and Internet-of-Things Laboratory, Prince Sultan University, Riyadh 12435, Saudi Arabia), Wadii Boulila (Robotics and Internet-of-Things Laboratory, Prince Sultan University, Riyadh 12435, Saudi Arabia), Zahra Moammeri (Cyberspace Research Institute, Shahid Beheshti University, Tehran, Iran), Ahmad Mahmoudi-Aznaveh (Cyberspace Research Institute, Shahid Beheshti University, Tehran, Iran), Ali Karbasi (Cyberspace Research Institute, Shahid Beheshti University, Tehran, Iran), Hossein Motamednia (Institute for Research in Fundamental Sciences, Tehran, Iran), Liangyan Li (McMaster University, Canada), Guanhua Zhao (McMaster University, Canada), Kevin Le (McMaster University, Canada), Yimo Ning (McMaster University, Canada), Haoxuan Huang (McMaster University, Canada), and Jun Chen (McMaster University, Canada)

NTIRE 2025 Challenge on Real-World Face Restoration: Methods and Results .....	1527
--	------

Zheng Chen (Shanghai Jiao Tong University), Jingkai Wang (Shanghai Jiao Tong University), Kai Liu (Shanghai Jiao Tong University), Jue Gong (Shanghai Jiao Tong University), Lei Sun (INSAIT), Zongwei Wu (University of Wurzburg), Radu Timofte (University of Wurzburg & ETH Zurich), Yulun Zhang (Shanghai Jiao Tong University), Jianxing Zhang (Samsung R&D Institute China - Beijing (SRC-B)), Jinlong Wu (Samsung R&D Institute China - Beijing (SRC-B)), Jun Wang (Samsung R&D Institute China - Beijing (SRC-B)), Zheng Xie (Samsung R&D Institute China - Beijing (SRC-B)), Hakjae Jeon (Samsung MX(Mobile eXperience) Business), Suejin Han (Samsung MX(Mobile eXperience) Business), Hyung-Ju Chun (Samsung MX(Mobile eXperience) Business), Hyunhee Park (Samsung MX(Mobile eXperience) Business), Zhicun Yin (Harbin Institute of Technology), Junjie Chen (Harbin Institute of Technology), Ming Liu (Harbin Institute of Technology), Xiaoming Li (Harbin Institute of Technology), Chao Zhou (Shanghai Transsion Co, Ltd), Wangmeng Zuo (Harbin Institute of Technology), Weixia Zhang (Shanghai Jiao Tong University), Dingquan Li (Pengcheng Laboratory), Kede Ma (City University of Hong Kong), Yun Zhang (Xiaomi Inc.), Zhuofan Zheng (Xiaomi Inc.), Yuyue Liu (Xiaomi Inc.), Shizhen Tang (Xiaomi Inc.), Zihao Zhang (Xiaomi Inc.), Yi Ning (Xiaomi Inc.), Hao Jiang (Xiaomi Inc.), Wenjie An (Harbin University of Technology), Kangmeng Yu (Harbin University of Technology), Chenyang Wang (Harbin University of Technology), Kui Jiang (Harbin University of Technology), Xianming Liu (Harbin University of Technology), Junjun Jiang (Harbin University of Technology), Yingfu Zhang (Xidian University), Gang He (Xidian University), Siqi Wang (Xidian University), Kepeng Xu (Xidian University), Zhenyang Liu (Xidian University), Changxin Zhou (Black Sesame Technologies (Singapore) Pte Ltd), Shanlan Shen (Black Sesame Technologies (Singapore) Pte Ltd), Yubo Duan (Black Sesame Technologies (Singapore) Pte Ltd), Yiang Chen (Harbin Institute of Technology), Jin Guo (Harbin Institute of Technology), Mengru Yang (Harbin Institute of Technology), Jen-Wei Lee (Institute of Data Science, National Cheng Kung University), Chia-Ming Lee (Institute of Data Science, National Cheng Kung University), Chih-Chung Hsu (Institute of Data Science, National Cheng Kung University), Hu Peng (Shenzhen International Graduate School, Tsinghua University), and Chunming He (Duke University)

Enhance Then Search: An Augmentation-Search Strategy with Foundation Models for Cross-Domain Few-Shot Object Detection .....	1539
--	------

Jiancheng Pan (Tsinghua University), Yanxing Liu (University of Chinese Academy of Sciences), Xiao He (Wuhan University), Long Peng (University of Science and Technology of China), Jiahao Li (Tsinghua University), Yuze Sun (Tsinghua University), and Xiaomeng Huang (Tsinghua University)

Multi-Dimensional Quality Assessment for UGC Videos via Modular Multi-Modal Vision-Language Models .....	1548
--	------

Weixia Zhang (Shanghai Jiao Tong University), Bingkun Zheng (Shanghai Jiao Tong University), Junlin Chen (Shanghai Jiao Tong University), and Zhihua Wang (City University of Hong Kong)

Instance Feature Caching for Cross-Domain Few-Shot Object Detection .....	1558
<i>Yali Huang (Zhengzhou University, China), Jie Mei (Zhengzhou University, China), Yiming Yang (Zhengzhou University, China), Mi Guo (Zhengzhou University, China), Mingyuan Jiu (Zhengzhou University, China), and Mingliang Xu (Zhengzhou University, China)</i>	

## Navigating the Future: Ensuring Trustworthiness in Multi-Modal Open-World Intelligence

Benchmarking Multi-modal Semantic Segmentation under Sensor Failures: Missing and Noisy Modality Robustness .....	1567
<i>Chenfei Liao (HKUST(GZ)), Kaiyu Lei (XJTU), Xu Zheng (HKUST(GZ) &amp; INSAIT, Sofia University "St. Kliment Ohridski"), Junha Moon (HKUST(GZ)), Zhixiong Wang (HKUST(GZ)), Yixuan Wang (HKUST(GZ)), Danda Pani Paudel (INSAIT, Sofia University "St. Kliment Ohridski"), Luc Van Gool (INSAIT, Sofia University "St. Kliment Ohridski"), and Xuming Hu (HKUST(GZ) &amp; CSE, HKUST)</i>	
A Survey of State of the Art Large Vision Language Models: Benchmark Evaluations and Challenges .....	1578
<i>Zongxia Li (University of Maryland, College Park), Xiyang Wu (University of Maryland, College Park), Hongyang Du (University of Maryland, College Park), Fuxiao Liu (University of Maryland, College Park), Huy Nghiem (University of Maryland, College Park), and Guangyao Shi (University of Southern California)</i>	
Attention-Guided Hierarchical Defense for Multimodal Attacks in Vision-Language Models .....	1598
<i>Long Chen (Guizhou University, China), Yuling Chen (Guizhou University, China), Yun Luo (Guizhou University, China), Hui Dou (Guizhou University, China), and Xinyang Zhong (Guizhou University, China)</i>	
On the Robustness of GUI Grounding Models Against Image Attacks .....	1609
<i>Haoren Zhao (Hangzhou Dianzi University, China), Tianyi Chen (Microsoft Research, United States), and Zhen Wang (Hangzhou Dianzi University, China)</i>	
IBD: Alleviating Hallucinations in Large Vision-Language Models via Image-Biased Decoding ....	1615
<i>Lanyun Zhu (Singapore University of Technology and Design), Deyi Ji (Alibaba), Tianrun Chen (Zhejiang University), Peng Xu (Alibaba), Jieping Ye (Alibaba), and Jun Liu (Lancaster University)</i>	
Multimodal Generalized Category Discovery .....	1625
<i>Yuchang Su (Independent Researcher), Renping Zhou (Independent Researcher), Siyu Huang (Clemson University), Xingjian Li (Carnegie Mellon University), Tianyang Wang (University of Alabama at Birmingham), Ziyue Wang (Carnegie Mellon University), and Min Xu (Carnegie Mellon University)</i>	

Machine Unlearning in Hyperbolic vs. Euclidean Multimodal Contrastive Learning: Adapting Alignment Calibration to MERU .....	1635
<i>Alex Pujol Vidal (Aalborg University, Denmark; Milestone Systems, Denmark), Kamal Nasrollahi (Aalborg University, Denmark; Milestone Systems, Denmark), Thomas B. Moeslund (Aalborg University, Denmark; Pioneer Center for Artificial Intelligence, Denmark), and Sergio Escalera (Aalborg University, Denmark; University of Barcelona and Computer Vision Center, Spain; Pioneer Center for Artificial Intelligence, Denmark)</i>	
Prompt the Missing: Prompt-Based Robust Audio-Visual Classification under Uncertain Modalities .....	1645
<i>Eunju Park (Ewha Womans University, South Korea)</i>	
HARMONY: Hidden Activation Representations and Model Output-Aware Uncertainty Estimation for Vision-Language Models .....	1654
<i>Erum Mushtaq (University of Southern California), Zalan Fabian (University of Southern California), Yavuz Faruk Bakman (University of Southern California), Anil Ramakrishna (Amazon), Mahdi Soltanolkotabi (University of Southern California), and Salman Avestimehr (University of Southern California)</i>	
Vision Language Models for Massive MIMO Semantic Communication .....	1660
<i>Stephen D. Liang (Hewlett Packard Enterprise)</i>	

## 8th International Workshop on Visual Odometry and Computer Vision Applications Based on Location Clues

Splat-SLAM: Globally Optimized RGB-only SLAM with 3D Gaussians .....	1671
<i>Erik Sandström (ETH Zürich), Ganlin Zhang (TUM), Keisuke Tateno (Google), Michael Oechsle (Google), Michael Niemeyer (Google), Youmin Zhang (Rock Universes), Manthan Patel (ETH Zurich), Luc Van Gool (INSAT), Martin Oswald (University of Amsterdam), and Federico Tombari (TUM)</i>	
Salient Object Detection with Dynamic Convolutions .....	1683
<i>Rohit Venkata Sai Dulam (University of Delaware, USA) and Chandra Kambhamettu (University of Delaware, USA)</i>	
Feature Matching in the Dark: Homography-Based RGB-IR Feature Transformation for Low-Light Vision .....	1694
<i>Kyle O'Donnell (University of Delaware) and Chandra Kambhamettu (University of Delaware)</i>	

## Uncertainty Quantification for Computer Vision

Improving Optical Flow and Stereo Depth Estimation by Leveraging Uncertainty-Based Learning Difficulties .....	1703
<i>Jisoo Jeong (Qualcomm AI Research), Hong Cai (Qualcomm AI Research), Jamie Menjay Lin (Qualcomm Technologies, Inc.), and Fatih Porikli (Qualcomm AI Research)</i>	

D-Feat Occlusions: Diffusion Features for Robustness to Partial Visual Occlusions in Object Recognition .....	1713
<i>Rupayan Mallick (Department of Computer Science, Georgetown University, USA), Sibong Dong (Department of Computer Science, Georgetown University, USA), Nataniel Ruiz (Google, USA), and Sarah Adel Bargal (Department of Computer Science, Georgetown University, USA)</i>	
WQLCP: Weighted Adaptive Conformal Prediction for Robust Uncertainty Quantification Under Distribution Shifts .....	1723
<i>Shadi Alijani (University of Victoria, Canada) and Homayoun Najjaran (University of Victoria, Canada)</i>	
The Surprising Utility of Group Partitioning in Improving Conformal Prediction of Visual Classifiers under Distributional Shifts .....	1733
<i>Kowshik Thopalli (Lawrence Livermore National Laboratory), Vivek Narayanaswamy (Lawrence Livermore National Laboratory), and Jayaraman J. Thiagarajan (Apple)</i>	
Uncertainty Quantification for Gradient-based Explanations in Neural Networks .....	1743
<i>Mihir Mulye (Heinrich-Heine Universität Düsseldorf, Germany) and Matias Valdenegro-Toro (University of Groningen, The Netherlands)</i>	

## The 4th Workshop on Federated Learning for Computer Vision

GPT-FL: Generative Pre-trained Model-Assisted Federated Learning .....	1752
<i>Tuo Zhang (University of Southern California, United States), Tiantian Feng (University of Southern California), Samiul Alam (The Ohio State University, United States), Dimitrios Dimitriadis (Amazon, United States), Sunwoo Lee (Inha University, South Korea), Mi Zhang (The Ohio State University, United States), Shrikanth S. Narayanan (University of Southern California, United States), and Salman Avestimehr (University of Southern California, United States)</i>	
FedSECA: Sign Election and Coordinate-wise Aggregation of Gradients for Byzantine Tolerant Federated Learning .....	1762
<i>Joseph Geo Benjamin (Michigan State University, MI, USA), Mothilal Asokan (Mohamed bin Zayed University of Artificial Intelligence (MBZUAI), Abu Dhabi, UAE), Mohammad Yaqub (Mohamed bin Zayed University of Artificial Intelligence (MBZUAI), Abu Dhabi, UAE), and Karthik Nandakumar (Michigan State University, MI, USA; Mohamed bin Zayed University of Artificial Intelligence (MBZUAI), Abu Dhabi, UAE)</i>	
FedCAPR: Federated Camera-Aware Unsupervised Person Re-Identification with Identity-Distributed Equalization for Decentralized Data Clustering .....	1772
<i>Yu-Syuan Tseng (Graduate Institute of Electronics Engineering, National Taiwan University, Taipei, Taiwan), Tzu-Chin Hsu (Graduate Institute of Electronics Engineering, National Taiwan University, Taipei, Taiwan), Chih-Ting Liu (Amazon), and Shao-Yi Chien (Graduate Institute of Electronics Engineering, National Taiwan University, Taipei, Taiwan)</i>	

Forget Less, Learn More: Contrastive-Based Federated Class Incremental Learning with a Low-Dimensional Projection Layer .....	1782
<i>Ensieh Khazaei (University of Toronto) and Dimitrios Hatzinakos (University of Toronto)</i>	
FedAlign: Federated Domain Generalization with Cross-Client Feature Alignment .....	1792
<i>Sunny Gupta (Indian Institute of Technology Bombay, India), Vinay Sutar (Indian Institute of Technology Bombay, India), Varunav Singh (Indian Institute of Technology Bombay, India), and Amit Sethi (Indian Institute of Technology Bombay, India)</i>	
FedDG-MoE: Test-Time Mixture-of-Experts Fusion for Federated Domain Generalization .....	1802
<i>Ahmed Radwan (The University of British Columbia), Mahmoud Soliman (The University of British Columbia), Omar Abdelaziz (The University of British Columbia), and Mohamed Shehata (The University of British Columbia)</i>	
FedCIAL: Federated Color-Invariant Adversarial Learning for Enhancing Fairness and Performance in Skin Lesion Classification .....	1812
<i>Rahmat Izwan Heroza (School of Computer Science and Electronic Engineering, University of Essex), John Q. Gan (School of Computer Science and Electronic Engineering, University of Essex), and Haider Raza (School of Computer Science and Electronic Engineering, University of Essex)</i>	
Sporadic Federated Learning Approach in Quantum Environment to Tackle Quantum Noise .....	1820
<i>Ratun Rahman (The University of Alabama in Huntsville, USA), Atit Pokharel (The University of Alabama in Huntsville, USA), and Dinh C. Nguyen (The University of Alabama in Huntsville, USA)</i>	

## Mobile AI workshop and associated challenges, 5th edition

ActNAS : Generating Efficient YOLO Models using Activation NAS .....	1830
<i>Sudhakar Sah (Deeplite), Ravish Kumar (Deeplite), Darshan C. Ganji (Deeplite), and Ehsan Saboori (Deeplite)</i>	
Robust 6DoF Pose Estimation Against Depth Noise and a Comprehensive Evaluation on a Mobile Dataset .....	1839
<i>Zixun Huang (UC Berkeley), Keling Yao (CMU), Zhihao Zhao (University of California, Los Angeles), Chuanyu Pan (UC Berkeley), and Allen Yang (UC Berkeley)</i>	
RepFC: Universal Structural Reparametrization Block for High Performance, Lightweight Deep Neural Networks. ....	1849
<i>Shambhavi Balamuthu Sampath (BMW Group), Judeson Anthony Fernando (Technical University of Munich), Moritz Thoma (BMW Group), Nael Fasfous (BMW Group), Lukas Frickenstein (BMW Group), Pierpaolo Mori (BMW), Manoj Rohit Vemparala (BMW Group), Alexander Frickenstein (BMW Group), Ulf Schlichtmann (Technical University of Munich), and Walter Stechele (Technical University of Munich)</i>	
PETAH: Parameter Efficient Task Adaptation for Hybrid Transformers .....	1858
<i>Maximilian Augustin (University of Tübingen), Syed Shakib Sarwar (Meta Inc.), Mostafa Elhoushi (Meta Inc.), Yuecheng Li (Meta Inc.), Sai Qian Zhang (New York University), and Barbara De Salvo (Meta Inc.)</i>	

Learned Lightweight Smartphone ISP with Unpaired Data .....	1869
<i>Andrei Arhire (UAIC, Romania) and Radu Timofte (University of Würzburg, Germany)</i>	
Cycle Training with Semi-Supervised Domain Adaptation: Bridging Accuracy and Efficiency for Real-Time Mobile Scene Detection .....	1879
<i>Huu-Phong Phan-Nguyen (Ho Chi Minh University of Information Technology, VNU-HCM, Vietnam), Anh Dao (Michigan State University, USA), Tien-Huy Nguyen (Ho Chi Minh University of Information Technology, VNU-HCM, Vietnam), Tuan Quang (LPL Financial, USA), Huu-Loc Tran (Ho Chi Minh University of Information Technology, VNU-HCM, Vietnam), Tinh-Anh Nguyen-Nhu (Ho Chi Minh University of Technology, VNU-HCM, Vietnam), Huy-Thach Pham (Georgia State University, USA), Quan Nguyen (Posts and Telecommunications Institute of Technology, Hanoi, Vietnam), Hoang Le (York University, Canada ), and Quang-Vinh Dinh (AI VIETNAM Lab)</i>	
FLAR-SVD: Fast and Latency-Aware Singular Value Decomposition for Model Compression .....	1889
<i>Moritz Thoma (BMW Group), Jorge Villasante (Technical University Munich), Emad Aghajanzadeh (BMW Group), Shambhavi Balamuthu Sampath (BMW Group), Pierpaolo Mori (BMW Group), Maximilian Groetzinger (Technical University Munich), Daniil Dylkin (Technical University Munich), Manoj-Rohit Vemparala (BMW Group), Nael Fafous (BMW Group), Alexander Frickenstein (BMW Group), Daniel Mueller-Gritschneider (Technical University Vienna), and Ulf Schlichtmann (Technical University Munich)</i>	
Quantized Image Super-Resolution on Mobile NPUs, Mobile AI 2025 Challenge: Report .....	1899
<i>Andrey Ignatov (ETH Zurich), Georgy Perevozchikov (University of Würzburg), Radu Timofte (University of Würzburg), Zhiyu Zhang, Tianxiao Gao, Yukun Yang, Shiai Zhu, Shihao Wang, Kihwan Yoon, Ganzorig Gankhuyag, Hyeon-Cheol Moon, Taehyun Jeong, Yumi Kim, Suhyeon Lee, Jaehun Baek, Jinwoo Jeong, Eunjun Park, Jun Lee, Heejun Lee, Sungjei Kim, Dafeng Zhang, Yong Yang, Heo Myeong Cheol, Yonghyun Park, Jooho Jeong, Wontae Kim, Kanghwan Lee, Diankai Zhang, Biao Wu, Chengjian Zheng, Shaoli Liu, Si Gao, Ning Wang, Mingshen Wang, Zhao Zhang, Suiyi Zhao, Jinhan Guan, Bo Wang, and Yan Luo</i>	
RGB Photo Enhancement on Mobile GPUs, Mobile AI 2025 Challenge: Report .....	1913
<i>Andrey Ignatov (ETH Zurich), Georgy Perevozchikov (University of Würzburg), Radu Timofte (University of Würzburg), Wu Pan, Song Wang, Dong Zhang, Zhao Ran, Xiaochen Li, Shichang Ju, Diankai Zhang, Biao Wu, Shaoli Liu, Si Gao, Chengjian Zheng, Ning Wang, Yi Feng, Cailu Wan, Xiangji Wu, Hailong Yan, Ao Li, Xiangtao Zhang, Zhe Liu, Ce Zhu, Le Zhang, Jinjie Zhou, Yang Lu, Feng Duo, Runhua Deng, Xuanyu Chen, Shuhui Xie, Guojie Xiao, Zhifeng Wang, Long Peng, and Aiwen Jiang</i>	

Learned Smartphone ISP on Mobile GPUs, Mobile AI 2025 Challenge: Report .....	1925
<i>Andrey Ignatov (ETH Zurich), Georgii Perevozchikov (University of Würzburg), Radu Timofte (University of Würzburg), Cheng Li, Lian Liu, Jun Cao, Heng Sun, Wu Pan, Song Wang, KeQiang Yu, Shuo Liu, HongQin He, ZhenHao Dong, JianKe Chen, DeJun Hao, Keqiang Yu, Tingniao Wang, Xiaoqing Zhou, Dong Zhang, Chunxia Zhang, Jianke Chen, Dejun Hao, Jianguang He, Hailong Yan, Ao Li, Xiangtao Zhang, Zhe Liu, Ce Zhu, Le Zhang, Andrei Arhire, Shuo Liu, Junpyo Seo, Zhenhao Dong, Jianke Chen, Fen Xie, Xiuzhi Fang, Chen Wu, Zhangsheng Wang, Pengbo Zhang, and Jiazi Huang</i>	
CDVS: Compressed Domain On Device Memory Efficient 8K Video SlowMo .....	1938
<i>Jing Li (Samsung Research America), Chengyu Wang (Samsung Research America), Hamid R. Sheikh (Samsung Research America), and SeokJun Lee (Samsung Research America)</i>	
Compressed Domain Multiframe Processing .....	1945
<i>Chengyu Wang (Samsung Research America), Jing Li (Samsung Research America), Saurabh Kumar (Samsung Research India-Bangalore), Seok-Jun Lee (Samsung Research America), and Hamid Sheikh (Samsung Research America)</i>	

## Data Driven Autonomous Driving Simulation

Learning to Drive from a World Model .....	1955
<i>Mitchell Goff (comma.ai), Greg Hogan (comma.ai), George Hotz (comma.ai), Armand du Parc Locmaria (comma.ai), Kacper Raczy (comma.ai), Harald Schäfer (comma.ai), Adeeb Shihadeh (comma.ai), Weixing Zhang (comma.ai), and Yassine Yousfi (comma.ai)</i>	
MOBI: Multimodal Object Inpainting Using Diffusion Models .....	1965
<i>Alexandru Buburuzan (Five AI, The University of Manchester), Anuj Sharma (Five AI), John Redford (Five AI), Puneet K. Dokania (Five AI, University of Oxford), and Romain Mueller (Five AI)</i>	

## 2nd Workshop on Urban Scene Modeling: Where Vision meets Photogrammetry and Graphics

NadirFloorNet: reconstructing multi-room floorplans from a small set of registered panoramic images .....	1976
<i>Giovanni Pintore (CRS4), Uzair Shah (HBKU), Marco Agus (HBKU), and Enrico Gobbetti (CRS4)</i>	
CityGen: Infinite and Controllable City Layout Generation .....	1986
<i>Jie Deng (Zhejiang University), Wenhao Chai (University of Washington), Jianshu Guo (Zhejiang University), Qixuan Huang (Zhejiang University), Junsheng Huang (Zhejiang University), Wenhao Hu (Zhejiang University), Shengyu Hao (Zhejiang University), Jenq-Neng Hwang (University of Washington), and Gaoang Wang (Zhejiang University)</i>	

Proc-GS: Procedural Building Generation for City Assembly with 3D Gaussians .....	1997
<i>Yixuan Li (The Chinese University of Hong Kong), Xingjian Ran (Shanghai Artificial Intelligence Laboratory), Linning Xu (The Chinese University of Hong Kong), Tao Lu (Brown University), Mulin Yu (Shanghai Artificial Intelligence Laboratory), Zhenzhi Wang (The Chinese University of Hong Kong), Yuanbo Xiangli (Cornell University), Dahua Lin (The Chinese University of Hong Kong), and Bo Dai (The University of Hong Kong)</i>	
Texture2LoD3: Enabling LoD3 Building Reconstruction With Panoramic Images .....	2007
<i>Wenzhao Tang (Technical University of Munich), Weihang Li (Technical University of Munich), Xiucheng Liang (National University of Singapore), Olaf Wysocki (Technical University of Munich), Filip Biljecki (National University of Singapore), Christoph Holst (Technical University of Munich), and Boris Jutzi (Technical University of Munich)</i>	
Turin3D: Evaluating Adaptation Strategies under Label Scarcity in Urban LiDAR Segmentation with Semi-Supervised Techniques .....	2018
<i>Luca Barco (Politecnico di Torino, Italy), Giacomo Blanco (LINKS Foundation, Italy), Gaetano Chiriaco (LINKS Foundation, Italy), Alessia Intini (Politecnico di Torino, Italy), Luigi La Riccia (Politecnico di Torino, Italy), Vittorio Scolamiero (Sapienza Università di Roma, Italy), Piero Boccardo (Politecnico di Torino, Italy), Paolo Garza (Politecnico di Torino, Italy), and Fabrizio Dominici (LINKS Foundation, Italy)</i>	
Near-incident detection in railroad environments: lateral distance estimation from train-mounted monocular camera .....	2027
<i>Yilei Wang (Eindhoven University of Technology), Giacomo D'Amicantonio (Eindhoven University of Technology), and Egor Bondarev (Eindhoven University of Technology)</i>	

## The 12th Workshop on Fine-Grained Visual Categorization

FungiTastic: A Multi-Modal Dataset and Benchmark for Image Categorization .....	2037
<i>Lukas Pícek (Inria / University of West Bohemia), Klara Janouskova (Czech Technical University in Prague), Vojtech Cermak (Czech Technical University in Prague), and Jiri Matas (Czech Technical University in Prague)</i>	
Fine-grained Few-Shot Classification with Part Matching .....	2048
<i>Samuel Black (Temple University) and Richard Souvenir (Temple University)</i>	
CYFLOD: Cyclic Filtering and Loss Damping for Alleviating Noisy Labels in Fine-grained Visual Classification .....	2059
<i>Nauman Ullah Gilal ( Hamad Bin Khalifa University), Khaled Al-Thelaya ( Hamad Bin Khalifa University), Fahad Majeed ( Hamad Bin Khalifa University), Zhihe Lu ( Hamad Bin Khalifa University), Sabri Boughorbel ( Hamad Bin Khalifa University), Jens Schneider ( Hamad Bin Khalifa University), and Marco Agus ( Hamad Bin Khalifa University)</i>	

Real-Time Ultra-Fine-Grained Surgical Instrument Classification .....	2070
<i>Md. Atabuzzaman (Virginia Tech), Gino DiMatteo (Virginia Tech), Hani Alomari (Virginia Tech), Chiawei Tang (Virginia Tech), Connor Hale (Carilion Clinic Innovations), Adam E. Goode (Carilion Clinic-Virginia Tech Carilion School of Medicine), David Ryan King (Carilion Clinic), and Chris Thomas (Virginia Tech)</i>	
A Visual RAG Pipeline for Few-Shot Fine-Grained Product Classification .....	2080
<i>Bianca Lamm (Offenburg University, Germany) and Janis Keuper (Offenburg University and University of Mannheim, Germany)</i>	
WildlifeReID-10k: Wildlife re-identification dataset with 10k individual animals .....	2090
<i>Lukas Adam (University of West Bohemia), Vojtech Cermak (Czech Technical University), Kostas Papafitsoros (Queen Mary University of London), and Lukas Pícek (Inria / University of West Bohemia)</i>	
Multi-entity Video Transformers for Fine-Grained Video Representation Learning .....	2101
<i>Matthew Walmer (University of Maryland, College Park), Rose Kanjirathinkal (Meta), Kai Sheng Tai (Meta), Keyur Muzumdar (Meta), Taipeng Tian (Meta), and Abhinav Shrivastava (University of Maryland, College Park)</i>	
Multi-aspect Knowledge Distillation with Large Language Model .....	2112
<i>Taegyeong Lee (UNIST), Jinsik Bang (UNIST), Soyeong Kwon (UNIST), and Taehwan Kim (UNIST)</i>	
Self-Supervised Pretraining for Fine-Grained Plankton Recognition .....	2122
<i>Joona Kareinen (LUT University, Finland), Tuomas Eerola (LUT University, Finland), Kaisa Kraft (Finnish Environment Institute, Finland), Lasse Lensu (LUT University, Finland), Sanna Suikkanen (Finnish Environment Institute, Finland), and Heikki Kälviäinen (LUT University, Finland)</i>	
Combining Vision-Language Models and Weak Supervision for Nuanced Vision Classification Tasks .....	2133
<i>Seyed Mohamad Ali Tousi (University of Missouri - Columbia), Jacket Demby's (University of Missouri - Columbia), Ramy Farag (University of Missouri - Columbia), Gbenga Omotara (University of Missouri - Columbia), and G. N. DeSouza (University of Missouri - Columbia)</i>	
Pseudo-labelling meets Label Smoothing for Noisy Partial Label Learning .....	2143
<i>Darshana Saravanan (IIIT Hyderabad), Naresh Manwani (IIIT Hyderabad), and Vineet Gandhi (IIIT Hyderabad)</i>	
A Fine-grained Artist Identification Method for Authentication and Attribution of Drawings using Hatching Lines .....	2153
<i>Shahrazad Ziaee (Art &amp; AI Lab, Department of Computer Science, Rutgers University), Ahmed Elgammal (Art &amp; AI Lab, Department of Computer Science, Rutgers University), and Marian Mazzone (Department of Art &amp; Architectural History, College of Charleston)</i>	
Predicting butterfly species presence from satellite imagery using soft contrastive regularisation .....	2165
<i>Thijs Van der Plas (Alan Turing Institute, the UK), Stephen Law (UCL, the UK), and Michael JO Pocock (UK Centre for Ecology &amp; Hydrology, the UK)</i>	

Hybrid AI-Physical Modeling for Penetration Bias Correction in X-band InSAR DEMs: A Greenland Case Study .....	2175
<i>Islam Mansour (German Aerospace Center (DLR), Germany, and ETH Zurich, Switzerland), Georg Fischer (German Aerospace Center (DLR), Germany), Ronny Hänsch (German Aerospace Center (DLR), Germany), and Irena Hajnsek (German Aerospace Center (DLR), Germany, and ETH Zurich, Switzerland)</i>	
CoDEX: Combining Domain Expertise for Spatial Generalization in Satellite Image Analysis .....	2185
<i>Abhishek Kuriyal (Ecole nationale des ponts et chaussées), Elliot Vincent (École des ponts ParisTech / Inria / IGN), Mathieu Aubry (Ecole nationale des ponts et chaussées), and Loic Landrieu (Ecole nationale des ponts et chaussées)</i>	
Panopticon: Advancing Any-Sensor Foundation Models for Earth Observation .....	2195
<i>Leonard Waldmann (Technical University of Munich), Ando Shah (University of California Berkeley), Yi Wang (Technical University of Munich), Nils Lehmann (Technical University of Munich), Adam Stewart (Technical University of Munich), Zhitong Xiong (Technical University of Munich), Xiao Xiang Zhu (Technical University of Munich), Stefan Bauer (Helmholtz Munich), and John Chuang (UC Berkeley)</i>	
Bridging Classical and Modern Computer Vision: PerceptiveNet for Tree Crown Semantic Segmentation .....	2206
<i>Georgios Voulgaris (University of Oxford)</i>	
SARFormer - An Acquisition Parameter Aware Vision Transformer for Synthetic Aperture Radar Data .....	2216
<i>Jonathan Prexl (University of the Bundeswehr Munich, Germany), Michael Recla (University of the Bundeswehr Munich, Germany), and Michael Schmitt (University of the Bundeswehr Munich, Germany)</i>	
LADI v2: Multi-label Dataset and Classifiers for Low-Altitude Disaster Imagery .....	2226
<i>Samuel Scheele (Massachusetts Institute of Technology Lincoln Lab), Katherine Picchione (NASA), and Jeffrey Liu (Massachusetts Institute of Technology Lincoln Lab)</i>	
Task-Informed Meta-Learning for Remote Sensing .....	2235
<i>Gabriel Tseng (McGill University), Hannah Kerner (Arizona State University), and David Rolnick (McGill University)</i>	
AerOSeg: Harnessing SAM for Open-Vocabulary Segmentation in Remote Sensing Images .....	2245
<i>Saikat Dutta (IITB-Monash Research Academy), Akhil Vasim (IIT Bombay), Siddhant Gole (IIT Bombay), Hamid Reza Tofighi (Monash University), and Biplab Banerjee (IIT Bombay)</i>	
Distribution Shifts at Scale: Out-of-distribution Detection in Earth Observation .....	2256
<i>Burak Ekim (University of the Bundeswehr), Girmaw Abebe Tadesse (Microsoft AI for Good Research Lab), Caleb Robinson (Microsoft AI for Good Research Lab), Gilles Hacheme (Microsoft AI for Good Research Lab), Michael Schmitt (University of the Bundeswehr Munich), Rahul Dodhia (Microsoft AI for Good Research Lab), and Juan M. Lavista Ferres (Microsoft AI for Good Research Lab)</i>	

EcoWikiRS: Learning Ecological Representation of Satellite Images from Weak Supervision with Species Observations and Wikipedia .....	2266
<i>Valerie Zermatten (EPFL, Switzerland), Javiera Castillo-Navarro (CNAM, France), Pallavi Jain (INRIA, France), Devis Tuia (EPFL, Switzerland), and Diego Marcos (INRIA, France)</i>	
Explainable Physical PolSAR Autoencoders for Soil Moisture Estimation .....	2277
<i>Nikita Basargin (German Aerospace Center (DLR)), Alberto Alonso-González (Technical University of Catalonia (UPC)), and Irena Hajnsek (German Aerospace Center (DLR))</i>	
Detecting Looted Archaeological Sites from Satellite Image Time Series .....	2287
<i>Elliot Vincent (LIGM, Ecole des Ponts, Univ Gustave Eiffel, CNRS, France; Inria Paris; LASTIG, Univ Gustave Eiffel, IGN-ENSG, 94160, Saint-Mande, France), Mehraïl Saroufim (Iconem), Jonathan Chemla (Iconem), Yves Ubelmann (Iconem), Philippe Marquis (DAFA, French archaeological delegation in Afghanistan), Jean Ponce (Department of Computer Science, Ecole normale supérieure (ENS-PSL, CNRS, Inria); Courant Institute of Mathematical Sciences and Center for Data Science, New York University), and Mathieu Aubry (LIGM, Ecole des Ponts, Univ Gustave Eiffel, CNRS, France)</i>	
Scale-Invariant Implicit Neural Representations For Object Counting .....	2299
<i>Siyuan Xu (Texas a&amp;m university), Yucheng Wang (Texas a&amp;m university), Xihaier Luo (Brookhaven National Laboratory), Byung-Jun Yoon (Texas a&amp;m university), and Xiaoning Qian (Texas a&amp;m university)</i>	
Visual Question Answering on Multiple Remote Sensing Image Modalities .....	2310
<i>Hichem Boussaid (Université Paris Cité (LIPADE)), Lucrezia Tosato (Université Paris Cité (LIPADE), DTIS Onera The french aerospace lab), Flora Weissgerber (DTIS Onera The french aerospace lab), Camille Kurtz (Université Paris Cité (LIPADE)), Laurent Wendling (Université Paris Cité (LIPADE)), and Sylvain Lobry (Université Paris Cité (LIPADE))</i>	
Better Coherence, Better Height: Fusing Physical Models and Deep Learning for Forest Height Estimation from Interferometric SAR Data .....	2320
<i>Ragini Bal Mahesh (German Aerospace Center - DLR) and Ronny Hänsch (German Aerospace Center - DLR)</i>	
s2p-hd: Gpu-Accelerated Binocular Stereo Pipeline for Large-Scale Same-Date Stereo .....	2330
<i>Tristan Amadei (ENS Paris-Saclay, Centre Borelli, France), Enric Meinhardt-Llopis (ENS Paris-Saclay, Centre Borelli, France), Carlo de Franchis (Kayrros, France), Jeremy Anger (ENS Paris-Saclay, Centre Borelli, France), Thibaud Ehret (AMIAD, France), and Gabriele Facciolo (ENS Paris-Saclay, Centre Borelli, France)</i>	
Mapping biodiversity at very-high resolution in Europe .....	2340
<i>Cesar Leblanc (INRIA), Lukas Picek (INRIA / University of West Bohemia), Rémi Palard (CIRAD), Benjamin Deneu (WSL), Maximilien Servajean (LIRMM), Pierre Bonnet (CIRAD), and Alexis Joly (INRIA)</i>	

FrogDogNet: Fourier frequency Retained visual prompt Output Guidance for Domain Generalization of CLIP in Remote Sensing .....	2350
<i>Hariseetharam Gunduboina (Indian Institute of Technology Bombay, India), Muhammad Haris Khan (Mohamed Bin Zayed University of Artificial Intelligence, UAE), and Biplab Banerjee (Indian Institute of Technology Bombay, India)</i>	
REJEPa: A Novel Joint-Embedding Predictive Architecture for Efficient Remote Sensing Image Retrieval .....	2364
<i>Shabnam Choudhury (Ph.D IIT Bombay), Yash Salunkhe (B.tech IIT Bombay), Sarthak Mehrotra (B.Tech IIT Bombay), and Biplab Banerjee (Professor IIT Bombay)</i>	
S-EO: A Large-Scale Dataset for Geometry-Aware Shadow Detection in Remote Sensing Applications .....	2374
<i>Elías Masquil (Universidad de la República, Uruguay), Roger Marí (Eurecat, Spain), Thibaud Ehret (AMIAD, France), Enric Meinhardt-Llopis (ENS Paris-Saclay, France), Pablo Musé (Universidad de la República, Uruguay), and Gabriele Facciolo (ENS Paris-Saclay, France)</i>	
TerraMesh: A Planetary Mosaic of Multimodal Earth Observation Data .....	2385
<i>Benedikt Blumenstiel (IBM Research Europe), Paolo Fraccaro (IBM Research Europe), Valerio Marsocci (European Space Agency Φ-lab), Johannes Jakubik (IBM Research Europe), Stefano Maurogiovanni (Forschungszentrum Jülich, University of Iceland), Mikolaj Czerkawski (European Space Agency Φ-lab), Rocco Sedona (Forschungszentrum Jülich), Gabriele Cavallaro (Forschungszentrum Jülich, University of Iceland), Thomas Brunschweiler (IBM Research Europe), Juan Bernabe Moreno (IBM Research Europe), and Nicolas Longepe (European Space Agency Φ-lab)</i>	
SSL4Eco: A Global Seasonal Dataset for Geospatial Foundation Models in Ecology .....	2394
<i>Elena Plekhanova (Swiss Federal Institute for Forest, Snow and Landscape Research, WSL), Damien Robert (University of Zurich), Johannes Dollinger (University of Zurich), Emilia Arens (University of Zurich), Philipp Brun (Swiss Federal Institute for Forest, Snow and Landscape Research, WSL), Jan Dirk Wegner (University of Zurich), and Niklaus E. Zimmermann (Swiss Federal Institute for Forest, Snow and Landscape Research, WSL)</i>	

## Workshop on Autonomous Driving

AttentiveGRU: Recurrent Spatio-Temporal Modeling for Advanced Radar-Based BEV Object Detection .....	2406
<i>Loveneet Saini (Aptiv), Mirko Meuter (Aptiv), Hasan Tercan (University of Wuppertal), and Tobias Meisen (University of Wuppertal)</i>	
Inferring Driving Maps by Deep Learning-based Trail Map Extraction .....	2416
<i>Michael Hubbertz (University of Wuppertal, Germany), Pascal Colling (Aptiv Services Deutschland GmbH, Germany), Qi Han (Aptiv Services Deutschland GmbH, Germany), and Tobias Meisen (University of Wuppertal, Germany)</i>	

LMFormer: Lane based Motion Prediction Transformer .....	2426
<i>Harsh Yadav (University of Wuppertal, Germany), Maximilian Schaefer (Aptiv Services Deutschland GmbH, Germany), Kun Zhao (Aptiv Services Deutschland GmbH, Germany), and Tobias Meisen (University of Wuppertal, Germany)</i>	
TB-Bench: Training and Testing Multi-Modal AI for Understanding Spatio-Temporal Traffic Behaviors from Dashcam Images/Videos .....	2436
<i>Korawat Charoenpitaks (Tohoku University), Van-Quang Nguyen (RIKEN Center for AIP), Masanori Suganuma (Tohoku University), Kentaro Arai (DENSO CORPORATION), Seiji Totsuka (DENSO CORPORATION), Hiroshi Ino (DENSO CORPORATION), and Takayuki Okatani (Tohoku University)</i>	
PatchContrast: Self-Supervised Pre-Training for 3D Object Detection .....	2447
<i>Oren Shtrout (Technion), Ori Nizan (Technion), Yizhak Ben-Shabat (Technion), and Ayellet Tal (Technion)</i>	
PS4PRO: Pixel-to-pixel Supervision for Photorealistic Rendering and Optimization .....	2458
<i>Yezhi Shen (Purdue University, USA), Qiuchen Zhai (Purdue University, USA), and Fengqing Zhu (Purdue University, USA)</i>	
Exploring Semi-Supervised Learning for Online Mapping .....	2468
<i>Adam Lilja (Zenseact), Erik Wallin (SAAB), Junsheng Fu (Zenseact), and Lars Hammarstrand (Chalmers University)</i>	
NeuRadar: Neural Radiance Fields for Automotive Radar Point Clouds .....	2479
<i>Mahan Rafidashti (Chalmers University of Technology), Ji Lan (Chalmers University of Technology), Maryam Fatemi (Zenseact), Junsheng Fu (Zenseact), Lars Hammarstrand (Chalmers University of Technology), and Lennart Svensson (Chalmers University of Technology)</i>	
Multimodal 3D Object Detection on Unseen Domains .....	2490
<i>Deepti Hegde (Johns Hopkins University), Suhas Lohit (MERL), Kuan-Chuan Peng (MERL), Michael Jones (MERL), and Vishal Patel (Johns Hopkins University)</i>	
DySS: Dynamic Queries and State-Space Learning for Efficient 3D Object Detection from Multi-Camera Videos .....	2501
<i>Rajeev Yasarla (Qualcomm AI Research), Shizhong Han (Qualcomm AI Research), Hong Cai (Qualcomm AI Research), and Fatih Porikli (Qualcomm AI Research)</i>	
Camera-Only 3D Panoptic Scene Completion for Autonomous Driving through Differentiable Object Shapes .....	2511
<i>Nicola Marinello (KU Leuven, Belgium), Simen Cassiman (KU Leuven, Belgium), Jonas Heylen (Trace vzw, Belgium), Marc Proesmans (KU Leuven, Belgium), and Luc Van Gool (KU Leuven, Belgium)</i>	
What is the Added Value of UDA in the VFM Era? .....	2521
<i>Brunó B. Englert (Eindhoven University of Technology), Tommie Keressies (Eindhoven University of Technology), and Gijs Dubbelman (Eindhoven University of Technology)</i>	
TrajGNAS: Heterogeneous Multiagent Trajectory Prediction Based on a Graph Neural Architecture Search .....	2532
<i>Yunheng Xu (Anhui University, China), Jie Chen (Anhui University, China), Shuoheng Wang (Anhui University, China), and Xinwen Wang (Anhui University, China)</i>	

CE-NPBG: Connectivity Enhanced Neural Point-Based Graphics for Novel View Synthesis in Autonomous Driving Scenes .....	2542
<i>Mohammad Altillawi (Huawei Munich Research Center), Fengyi Shen (Technical University of Munich), Liudi Yang (Huawei Group), Sai Manoj Prakhya (Huawei Group), and Ziyuan Liu (Huawei Group)</i>	
DuoSpaceNet: Leveraging Both Bird's-Eye-View and Perspective View Representations for 3D Object Detection .....	2551
<i>Zhe Huang (Carnegie Mellon University), Yizhe Zhao (University of California, San Diego), Hao Xiao (University of Washington, Seattle), Chenyan Wu (The Pennsylvania State University), and Lingting Ge (University of California, San Diego)</i>	
Data Scaling Laws for End-to-End Autonomous Driving .....	2562
<i>Alexander Naumann (NVIDIA), Xunjiang Gu (University of Toronto), Tolga Dimlioglu (New York University), Mariusz Bojarski (NVIDIA), Alperen Degirmenci (NVIDIA), Alexander Popov (NVIDIA), Devansh Bisla (NVIDIA), Marco Pavone (NVIDIA), Urs Muller (NVIDIA), and Boris Ivanovic (NVIDIA)</i>	
Nexar Dashcam Collision Prediction Dataset and Challenge .....	2574
<i>Daniel Moura (Nexar Inc.), Shizhan Zhu (Nexar Inc.), and Orly Zvitia (Nexar Inc.)</i>	

## Pixel-level Video Understanding in the Wild Challenge

REEF: Relevance-Aware and Efficient LLM Adapter for Video Understanding .....	2583
<i>Sakib Reza (Northeastern University, USA), Xiyun Song (Futurewei Technologies Inc, USA), Heather Yu (Futurewei Technologies Inc, USA), Zongfang Lin (Futurewei Technologies Inc, USA), Mohsen Moghaddam (Georgia Institute of Technology, USA), and Octavia Camps (Northeastern University, USA)</i>	
M-Adaptor: Text-driven Whole-body Human Motion Generation .....	2595
<i>Alicia Li (Horace Mann School, America), Xiaodong Chen (University of Science and Technology of China, China), Bohao Liang (University of New South Wales, Australia), Qian Bao (JD Explore Academy, China), and Wu Liu (University of Science and Technology of China, China)</i>	
Adversarially Domain-adaptive Latent Diffusion for Unsupervised Semantic Segmentation .....	2605
<i>Jongmin Yu (ProjectG.AI &amp; Department of Applied Mathematics and Theoretical Physics, University of Cambridge), Zhongtian Sun (Computer Science, University of Kent, and Department of Statistics, University of Oxford), Chi Bene Chen (Department of Automation, Tsinghua University), Jinhong Yang (Department of Healthcare and Information Technology, Inje University), and Shan Luo (Department of Engineering, King's College London)</i>	

OpenTAD: A Unified Framework and Comprehensive Study of Temporal Action Detection .....	2616
<i>Shuming Liu (King Abdullah University of Science and Technology, Saudi Arabia), Chen Zhao (King Abdullah University of Science and Technology, Saudi Arabia), Fatimah Zohra (King Abdullah University of Science and Technology, Saudi Arabia), Mattia Soldan (King Abdullah University of Science and Technology, Saudi Arabia), Alejandro Pardo (King Abdullah University of Science and Technology, Saudi Arabia), Mengmeng Xu (King Abdullah University of Science and Technology, Saudi Arabia), Lama Alssum (King Abdullah University of Science and Technology, Saudi Arabia), Merey Ramazanova (King Abdullah University of Science and Technology, Saudi Arabia), Juan León Alcázar (King Abdullah University of Science and Technology, Saudi Arabia), Anthony Cioppa (King Abdullah University of Science and Technology, Saudi Arabia), Silvio Giancola (King Abdullah University of Science and Technology, Saudi Arabia), Carlos Hinojosa (King Abdullah University of Science and Technology, Saudi Arabia), and Bernard Ghanem (King Abdullah University of Science and Technology, Saudi Arabia)</i>	
Studying Image Diffusion Features for Zero-Shot Video Object Segmentation .....	2627
<i>Thanos Delatolas (Technical University of Denmark &amp; Pioneer Center for AI), Vicky Kalogeiton (LIX, École Polytechnique, CNRS, IP Paris, France), and Dim Papadopoulos (Technical University of Denmark &amp; Pioneer Center for AI)</i>	
MTA-VPS: A Large-scale Benchmark for Video-Based Person Search .....	2639
<i>Ding Qi (Tongji University), Shuguang Dou (Huawei Inc.), Jian Liu (ShanDong Energy Group CO. LTD.), Huaixuan Cao (Yunding Technology Co. Ltd.), Hao Zhang (Yunding Technology Co. Ltd.), Dongsheng Jiang (Huawei Inc.), and Cairong Zhao (Tongji University)</i>	
Efficient VideoMAE via Temporal Progressive Training .....	2650
<i>Xianhang Li (University of California, Santa Cruz), Peng Wang (Bytedance Inc.), Xinyu Li (Amazon Inc.), Heng Wang (NVIDIA Inc.), Hongru Zhu (Johns Hopkins University), and Cihang Xie (University of California, Santa Cruz)</i>	
PVUW 2025 Challenge Report: Advances in Pixel-level Understanding of Complex Videos in the Wild .....	2660
<i>Henghui Ding (Fudan University, China), Chang Liu (ByteDance, Singapore), Nikhila Ravi (Meta AI, USA), Shuting He (Shanghai University of Finance and Economics, China), Yunchao Wei (Beijing Jiaotong University, China), Song Bai (ByteDance, Singapore), and Philip Torr (University of Oxford, UK)</i>	

## The 4th Explainable AI for Computer Vision (XAI4CV) Workshop

Disentangling Visual Transformers: Patch-level Interpretability for Image Classification .....	2670
<i>Guillaume Jeanneret (ISIR Laboratory - Sorbonne Université), Loic Simon (GREYC Laboratory - ENSICAEN, UNICAEN), and Frederic Jurie (GREYC Laboratory - ENSICAEN, UNICAEN)</i>	

PCBEAR: Pose Concept Bottleneck for Explainable Action Recognition .....	2681
<i>Jongseo Lee (Kyung Hee University), Wooil Lee (Kyung Hee University), Gyeong-Moon Park (Korea University), Seong Tae Kim (Kyung Hee University), and Jinwoo Choi (Kyung Hee University)</i>	
ScoreCAM++: Gated Score-Weighted Visual Explanations for CNNs .....	2691
<i>Soham Mitra (Indian Institute of Information Technology Kalyani), Atri Sukul (Jalpaiguri Government Engineering College), Swalpa Kumar Roy (Alipurduar Government Engineering and Management College), Pravendra Singh (IIT Roorkee), and Vinay Kumar Verma (Amazon)</i>	
How does the Machine Perceive Depth for Indoor Single Images with CNN? .....	2701
<i>Yihong Wu (University of Southampton), Yuwen Heng (University of Southampton), Mahesan Niranjan (University of Southampton), and Hansung Kim (University of Southampton)</i>	
Towards Synthetic Concept Activation Vectors via Generative Models .....	2711
<i>Riccardo Campi (Politecnico di Milano, DEIB, Italy), Santiago Borrego (Massachusetts Institute of Technology, CSAIL, USA), Antonio De Santis (Politecnico di Milano, DEIB, Italy), Matteo Bianchi (Politecnico di Milano, DEIB, Italy), Andrea Tocchetti (Politecnico di Milano, DEIB, Italy), and Marco Brambilla (Politecnico di Milano, DEIB, Italy)</i>	
X-Edit: Detecting and Localizing Edits in Images Altered by Text-Guided Diffusion Models .....	2720
<i>Valentina Bazyleva (Reality Defender), Nicolò Bonettini (Reality Defender), and Gaurav Bharaj (Reality Defender)</i>	
PoseGuru: Landmarks for Explainable Pose Correction using Exemplar-Guided Algorithmic Recourse .....	2731
<i>Bhat Dittakavi (Indian Institute of Technology, Hyderabad), Bharathi Callepalli (Variance AI), Swarnim Maheshwari (Indian Institute of Technology, Hyderabad), and Vineeth Balasubramanian (Indian Institute of Technology, Hyderabad)</i>	
ExaM: Unsupervised Concept-Based Representation Learning to Better Explain Models in Vision Tasks .....	2741
<i>Maguelonne Heritier (GENETEC), Djebril Mekhazni (GENETEC), Cedric Leblond-Menard (GENETEC), Benoit Godbout (GENETEC), Nathan Guilbaud (GENETEC), Mahdi Alehdaghi (ETS), and Eric Granger (ETS)</i>	
Visually Interpretable Subtask Reasoning for Visual Question Answering .....	2751
<i>Yu Cheng (The University of Edinburgh), Arushi Goel (NVIDIA), and Hakan Bilen (The University of Edinburgh)</i>	
gMINT: Gradient-based Membership Inference Test applied to Image Models. ....	2772
<i>Daniel DeAlcala (Universidad Autónoma de Madrid), Aythami Morales (Universidad Autónoma de Madrid), Julian Fierrez (Universidad Autónoma de Madrid), Gonzalo Mancera (Universidad Autónoma de Madrid), and Ruben Tolosana (Universidad Autónoma de Madrid)</i>	
Explaining 3D Point Cloud Semantic Segmentation Models Through Adversarial Attacks .....	2782
<i>Jorge Francisco Ciprián-Sánchez (University of Potsdam, Hasso Plattner Institute, Germany), Josafat-Mattias Burmeister (University of Potsdam, Germany), Rico Richter (University of Potsdam, Germany), and Jürgen Döllner (University of Potsdam, Hasso Plattner Institute, Germany)</i>	

Rel-SA: Alzheimer’s Disease Detection using Relevance-augmented Self Attention by Inducing Domain Priors in Vision Transformers .....	2792
<i>Madhumitha V (Indian Institute of Technology Hyderabad), Sunayna Padhye (Indian Institute of Technology Hyderabad), Shanawaj S Madarkar (Indian Institute of Technology Hyderabad), Susmit Agrawal (Indian Institute of Technology Hyderabad), and Konda Reddy Mopuri (Indian Institute of Technology Hyderabad)</i>	

## Image Matching: Local Features and Beyond

Less Biased Noise Scale Estimation for Threshold-Robust RANSAC .....	2802
<i>Johan Edstedt (Linköping University)</i>	
Outlier-Robust Multi-Model Fitting on Quantum Annealers .....	2812
<i>Saurabh Pandey (Saarland University), Luca Magri (Politecnico di Milano), Federica Arrigoni (Politecnico di Milano), and Vladislav Golyanik (MPI for Informatics)</i>	
Learning Optical Flow Field via Neural Ordinary Differential Equation .....	2822
<i>Leyla Mirvakhabova (Qualcomm AI Research), Hong Cai (Qualcomm AI Research), Jisoo Jeong (Qualcomm AI Research), Hanno Ackermann (Qualcomm AI Research), Farhad Zanjani (Qualcomm AI Research), and Fatih Porikli (Qualcomm AI Research)</i>	
Detector-free Image Matching with Lightweight Backbone and Feature Filtering .....	2832
<i>Xiaolong Guo (EEIS Department, University of Science and Technology of China), Min Wang (IAI-ATCDI Joint R&amp;D Center (Laboratory) on Intelligent Technologies for Comprehensive Transportation, Institute of Artificial Intelligence, Hefei Comprehensive National Science Center), Hui Wu (EEIS Department, University of Science and Technology of China), Wengang Zhou (EEIS Department, University of Science and Technology of China; IAI-ATCDI Joint R&amp;D Center (Laboratory) on Intelligent Technologies for Comprehensive Transportation, Institute of Artificial Intelligence, Hefei Comprehensive National Science Center), and Houqiang Li (EEIS Department, University of Science and Technology of China; IAI-ATCDI Joint R&amp;D Center (Laboratory) on Intelligent Technologies for Comprehensive Transportation, Institute of Artificial Intelligence, Hefei Comprehensive National Science Center)</i>	
To Match or Not to Match: Revisiting Image Matching for Reliable Visual Place Recognition .....	2840
<i>Davide Sferrazza (Politecnico di Torino), Gabriele Berton (Politecnico di Torino), Gabriele Trivigno (Politecnico di Torino), and Carlo Masone (Politecnico di Torino)</i>	
MegaLoc: One Retrieval to Place Them All .....	2852
<i>Gabriele Berton (Politecnico di Torino) and Carlo Masone (Politecnico di Torino)</i>	

## 2nd Workshop on Human Motion Generation

Dyadic Mamba: Long-term Dyadic Human Motion Synthesis .....	2859
<i>Julian Tanke (Sony AI), Takashi Shibuya (Sony AI), Kengo Uchida (Sony AI), Koichi Saito (Sony AI), and Yuki Mitsufuji (Sony AI)</i>	
HOI-Diff: Text-Driven Synthesis of 3D Human-Object Interactions using Diffusion Models .....	2869
<i>Xiaogang Peng (Northeastern University), Yiming Xie (Northeastern University), Zizhao Wu (Hangzhou Dianzi University), Varun Jampani (Stability AI), Deqing Sun (Google Research), and Huaizu Jiang (Northeastern University)</i>	
MDMP: Multi-modal Diffusion for supervised Motion Predictions with uncertainty .....	2880
<i>Leo Bringer (University of Michigan - Ann Arbor), Joey Wilson (University of Michigan - Ann Arbor), Kira Barton (University of Michigan - Ann Arbor), and Maani Ghaffari (University of Michigan - Ann Arbor)</i>	
SILK: Smooth InterpoLation frameworK for motion in-betweening .....	2891
<i>Elly Akhouni (SEED (Electronic Arts)), Hung Yu Ling (Electronic Arts), Anup Anand Deshmukh (Electronic Arts), and Judith Bütepage (SEED (Electronic Arts))</i>	
MoLA: Motion Generation and Editing with Latent Diffusion Enhanced by Adversarial Training. ....	2901
<i>Kengo Uchida (Sony AI), Takashi Shibuya (Sony AI), Yuhta Takida (Sony AI), Naoki Murata (Sony AI), Julian Tanke (Sony AI), Shusuke Takahashi (Sony Group Corporation), and Yuki Mitsufuji (Sony AI, Sony Group Corporation)</i>	
Goal-Driven Human Motion Synthesis in Diverse Task .....	2911
<i>Inwoo Hwang (Seoul National University, South Korea), Jinseok Bae (Seoul National University, South Korea), Donggeun Lim (Seoul National University, South Korea), and Young Min Kim (Seoul National University, South Korea)</i>	
MoCLIP Motion-Aware Fine-Tuning and Distillation of CLIP for Human Motion Generation .....	2922
<i>Gabriel Maldonado (Department of Electrical and Computer Engineering, UNC Charlotte), Armin Danesh Pazho (Department of Electrical and Computer Engineering, UNC Charlotte), Ghazal Alinezhad Noghre (Department of Electrical and Computer Engineering, UNC Charlotte), Vinit Katariya (Department of Electrical and Computer Engineering, UNC Charlotte), and Hamed Tabkhi (Department of Electrical and Computer Engineering, UNC Charlotte)</i>	

## Multimodal Algorithmic Reasoning Workshop

Exemplar Masking for Multimodal Incremental Learning .....	2933
<i>Yi-Lun Lee (National Yang Ming Chiao Tung University; Atmanity), Chen-Yu Lee (Google), Wei-Chen Chiu (National Yang Ming Chiao Tung University), and Yi-Hsuan Tsai (Atmanity)</i>	

Human vs. Machine Minds: Ego-Centric Action Recognition Compared .....	2943
<i>Sadegh Rahmaniboldaji (University of Surrey), Filip Rybansky (University of Newcastle), Quoc Vuong (University of Newcastle), Frank Guerin (University of Surrey), and Andrew Gilbert (University of Surrey)</i>	
Autonomous Multimodal Reasoning via Implicit Chain-of-Vision .....	2954
<i>Yiqiao Huang (Harvard University), Qi He (George Mason University), Zhaorun Chen (University of Chicago), Haopeng Zhang (University of Hawaii), Hanchao Yu (Meta AI), and Zhuokai Zhao (Meta AI)</i>	
Comparison Visual Instruction Tuning .....	2964
<i>Wei Lin (Johannes Kepler University Linz), Muhammad Jehanzeb Mirza (MIT), Sivan Doreh (Google), Rogerio Feris (MIT-IBM Watson AI Lab, IBM Research), Raja Giryes (Tel Aviv University), Sepp Hochreiter (Johannes Kepler University Linz), and Leonid Karlinsky (Xero)</i>	
SilVar-Med: A Speech-Driven Visual Language Model for Explainable Abnormality Detection in Medical Imaging .....	2975
<i>Tan-Hanh Pham (Florida Institute of Technology), Trong-Duong Bui (Vietnam Military Medical University), Minh Luu Quang (Cardiovascular Intensive Care Unit, Heart Institute, 108 Military Central Hospital), Tan Huong Pham (Can Tho University of Medicine and Pharmacy), Chris Ngo (Knovel Engineering Lab), and Truong Son Hy (University of Alabama at Birmingham)</i>	
Illusory VQA: Benchmarking and Enhancing Multimodal Models on Visual Illusions .....	2986
<i>Mohammadmostafa Rostamkhani (Iran University of Science and Technology (IUST)), Baktash Ansari (Iran University of Science and Technology (IUST)), Hoorieh Sabzevari (Iran University of Science and Technology (IUST)), Farzan Rahmani (Iran University of Science and Technology (IUST)), and Sauleh Eetemadi (Iran University of Science and Technology (IUST))</i>	

## Workshop on Foundation and Large Vision Models in Remote Sensing

PAN-RSVQA: Vision Foundation Models as Pseudo-ANnotators for Remote Sensing Visual Question Answering .....	2996
<i>Christel Chappuis (Ecole Polytechnique Fédérale de Lausanne, Switzerland), Gencer Sümbül (Ecole Polytechnique Fédérale de Lausanne, Switzerland), Syrielle Montariol (Ecole Polytechnique Fédérale de Lausanne, Switzerland), Sylvain Lobry (LIPADE, Université Paris Cité, France), and Devis Tuia (Ecole Polytechnique Fédérale de Lausanne, Switzerland)</i>	
Efficient Self-Supervised Learning for Earth Observation via Dynamic Dataset Curation .....	3008
<i>Thomas Kerdreux (Galeio), Alexandre Tuel (Galeio), Quentin Febvre (Ifremer, LOPS, Brest, France), Alexis Mouche (Ifremer, LOPS, Brest, France), and Bertrand Chapron (Ifremer, LOPS, Brest, France)</i>	

Foundation Models for Remote Sensing: An Analysis of MLLMs for Object Localization. ....	3019
<i>Darryl Hannan (Pacific Northwest National Laboratory), John Cooper (Pacific Northwest National Laboratory), Dylan White (Pacific Northwest National Laboratory), Timothy Doster (Pacific Northwest National Laboratory), Henry Kvinge (Pacific Northwest National Laboratory), and Yijing Watkins (Pacific Northwest National Laboratory)</i>	
Deep Diffusion Models and Unsupervised Hyperspectral Unmixing for Realistic Abundance Map Synthesis .....	3029
<i>Martina Pastorino (DITEN, University of Genoa), Michael Alibani (DII, University of Pisa), Nicola Acito (DII, University of Pisa), and Gabriele Moser (DITEN, University of Genoa)</i>	
A Sensor Agnostic Domain Generalization Framework for Leveraging Geospatial Foundation Models: Enhancing Semantic Segmentation via Synergistic Pseudo-Labeling and Generative Learning .....	3038
<i>Anan Yaghmour (University of Houston, USA), Melba Crawford (Purdue University, USA), and Saurabh Prasad (University of Houston, USA)</i>	
Bridging the Modality Gap: Training-free Adaptation of Vision-Language Models for Remote Sensing via Visual Prototypes .....	3048
<i>Clément Barbier (ONERA, France), Baptiste Abeloss (ONERA, France), and Stéphane Herbin (ONERA, France)</i>	
MESA: Text-Driven Terrain Generation Using Latent Diffusion and Global Copernicus Data .....	3058
<i>Paul Borne--Pons (Adobe Research, European Space Agency (ESA)), Mikolaj Czerkawski (Asterisk Labs, European Space Agency (ESA)), Rosalie Martin (Adobe Research), and Romain Rouffet (Adobe Research)</i>	
Dynamic State-Control Modeling for Generalized Remote Sensing Image Super-Resolution .....	3067
<i>Chenyu Li (Southeast University), Zhaojie Pan (Southeast University), and Danfeng Hong (Chinese Academy of Sciences)</i>	
COP-GEN-Beta: Unified Generative Modelling of COPernicus Imagery Thumbnails .....	3076
<i>Miguel Espinosa (University of Edinburgh), Valerio Marsocci (European Space Agency (ESA)), Yuru Jia (KU Leuven), Elliot Crowley (University of Edinburgh), and Mikolaj Czerkawski (Asterisk Labs)</i>	
Towards Efficient Benchmarking of Foundation Models in Remote Sensing: A Capabilities Encoding Approach .....	3087
<i>Pierre Adorni (IRISA, Université Bretagne Sud, UMR 6074, Vannes, France), Minh-Tan Pham (IRISA, Université Bretagne Sud, UMR 6074, Vannes, France), Stéphane May (Centre National d'Études Spatiales (CNES), Toulouse, France), and Sébastien Lefèvre (IRISA, Université Bretagne Sud, UMR 6074, Vannes, France; UiT The Arctic University of Norway, Tromsø, Norway)</i>	

# SyntaGen: 2nd Workshop on Harnessing Generative Models for Synthetic Visual Datasets

NeIn: Telling What You Don't Want .....	3098
<i>Nhat-Tan Bui (University of Arkansas, USA), Dinh-Hieu Hoang (University of Science, VNU-HCM, Vietnam), Quoc-Huy Trinh (Aalto University, Finland &amp; SpexAI GmbH, Germany), Minh-Triet Tran (University of Science, VNU-HCM, Vietnam), Truong Nguyen (University of California, San Diego, USA), and Susan Gauch (University of Arkansas, USA)</i>	
Noise Consistency Regularization for Improved Subject-Driven Image Synthesis .....	3107
<i>Yao Ni (Australian National University), Song Wen (Rutgers University), Piotr Koniusz (Australian National University), and Anoop Cherian (Mitsubishi Electric Research Labs (MERL))</i>	
AnomalyHybrid: A Domain-agnostic Generative Framework for General Anomaly Detection .....	3118
<i>YING ZHAO (Ricoh Software Research Center (Beijing) Co., Ltd.)</i>	
SVAD: From Single Image to 3D Avatar via Synthetic Data Generation with Video Diffusion and Data Augmentation .....	3128
<i>Yonwoo Choi (SECERN AI)</i>	
good4cir: Generating Detailed Synthetic Captions for Composed Image Retrieval .....	3139
<i>Pranavi Kolouju (Saint Louis University), Eric Xing (Washington University in St. Louis), Robert Pless (George Washington University), Nathan Jacobs (Washington University in St. Louis), and Abby Stylianou (Saint Louis University)</i>	
Syn3DTxt: Embedding 3D Cues for Scene Text Generation .....	3149
<i>Li-Syun Hsiung (National Taiwan University of Science and Technology), Jun-Kai Tu (National Taiwan University of Science and Technology), Kuan-Wu Chu (National Chengchi University), Yu-Hsuan Chiu (National Taiwan University of Science and Technology), Yan-Tsung Peng (National Chengchi University), Sheng-Luen Chung (National Taiwan University of Science and Technology), and Gee-Sern Hsu (National Taiwan University of Science and Technology)</i>	
Eyes Tell the Truth: GazeVal Highlights Shortcomings of Generative AI in Medical Imaging .....	3158
<i>David Wong (Northwestern University), Bin Wang (Northwestern University), Gorkem Durak (Northwestern University), Marouane Tliba (University of Illinois at Chicago), Akshay Chaudhari (Stanford University), Aladine Chetouani (Université Sorbonne Paris Nord), Ahmet Enis Cetin (University of Illinois at Chicago), Cagdas Topel (Northwestern University), Nicolo Gennaro (Northwestern University), Camila Lopes Vendrami (Northwestern University), Tugce Agirlar Trabzonlu (Northwestern University), Amir Ali Rahsepar (Northwestern University), Laetitia Perronne (Northwestern University), Matthew Antalek (Northwestern University), Onural Ozturk (Northwestern University), Gokcan Okur (Loyola University Chicago), Andrew C Gordon (Northwestern University), Ayis Pyrros (DuPage Medical Group), Frank H Miller (Northwestern University), Amir Borhani (Northwestern University), Hatice Savas (Northwestern University), Eric Hart (Northwestern University), Drew Torigian (University of Pennsylvania), Jayaram K udupa (University of Pennsylvania), Elizabeth Krupinski (Emory University), and Ulas Bagci (Northwestern University)</i>	

## 2nd Workshop on Efficient Large Vision Models

ViDROP: Video Dense Representation through Spatio-Temporal Sparsity .....	3167
<i>Sepehr Sameni (Independent), Simon Jenni (Adobe Research), and Paolo Favaro (University of Bern)</i>	
Window Token Concatenation for Efficient Visual Large Language Models .....	3178
<i>Yifan Li (Michigan State University), Wentao Bao, Botao Ye, Zhen Tan, Tianlong Chen, Huan Liu, and Yu Kong</i>	
Prompt Categories Cluster for Weakly Supervised Semantic Segmentation .....	3189
<i>Wangyu Wu (The University of Liverpool), Xianglin Qiu (The University of Liverpool), Siqi Song (The University of Liverpool), Zhenhong Chen (Microsoft), Xiaowei Huang (The University of Liverpool), Fei Ma (Xi'an Jiaotong-Liverpool University), and Jimin Xiao (Xi'an Jiaotong-Liverpool University)</i>	
From Data to Design: Leveraging Frequency Statistics for Efficient Neural Network Architectures .....	3199
<i>Mustafa Munir (The University of Texas at Austin), Guihong Li (The University of Texas at Austin), Md Mostafijur Rahman (The University of Texas at Austin), Alex Zhang (The University of Texas at Austin), and Radu Marculescu (The University of Texas at Austin)</i>	
Rethinking the Role of Spatial Mixing .....	3210
<i>George Cazenavette (MIT), Joel Julin (Carnegie Mellon University), and Simon Lucey (University of Adelaide)</i>	
SmoothCache: A Universal Inference Acceleration Technique for Diffusion Transformers .....	3220
<i>Joseph Liu (Roblox), Joshua Geddes (Queen's University), Ziyu Guo (Roblox), Haomiao Jiang (Roblox), and Mahesh Kumar Nandwana (Roblox)</i>	
Efficient Image Generation with Variadic Attention Heads .....	3230
<i>Steven Walton (University of Oregon, USA), Ali Hassani (Georgia Institute of Technology, USA), Xingqian Xu (University of Illinois Urbana-Champaign, USA), Zhangyang Wang (University of Texas at Austin), and Humphrey Shi (Georgia Institute of Technology, USA)</i>	
U-Shape Mamba: State Space Model for faster diffusion .....	3242
<i>Alex Ergasti (University Of Parma), Filippo Botti (University Of Parma), Tomaso Fontanini (University Of Parma), Claudio Ferrari (University Of Siena), Massimo Bertozzi (University Of Parma), and Andrea Prati (University Of Parma)</i>	
Dual Precision Quantization for Efficient and Accurate Deep Neural Networks Inference .....	3250
<i>Tomer Gafni (Intel), Asaf Karnieli (Intel), and Yair Hanani (Intel)</i>	
VisionCube: 3D-Aware Vision-Language Model for Multi-Step Spatial Reasoning .....	3261
<i>Feiyang Wang (Central South University), Nan Luo (Northeastern University), and Wangyu Wu (The University of Liverpool)</i>	
Mix-QSAM: Mixed-Precision Quantization of the Segment Anything Model .....	3271
<i>Navin Ranjan (Rochester Institute of Technology) and Andreas Savakis (Rochester Institute of Technology)</i>	

Effectiveness of Max-Pooling for Fine-Tuning CLIP on Videos .....	3282
<i>Fatimah Zohra (KAUST), Chen Zhao (KAUST), Shuming Liu (KAUST), and Bernard Ghanem (KAUST)</i>	
Efficiently Mitigating Video Content Misalignment on Large Vision Model with Time-Series Data Alignment .....	3292
<i>Hanchen Xie (University of Southern California), Rose Ma (Bosch Research North America and Bosch Center for Artificial Intelligence (BCAI)), Jiageng Zhu (University of Southern California), Zheda Mai (Ohio State University), Wael Abd-Almageed (Clemson University), and Zubin Abraham (Bosch Research North America and Bosch Center for Artificial Intelligence (BCAI))</i>	
DEFT-VTON: Efficient Virtual Try-On with Consistent Generalised H-Transform .....	3299
<i>Xingzi Xu (Duke University, United States), Qi Li (Amazon), Shuwen Qiu (UCLA, United States), Julien Han (Amazon), and Karim Bouyarmane (Amazon)</i>	
CARN: Complexity-Aware Routing Network for Efficient and Adaptive Inference .....	3309
<i>Rebati Gaire (University of Illinois Chicago) and Arman Roohi (University of Illinois Chicago)</i>	
SimCache: Similarity Caching for Efficient VLM-based Scene Understanding .....	3318
<i>Surya Selvam (Purdue University, USA), Ravi K. Rajendran (NEC Laboratories America, USA), Murugan Sankaradas (NEC Laboratories America, USA), Anand Raghunathan (Purdue University, USA), and Srimat T. Chakradhar (NEC Laboratories America, USA)</i>	
Distilling Normalizing Flows .....	3328
<i>Steven Walton (University of Oregon, USA), Valeriy Klyukin (HSE University, Russia), Maksim Artemev (Picsart AI Research, Armenia), Denis Derkach (HSE University, Russia), Nikita Orlov (Picsart AI Research, Armenia), and Humphrey Shi (Georgia Institute of Technology, USA)</i>	
Video, How Do Your Tokens Merge? .....	3338
<i>Sam Pollard (University of Bristol, United Kingdom) and Michael Wray (University of Bristol, United Kingdom)</i>	

## Test-time Scaling for Computer Vision

On the Suitability of Reinforcement Fine-Tuning to Visual Tasks .....	3348
<i>Xiaxu Chen (Beijing institute of technology, China), Wei Li (SenseTime Research), Chunxu Liu (Nanjing University, China), Chi Xie (Tongji University, China), Xiaoyan Hu (SenseTime Research), Chengqian Ma (SenseTime Research), Feng Zhu (SenseTime Research), and Rui Zhao (SenseTime Research)</i>	
TTGen: Incorporating Test-time Scaling to Diffusion Models .....	3353
<i>Yuming Qiao (OPPO Research Institute), Yuechen Wang (OPPO Research Institute), Xudong Zhang (OPPO Research Institute), and Dan Meng (OPPO Research Institute)</i>	

Scaling Test-Time Compute Can Outperform Larger Architectures in Computer Vision .....	3358
<i>Erfan Darzi (Harvard University, Massachusetts Institute of Technology, USA), Dylan Nguyen (Massachusetts Institute of Technology, USA), and George Cheng (Massachusetts Institute of Technology, USA)</i>	
Get a GRIP on Test Time Adaptation! - Group Robust Inference-Time Policy Optimization for Vision Models .....	3367
<i>Prabhav Sanga (University College London, UK), Jaskaran Singh (University of Nottingham, UK), and Tapabrata Chakraborti (University College London, UK)</i>	

## Women in Computer Vision

Enhanced Multi-View Pedestrian Detection Using Probabilistic Occupancy Volume .....	3377
<i>Reef Alturki (University of Surrey, UK.), Adrian Hilton (University of Surrey, UK.), and Jean-Yves Guillemaut (University of Surrey, UK.)</i>	
Document Image Rectification using Stable Diffusion Transformer .....	3387
<i>Pooja Kumari (IIT Madras) and Sukhendu Das (IIT Madras)</i>	
Uncertainty-guided Style-aware Perceptual Quality Assessment for AI-Generated Images .....	3397
<i>Tushar Shinde (IIT Madras Zanzibar) and Shivaanee Eswaran (IIT Madras Zanzibar)</i>	
LLaVA-SCo: Teach Vision Language Models to Self-Correct .....	3406
<i>Zixuan Liu (Tulane University), Guangkai Jiang (University of Helsinki), and Siavash Khajavi (Aalto University)</i>	
Exploiting Adversarial Learning and Topology Augmentation for Open-Set Visual Recognition ..	3416
<i>Rosa Zuccarà (University of Catania), Georgia Fargetta (University of Catania), Alessandro Ortis (University of Catania), and Sebastiano Battiato (University of Catania)</i>	
Improved Out-of-Distribution Detection with Additive Angular Margin Loss .....	3425
<i>Deepak Ravikumar (Purdue University, USA), Efstathia Soufleri (Archimedes, Athena Research Center, Greece), and Kaushik Roy (Purdue University, USA)</i>	
Domain Adaptation for Skin Lesion: Evaluating Real-World Generalisation .....	3433
<i>Nurjahan Sultana (Manchester Metropolitan University, UK), Wenqi Lu (Manchester Metropolitan University, UK), Xinqi Fan (Manchester Metropolitan University), and Moi Hoon Yap (Manchester Metropolitan University)</i>	
Low-Resource Video Super-Resolution using Memory, Wavelets, and Deformable Convolutions .	3444
<i>Kavitha Viswanathan (Department of Electrical Engineering, IIT Bombay, India), Amit Sethi ( Department of Electrical Engineering, IIT Bombay, India), Shashwat Pathak ( Department of Electrical Engineering, IIT Bombay, India), Piyush Bharambe (Department of Electrical Engineering, IIT Bombay, India), and Harsh Choudhary (Department of Electrical Engineering, IIT Bombay, India)</i>	
IdolDanceNet:Indian Heritage idol Dance Pose Classification .....	3454
<i>S Kanimozhi (Vellore Institute of Technology,India), Sabari Nathan (Couger Inc , Japan), and Sasithradevi A (Vellore Institute of Technology,India)</i>	

Leveraging Intermediate Features of Vision Transformer for Face Anti-Spoofing .....	3464
<i>Mika Feng (Tohoku University, Japan), Koichi Ito (Tohoku University, Japan), Takafumi Aoki (Tohoku University, Japan), Tetsushi Ohki (Shizuoka University, Japan), and Masakatsu Nishigaki (Shizuoka University, Japan)</i>	
Knowledge Distillation Approach for SOS Fusion Staging: Towards Fully Automated Skeletal Maturity Assessment .....	3473
<i>Omid Halimi Milani (Department of Electrical and Computer Engineering, University of Illinois Chicago, USA), Amanda N Nikho (Department of Orthodontics, College of Dentistry, University of Illinois Chicago, USA), Marouane Tliba (University of Orleans), Lauren Mills (Department of Orthodontics, College of Dentistry, University of Illinois Chicago, USA), Ahmet Enis Cetin (Department of Electrical and Computer Engineering, University of Illinois Chicago, USA), and Mohammed H Elnagar (Department of Orthodontics, College of Dentistry, University of Illinois Chicago, USA)</i>	
Dust to Detail: Restoring Sand-dust Images with Frequency-Guided Attention and Multi-Scale Features .....	3481
<i>Romala Mishra (National Institute of Technology Rourkela, India) and Sobhan Dhara (National Institute of Technology Rourkela, India)</i>	
Leveraging Fixed and Dynamic Pseudo-Labels in Cross-Supervision Framework for Semi-Supervised Medical Image Segmentation .....	3491
<i>Suruchi Kumari (Indian Institute of Technology Roorkee, India) and Pravendra Singh (Indian Institute of Technology Roorkee, India)</i>	
Training Data Reconstruction: Privacy due to Uncertainty? .....	3502
<i>Christina Runkel (University of Cambridge), Kanchana Vaishnavi Gandikota (University of Siegen), Jonas Geiping (ELLIS Institute Tübingen, MPI Intelligent Systems), Carola-Bibiane Schönlieb (University of Cambridge), and Michael Moeller (University of Siegen)</i>	

## **The 5th Workshop of Adversarial Machine Learning on Computer Vision: Foundation Models + X**

Trustworthy Multi-UAV Collaboration: A Self-Supervised Framework for Explainable and Adversarially Robust Decision-Making .....	3511
<i>Yuwei Chen (Aviation Industry Development Research Center of China, China) and Shiyong Chu (Aviation Industry Development Research Center of China, China)</i>	
Defending Against Frequency-Based Attacks with Diffusion Models .....	3523
<i>Fatemeh Amerehi (University of Limerick) and Patrick Healy (University of Limerick)</i>	
Attacking Attention of Foundation Models Disrupts Downstream Tasks .....	3534
<i>Hondamunige Prasanna Silva (University of Florence), Federico Becattini (University of Siena), and Lorenzo Seidenari (University of Florence)</i>	

Towards Evaluating the Robustness of Visual State Space Models .....	3544
<i>Hashmat Shadab Malik (Mohamed Bin Zayed University of AI, UAE), Fahad Shamshad (Mohamed Bin Zayed University of AI, UAE), Muzammal Naseer (Khalifa University, UAE), Karthik Nandakumar (Michigan State University), Fahad Shahbaz Khan (Mohamed Bin Zayed University of AI, UAE and Linkoping University, Sweden ), and Salman Khan (Mohamed Bin Zayed University of AI, UAE and Australian National University, Australia)</i>	
FullCycle: Full Stage Adversarial Attack For Reinforcement Learning Robustness Evaluation .....	3554
<i>Zhenshu Ma (State Key Laboratory of Intelligent Transportation Systems, Beijing), Xuan Cai (State Key Laboratory of Intelligent Transportation Systems, Beijing), Changhang Tian (State Key Laboratory of Intelligent Transportation Systems, Beijing), Yuqi Fan (State Key Laboratory of Intelligent Transportation Systems, Beijing), Kemou Jiang (State Key Laboratory of Intelligent Transportation Systems, Beijing), Gangfu Liu (State Key Laboratory of Intelligent Transportation Systems, Beijing), Xuesong Bai (State Key Laboratory of Intelligent Transportation Systems, Beijing), Aoyong Li (State Key Laboratory of Intelligent Transportation Systems, Beijing), Yilong Ren (State Key Laboratory of Intelligent Transportation Systems, Beijing), and Haiyang Yu (hyyu@buaa.edu.cn)</i>	
Probing Vulnerabilities of Vision-LiDAR Based Autonomous Driving Systems .....	3561
<i>Siwei Yang (UC Santa Cruz), Zeyu Wang (UC Santa Cruz), Diego Ortiz (UC Santa Cruz), Luis Burbano (UC Santa Cruz), Murat Kantarcioglu (Virginia Tech), Alvaro Cardenas (UC Santa Cruz), and Cihang Xie (UC Santa Cruz)</i>	
Task-Agnostic Attacks Against Vision Foundation Models .....	3570
<i>Brian Pulfer (University of Geneva), Yury Belousov (University of Geneva), Vitaliy Kinakh (University of Geneva), Teddy Furon (University of Rennes, Inria, CNRS, IRISA), and Slava Voloshynovskiy (University of Geneva)</i>	
EL-Attack: Explicit and Latent Space Hybrid Optimization based General and Effective Attack for Autonomous Driving Trajectory Prediction .....	3582
<i>Xuesong Bai (School of transportation science and engineering, Beihang University), Changhang Tian (State Key Laboratory of Intelligent Transportation Systems, Beijing), Wei Xia (State Key Laboratory of Intelligent Transportation Systems, Beijing), Zhenshu Ma (School of transportation science and engineering, Beihang University), Haiyang Yu (School of transportation science and engineering, Beihang University), and Yilong Ren (School of transportation science and engineering, Beihang University)</i>	
Attention-Aware Temporal Adversarial Shadows on Traffic Sign Sequences .....	3591
<i>Pedram MohajerAnsari (Clemson University, USA), Amir Salarpour (Clemson University, USA), David Fernandez (Clemson University, USA), Cigdem Kokenoz (Clemson University, USA), Bing Li (Clemson University, USA), and Mert D. Pesé (Clemson University, USA)</i>	

# The 3rd Workshop on What is Next in Multimodal Foundation Models?

UniToken: Harmonizing Multimodal Understanding and Generation through Unified Visual Encoding .....	3600
<i>Yang Jiao (Fudan University), Haibo Qiu (Meituan), Zequn Jie (Meituan), Shaoxiang Chen (Meituan), Jingjing Chen (Fudan University), Lin Ma (Meituan), and Yu-Gang Jiang (Fudan University)</i>	
Understanding Depth and Height Perception in Large Visual-Language Models .....	3611
<i>Shehreen Azad (University of Central Florida), Yash Jain (Microsoft Research), Rishit Garg ( Indian Institute of Technology Kharagpur), Vibhav Vineet (Microsoft Research), and Yogesh Rawat (University of Central Florida)</i>	
Repurposing SAM for User-Defined Semantics Aware Segmentation .....	3621
<i>Rohit Kundu (University of California, Riverside), Sudipta Paul (Samsung Research America), Arindam Dutta (University of California, Riverside), and Amit Roy-Chowdhury (University of California, Riverside)</i>	
PLVM: A tuning-free approach for Personalized Large Vision-Language Model .....	3632
<i>Chau Pham (University at Buffalo), Hoang Phan (New York University), David Doermann (University at Buffalo), and Yunjie Tian (University at Buffalo)</i>	
How Good is my Video-LMM? Complex Video Reasoning and Robustness Evaluation Suite for Video-LMMs .....	3642
<i>Muhammad Uzair Khattak (Mohamed Bin Zayed University of AI, Abu-Dhabi), Muhammad Ferjad Naeem (Google Zurich), Jameel Hassan (Mohamed Bin Zayed University of AI, Abu-Dhabi), Muzammal Naseer (Khalifa University, Abu-Dhabi), Federico Tombari (Google Zurich), Fahad Shahbaz Khan (Mohamed Bin Zayed University of AI, Abu-Dhabi), and Salman Khan (Mohamed Bin Zayed University of AI, Abu-Dhabi)</i>	
An Interactive Agent Foundation Model .....	3652
<i>Zane Durante (Stanford University), Ran Gong (University of California, Los Angeles), Bidipta Sarkar (Stanford University), Naoki Wake (Microsoft Research, Redmond), Rohan Taori (Stanford University), Paul Tang (Stanford University), Shrinidhi Lakshmikanth (Stanford University), Kevin Schulman (Stanford University), Arnold Milstein (Stanford University), Hoi Vo (Microsoft Gaming), Ehsan Adeli (Stanford University), Demetri Terzopoulos (University of California, Los Angeles), Li Fei-Fei (Stanford University), and Jianfeng Gao (Microsoft Research, Redmond)</i>	
Understanding the Effect of using Semantically Meaningful Tokens for Visual Representation Learning .....	3663
<i>Neha Kalibhat (University of Maryland), Priyatham Kattakinda (University of Maryland), Sumit Nawathe (University of Maryland), Arman Zarei (University of Maryland), Nikita Seleznev (CapitalOne Research), Samuel Sharpe (CapitalOne Research), Senthil Kumar (CapitalOne Research), and Soheil Feizi (University of Maryland)</i>	

# 1st International Workshop on Interactive Video Search and Exploration

Overview of the 1st International Workshop on Interactive Video Search and Exploration .....	3673
<i>Luca Rossetto (Dublin City University), George Awad (NIST), Werner Bailer (Joanneum Research), Cathal Gurrin (Dublin City University), Björn Þór Jónsson (Reykjavik University), Jakub Lokoč (Charles University), Stevan Rudinac (University of Amsterdam), and Klaus Schoeffmann (Klagenfurt University)</i>	
CadenceRAG: Context-Aware and Dependency-Enhanced Retrieval Augmented Generation for Holistic Video Understanding .....	3679
<i>Heng Liu (Wuhan University of Technology), Siru Jiang (Wuhan University), Fangyun Duan (Wuhan University), Yongzhe Lyu (Wuhan University), Xiusong Wang (Wuhan University), Hanlin Ge (Wuhan University), and Chao Liang (Wuhan University)</i>	
VRAG: Retrieval-Augmented Video Question Answering for Long-Form Videos .....	3689
<i>Bao Tran Gia (University of Information Technology - Vietnam National University), Khiem Le (University of Information Technology, Vietnam National University), Tien Do (University of Information Technology, Vietnam National University), Tien-Dung Mai (University of Information Technology, Vietnam National University), Thanh Duc Ngo (University of Information Technology, Vietnam National University), Duy-Dinh Le (University of Information Technology, Vietnam National University), and Shin'ichi Satoh (National Institute of Informatics)</i>	
Toward Automation in Text-based Video Retrieval with LLM Assistance .....	3699
<i>Khanh-An C. Quan (University of Information Technology, Vietnam National University), Qui Ngoc Nguyen (University of Information Technology, Vietnam National University), and Duc-Tuan Luu (University of Information Technology, Vietnam National University)</i>	
A Lightweight Moment Retrieval System with Global Re-Ranking and Robust Adaptive Bidirectional Temporal Search .....	3708
<i>Tinh-Anh Nguyen-Nhu (Ho Chi Minh University of Technology, VNU-HCM, Vietnam), Huu-Loc Tran (University of Information Technology, VNU-HCM, Vietnam), Nguyen-Khang Le (Ho Chi Minh University of Science, VNU-HCM, Vietnam), Minh-Nhat Nguyen (University of Economics Ho Chi Minh City, Ho Chi Minh, Vietnam), Tien-Huy Nguyen (University of Information Technology, VNU-HCM, Vietnam), Hoang-Long Nguyen-Huu (University of Information Technology, VNU-HCM, Vietnam), Huu-Phong Phan-Nguyen (University of Information Technology, VNU-HCM, Vietnam), Huy-Thach Pham (Georgia State University, USA), Quan Nguyen (Posts and Telecommunications Institute of Technology, Hanoi, Vietnam), Hoang M. Le (York University, Canada), and Quang-Vinh Dinh (AI VIETNAM Lab)</i>	

Towards Efficient and Robust Moment Retrieval System: A Unified Framework for Multi-Granularity Models and Temporal Reranking .....	3719
<i>Huu-Loc Tran (University of Information Technology, VNU-HCM, Vietnam), Tinh-Anh Nguyen-Nhu (Ho Chi Minh University of Technology, VNU-HCM, Vietnam), Huu-Phong Phan-Nguyen (University of Information Technology, VNU-HCM, Vietnam), Tien-Huy Nguyen (University of Information Technology, VNU-HCM, Vietnam), Nhat-Minh Nguyen-Dich (Hanoi University of Science and Technology, Hanoi, Vietnam), Anh Dao (Michigan State University, USA), Huy-Duc Do (National Economics University, Hanoi, Vietnam), Quan Nguyen (Posts and Telecommunications Institute of Technology, Hanoi, Vietnam), Hoang M. Le (York University, Canada ), and Quang-Vinh Dinh (AI VIETNAM Lab)</i>	
An LLM Framework for Long-form Video Retrieval and Audio-Visual Question Answering Using Qwen2/2.5 .....	3730
<i>Damianos Galanopoulos (CERTH/ITI), Andreas Goulas (CERTH/ITI), Antonios Leventakis (CERTH/ITI), Ioannis Patras (Queen Mary University of London), and Vasileios Mezaris (CERTH/ITI)</i>	
Can Relevance Feedback, Conversational Search and Foundation Models Work Together for Interactive Video Search and Exploration? .....	3740
<i>Ujjwal Sharma (University of Amsterdam), Omar Shahbaz Khan (Reykjavík University), Stevan Rudinac (University of Amsterdam), and Björn Þór Jónsson (Reykjavík University)</i>	
AI-based Video Content Understanding for Automatic and Interactive Multimedia Retrieval .....	3750
<i>Klaus Schoeffmann (Klagenfurt University) and Mario Leopold (Klagenfurt University)</i>	

## Workshop on Distillation of Foundation Models for Autonomous Driving

Harmonizing Attention Fields with Knowledge Distillation for Multi-View 3D Object Detection .....	3759
<i>Yafei Qi (Central South University, China), Menghao Yang (Changsha University of Science and Technology, China), Fan Wu (Central South University, China), Chen Wang (Shaoxing University, China), and Yongmin Zhang (Central South University, China)</i>	
ZFusion: An Effective Fuser of Camera and 4D Radar for 3D Object Perception in Autonomous Driving .....	3768
<i>Sheng Yang (School of Data Science, Fudan University), Tong Zhan (School of Data Science, Fudan University), Shichen Qiao (School of Data Science, Fudan University), Jicheng Gong (ZF (China) Investment Co., Ltd.), Qing Yang (ZF (China) Investment Co., Ltd.), Jian Wang (School of Data Science, Fudan University), and Yanfeng Lu (Institute of Automation, Chinese Academy of Sciences)</i>	
PF3Det: A Prompted Foundation Feature Assisted Visual LiDAR 3D Detector .....	3778
<i>Kaidong Li (University of Kansas), Tianxiao Zhang (University of Kansas), Kuan-Chuan Peng (Mitsubishi Electric Research Laboratories), and Guanghui Wang (Toronto Metropolitan University)</i>	

Talk2Traffic: Interactive and Editable Traffic Scenario Generation for Autonomous Driving with Multimodal Large Language Model .....	3788
<i>Zihao Sheng (University of Wisconsin - Madison), Zilin Huang (University of Wisconsin - Madison), Yansong Qu (Purdue University), Yue Leng (Google), and Sikai Chen (University of Wisconsin - Madison)</i>	
CleanMAP: Distilling Multimodal LLMs for Confidence-Driven Crowdsourced HD Map Updates .....	3798
<i>Ankit Kumar Shaw (Tsinghua University), Kun Jiang (Tsinghua University), Tuopu Wen (Tsinghua University), Chandan Kumar Sah (Beihang University), Yining Shi (Tsinghua University), Mengmeng Yang (Tsinghua University), Diange Yang (Tsinghua University), and Xiaoli Lian (Beihang University)</i>	
Fairness-Aware Boosting Model for Imbalanced 3D Point Cloud Segmentation in Autonomous Driving .....	3808
<i>Elahe Yahyapour (University of Massachusetts at Amherst, USA) and Chengbo Ai (University of Massachusetts, Amherst, USA)</i>	
Robusto-1 Dataset: Comparing Humans and VLMs on real out-of-distribution Autonomous Driving VQA from Peru .....	3817
<i>Dunant Cusipuma (Artificio), David Ortega (Artificio), Victor Flores-Benites (Artificio &amp; UTEC), and Arturo Deza (Artificio &amp; UTEC)</i>	
Clip4Retrofit: Enabling Real-Time Image Labeling on Edge Devices via Cross-Architecture CLIP Distillation .....	3829
<i>Li Zhong (Robert Bosch GmbH), Ahmed Ghazal (Robert Bosch GmbH), Jun-Jun Wan (Robert Bosch GmbH), Frederik Zilly (Robert Bosch GmbH), Patrick Mackens (Robert Bosch GmbH), Joachim Vollrath (Robert Bosch GmbH), and Bogdan Coseriu (Robert Bosch GmbH)</i>	
RAD: Retrieval-Augmented Decision-Making of Meta-Actions with Vision-Language Models in Autonomous Driving .....	3838
<i>Yujin Wang (School of Automotive Studies, Tongji University), Quanfeng Liu (School of Automotive Studies, Tongji University), Zhengxin Jiang (Shanghai Research Institute for Intelligent Autonomous Systems, Tongji University), Tianyi Wang (Yale University), Junfeng Jiao (University of Texas at Austin), Hongqing Chu (School of Automotive Studies, Tongji University), Bingzhao Gao (School of Automotive Studies, Tongji University), and Hong Chen (College of Electronic and Information Engineering, Tongji University)</i>	
ReferGPT: Towards Zero-Shot Referring Multi-Object Tracking .....	3849
<i>Tzoulis Chamiti (Vrije Universiteit Brussel, Belgium ), Leandro Di Bella (Vrije Universiteit Brussel, Belgium ), Adrian Munteanu (Vrije Universiteit Brussel, Belgium ), and Nikos Deligiannis (Vrije Universiteit Brussel, Belgium )</i>	

Drive4C: A Closed-Loop Benchmark on What Foundation Models Really Need to Be Capable of for Language-Guided Autonomous Driving .....	3859
<i>Tin Stribor Sohn (Dr. Ing. h.c. F. Porsche AG, Germany), Maximilian Dillitzer (Dr. Ing. h.c. F. Porsche AG, Germany), Johannes Bach (Dr. Ing. h.c. F. Porsche AG, Germany), Jason J. Corso (University of Michigan, USA), Tim Brühl (Dr. Ing. h.c. F. Porsche AG, Germany), Robin Schwager (Dr. Ing. h.c. F. Porsche AG, Germany), Tim Dieter Eberhardt (Dr. Ing. h.c. F. Porsche AG, Germany), and Eric Sax (Karlsruhe Institute of Technology, Germany)</i>	
ReasonDrive: Efficient Visual Question Answering for Autonomous Vehicles with Reasoning-Enhanced Small Vision-Language Models .....	3870
<i>Amirhosein Chahe (Drexel University, USA) and Lifeng Zhou (Drexel University, USA)</i>	

## **Rhobin2025: The Third Rhobin Challenge on Reconstruction of Human-Object Interaction**

CondiMen: Conditional Multi-Person Mesh Recovery .....	3880
<i>Romain Brégier (Naver Labs Europe), Fabien Baradel (Naver Labs Europe), Thomas Lucas (Naver Labs Europe), Salma Galaaoui (Valeo, ENPC, Ecole Nationale des Ponts et Chausees), Matthieu Armando (Naver Labs Europe), Philippe Weinzaepfel (Naver Labs Europe), and Grégory Rogez (Naver Labs Europe)</i>	
Physics-based Human Pose Estimation from a Single Moving RGB Camera .....	3891
<i>Ayce Idil Aytekin (Max Planck Institute for Informatics), Chuqiao Li (University of Tübingen), Diogo Luvizon (Max Planck Institute for Informatics), Rishabh Dabral (Max Planck Institute for Informatics), Martin Oswald (University of Amsterdam), Marc Habermann (Max Planck Institute for Informatics), and Christian Theobalt (Max Planck Institute for Informatics)</i>	
Short-term 3D Human Mesh Recovery with Virtual Markers Disentanglement .....	3901
<i>Xiyuan Kang (University of Surrey), Yi Yuan (University of Surrey), Xu Dong (University of Surrey), Muhammad Awais (University of Surrey), Lilian Tang (University of Surrey), Josef Kittler (University of Surrey), and Zhenhua Feng (Jiangnan University)</i>	
PoseSynViT: Lightweight and Scalable Vision Transformers for Human Pose Estimation .....	3912
<i>Sonain Jamil (Faculty of Science and Technology (FST), University of Jean Monnet (UJM), Saint-Etienne, France)</i>	

## **VAND: Visual Anomaly and Novelty Detection - 3rd Edition**

Feature Attenuation of Defective Representation Can Resolve Incomplete Masking on Anomaly Detection .....	3922
<i>YeongHyeon Park (SungKyunKwan University, SK Planet), Sungho Kang (SungKyunKwan University), Myung Jin Kim (SK Planet), Hyeong Seok Kim (SK Planet), and Juneho Yi (SungKyunKwan University)</i>	
Multi-Flow: Multi-View-Enriched Normalizing Flows for Industrial Anomaly Detection .....	3933
<i>Mathis Kruse (Leibniz University Hannover) and Bodo Rosenhahn (Leibniz University Hannover)</i>	

SK-RD4AD : Skip-Connected Reverse Distillation For Robust One-Class Anomaly Detection .....	3945
<i>EunJu Park (Ewha Womans University, South Korea), Taekyung Kim (Dongguk University, South Korea), Minju Kim (Chung-Ang University, South Korea), Hojun Lee (Keimyung University, South Korea), and Gil-Jun Lee (SK On, South Korea)</i>	
Multi-layer Radial Basis Function Networks for Out-of-distribution Detection .....	3954
<i>Amol Khanna (Booz Allen Hamilton), Chenyi Ling (Booz Allen Hamilton), Derek Everett (Booz Allen Hamilton), Edward Raff (Booz Allen Hamilton), and Nathan Inkawhich (U.S. Air Force Research Laboratory)</i>	
Automated Essential Concept Discovery for Few-Shot Out-of-Distribution Detection .....	3964
<i>Guangyao Chen (Cornell University), Kai Horstmann (Cornell University), Zhilong Wang (Cornell University), and Fengqi You (Cornell University)</i>	
SmartHome-Bench: A Comprehensive Benchmark for Video Anomaly Detection in Smart Homes Using Multi-Modal Large Language Models .....	3975
<i>Xinyi Zhao (University of Washington), Congjing Zhang (University of Washington), Pei Guo (Wyze Labs, Inc.), Wei Li (Wyze Labs, Inc.), Lin Chen (Wyze Labs, Inc.), Chaoyue Zhao (University of Washington), and Shuai Huang (University of Washington)</i>	
No-MambAAD: Revitalizing Conv-Only Networks for Unsupervised Anomaly Detection .....	3986
<i>Masud An Nur Islam Fahim (University of Vaasa) and Jani Boutellier (University of Vaasa)</i>	
Scene-Specific Anomalous Relationship Detection Using Scene Graph Summarization .....	3995
<i>Yu-Chen Lai (NCKU), Motoharu Sonogashira (RIKEN), Itthisak Phueaksri (RIKEN), and Yasutomo Kawanishi (RIKEN)</i>	
Semi-supervised Object-Wise Anomaly Detection for Firearm and Firearm Component Detection in X-ray Security Imagery .....	4004
<i>Yona Falinie A. Gaus (Durham University), Brian K.S. Isaac Medina (Durham University), Neelanjan Bhowmik (Durham University), Yam T. Lee (Durham University), and Toby Breckon (Durham University)</i>	
Beyond Academic Benchmarks: Critical Analysis and Best Practices for Visual Industrial Anomaly Detection .....	4015
<i>Aimira Baitieva (Valeo), Yacine Bouaouni (Valeo), Alexandre Briot (Valeo), Dick Ameln (Intel), Souhaïel Khalfaoui (Valeo), and Samet Akcay (Intel)</i>	
PaSTe: Improving the Efficiency of Visual Anomaly Detection at the Edge .....	4026
<i>Manuel Barusco (University of Padova, Italy), Francesco Borsatti (University of Padova, Italy), Davide Dalle Pezze (University of Padova, Italy), Francesco Paissan (Fondazione Bruno Kessler, Italy), Elisabetta Farella (Fondazione Bruno Kessler, Italy), and Gian Antonio Susto (University of Padova, Italy)</i>	
FusedVision: A Knowledge-Infusing Approach for Practical Anomaly Detection in Real-world Surveillance Videos .....	4036
<i>Khaled Dawoud (MBZUAI), Zaigham Zaheer (MBZUAI), Mustaqeem Khan (MBZUAI), Karthik Nandakumar (Michigan State University), Abdulmotaleb Elsadik (University of Ottawa), and Muhammad Haris Khan (MBZUAI)</i>	

Robust AD: A Real World Benchmark Dataset For Robustness in Industrial Anomaly Detection ..	4047
<i>Latha Pemula (AWS AI Labs), Dongqing Zhang (Amazon), and Onkar Dabeer (AWS AI Labs)</i>	
Detect, Classify, Act: Categorizing Industrial Anomalies with Multi-Modal Large Language Models .....	4058
<i>Sassan Mokhtar (University of Bonn), Arian Mousakhan (University of Freiburg), Silvio Galesso (University of Freiburg), Jawad Tayyub (Endress+Hauser), and Thomas Brox (University of Freiburg)</i>	
When Textures Deceive: Weakly Supervised Industrial Anomaly Detection with Adapted-Loss CycleGAN .....	4068
<i>Tapan Ganatma Nakkina (Texas A&amp;M University), Yuhao Zhong (Texas A&amp;M University), Pete Sumethasorn (Texas A&amp;M University), Haopeng Tian (Texas A&amp;M University), and Satish Bukkapatnam (Texas A&amp;M University)</i>	

## **The 3rd Workshop on Sign Language Recognition, Translation and Production**

Cross-Modal Consistency Learning for Sign Language Recognition .....	4078
<i>Kepeng Wu (University of Science and Technology of China), Zecheng Li (University of Science and Technology of China), Weichao Zhao (University of Science and Technology of China), Hezhen Hu (University of Texas at Austin), Wengang Zhou (University of Science and Technology of China), and Houqiang Li (University of Science and Technology of China)</i>	
Diffusion-Based Continuous Sign Language Generation with Cluster-Specific Fine-Tuning and Motion-Adapted Transformer .....	4088
<i>Razieh Rastgoo (Semnan University), Kourosh Kiani (Semnan University), and Sergio Escalera (Barcelona University)</i>	
CLIP-SLA: Parameter-Efficient CLIP Adaptation for Continuous Sign Language Recognition .....	4098
<i>Sarah Alyami (Information and Computer Science Department, King Fahd University of Petroleum and Minerals and Computing Department, Imam Abdulrahman Bin Faisal University) and Hamzah Luqman (Information and Computer Science Department, King Fahd University of Petroleum and Minerals and 3SDAIA–KFUPM Joint Research Center for Artificial Intelligence)</i>	

SLRTP2025 Sign Language Production Challenge: Methodology, Results and Future Work .....	4109
<i>Harry Walsh (University of Surrey), Ed Fish (University of Surrey),  Ozge Mercanoglu Sincan (University of Surrey), Mohamed Ilyes Lakhel  (University of Surrey), Richard Bowden (University of Surrey), Neil  Fox (University of Birmingham), Bencie Woll (University College  London), Kepeng Wu (University of Science and Technology of China),  Zecheng Li (University of Science and Technology of China), Weichao  Zhao (University of Science and Technology of China), Haodong Wang  (University of Science and Technology of China), Wengang Zhou  (University of Science and Technology of China), Houqiang Li  (University of Science and Technology of China), Shengeng Tang (Hefei  University of Technology), Jiayi He (Hefei University of Technology),  Xu Wang (Hefei University of Technology), Ruobei Zhang (Hefei  University of Technology), Yaxiong Wang (Hefei University of  Technology), Lechao Cheng (Hefei University of Technology), Meryem  Tasyurek (Hacettepe University), Tugce Kiziltepe (Hacettepe  University), and Hacer Yalim Keles (Hacettepe University)</i>	

## Workshop on Pixel-level understanding with Vision Foundation Models

Hierarchical Semantic Segmentation with Autoregressive Language Modeling .....	4120
<i>Josh Myers-Dean (University of Colorado Boulder), Brian Price (Adobe),  Yifei Fan (Adobe), and Danna Gurari (University of Colorado Boulder)</i>	
Prompt-Guided Attention Head Selection for Focus-Oriented Image Retrieval .....	4131
<i>Yuji Nozawa (Kioxia Corporation), Yu-Chieh Lin (Kioxia Corporation),  Kazumoto Nakamura (Kioxia Corporation), and Youyang Ng (Kioxia  Corporation)</i>	
ITACLIP: Boosting Training-Free Semantic Segmentation with Image, Text, and Architectural Enhancements .....	4142
<i>M. Arda Aydın (Bilkent University), Efe Mert Çırpar (RWTH Aachen  University), Elvin Abidinli (Technical University of Munich), Gozde  Unal (Istanbul Technical University), and Yusuf H. Sahin (Istanbul  Technical University)</i>	
Show or Tell? A Benchmark To Evaluate Visual and Textual Prompts in Semantic Segmentation ..	4153
<i>Gabriele Rosi (Politecnico di Torino) and Fabio Cermelli (Focoos AI)</i>	

## 2nd Workshop on Multi-Agent Embodied Intelligent Systems Meet Generative-AI Era: Opportunities, Challenges and Futures

SwarmDiff: Swarm Robotic Trajectory Planning in Cluttered Environments via Diffusion Transformer .....	4164
<i>Kang Ding (Peking University), Chunxuan Jiao (Peking University),  Yunze Hu (Peking University), Kangjie Zhou (Peking University),  Pengying Wu (Peking University, China), Yao Mu (The University of Hong  Kong), and Chang Liu (Peking University)</i>	

Efficient Task-specific Conditional Diffusion Policies: Shortcut Model Acceleration and SO(3) Optimization .....	4174
<i>Haiyong Yu (D-Robotics, China), Yanqiong Jin (D-Robotics, China), Yonghao He (D-Robotics, China), and Wei Sui (D-Robotics, China)</i>	
AI Hiring with LLMs: A Context-Aware and Explainable Multi-Agent Framework for Resume Screening .....	4184
<i>Frank P.-W. Lo (Imperial College London), Jianing Qiu (The Chinese University of Hong Kong), Zeyu Wang (Imperial College London), Haibao Yu (The University of Hong Kong), Yeming Chen (Wedon Education Technologies), Gao Zhang (Brest Business School), and Benny Lo (Imperial College London)</i>	
Multi-Agent Systems for Robotic Autonomy with LLMs .....	4194
<i>Junhong Chen (Imperial College London), Ziqi Yang (Imperial College London), Haoyuan G Xu (Imperial College London), Dandan Zhang (Imperial College London), and George Mylonas (Imperial College London)</i>	
An LLM-enabled Multi-Agent Autonomous Mechatronics Design Framework .....	4205
<i>Zeyu Wang (Imperial College London), Frank Po Wen Lo (Imperial College London), Qian Chen (Imperial College London), Yongqi Zhang (Imperial College London), Chen Lin (Imperial College London), Xu Chen (Imperial College London), Zhenhua Yu (University of Aberdeen), Alexander J. Thompson (Imperial College London), Eric M. Yeatman (Imperial College London), and Benny P. L. Lo (Imperial College London)</i>	
Deciding the Path: Leveraging Multi-Agent Systems for Solving Complex Tasks .....	4216
<i>Iman Abbasnejad (Amazon), Xuefeng Liu (Amazon), and Atanu Roy (Amazon)</i>	
LangCoop: Collaborative Driving with Language .....	4226
<i>Xiangbo Gao (Texas A&amp;M University), Yuheng Wu (Korea Advanced Institute of Science &amp; Technology), Rujia Wang (Texas A&amp;M University), Chenxi Liu (The University of Utah), Yang Zhou (Texas A&amp;M University), and Zhengzhong Tu (Texas A&amp;M University)</i>	

## Computer Vision for Drug Discovery Where Are We and What is Beyond?

LNTransformer: Lung Nodule Transformer for Sparse CT Segmentation .....	4238
<i>Hooman Ramezani (University of Toronto), Charlotte Vadrines (University of Toronto), Dionne Aleman (University of Toronto), and Daniel Letourneau (University of Toronto, Canada)</i>	
HCS-DFC: A Diffusion Classifier for Mode of Action Prediction Using Morphological Profiles.....	4246
<i>Jakub Kościukiewicz (Jagiellonian University, Poland; Ardigen S.A., Poland), Dawid Rymarczyk (Jagiellonian University, Poland; Ardigen S.A., Poland), and Bartosz Zieliński (Jagiellonian University)</i>	
Revolutionizing Drug Discovery: Integrating Spatial Transcriptomics with Advanced Computer Vision Techniques .....	4252
<i>Zichao Li (Canoakbit Alliance Inc.), Shiqing Qiu (Chengdu University of Technology), and Zong Ke (National University of Singapore)</i>	

Towards exploring continual learning for toxicologic pathology in pharmaceutical drug discovery .....	4259
<i>Arijit Patra (UCB Pharma, University of Glasgow), Jing Wu (University College London), Honghan Wu (University College London), and Anshul Thakur (University of Oxford)</i>	
Drug Discovery Agent: An Automated Vision Detection System for Drug-Cell Interactions .....	4269
<i>Adib Bazgir (University of Missouri, USA) and Yuwen Zhang (University of Missouri, USA)</i>	
Bridging Self-Supervision and Mechanism of Action Discovery in Morphological Profiling .....	4278
<i>Syed Sameed Husain (University of Surrey, UK), Jan Bober (ForecomAI, UK), Amaia Irizar (ForecomAI, UK), and Miroslaw Bober (University of Surrey, UK)</i>	
Segment AnyNeuron .....	4286
<i>Taha Razzaq (Tibbling Technologies), Ahmed Rashid Qazi (Tibbling Technologies), and Asim Iqbal (Tibbling Technologies)</i>	
Bridging Morphology and Molecular Signatures: Multi-Task Deep Learning for Multi-Omics Prediction from Histopathology .....	4294
<i>Fatemeh Dashti Ahangar (Department of Electrical and Computer Engineering, University of Central Florida) and Jiann-Shiun Yuan (Department of Electrical and Computer Engineering, University of Central Florida)</i>	
CellRep: A Multichannel Image Representation Learning Model .....	4303
<i>Lawrence Phillips (Novo Nordisk ) and Rory Donovan-Maiye (Novo Nordisk )</i>	

## 7th Safe Artificial Intelligence for All Domains

Out-of-Distribution Segmentation in Autonomous Driving: Problems and State of the Art .....	4310
<i>Youssef Shoeb (Continental AG), Azarm Nowzad (Continental AG), and Hanno Gottschalk (Technische Universität Berlin)</i>	
Improving Weather-based OOD Generalisation in Lidar-based Object Detection Models via Adversarial Training .....	4321
<i>Ben Batten (Imperial College London) and Alessio Lomuscio (Imperial College London)</i>	
Universal Shape of Strong Remote Adversarial Patches for Object Detection with Convolutional Neural Networks .....	4330
<i>Kento Oonishi (Mitsubishi Electric Corporation, Japan) and Tsunato Nakai (Mitsubishi Electric Corporation, Japan)</i>	
Detecting Localized Deepfake Manipulations Using Action Unit-Guided Video Representations ..	4341
<i>Tharun Anand (Indian Institute Of Technology Madras), Siva Sankar (Indian Institute Of Technology Madras), and Pravin Nair (Indian Institute Of Technology Madras)</i>	
CTC: Contribution to Classification of Complex Features .....	4352
<i>Sophia Kalanovska (King's College London), Michael Luck (University of Sussex), and Christopher Hampson (King's College London)</i>	

Robustness Evaluation for Video Models with Reinforcement Learning .....	4362
<i>Ashwin Ramesh Babu (Hewlett Packard Enterprise Labs), Sajad Mousavi (Hewlett Packard Enterprise Labs), Vineet Gundecha (Hewlett Packard Enterprise Labs), Sahand Ghorbanpour (Hewlett Packard Enterprise Labs), Avisek Naug (Hewlett Packard Enterprise Labs), Antonio Guillen (Hewlett Packard Enterprise Labs), Ricardo Luna (Hewlett Packard Enterprise Labs), and Soumyendu Sarkar (Hewlett Packard Enterprise Labs)</i>	
How Much Noise is there in Labels Generated by Humans? A Method to Validate Automatically Generated Bounding Boxes. ....	4371
<i>Mariusz Nowak (Aptiv), Jacek Cyranka (Aptiv), Natalia Maslany (Aptiv, Jagiellonian University), Aleksander Kostuch (Aptiv, AGH University of Krakow), Jakub Derbisz (Aptiv), Mateusz Komorkiewicz (Aptiv, AGH University of Krakow), Patryk Siwek (Aptiv, AGH University of Krakow), Mateusz Wójcik (Aptiv), Dariusz Marchewka (Aptiv, AGH University of Krakow), and Paweł Skruch (Aptiv, AGH University of Krakow)</i>	
Traffic Sign Recognition Under Visual Perturbations: Shadows, Light Patches, and Simulated Obstructions .....	4381
<i>Muneeb A. Khan (Sangmyung University), Yujin Choi (Sangmyung University), Jiho Eum (Sangmyung University), and Heemin Park (Sangmyung University)</i>	
Out-of-Distribution Detection with Adversarial Outlier Exposure .....	4391
<i>Thomas Botschen (OvGU Magdeburg), Konstantin Kirchheim (OvGU Magdeburg), and Frank Ortmeier (OvGU Magdeburg)</i>	
Uncertainty Aware Training to Improve Uncertainty Active Learning for Semantic Segmentation .....	4401
<i>Moritz Thoma (BMW Group, Germany), Tobias Preintner (BMW Group, Germany), Emad Aghajanzadeh (BMW Group, Germany), Shambhavi Balamuthu Sampath (BMW Group, Germany), Pierpaolo Mori (BMW Group, Germany), Nael Fafous (BMW Group, Germany), Manoj-Rohit Vemparala (BMW Group, Germany), Alexander Frickenstein (BMW Group, Germany), Daniel Mueller-Gritschneider (TU-Vienna, Austria), and Ulf Schlichtmann (TU-Munich, Germany)</i>	

## 21th Workshop on Perception Beyond the Visible Spectrum (PBVS'2025)

A True Hyperspectral Image Super-Resolution Dataset .....	4412
<i>Alexander Ulrichsen (University of Strathclyde), Thomas De Kerf (University of Antwerp), David Dunphy (University of Strathclyde), Paul Murray (University of Strathclyde), Steve Vanlanduit (University of Antwerp), and Stephen Marshall (University of Strathclyde)</i>	

3D Face Reconstruction From Radar Images .....	4422
<i>Valentin Braeutigam (Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany), Vanessa Wirth (Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany), Ingrid Ullmann (Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany), Christian Schüßler (Friedrich-Alexander Universität Erlangen-Nürnberg, Germany), Martin Vossiek (Friedrich-Alexander Universität Erlangen-Nürnberg, Germany), Matthias Berking (Friedrich-Alexander Universität Erlangen-Nürnberg, Germany), and Bernhard Egger (Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany)</i>	
Fusion or Confusion? A Look at Dataset Pooling for Infrared Object Detection .....	4432
<i>Stefan Becker (Fraunhofer IOSB), Ann-Kristin Grosselfinger (Fraunhofer IOSB), Jens Bayer (Fraunhofer IOSB), David Münch (Fraunhofer IOSB), Wolfgang Hübner (Fraunhofer IOSB), and Michael Arens (Fraunhofer IOSB)</i>	
ToF-360 - A Panoramic Time-of-flight RGB-D Dataset for Single Capture Indoor Semantic 3D Reconstruction .....	4442
<i>Hideaki Kanayama (Ricoh Company, Ltd. Japan), Mahdi Chamseddine (German Research Center for Artificial Intelligence, DFKI, Germany), Suresh Guttikonda (German Research Center for Artificial Intelligence, DFKI, Germany), So Okumura (Ricoh Company, Ltd. Japan), Soichiro Yokota (Ricoh Company, Ltd. Japan), Didier Stricker (German Research Center for Artificial Intelligence, DFKI, Germany), and Jason Rambach (German Research Center for Artificial Intelligence, DFKI, Germany)</i>	
RADLER: Radar Object Detection Leveraging Semantic 3D City Models and Self-Supervised Radar-Image Learning .....	4452
<i>Yuan Luo (GPP Communication GmbH, Germany), Rudolf Hoffmann (GPP Communication GmbH, Germany), Yan Xia (Technical University of Munich, Germany), Olaf Wysocki (Technical University of Munich, Germany), Benedikt Schwab (Technical University of Munich, Germany), Thomas H. Kolbe (Technical University of Munich, Germany), and Daniel Cremers (Technical University of Munich, Germany)</i>	
Aerial Infrared Health Monitoring of Solar Photovoltaic Farms at Scale .....	4462
<i>Isaac Corley (Zeitview), Conor Wallace (Zeitview), Sourav Agrawal (Zeitview), Burton Putrah (Zeitview), and Jonathan Lwowski (Zeitview)</i>	
GaussianVideo: Efficient Video Representation and Compression by Gaussian Splatting .....	4471
<i>Inseo Lee (Seoul National University, South Korea), Youngyoon Choi (Seoul National University, South Korea), and Joonseok Lee (Seoul National University, South Korea)</i>	
XiEff Representation for Interpretable Near-Field Imaging .....	4481
<i>Vasyl Vasylenko (Individual Researcher), Ihor Tymchyshyn (Taras Shevchenko National University of Kyiv), and Vitalii Tymchyshyn (Bogolyubov Institute for Theoretical Physics; Kyiv Academic University)</i>	
LangGas: Introducing Language in Selective Zero-Shot Background Subtraction for Semi-Transparent Gas Leak Detection with a New Dataset .....	4490
<i>Wenqi Guo (University of British Columbia, Canada), Yiyang Du (University of British Columbia, Canada), and Shan Du (University of British Columbia, Canada)</i>	

Define, Refine, Align: Correspondence-free 3D Line Alignment with Attentional, Equivariant and Rotational Layers .....	4501
<i>Alberto Pepe (University of Cambridge), Yuxin Yao (University of Cambridge), and Joan Lasenby (University of Cambridge)</i>	
Enhancing Multi-modal Automatic Target Recognition using Out-of-Distribution Exploitation (MATRODE) .....	4512
<i>Hongzhi Guo (University of Nebraska-Lincoln), Paul T. Schrader (Information Directorate (AFRL/RIGC), Air Force Research Laboratory), and Erik Blasch (Air Force Office of Scientific Research)</i>	
SmallGS: Gaussian Splatting-based Camera Pose Estimation for Small-Baseline Videos .....	4521
<i>Yuxin Yao (University of Cambridge, UK), Yan Zhang (Meshcapade, Germany), Zhening Huang (University of Cambridge, UK), and Joan Lasenby (University of Cambridge, UK)</i>	
S-Band SAR Target Classification via 2D and 3D Deep Learning Methods .....	4531
<i>Tyler Rust (University of Delaware), Michael Pergeorelis (University of Delaware), Chandra Kambhamettu (University of Delaware), and Colin Kelly (Army Research Laboratory)</i>	
A Strong Baseline for Multi-Person Tracking in Thermal Infrared Imagery .....	4540
<i>Daniel Stadler (Fraunhofer IOSB) and Andreas Specker (Fraunhofer IOSB)</i>	
UPPET: Unified Pedestrian Pose Estimation in Thermal Imaging .....	4551
<i>Mickael Cormier (Fraunhofer IOSB; Karlsruhe Institute of Technology), Andreas Specker (Fraunhofer IOSB), and Jürgen Beyerer (Fraunhofer IOSB; Karlsruhe Institute of Technology)</i>	
Multi-Spectral Imaging and Data Fusion for Real-Time Bleeding Detection .....	4561
<i>Ghazal Rouhafzay (Université de Moncton), Stephen Rowlands (Spectronix Inc.), Angel J. Valencia (University of Ottawa), Shengsong Yang (University of Ottawa), Pierre Payeur (University of Ottawa), Haitao Tian (University of Ottawa), and James Dickens (University of Ottawa)</i>	
Cross-Spectral Body Recognition with Side Information Embedding: Benchmarks on LLCM and Analyzing Range-Induced Occlusions on IJB-MDF .....	4569
<i>Anirudh Nanduri (University of Maryland, College Park, MD, USA), Siyuan Huang (Johns Hopkins University, Baltimore, MD, USA), and Rama Chellappa (Johns Hopkins University, Baltimore, MD, USA)</i>	
SAGA: Semantic-Aware Gray color Augmentation for Visible-to-Thermal Domain Adaptation across Multi-View Drone and Ground-Based Vision Systems .....	4578
<i>Manjunath D (Department of Aerospace Engineering, Indian Institute of Science, Bengaluru, India), Aniruddh Sikdar (Robert Bosch Centre for Cyber Physical Systems, Indian Institute of Science, Bengaluru, India), Prajwal Gurunath (Department of Aerospace Engineering, Indian Institute of Science, Bengaluru, India), Sumanth Udupa (Department of Aerospace Engineering, Indian Institute of Science, Bengaluru, India), and Suresh Sundaram (Department of Aerospace Engineering, Indian Institute of Science, Bengaluru, India)</i>	

SwinPaste: A Swin Transformer-Based Framework for RGB-Guided Thermal Image Super-Resolution .....	4589
<i>Hang Zhong (School of Computer Science and Technology, Tongji University, China), Yu Wang (School of Computer Science and Technology, Tongji University, China), and Shengjie Zhao (School of Computer Science and Technology, Tongji University, China)</i>	
Open Dataset and Enhancement Method for Long-wave Thermal Diurnal Material Classification .....	4595
<i>Michael Pergeorelis (University of Delaware), Tyler Rust (University of Delaware), and Chandra Kambhamettu (University of Delaware)</i>	
Thermal Pedestrian Multiple Object Tracking Challenge (TP-MOT) .....	4602
<i>Wassim El Ahmar (University of Ottawa), Angel Sappa (Computer Vision Center, ESPOL Polytechnic University), and Riad Hammoud (PlusAi Inc)</i>	
Dual-Input Frequency-Aware Network for High-Quality Thermal Image Super-Resolution .....	4610
<i>Priya Kansal (Couger Inc,Japan) and Sabari Nathan (Couger Inc,Japan)</i>	
Thermal Image Super-Resolution Challenge Results - PBVS 2025 .....	4621
<i>Rafael E. Rivadeneira (Escuela Superior Politécnica del Litoral, Ecuador), Angel D. Sappa (Computer Vision Center, Spain), Riad Hammoud (Plus.ai, USA), Jiyong Rao (Tongji University, China), Hang Zhong (Tongji University, China), Yu Wang (Tongji University, China), Shengjie Zhao (Tongji University, China), Zhiwei Zhong (City University of Hong Kong, Hong Kong), Yung-Hui Li (Hon Hai Research Institute, Taiwan), Shiqi Wang (City University of Hong Kong, Hong Kong), Qiangqiang Shen (Harbin Institute of Technology, Shenzhen), Hanzhang Wang (Harbin Institute of Technology, Shenzhen), and Xuanqi Zhang (University of British Columbia, Canada)</i>	
4th Multi-modal Aerial View Image Challenge: SAR Classification - PBVS 2025 .....	4631
<i>Nathan Inkawhich (AFRL, USA), Claire Thorp (AFRL, USA), Justice Wheelwright (AFRL, USA), Oliver Nina (AFRL, USA), Dylan Bowald (AFRL, USA), Angel Sappa (ESPOL Polytechnic University, Ecuador), and Erik Blasch (AFRL, USA)</i>	
3rd Multi-modal Aerial View Image Challenge: Sensor Domain Translation - PBVS 2025 .....	4640
<i>Dylan Bowald (AFRL, USA), Justice Wheelwright (AFRL, USA), Oliver Nina (AFRL, USA), Angel Sappa (ESPOL Polytechnic University, Ecuador), Riad Hammoud (PlusAI Inc., USA), Erik Blasch (AFRL, USA), and Nathan Inkawhich (AFRL, USA)</i>	
Probabilistic Perspective-n-lines for Indoor Camera Pose Estimation .....	4650
<i>Xiaowei Chen (University of Central Arkansas) and Guoliang Fan (Oklahoma State University)</i>	
CSRN: Cross-Sensor Robust Recognition Network for Multi-modal Aerial View Object Classification .....	4660
<i>Hongli Liu (Tongji University), Wang Yu (Tongji University), and Shengjie Zhao (Tongji University)</i>	

# Workshop on Computer Vision for Microscopy Image Analysis

Fast Sphericity and Roundness approximation in 2D and 3D using Local Thickness .....	4667
<i>Pawel Tomasz Pieta (Technical University of Denmark), Peter Winkel Rasmussen (Technical University of Denmark), Anders Bjorholm Dahl (Technical University of Denmark), and Anders Nymark Christensen (Technical University of Denmark)</i>	
SAM4EM: Efficient memory-based two stage prompt-free segment anything model adapter for complex 3D neuroscience electron microscopy stacks .....	4678
<i>Uzair Shah (CSE, HBKU), Marco Agus (CSE, HBKU), Daniya Boges (CEMSE, KAUST), Vanessa Chiappini (University of Turin), Mahmood Alzubaidi (CSE, HBKU), Jens Schneider (CSE, HBKU), Markus Hadwiger (CEMSE, KAUST), Pierre J. Magistretti (BESE, KAUST), Mowafa Househ (CSE, HBKU), and Corrado Calì (University of Turin)</i>	
Prompt-Tuning SAM: From Generalist to Specialist with only 2048 Parameters and 16 Training Images .....	4688
<i>Tristan Piater (Carl Zeiss AG), Björn Barz (Carl Zeiss AG), and Alexander Freytag (Carl Zeiss AG)</i>	
A Novel 3D Decoder with Weighted and Learnable Triple Attention for 3D Microscopy Image Segmentation .....	4699
<i>Siyavash Shabani (university of nevada, Reno), Sahar Mohammed (University of Nevada, Reno), and Bahram Parvin (University of Nevada, Reno)</i>	
Generalizable Unsupervised Microscopy Video Denoising via Weighted SpatioTemporal Sampling.....	4709
<i>Mary Aiyetigbo (Clemson University), Wanqi Yuan (Clemson University), Feng Luo (Clemson University), Xin Li (University at Albany - State University of New York), Tong Ye (Clemson University), and Nianyi Li (Clemson University)</i>	
Zero-Shot Denoising for Fluorescence Lifetime Imaging Microscopy with Intensity-Guided Learning .....	4719
<i>Hao Chen (University of Notre Dame, USA), Julian Najera (University of Notre Dame, USA), Dagmawit Geresu (University of Notre Dame, USA), Meenal Datta (University of Notre Dame, USA), Cody Smith (University of Notre Dame, USA), and Scott Howard (University of Notre Dame, USA)</i>	
Low-Frame-Rate Cell Tracking: Unmet Needs and Future Directions .....	4729
<i>Mina Gachloo (Clemson University), Akhila Nangineedi (Clemson University), Mahsa Partovi (Clemson University), Fardifa Fathmiul Alam (Clemson University), Tzu-Yu Chu (Clemson University), James Schvaneveldt (Clemson University), Xiaoming Lu (Clemson University), Tirthankar Biswas (Clemson University), Marc Russel Birtwistle (Clemson University), and Federico Iuricich (Clemson University)</i>	
IAUNet: Instance-Aware U-Net .....	4739
<i>Yaroslav Prytula (University of Tartu, Estonia), Illia Tsiporenko (University of Tartu, Estonia), Ali Zeynalli (University of Tartu, Estonia), and Dmytro Fishman (University of Tartu, Estonia)</i>	

CytoFM: The first cytology foundation model .....	4749
<i>Vedrana Ivezic (UCLA), Ashwath Radhachandran (UCLA), Ekaterina Redekop (UCLA), Shreeram Athreya (UCLA), Dongwoo Lee (UCLA), Vivek Sant (UT Southwestern), Corey Arnold (UCLA), and William Speier (UCLA)</i>	
Beyond Neurofibrillary Tangles: Explainable AI for Microscopic Tauopathy Classification in Immunofluorescence Imaging .....	4758
<i>Jesus Dassaef López-Barrios (Tecnologico de Monterrey, Mexico), Miguel Angel Ontiveros-Torres (Tecnologico de Monterrey, Mexico; National Dementia BioBank, AMPAEYDEN A.C., Federación Mexicana de Alzheimer, Mexico), and Jose Antonio Cantoral-Ceballos (Tecnologico de Monterrey, Mexico)</i>	

## Foundation Models for V2X-Based Cooperative Autonomous Driving

V3LMA: Visual 3D-enhanced Language Model for Autonomous Driving .....	4769
<i>Jannik Lübberstedt (Technical University of Munich), Esteban Rivera Guerrero (Technical University of Munich), Nico Uhlemann (Technical University of Munich), and Markus Lienkamp (Technical University of Munich)</i>	
Can Vision-Language Models Understand and Interpret Dynamic Gestures from Pedestrians? Pilot Datasets and Exploration Towards Instructive Nonverbal Commands for Cooperative Autonomous Vehicles .....	4779
<i>Tonko E. W. Bossen (University of California Merced, US), Andreas Møgelmoose (Aalborg University, Denmark), and Ross Greer (University of California Merced, US)</i>	
Exploring Modality Guidance to Enhance VFM-based Feature Fusion for UDA in 3D Semantic Segmentation .....	4789
<i>Johannes Spoecklberger (Graz University of Technology), Wei Lin (JKU LINZ), Pedro Hermosilla (TU Wien), Sivan Doveh (IBM Research), Horst Possegger (Institute of Visual Computing), and M. Jehanzeb Mirza (MIT CSAIL)</i>	

## Mechanistic Interpretability for Vision

Disentangling Polysemantic Channels in Convolutional Neural Networks .....	4799
<i>Robin Hesse (Technical University of Darmstadt), Jonas Fischer (Max Planck Institute for Informatics), Simone Schaub-Meyer (Technical University of Darmstadt, hessian.AI), and Stefan Roth (Technical University of Darmstadt, hessian.AI)</i>	
Naturally Computed Scale Invariance in the Residual Stream of ResNet18 .....	4804
<i>André Longon (Independent)</i>	
Uncovering Branch-specialization in InceptionV1 using k sparse autoencoders .....	4809
<i>Matthew Bozoukov (Miramar college)</i>	

Embedding Shift Dissection on CLIP: Effects of Augmentations on VLM’s Representation Learning .....	4814
<i>Ashim Dahal (University of Southern Mississippi, Hattiesburg, Mississippi, USA), Saydul Akbar Murad (University of Southern Mississippi, Hattiesburg, Mississippi, USA), and Nick Rahimi (University of Southern Mississippi, Hattiesburg, Mississippi, USA)</i>	
Decoding Vision Transformers: the Diffusion Steering Lens .....	4819
<i>Ryota Takatsuki (Araya Inc.; AI Alignment Network; The University of Tokyo), Sonia Joseph (Mila - Quebec Artificial Intelligence Institute; McGill University), Ipppei Fujisawa (Araya Inc.; AI Alignment Network), and Ryota Kanai (Araya Inc.; AI Alignment Network)</i>	
Visualizing and Controlling Cortical Responses Using Voxel-Weighted Activation Maximization .....	4825
<i>Matthew Shinkle (University of Nevada, Reno) and Mark D. Lescroart (University of Nevada, Reno)</i>	
Wavelet-Based Mechanistic Interpretability of Vision Transformers via Frequency-Aware Ablations .....	4830
<i>Sophia J. Abraham (University of Notre Dame), Jonathan D. Hauenstein (University of Notre Dame), and Walter J. Scheirer (University of Notre Dame)</i>	
Analyzing Hierarchical Structure in Vision Models with Sparse Autoencoders .....	4835
<i>Matthew Lyle Olson (Intel Labs), Musashi Hinck (Intel Labs), Neale Ratzlaff (Intel Labs), Changbai Li (Oregon State University), Phillip Howard (Intel Labs), Vasudev Lal (Intel Labs), and Shao-Yen Tseng (Intel Labs)</i>	
Leveraging Vision-Language Foundation Models to Reveal Hidden Image-Attribute Relationships in Medical Imaging .....	4840
<i>Amar Kumar (McGill University, Canada), Anita Kriz (McGill University, Canada), Barak Pertzov (McMaster University, Canada), and Tal Arbel (McGill University, Canada)</i>	
Language-Guided Trajectory Traversal in Disentangled Stable Diffusion Latent Space for Factorized Medical Image Generation .....	4846
<i>Zahra TehraniNasab (McGill University, Canada), Amar Kumar (McGill University, Canada), and Tal Arbel (McGill University, Canada)</i>	
Investigating Mechanisms for In-Context Vision Language Binding .....	4852
<i>Darshana Saravanan (IIIT Hyderabad), Makarand Tapaswi (IIIT Hyderabad), and Vineet Gandhi (IIIT Hyderabad)</i>	

## 5th International Workshop on Event-Based Vision

Probabilistic Online Event Downsampling .....	4857
<i>Andreu Girbau-Xalabarder (Denso IT Laboratory), Jun Nagata (Denso IT Laboratory), and Shinichi Sumiyoshi (Denso IT Laboratory)</i>	

Dynamic EventNeRF: Reconstructing General Dynamic Scenes from Multi-view RGB and Event Streams .....	4866
<i>Viktor Rudnev (Max Planck Institute for Informatics), Gereon Fox (Max Planck Institute for Informatics), Mohamed Elgharib (Max Planck Institute for Informatics), Christian Theobalt (Max Planck Institute for Informatics), and Vladislav Golyanik (Max Planck Institute for Informatics)</i>	
E-VLC: A Real-World Dataset for Event-based Visible Light Communication And Localization ...	4877
<i>Shintaro Shiba (Woven by Toyota), Quan Kong (Woven by Toyota), and Norimasa Kobori (Woven by Toyota)</i>	
Augmented Reality Applications Using Active Markers With An Event Camera .....	4887
<i>Shintaro Shiba (Woven by Toyota), Quan Kong (Woven by Toyota), and Norimasa Kobori (Woven by Toyota)</i>	
Quadrocular, Neuromorphic Stereo Triangulation and Asynchronous Data Fusion for 3D Object Tracking .....	4889
<i>Jonah Sengupta (DEVCOM Army Research Laboratory)</i>	
DELTA: Dense Depth from Events and LiDAR using Transformer's Attention .....	4898
<i>Vincent Brebion (Centre for Environmental and Climate Science, Lund University, Sweden), Julien Moreau (Université de technologie de Compiègne, CNRS, Heudiasyc, France), and Franck Davoine (CNRS, INSA Lyon, UCBL, LIRIS, UMR5205, France)</i>	
Best Linear Unbiased Estimation for 2D and 3D Flow with Event-based Cameras .....	4908
<i>Juan L. Valerdi (Ikerlan) and Xabier Iturbe (Ikerlan)</i>	
Iterative Event-based Motion Segmentation by Variational Contrast Maximization .....	4918
<i>Ryo Yamaki (Keio University), Shintaro Shiba (Keio University), Gallego Guillermo (TU Berlin), and Yoshimitsu Aoki (Keio University )</i>	
Nanoparticle Diameter Measurements With Event Camera Tracking .....	4928
<i>Michael Daugherty (National Institute of Standards and Technology), Matthew DiSalvo (National Institute of Standards and Technology), Aaron Goldfain (National Institute of Standards and Technology), Alexander Peterson (National Institute of Standards and Technology), Edward Kwee (National Institute of Standards and Technology), Thomas Germer (National Institute of Standards and Technology), Gregory Cooksey (National Institute of Standards and Technology), Jagat Budhathoki (National Institute of Standards and Technology), and Peter Bajcsy (National Institute of Standards and Technology)</i>	
Towards Low-Latency Event-based Obstacle Avoidance on a FPGA-Drone .....	4938
<i>Pietro Bonazzi (ETH Zurich, Switzerland), Christian Vogt (ETH Zurich, Switzerland), Michael Jost (ETH Zurich, Switzerland), Lyes Khacef (Sony Semiconductor Solutions Europe, Sony Europe B.V., Zürich), Federico Paredes-Valles (Sony Semiconductor Solutions Europe, Sony Europe B.V., Zürich), and Michele Magno (ETH Zurich, Switzerland)</i>	
EV-Flying: an Event-based Dataset for In-The-Wild Recognition of Flying Objects .....	4947
<i>Gabriele Magrini (University of Florence), Federico Becattini (University of Siena), Giovanni Colombo (University of Florence), and Pietro Pala (University of Florence)</i>	

Demo : Point-Feature Tracking for Pixel Processor Arrays .....	4956
<i>Laurie Bose (VisionChip) and Piotr Dudek (The University Of Manchester)</i>	
Real-Time Pedestrian Detection at the Edge on a Fully Asynchronous Neuromorphic System .....	4958
<i>Hugo Bulzomi (Université Côte d'Azur, IMRA Europe), Alimatou Sadia Memudu (Université Côte d'Azur, IMRA Europe), Yuta Nakano (IMRA Europe), and Jean Martinet (Université Côte d'Azur)</i>	
Event-based Continuous Color Video Decompression from Single Frames .....	4968
<i>Ziyun Wang (University of Pennsylvania, USA), Friedhelm Hamann (TU Berlin, Germany), Kenneth Chaney (University of Pennsylvania, USA), Wen Jiang (University of Pennsylvania, USA), Guillermo Gallego (TU Berlin, ECDF Berlin, SCIOI Cluster and RIG, Germany), and Kostas Daniilidis (University of Pennsylvania, USA; Archimedes, Athena RC)</i>	
Looking into the Shadow: Recording a Total Solar Eclipse with High-resolution Event Cameras .....	4979
<i>Fernando Cladera (University of Pennsylvania), Kenneth Chaney (University of Pennsylvania), Caroline Pritchard (The RNA Institute &amp; Department of Biological Sciences, University at Albany SUNY), M. Ani Hsieh (University of Pennsylvania), Vijay Kumar (University of Pennsylvania), Camillo J. Taylor (University of Pennsylvania), and Kostas Daniilidis (University of Pennsylvania)</i>	
Seeing like a Cephalopod: Colour Vision with a Monochrome Event Camera .....	4984
<i>Sami Arja (Western Sydney University ), Nimrod Kruger (Western Sydney university ), Alexander Marcireau (Western Sydney university ), Nicholas Owen Ralph (Western Sydney University ), Saeed Afshar (Western Sydney university ), and Gregory Cohen (Western Sydney university )</i>	
BiasBench: A reproducible benchmark for tuning the biases of event cameras .....	4994
<i>Andreas Ziegler (University of Tübingen), David Joseph (University of Tübingen), Thomas Gossard (University of Tübingen), Emil Moldovan (University of Tübingen), and Andreas Zell (University of Tübingen)</i>	
Human-Robot Navigation using Event-based Cameras and Reinforcement Learning .....	5004
<i>Ignacio Bugueno-Cordova (Department of Electrical Engineering, Universidad de Chile. Institute of Engineering Sciences, Universidad de O'Higgins), Javier Ruiz-del-Solar (Department of Electrical Engineering, Universidad de Chile. Advanced Mining Technology Center (AMTC), Universidad de Chile), and Rodrigo Verschae (Institute of Engineering Sciences, Universidad de O'Higgins)</i>	
E-BARF: Bundle Adjusting Neural Radiance Fields from a Moving Event Camera .....	5013
<i>Zhipeng Tang (University of Massachusetts Amherst, USA), Shifan Zhu (University of Massachusetts Amherst, USA), Zezhou Cheng (University of Virginia, USA), Donghyun Kim (University of Massachusetts Amherst, USA), and Erik Learned-Miller (University of Massachusetts Amherst, USA)</i>	
Live Demonstration: NeuroTouch - A Neuromorphic Vision-based Tactile Sensor for Real-Time Gesture Recognition .....	5023
<i>Victor Hoffmann (Sony Europe B.V.), Valentina Cavinato (Sony Europe B.V.), and Kirk Scheper (Sony Europe B.V.)</i>	

Egocentric Event-Based Vision for Ping Pong Ball Trajectory Prediction .....	5025
<i>Ivan Alberico (University of Zürich), Marco Cannici (University of Zürich), Giovanni Cioffi (University of Zürich), and Davide Scaramuzza (University of Zürich)</i>	
Reading in the Dark with Foveated Event Vision .....	5035
<i>Carl Brander (University of Zurich, Switzerland), Giovanni Cioffi (University of Zurich, Switzerland), Nico Messikommer (University of Zurich, Switzerland), and Davide Scaramuzza (University of Zurich, Switzerland)</i>	
Making Every Event Count: Balancing Data Efficiency and Accuracy in Event Camera Subsampling .....	5044
<i>Hesam Araghi (Delft University of Technology, Netherlands), Jan van Gemert (Delft University of Technology, Netherlands), and Nergis Tomen (Delft University of Technology, Netherlands)</i>	
Event-Driven Dynamic Attention for Multi-Object Tracking on Neuromorphic Hardware .....	5055
<i>Muhammad Aitsam (Sheffield Hallam University), Sergio Davies (Manchester Metropolitan University), and Alessandro Di Nuovo (Sheffield Hallam University)</i>	
MTevent: A Multi-Task Event Camera Dataset for 6D Pose Estimation and Moving Object Detection .....	5063
<i>Shrutarov Awasthi (TU Dortmund), Anas Gouda (TU Dortmund), Sven Franke (TU Dortmund), Jérôme Rutinowski (TU Dortmund), Frank Hoffmann (TU Dortmund), and Moritz Roidl (TU Dortmund)</i>	
Perturbed State Space Feature Encoders for Optical Flow with Event Cameras .....	5072
<i>Gokul Raju Govinda Raju (ETH Zurich), Nikola Zubic (University of Zurich), Marco Cannici (University of Zurich), and Davide Scaramuzza (University of Zurich)</i>	
Spatio-Temporal State Space Model For Efficient Event-Based Optical Flow .....	5082
<i>Muhammad Ahmed Humais (Khalifa University), Xiaoqian Huang (Khalifa University), Hussain Sajwani (Khalifa University), Sajid Javed (Khalifa University), and Yahya Zweiri (Khalifa University)</i>	
Learning from Noise: Enhancing DNNs for Event-Based Vision through Controlled Noise Injection .....	5092
<i>Marcin Kowalczyk (AGH University of Krakow, Poland), Kamil Jeziorek (AGH University of Krakow, Poland), and Tomasz Kryjak (AGH University of Krakow, Poland)</i>	
Live Demonstration: Real-time event-data processing with Graph Convolutional Neural Networks and SoC FPGA .....	5103
<i>Piotr Wzorek (AGH University of Krakow), Krzysztof Blachut (AGH University of Krakow), Kamil Jeziorek (AGH University of Krakow), and Tomasz Kryjak (AGH University of Krakow)</i>	
Event Quality Score (EQS): Assessing the Realism of Simulated Event Camera Streams via Distance in Latent Space .....	5105
<i>Kaustav Chanda (Arizona State University), Aayush Verma (Arizona State University), Arpitsinh Vaghela (Arizona State University), Yezhou Yang (Arizona State University), and Bharatesh Chakravarthi (Arizona State University)</i>	

Event-based Tracking and Imaging of Randomly Moving Objects in Dense Dynamical Scattering Media .....	5114
<i>Ning Zhang (Brown University), Timothy Shea (Intel Lab), and Arto Nurmikko (Brown University)</i>	
EV-LayerSegNet: Self-supervised Motion Segmentation using Event Cameras .....	5126
<i>Youssef Farah (Khalifa University), Federico Paredes-Valles (Delft University of Technology), Guido De Croon (Delft University of Technology), Muhammad Ahmed Humais (Khalifa University), Hussain Sajwani (Khalifa University), and Yahya Zweiri (Khalifa University)</i>	
BRAT: Bidirectional Relative Positional Attention Transformer for Event-based Eye tracking.....	5136
<i>Yuliang Wu (University of Science and Technology of China), Han Han (University of Science and Technology of China), Jinze Chen (University of Science and Technology of China), Wei Zhai (University of Science and Technology of China), Yang Cao (University of Science and Technology of China), and Zheng-jun Zha (University of Science and Technology of China)</i>	
Exploring Temporal Dynamics in Event-based Eye Tracker .....	5145
<i>Hongxiang Huang (The Hong Kong University of Science and Technology (Guangzhou)), Xiaopeng Lin (The Hong Kong University of Science and Technology (Guangzhou)), Hongwei Ren (The Hong Kong University of Science and Technology (Guangzhou)), Yue Zhou (The Hong Kong University of Science and Technology (Guangzhou)), and Bojun Cheng (The Hong Kong University of Science and Technology (Guangzhou))</i>	
Dual-Path Enhancements in Event-Based Eye Tracking: Augmented Robustness and Adaptive Temporal Modeling .....	5155
<i>Hoang M. Truong (University of Science, VNU-HCM, Ho Chi Minh City, Vietnam), Vinh-Thuan Ly (University of Science, VNU-HCM, Ho Chi Minh City, Vietnam), Huy G. Tran (University of Science, VNU-HCM, Ho Chi Minh City, Vietnam), Thuan-Phat Nguyen (University of Science, VNU-HCM, Ho Chi Minh City, Vietnam), and Tram T. Doan (University of Science, VNU-HCM, Ho Chi Minh City, Vietnam)</i>	

Event-based eye tracking. Event-based Vision Workshop 2025 .....	5164
<i>Qinyu Chen (Leiden University), Chang Gao (TU Delft ), Min Liu (DVSense), Daniele Perrone (Prophesee), Yan Ru Pei (Nvidia), Zuowen Wang (Institute of Neuroinformatics, UZH/ETH Zurich ), Zhuo Zou (Fudan University ), Shihang Tan (Fudan University), Tao Han (Delft University of Technology), Guorui Lu (Leiden University), Xu Zhen (Leiden University ), Junyuan Ding (DVSense), Ziteng Wang (DVSense), Zongwei Wu (University of Würzburg), Han Han (University of Science and Technology of China), Yuliang Wu (University of Science and Technology of China), Jinze Chen (University of Science and Technology of China), Wei Zhai (University of Science and Technology of China), Yang Cao (University of Science and Technology of China), Zheng-jun Zha (University of Science and Technology of China), Nuwan Bandara (Singapore Management University), Thivya Kandappu (Singapore Management University), Archan Misra (Singapore Management University), Xiaopeng Lin (The Hong Kong University of Science and Technology (Guangzhou)), Hongxiang Huang (The Hong Kong University of Science and Technology (Guangzhou)), Hongwei Ren (The Hong Kong University of Science and Technology (Guangzhou)), Bojun Cheng (The Hong Kong University of Science and Technology (Guangzhou)), Hoang M. Truong (University of Science, VNU-HCM &amp; Vietnam National University, Ho Chi Minh City, Vietnam), Vinh-Thuan Ly (University of Science, VNU-HCM &amp; Vietnam National University, Ho Chi Minh City, Vietnam), Huy G. Tran (University of Science, VNU-HCM &amp; Vietnam National University, Ho Chi Minh City, Vietnam), Thuan-Phat Nguyen (University of Science, VNU-HCM &amp; Vietnam National University, Ho Chi Minh City, Vietnam), and Tram T. Doan (University of Science, VNU-HCM &amp; Vietnam National University, Ho Chi Minh City, Vietnam)</i>	

## Exploring the Next Generation of Data

CACP: Context-Aware Copy-Paste to Enrich Image Content for Data Augmentation .....	5177
<i>Qiushi Guo (csrd), Shaoxiang Wang (DFKI), Chunpeng Chang (DFKI), and Jason Rambach (DFKI)</i>	
Prototype-Guided Diffusion for Digital Pathology: Achieving Foundation Model Performance with Minimal Clinical Data .....	5187
<i>Ekaterina Redekop (University of California, Los Angeles, USA), Mara Pleasure (University of California, Los Angeles, USA), Vedrana Ivezic (University of California, Los Angeles, USA), Zichen Wang (University of California, Los Angeles, USA), Kimberly Flores (University of California, Los Angeles, USA), Anthony Sisk (University of California, Los Angeles, USA), William Speier (University of California, Los Angeles, USA), and Corey Arnold (University of California, Los Angeles, USA)</i>	

Why We Feel: Breaking Boundaries in Emotional Reasoning with Multimodal Large Language Models .....	5196
<i>Yuxiang Lin (Georgia Institute of Technology), Jingdong Sun (Carnegie Mellon University), Zhi-Qi Cheng (University of Washington), Jue Wang (Shenzhen Institute of Advanced Technology), Haomin Liang (Shenzhen Technology University), Zebang Cheng (Shenzhen Technology University), Yifei Dong (University of Washington), Jun-Yan He (Alibaba), Xiaojiang Peng (Shenzhen Technology University), and Xian-Sheng Hua (Tongji University)</i>	

## Open-World 3D Scene Understanding with Foundation Models

Enforcing View-Consistency in Class-Agnostic 3D Segmentation Fields .....	5207
<i>Corentin Dumery (EPFL), Aoxiang Fan (EPFL), Ren Li (EPF Lausanne), Nicolas Talabot (EPFL), and Pascal Fua (EPFL)</i>	
DCSEG: Decoupled 3D Open-Set Segmentation using Gaussian Splatting .....	5217
<i>Luis Wiedmann (Technical University of Munich), Luca Wiehe (Technical University of Munich), and David Rozenberszki (Technical University of Munich)</i>	
Segment Any Primitive: Zero-Shot 3D Primitive Segmentation from Point Cloud .....	5227
<i>Yushan Bai (Institute of Automation, Chinese Academy of Sciences, China), Shaohu Wang (Institute of Automation, Chinese Academy of Sciences, China), Rongtao Xu (Institute of Automation, Chinese Academy of Sciences, China), Yuchuang Tong (Institute of Automation, Chinese Academy of Sciences, China), Chaoran Xu (Institute of Automation, Chinese Academy of Sciences, China), and Zhengtao Zhang (Institute of Automation, Chinese Academy of Sciences, China)</i>	
ForesightNav: Learning Scene Imagination for Efficient Exploration .....	5236
<i>Hardik Shah (ETH Zurich), Jiaxu Xing (University of Zurich), Nico Messikommer (University of Zurich), Boyang Sun (ETHZ - ETH Zurich), Marc Pollefeys (ETH Zurich), and Davide Scaramuzza (University of Zurich)</i>	
OpenSplat3D: Open-Vocabulary 3D Instance Segmentation using Gaussian Splatting .....	5246
<i>Jens Piekenbrinck (RWTH Aachen University), Christian Schmidt (RWTH Aachen University), Alexander Hermans (RWTH Aachen University), Narunas Vaskevicius (Robert Bosch GmbH), Timm Linder (Robert Bosch GmbH), and Bastian Leibe (RWTH Aachen University)</i>	
HAECcity: Open-Vocabulary Scene Understanding of City-Scale Point Clouds with Superpoint Graph Clustering .....	5256
<i>Alexander M. Rusnak (EPF Lausanne) and Frederic Kaplan (EPF Lausanne)</i>	

## ReGenAI: Second Workshop on Responsible Generative AI

ECO-AI – Energy-Conscious Optimization for AI Training .....	5266
<i>János Horváth (Purdue University)</i>	

Dynamic watermarks in images generated by diffusion models .....	5271
<i>Yunzhuo Chen (The University of Western Australia), Jordan Vice (The University of Western Australia), Naveed Akhtar (The University of Melbourne ), Nur Haldar (The University of Western Australia), and Ajmal Mian (The University of Western Australia)</i>	
Understanding and Mitigating Toxicity in Image-Text Pretraining Datasets: A Case Study on LLaVA .....	5278
<i>Nahid Alam (Cisco Meraki), Karthik Reddy Kanjula (Cohere for AI community), Surya Guthikonda (ndiana University Bloomington), and Shayakh Islam (BUET)</i>	

## The 7th Workshop on Precognition: Seeing through the Future

SRVP: Strong Recollection Video Prediction Model Using Attention-Based Spatiotemporal Correlation Fusion .....	5283
<i>Yuseon Kim (Korea Institute of Science and Technology Information (KISTI), South Korea / University of Science and Technology (UST), South Korea) and Kyongseok Park (Korea Institute of Science and Technology Information (KISTI), South Korea / University of Science and Technology (UST), South Korea)</i>	
PRIMEDrive-CoT: A Precognitive Chain-of-Thought Framework for Uncertainty-Aware Object Interaction in Driving Scene Scenario .....	5293
<i>Sriram Mandalika (SRM Institute of Science and Technology, Kattankulathur, India), Lalitha V (SRM Institute of Science and Technology, Kattankulathur, India), and Athira Nambiar (SRM Institute of Science and Technology, Kattankulathur, India)</i>	
BIMA: Bijective Maximum Likelihood Learning Approach to Hallucination Prediction and Mitigation in Large Vision-Language Models .....	5302
<i>Huu-Thien Tran (University of Arkansas, US), Thanh-Dat Truong (University of Arkansas, US), and Khoa Luu (University of Arkansas, US)</i>	
IGL-DT: Iterative Global-Local Feature Learning with Dual-Teacher Semantic Segmentation Framework under Limited Annotation Scheme .....	5312
<i>Quan Tran (National Chung Cheng University, Taiwan), Hoang-Thien Nguyen (Posts and Telecommunications Institute of Technology, Ho Chi Minh, Vietnam), Thanh-Huy Nguyen (Université de Bourgogne Europe, France), Gia-Van To (Institut de Science Financière et d'Assurances, France), Tien-Huy Nguyen (University of Information Technology, Vietnam), and Quan Nguyen (Posts and Telecommunications Institute of Technology, Ha Noi, Vietnam)</i>	

HDC: Hierarchical Distillation for Multi-level Noisy Consistency in Semi-Supervised Fetal Ultrasound Segmentation .....	5322
<i>Tran Quoc Khanh Le (University of Information Technology, Ho Chi Minh City, Vietnam), Nguye Lan Vi Vu (Ho Chi Minh University of Technology, Ho Chi Minh city, Vietnam), Ha-Hieu Pham (University of Science, Ho Chi Minh city, Vietnam), Xuan-Loc Huynh (Boston University, MA, USA), Tien-Huy Nguyen (University of Information Technology, Ho Chi Minh City city, Vietnam), Minh Huu Nhat Le (Taipei Medical University, Taipei, Taiwan), Quan Nguyen (Posts and Telecommunications Institute of Technology, Hanoi), and Hien D. Nguyen (University of Information Technology, Ho Chi Minh city, Viet Nam)</i>	

## **The 6th Annual International Workshop And Prize Challenge on Agriculture-Vision: Challenges and Opportunities for Computer Vision in Agriculture**

Few-Shot Adaptation of Grounding DINO for Agricultural Domain .....	5332
<i>Rajhans Singh (Corteva Agriscience), Rafael Bidese (Corteva Agriscience), Kshitiz Dhakal (Corteva Agriscience), and Sudhir Sornapudi (Corteva Agriscience)</i>	
Vit4V: a Video Classification Method for the Detection of Varroa Destructor from Honeybees.....	5343
<i>Luca Giovannesi (Sapienza University of Rome), Paolo Russo (Sapienza University of Rome), and Roberto Beraldi (Sapienza University of Rome)</i>	
Multiple Instance Learning for Visual Grain Quality Analysis Without Instance-level Annotation .....	5352
<i>Bradley Ezard (Curtin University, Australia), Ling Li (Curtin University, Australia), and Senjian An (Curtin University, Australia)</i>	
Wheat3DGS: In-field 3D Reconstruction, Instance Segmentation and Phenotyping of Wheat Heads with Gaussian Splatting .....	5360
<i>Daiwei Zhang (ETH Zurich), Joaquin Gajardo (ETH Zurich), Tomislav Medic (ETH Zurich), Isinsu Katircioglu (ETH Zurich), Mike Boss (ETH Zurich), Norbert Kirchgessner (ETH Zurich), Achim Walter (ETH Zurich), and Lukas Roth (ETH Zurich)</i>	
PhytoSynth: Leveraging Multi-modal Generative Model for Crop Disease Data Generation with Novel Benchmarking and Prompt Engineering Approach .....	5371
<i>Nitin Rai (University of Florida), Arnold Schumann (University of Florida), and Nathan Boyd (University of Florida)</i>	
A Dataset for Semantic and Instance Segmentation of Modern Fruit Orchards .....	5381
<i>Tieqiao Wang (Oregon State University), Abhinav Jain (Oregon State University), Liqiang He (Oregon State University), Cindy Grimm (Oregon State University), and Sinisa Todorovic (Oregon State University)</i>	
AGILE: A Diffusion-Based Attention-Guided Image and Label Translation for Efficient Cross-Domain Plant Trait Identification .....	5392
<i>Earl Ranario (UC Davis), Lars Lundqvist (UC Davis), Heesup Yun (UC Davis), Brian N. Bailey (UC Davis), and J. Mason Earles (UC Davis)</i>	

Maize ear sensing for on-farm yield predictions .....	5402
<i>Pedro Cisdeli (Purdue University), Gustavo Nocera Santiago (Purdue University), German Mandrini (Purdue University), and Ignacio Ciampitti (Purdue University)</i>	
A Semi-Self-Supervised Approach for Dense-Pattern Video Object Segmentation .....	5412
<i>Keyhan Najafian (University of Saskatchewan, Canada), Farhad Maleki (University of Calgary, Canada), Lingling Jin (University of Saskatchewan, Canada), and Ian Stavness (University of Saskatchewan, Canada)</i>	
MaskAdapt: Unsupervised Geometry-Aware Domain Adaptation Using Multimodal Contextual Learning and RGB-Depth Masking .....	5422
<i>Numair Nadeem (University of Regina, Canada), Muhammad Hamza Asad (University Canada West, Canada), Saeed Anwar (The University of Western Australia, Australia), and Abdul Bais (University of Regina, Canada)</i>	
SemanticSugarBeets: A Multi-Task Framework and Dataset for Inspecting Harvest and Storage Characteristics of Sugar Beets .....	5433
<i>Gerardus Croonen (AIT Austrian Institute of Technology), Andreas Trondl (AIT Austrian Institute of Technology), Julia Simon (AIT Austrian Institute of Technology), and Daniel Steininger (AIT Austrian Institute of Technology)</i>	
AppleGrowthVision: A large-scale stereo dataset for phenological analysis, fruit detection, and 3D reconstruction in apple orchards .....	5443
<i>Laura-Sophia von Hirschhausen (Fraunhofer HHI), Jannes S. Magnusson (Fraunhofer HHI), Mykyta Kovalenko (Fraunhofer HHI), Fredrik Boye (Fraunhofer IVI), Tanay Rawat (Fraunhofer IVI), Peter Eisert (Fraunhofer HHI), Anna Hilsmann (Fraunhofer HHI), Sebastian Pretzsch (Fraunhofer IVI), and Sebastian Bosse (Fraunhofer HHI)</i>	
Effectiveness of Training with Procedurally Generated Synthetic Images of Crop Plants .....	5451
<i>Nazifa Khan (University of Saskatchewan), Mikolaj Cieslak (University of Calgary), Mark Eramian (University of Saskatchewan), and Ian McQuillan (University of Saskatchewan)</i>	
Weakly Supervised Panoptic Segmentation for Defect-Based Grading of Fresh Produce .....	5462
<i>Manuel Knott (Empa, Swiss Federal Laboratories for Materials Science and Technology, St. Gallen, Switzerland), Divinefavour Odion (Constructor University, Bremen, Germany), Sameer Sontakke (Innoterra BioScience Private Limited, Mumbai, India), Anup Karwa (Innoterra BioScience Private Limited, Mumbai, India), and Thijs Defraeye (Empa, Swiss Federal Laboratories for Materials Science and Technology, St. Gallen, Switzerland)</i>	
SC-NeRF: NeRF-based Point Cloud Reconstruction using a Stationary Camera for Agricultural Applications .....	5472
<i>Kibon Ku (Iowa State University, USA), Talukder Z. Jubery (Iowa State University, USA), Elijah Rodriguez (Iowa State University, USA), Aditya Balu (Iowa State University, USA), Soumik Sarkar (Iowa State University, USA), Adarsh Krishnamurthy (Iowa State University, USA), and Baskar Ganapathysubramanian (Iowa State University, USA)</i>	

Maximizing aerial detection of organic objects in non-exhaustively searchable survey area .....	5482
<i>Amir Niaraki (Iowa State University), Jansel Herrera-Gerena (Iowa State University), Jeremy Roghair (Amazon Web Services (AWS)), and Ali Jannesari (Iowa State University)</i>	
California Crop Yield Benchmark: Combining Satellite Image, Climate, Evapotranspiration, and Soil Data Layers for County-Level Yield Forecasting of Over 70 Crops .....	5491
<i>Hamid Kamangir (University of California Davis), Mona Hajiesmaeli (University of California Davis), and J. Mason Earles (University of California Davis)</i>	
iNatAg: Multi-Class Classification Models Enabled by a Large-Scale Benchmark Dataset with 4.7M Images of 2,959 Crop and Weed Species .....	5501
<i>Naitik Jain (University of California, Davis), Amogh Joshi (University of California, Davis, Princeton University, AI Institute for Food Systems), and Mason Earles (University of California, Davis, AI Institute for Food Systems)</i>	
Agri-FM+: A Self-Supervised Foundation Model for Agricultural Vision .....	5511
<i>Md Jaber Al Nahian (University of Calgary, Canada), Tapotosh Ghosh (University of Calgary, Canada), Farnaz Sheikhi (University of Calgary, Canada), and Farhad Maleki (University of Calgary, Canada)</i>	
SoyStageNet: Balancing Accuracy and Efficiency for Real-Time Soybean Growth Stage Detection .....	5524
<i>Abdellah Lakhssassi (Southern Illinois University Carbondale, USA), Toqi Tahamid Sarker (Southern Illinois University Carbondale, USA), Khaled Ahmed (Southern Illinois University Carbondale, USA), Naoufal Lakhssassi (Hampton University, USA), and Khalid Meksem (Southern Illinois University Carbondale, USA)</i>	
An Efficient and Scalable Framework for Lightweight Crop Disease Recognition in Low-Resource Settings .....	5534
<i>Tushar Shinde (IIT Madras Zanzibar)</i>	

## 8th Workshop and Competition on Affective & Behavior Analysis in-the-wild

Learning Pose-aware Representations in Vision Transformers for Understanding Activities of Daily Living .....	5542
<i>Dominick Reilly (University of North Carolina at Charlotte), Srijita Das (University of Michigan-Dearborn), and Srijan Das (University of North Carolina at Charlotte)</i>	
Decoupling Identity Confounders for Enhanced Facial Expression Recognition: An Information-Theoretic Approach .....	5552
<i>Mohd Aquib (IIT Kanpur, India), Nishchal K. Verma (IIT Kanpur, India), and M. Jaleel Akhtar (IIT Kanpur, India)</i>	
Leveraging Synthetic Adult Datasets for Unsupervised Infant Pose Estimation .....	5562
<i>Sarosij Bose (University of California), Hannah Dela Cruz (University of California), Arindam Dutta (University of California), Elena Kokkoni (University of California), Konstantinos Karydis (University of California), and Amit Kumar Roy Chowdhury (University of California)</i>	

Advancements in Affective and Behavior Analysis: The 8th ABAW Workshop and Competition ..	5572
<i>Dimitrios Kollias (Queen Mary University of London), Panagiotis Tzirakis (Hume AI), Alan Cowen (Hume AI), Stefanos Zafeiriou (Imperial College London), Irene Kotsia (Cogitat), Eric Granger (ETS Montreal), Marco Pedersoli (ETS Montreal), Simon Bacon (Concordia University), Alice Baird (Hume AI), Chris Gagne (Hume AI), Chunchang Shao (Queen Mary University of London), Guanyu Hu (Queen Mary University of London; Xi'an Jiaotong University), Soufiane Belharbi (ETS Montreal), and Muhammad Haseeb Aslam (ETS Montreal)</i>	
Face Reconstruction from Face Embeddings using Adapter to a Face Foundation Model .....	5584
<i>Hatef Otroschi Shahreza (Idiap Research Institute), Anjith George (Idiap Research Institute), and Sébastien Marcel (Idiap Research Institute)</i>	
DAF: Distillation, Augmentation and Filtering based Framework for Efficient Smartphone Human Activity Recognition .....	5594
<i>Ujjal Kr Dutta (Dolby Laboratories) and Guan-Ming Su (Dolby Laboratories)</i>	
Is Multi-Person Gait Recognition Feasible under Mutual Occlusion? A Human Model Regression-based Approach .....	5603
<i>Ziruo Li (The University of Osaka, Japan), Chi Xu (The University of Osaka, Japan), Xiang Li (The University of Osaka, Japan), Shuqiong Wu (The University of Osaka, Japan), and Yasushi Yagi (The University of Osaka, Japan)</i>	
Advancing Facial Age Progression for Occluded Faces .....	5614
<i>Ankit Birla (IISER Bhopal) and Akshay Agarwal (IISER Bhopal)</i>	
VRU-CIPI: Crossing Intention Prediction at Intersections for Improving Vulnerable Road Users Safety .....	5623
<i>Ahmed S. Abdelrahman (University of Central Florida), Mohamed Abdel-Aty (University of Central Florida), and Quoc Dai Tran (University of Central Florida)</i>	
EmoVLM-KD: Fusing Distilled Expertise with Vision-Language Models for Visual Emotion Analysis .....	5633
<i>SangEun Lee (Electronics and Telecommunications Research Institute), Yubeen Lee (Sungkyunkwan University), and Eunil Park (Sungkyunkwan University)</i>	
V-NAW: Video-based Noise-aware Adaptive Weighting for Facial Expression Recognition .....	5643
<i>JunGyu Lee (Korea Institute of Science and Technology, University of Science and Technology(UST)), Kunyoung Lee (Korea Institute of Science and Technology), Haesol Park (Korea Institute of Science and Technology), Ig-Jae Kim (Korea Institute of Science and Technology, University of Science and Technology(UST)), and Gi Pyo Nam (Korea Institute of Science and Technology, University of Science and Technology(UST))</i>	

Mamba-VA: A Mamba-based Approach for Continuous Emotion Recognition in Valence-Arousal Space .....	5651
<i>Yuheng Liang (Nanjing University of Posts and Telecommunications), Zheyu Wang (Nanjing University of Posts and Telecommunications), Feng Liu (Nanjing University of Posts and Telecommunications), Mingzhou Liu (Nanjing University of Posts and Telecommunications), and Yu Yao (Nanjing University of Posts and Telecommunications)</i>	
Datasets for Valence and Arousal Inference: A Survey .....	5657
<i>Helen Schneider (Karlsruhe Institute of Technology), Svetlana Pavlotska (FZI Research Center for Information Technology), Helen Gremmelmaier (FZI Research Center for Information Technology), and Marius Zöllner (FZI Research Center for Information Technology)</i>	
Enhancing Facial Expression Recognition with LSTM through Dual-Direction Attention Mixed Feature Networks and CLIP .....	5665
<i>Josep Cabacas-Maso (Universitat Oberta de Catalunya), Elena Ortega-Beltrán (Universitat Oberta de Catalunya), Ismael Benito-Altamirano (Universitat Oberta de Catalunya), and Carles Ventura (Universitat Oberta de Catalunya)</i>	
Read My Ears! Horse Ear Movement Detection for Equine Affective State Assessment .....	5672
<i>Joao Alves (Aalborg University, Denmark), Pia Haubro Andersen (University of Copenhagen, Denmark), and Rikke Gade (Aalborg University, Denmark)</i>	
Multimodal Emotion Prediction in Interpersonal Videos Integrating Facial and Speech Cues .....	5681
<i>Hajer Guerdelli (University of Florence), Claudio Ferrari (University of Siena), Stefano Berretti (University of Florence), and Alberto Del Bimbo (University of Florence)</i>	
MMDrive: Multi-modal Remote Physiological Signal Measurement Dataset for Driver Status Monitoring .....	5691
<i>Jiho Choi (Jeonbuk National University, Republic of Korea) and Sang Jun Lee (Jeonbuk National University, Republic of Korea)</i>	
Interactive Multimodal Framework with Temporal Modeling for Emotion Recognition .....	5699
<i>Jun Yu (University of Science and Technology of China), Yongqi Wang (University of Science and Technology of China), Lei Wang (University of Science and Technology of China), Yang Zheng (University of Science and Technology of China), and Shengfan Xu (University of Science and Technology of China)</i>	
Cross-Modal Facial Expression Recognition with Global Channel-Spatial Attention: Modal Enhancement and Proportional Criterion Fusion .....	5707
<i>Jun Yu (University of Science and Technology of China), Yang Zheng (University of Science and Technology of China), Lei Wang (University of Science and Technology of China), Yongqi Wang (University of Science and Technology of China), and Shengfan Xu (Macau University of Science and Technology)</i>	
Semantic Matters: Multimodal Features for Affective Analysis .....	5715
<i>Tobias Hallmen (University of Augsburg), Robin-Nico Kampa (University of the Bundeswehr Munich), Fabian Deuser (University of the Bundeswehr Munich), Norbert Oswald (University of the Bundeswehr Munich), and Elisabeth André (University of Augsburg)</i>	

Towards Robust Multimodal AU Detection: STN-Enhanced Visual Encoding and Audio-Visual Spatial-Temporal Alignment .....	5725
<i>Jun Yu (University of Science and Technology of China), Yunxiang Zhang (University of Science and Technology of China), Fengzhao Sun (University of Science and Technology of China), Leilei Wang (University of Science and Technology of China), Renjie Lu (University of Science and Technology of China), Lingsi Zhu (University of Science and Technology of China), Xilong Lu (University of Science and Technology of China), Yang Zheng (University of Science and Technology of China), and Yongqi Wang (University of Science and Technology of China)</i>	
Dual-Stage Cross-Modal Network with Dynamic Feature Fusion for Emotional Mimicry Intensity Estimation .....	5733
<i>Jun Yu (University of Science and Technology of China), Lingsi Zhu (University of Science and Technology of China), Yanjun Chi (University of Science and Technology of China), Yunxiang Zhang (University of Science and Technology of China), Yang Zhen (University of Science and Technology of China), Yongqi Wang (University of Science and Technology of China), and Xilong Lu (University of Science and Technology of China)</i>	
United we stand, Divided we fall: Handling Weak Complementarity for Audio-Visual Emotion Recognition in Valence-Arousal Space .....	5741
<i>Gnana Praveen Rajasekhar (Computer Research Institute of Montreal), Jahangir Alam (Computer Research Institute of Montreal), and Eric Charton (Computer Research Institute of Montreal)</i>	
Shopformer: Transformer-Based Framework for Detecting Shoplifting via Human Pose .....	5752
<i>Narges Rashvand (University of North Carolina at Charlotte), Ghazal Alinezhad Noghre (University of North Carolina at Charlotte), Armin Danesh Pazho (University of North Carolina at Charlotte), Babak Rahimi Ardabili (University of North Carolina at Charlotte), and Hamed Tabkhi (University of North Carolina at Charlotte)</i>	
Neighbor-Based Feature and Index Enhancement for Person Re-Identification .....	5762
<i>Chao Yuan (School of Computer Science and Engineering, Beihang University), Tianyi Zhang (School of Artificial Intelligence, Beihang University), and Guanglin Niu (School of Artificial Intelligence, Beihang University)</i>	
Robust Stage-Wise LVLM Adaptation: Multi-Phase Prompt Lora Fine-tuning for Compound Expression Recognition .....	5770
<i>Xilong Lu (University of Science and Technology of China, China), Jun Yu (University of Science and Technology of China, China), Yunxiang Zhang (University of Science and Technology of China, China), Lingsi Zhu (University of Science and Technology of China, China), Yang Zheng (University of Science and Technology of China, China), Yongqi Wang (University of Science and Technology of China, China), and Qiang Ling (University of Science and Technology of China, China)</i>	
Leveraging Lightweight Facial Models and Textual Modality in Audio-visual Emotional Understanding in-the-Wild .....	5778
<i>Andrey Savchenko (Sber AI Lab) and Lyudmila Savchenko (HSE University)</i>	

MAVEN: Multi-modal Attention for Valence-Arousal Emotion Network .....	5789
<i>Vrushank Ahire (Indian Institute of Technology Ropar), Kunal Shah (Indian Institute of Technology Ropar), Mudassir Khan (Indian Institute of Technology Ropar), Nikhil Pakhale (Indian Institute of Technology Ropar), Lownish Sookha (Indian Institute of Technology Ropar), Mudassir Ganaie (Indian Institute of Technology Ropar), and Abhinav Dhall (Monash University)</i>	
Selective Test-time Domain Adaptation Using Fisher Information for Robust Facial Expression Recognition In-the-wild .....	5800
<i>Mohammadmahdi Honarmand (Stanford University), Onur Cezmi Mutlu (Stanford University), Parnian Azizian (Stanford University), Saimourya Surabhi (Stanford University), and Dennis Wall (Stanford University)</i>	
Safe-Construct: Redefining Construction Safety Violation Recognition as 3D Multi-View Engagement Task .....	5811
<i>Aviral Chharia (Carnegie Mellon University), Tianyu Ren (University of Illinois Urbana-Champaign), Tomotake Furuhashi (Carnegie Mellon University), and Kenji Shimada (Carnegie Mellon University)</i>	

## 11th IEEE International Workshop on Computer Vision in Sports

TT3D: Table Tennis 3D Reconstruction .....	5821
<i>Thomas Gossard (Cognitive systems - University Tuebingen), Andreas Ziegler (Cognitive systems - University Tuebingen), and Andreas Zell (Cognitive systems - University Tuebingen)</i>	
Towards Ball Spin and Trajectory Analysis in Table Tennis Broadcast Videos via Physically Grounded Synthetic-to-Real Transfer .....	5832
<i>Daniel Kienzle (University of Augsburg), Robin Schön (University of Augsburg), Rainer Lienhart (University of Augsburg), and Shin'Ichi Satoh (National Institute of Informatics)</i>	
Efficient 2D to Full 3D Human Pose Uplifting including Joint Rotations .....	5842
<i>Katja Ludwig (University of Augsburg), Yuliia Oksymets (University of Augsburg), Robin Schön (University of Augsburg), Daniel Kienzle (University of Augsburg), and Rainer Lienhart (University of Augsburg)</i>	
VNL-STES: A Benchmark Dataset and Model for Spatiotemporal Event Spotting in Volleyball Analytics .....	5852
<i>Hoang Nguyen (AI-Robotics, KIST School, University of Science and Technology, Seoul 02792, Republic of Korea), Ankhzaya Jamsrandorj (Korea Institute of Science and Technology, Seoul 02792, South Korea), Vanyi Chao (AI-Robotics, KIST School, University of Science and Technology, Seoul 02792, Republic of Korea), Yin May Oo (AI-Robotics, KIST School, University of Science and Technology, Seoul 02792, Republic of Korea), Muhammad Amrulloh Robbani (AI-Robotics, KIST School, University of Science and Technology, Seoul 02792, Republic of Korea), Kyung-Ryoul Mun (Korea Institute of Science and Technology, Seoul 02792, South Korea), and Jinwook Kim (Korea Institute of Science and Technology, Seoul 02792, South Korea)</i>	

Leveraging Anthropometric Measurements to Improve Human Mesh Estimation and Ensure Consistent Body Shapes .....	5862
<i>Katja Ludwig (University of Augsburg), Julian Lorenz (University of Augsburg), Daniel Kienzle (University of Augsburg), Tuan Bui (University of Augsburg), and Rainer Lienhart (University of Augsburg)</i>	
An End-to-End Pipeline for Virtual Banner Replacement in Football Broadcasts .....	5872
<i>Victor Gaspar (University of Liège), Anthony Cioppa (University of Liège), Jan Held (University of Liège), Silvio Giancola (King Abdullah University of Science and Technology), Marc Braham (University of Liège), Adrien Delière (University of Liège), Bernard Ghanem (King Abdullah University of Science and Technology), and Marc Van Droogenbroeck (University of Liège)</i>	
FieldMOT: A Field-Registered Multi-Object Tracking for Sports Videos .....	5884
<i>Hong-Qi Chen (National Yang Ming Chiao Tung University, Taiwan), Chao-Chi Liao (National Yang Ming Chiao Tung University, Taiwan), Yuan-Heng Sun (National Yang Ming Chiao Tung University, Taiwan), Cheng-Kuan Lin (National Yang Ming Chiao Tung University, Taiwan), and Yu-Chee Tseng (National Yang Ming Chiao Tung University, Taiwan)</i>	
From Beats to Scores: A Multi-Modal Framework for Comprehensive Figure Skating Assessment .....	5895
<i>Fengshun Wang (Capital University of Physical Education and Sports), Qiurui Wang (Capital University of Physical Education and Sports), and Dan Chen (Capital University of Physical Education and Sports)</i>	
Two Views Are Better than One: Monocular 3D Pose Estimation with Multiview Consistency .....	5905
<i>Christian Keilstrup Ingwersen (TrackMan), Rasmus Tirsgaard (Technical University of Denmark), Rasmus Nylander (TrackMan), Janus Nortoft Jensen (3Shape), Anders Bjorholm Dahl (Technical University of Denmark), and Morten Rieger Hannemose (Technical University of Denmark)</i>	
Virtual Pose Coach: A Motion-Retargeting Approach for Pose Training .....	5916
<i>Tzu-Chun Chiu (National Yang Ming Chiao Tung University), Ming-Han Lee (National Yang Ming Chiao Tung University), Kun-Ru Wu (National Yang Ming Chiao Tung University), Yu-Shuen Wang (National Yang Ming Chiao Tung University), and Yu-Chee Tseng (National Yang Ming Chiao Tung University)</i>	
Jump-Aware: Player Position Rectification and Identification in Dynamic Sports Using Jump Event Spotting .....	5925
<i>Yin May Oo (Korea Institute of Science and Technology), Ankhzaya Jamsrandorj (Korea Institute of Science and Technology), Vanyi Chao (Korea Institute of Science and Technology), Hoang Quoc Nguyen (Korea Institute of Science and Technology), Yewon Hwang (Korea Institute of Science and Technology), Kyung-Ryoul Mun (Korea Institute of Science and Technology), and Jinwook Kim (Korea Institute of Science and Technology)</i>	

AthletePose3D: A Benchmark Dataset for 3D Human Pose Estimation and Kinematic Validation in Athletic Movements .....	5935
<i>Calvin Yeung (Nagoya University, Japan), Tomohiro Suzuki (Nagoya University, Japan), Ryota Tanaka (Nagoya University, Japan), Zhuoer Yin (Nagoya University, Japan), and Keisuke Fujii (Nagoya University, Japan)</i>	
Towards fine-grained spatial control for soccer game image generation .....	5947
<i>Amadou S. Sangare (Université Paris-Saclay, CEA, List, F-91120, Palaiseau, France), Adrien Maglo (Université Paris-Saclay, CEA, List, F-91120, Palaiseau, France), Baptiste Engel (Université Paris-Saclay, CEA, List, F-91120, Palaiseau, France), and Mohamed Chaouch (Université Paris-Saclay, CEA, List, F-91120, Palaiseau, France)</i>	
Skor-xG: SKEleton-ORiented Expected Goal Estimation in Soccer .....	5957
<i>Yizhou Xu (KTH Royal Institute of Technology, Sweden), Lars Bretzner (Tracab, Sweden), Tiesheng Wang (Tracab, Sweden), and Atsuto Maki (KTH Royal Institute of Technology, Sweden)</i>	
SoccerNet-v3D: Leveraging Sports Broadcast Replays for 3D Scene Understanding .....	5968
<i>Marc Gutiérrez-Pérez (Institut de Robòtica i Informàtica Industrial, CSIC-UPC, Spain) and Antonio Agudo (Institut de Robòtica i Informàtica Industrial, CSIC-UPC, Spain)</i>	
CaddieSet: A Golf Swing Dataset with Human Joint Features and Ball Information .....	5978
<i>Seunghyeon Jung (Dongguk University, Seoul, Republic of Korea), Seoyoung Hong (Kimcaddie Inc, Seoul, Republic of Korea), Jiwoo Jeong (Dongguk University, Seoul, Republic of Korea), Seungwon Jeong (Dongguk University, Seoul, Republic of Korea), Jaerim Choi (Kimcaddie Inc, Seoul, Republic of Korea), Hoki Kim (Chung-Ang University, Seoul, Republic of Korea), and Woojin Lee (Dongguk University, Seoul, Republic of Korea)</i>	
Sport Field Calibration with NeRF-guided Camera Optimization from a Single Image .....	5987
<i>Liang Fan (Loughborough University), Xiaoqian Liu (Beijing Jiaotong University), and Malcolm Roberts (Irish Football Association)</i>	
GST: Precise 3D Human Body from a Single Image with Gaussian Splatting Transformers .....	5997
<i>Lorenza Prospero (University of Oxford), Abdullah Hamdi (University of Oxford), Joao F. Henriques (University of Oxford), and Christian Rupprecht (University of Oxford)</i>	
FineCausal: A Causal-Based Framework for Interpretable Fine-Grained Action Quality Assessment .....	6008
<i>Ruisheng Han (Department of Computer Science, Durham University), Kanglei Zhou (Beihang University), Amir Atapour-Abarghouei (Department of Computer Science, Durham University), Xiaohui Liang (Beihang University), and Hubert P. H. Shum (Department of Computer Science, Durham University)</i>	

From Broadcast to Minimap: Achieving State-of-the-Art SoccerNet Game State Reconstruction ....	6018
<i>Vladimir Golovkin (Constructor Tech, Bulgaria), Nikolay Nemtsev (Constructor Tech, Bulgaria), Vasyi Shandyba (Constructor Tech, Bulgaria), Oleg Udin (Constructor Tech, Bulgaria), Nikita Kasatkin (Constructor Tech, Bulgaria), Pavel Kononov (Constructor Tech, Bulgaria), Anton Afanasiev (Constructor Tech, Singapore), Sergey Ulasen (Constructor Tech, Singapore), and Andrei Boiarov (Constructor Tech, Bulgaria)</i>	
No Train Yet Gain: Towards Generic Multi-Object Tracking in Sports and Beyond .....	6029
<i>Tomasz Stanczyk (Inria, France), Seongro Yoon (Inria, France), and Francois Bremond (Inria, France)</i>	
Human Mesh Reconstruction of Sports Players with Multiple Dynamic Cameras .....	6039
<i>Yamato Hokari (Keio University, Japan), Ryosuke Hori (Keio University, Japan), and Hideo Saito (Keio University, Japan)</i>	
Ice Hockey Puck Localization Using Contextual Cues .....	6050
<i>Liam Salass (University of Waterloo), Jerrin Bright (University of Waterloo), Amir Nazemi (University of Waterloo), Yuhao Chen (University of Waterloo), John Zelek (University of Waterloo), and David Clausi (University of Waterloo)</i>	
Can Geometry Save Central Views for Sports Field Registration? .....	6060
<i>Floriane Magera (EVS Broadcast Equipment, Belgium), Thomas Hoyoux (EVS Broadcast Equipment, Belgium), Martin Castin (EVS Broadcast Equipment, Belgium), Olivier Barnich (EVS Broadcast Equipment, Belgium), Anthony Cioppa (University of Liege, Belgium), and Marc Van Droogenbroeck (University of Liege, Belgium)</i>	
Action Anticipation from SoccerNet Football Video Broadcasts .....	6070
<i>Mohamad Dalal (Aalborg University, Denmark), Artur Xarles (Universitat de Barcelona, Spain), Anthony Cioppa (University of Liège, Belgium), Silvio Giancola (King Abdullah University of Science and Technology (KAUST), Saudi Arabia), Marc Van Droogenbroeck (University of Liège, Belgium), Bernard Ghanem (King Abdullah University of Science and Technology (KAUST), Saudi Arabia), Albert Clapés (Universitat de Barcelona, Spain), Sergio Escalera (Universitat de Barcelona, Spain), and Thomas B. Moeslund (Aalborg University, Spain)</i>	
Pose-to-Pose: A New Task and Benchmark for Human Pose Transition in Yoga .....	6082
<i>Bhat Dittakavi (Indian Institute of Technology Hyderabad), Swarnim Maheshwari (Indian Institute of Technology Hyderabad), and Vineeth N. Balasubramanian (Indian Institute of Technology Hyderabad)</i>	
Single-Stage Uncertainty-Aware Jersey Number Recognition in Soccer .....	6092
<i>Lukasz Grad (University of Warsaw)</i>	
Domain Adaptation of VLM for Soccer Video Understanding .....	6101
<i>Tiancheng Jiang (Massachusetts Institute of Technology), Henry Wang (Amazon Web Services), Md Sirajus Salekin (Amazon Web Services), Parmida Atighehchian (Amazon Web Services), and Shinan Zhang (Amazon Web Services)</i>	
Where Is The Ball: 3D Ball Trajectory Estimation From 2D Monocular Tracking .....	6112
<i>Puntawat Ponglertnapakorn (VISTEC, Thailand) and Supasorn Suwajanakorn (VISTEC, Thailand)</i>	

Action Valuation in Sports: A Survey .....	6122
<i>Artur Xarles (Universitat de Barcelona, Spain), Sergio Escalera (Universitat de Barcelona, Spain), Thomas B. Moeslund (Aalborg University, Denmark), and Albert Clapés (Universitat de Barcelona, Spain)</i>	
SportMamba: Adaptive Non-Linear Multi-Object Tracking with State Space Models for Team Sports .....	6133
<i>Dheeraj Khanna (University of Waterloo), Jerrin Bright (University of Waterloo), Yuhao Chen (University of Waterloo), and John Zelek (University of Waterloo)</i>	
Polar Coordinate-Based 2D Pose Prior with Neural Distance Field .....	6144
<i>Qi Gan (Télécom Paris, Institut Polytechnique de Paris, France), Sao Mai Nguyen (ENSTA Paris, Institut Polytechnique de Paris, France), Eric Fenaux (Ef-e-science, France), Stephan Cléménçon (Télécom Paris, Institut Polytechnique de Paris, France), and Mounim El Yacoubi (Télécom SudParis, Institut Polytechnique de Paris, France)</i>	
The Way Up: A Dataset for Hold Usage Detection in Sport Climbing .....	6153
<i>Anna Maschek (University of Applied Sciences Upper Austria, Austria) and David C. Schedl (University of Applied Sciences Upper Austria, Austria)</i>	
Towards Unconstrained 2D Pose Estimation of the Human Spine .....	6162
<i>Muhammad Saif Ullah Khan (German Research Center for Artificial Intelligence (DFKI)), Stephan Krauß (German Research Center for Artificial Intelligence (DFKI)), and Didier Stricker (German Research Center for Artificial Intelligence (DFKI))</i>	

## 4th Monocular Depth Estimation Challenge

The Fourth Monocular Depth Estimation Challenge ..... 6172

*Anton Obukhov (Huawei Research Zürich), Matteo Poggi (University of Bologna), Fabio Tosi (University of Bologna), Ripudaman Singh Arora (Blue River Technology), Jaime Spencer (Independent Researcher), Chris Russel (Oxford Internet Institute, UK), Simon Hadfield (University of Surrey, UK), Richard Bowden (University of Surrey, UK), Shuaihang Wang (Hikvision Research Institute), Zhenxin Ma (Hikvision Research Institute), Weijie Chen (Hikvision Research Institute), Baobei Xu (Hikvision Research Institute), Fengyu Sun (Hikvision Research Institute), Di Xie (Hikvision Research Institute), Jiang Zhu (Hikvision Research Institute), Mykola Lavreniuk (Space Research Institute NASU-SSAU, Kyiv, Ukraine), Haining Guan (Megvii), Qun Wu (Megvii), Yupei Zeng (Megvii), Chao Lu (Megvii), Huanran Wang (Megvii), Guangyuan Zhou (Independent Researcher), Haotian Zhang (Independent Researcher), Jianxiong Wang (Independent Researcher), Qiang Rao (Independent Researcher), Chunjie Wang (Independent Researcher), Xiao Liu (Independent Researcher), Zhiqiang Lou (Insta360), Hualie Jiang (Insta360), Yihao Chen (Insta360), Rui Xu (Insta360), Minglang Tan (Insta360), Zihan Qin (Harbin Institute of Technology), Yifan Mao (Harbin Institute of Technology), Jiayang Liu (Harbin Institute of Technology), Jialei Xu (Harbin Institute of Technology), Yifan Yang (Harbin Institute of Technology), Wenbo Zhao (Harbin Institute of Technology), Junjun Jiang (Harbin Institute of Technology), Xianming Liu (Harbin Institute of Technology), Mingshuai Zhao (Beijing University of Posts and Telecommunications), Anlong Ming (Beijing University of Posts and Telecommunications), Wu Chen (Beijing University of Posts and Telecommunications), Feng Xue (University of Trento), Mengying Yu (Beijing University of Posts and Telecommunications), Shida Gao (Beijing University of Posts and Telecommunications), Xiangfeng Wang (Beijing University of Posts and Telecommunications), Gbenga Omotara (University of Missouri), Ramy Farag (University of Missouri), Jacket Demby's (University of Missouri), Seyed Mohamad Ali Tousi (University of Missouri), Guilherme N. DeSouza (University of Missouri), Tuấn Anh Yang (University of Science, VNU-HCM, Vietnam and Vietnam National University, Ho Chi Minh City, Vietnam), Minh Quang Nguyễn (University of Science, VNU-HCM, Vietnam and Vietnam National University, Ho Chi Minh City, Vietnam), Thiên Phúc Trần (University of Science, VNU-HCM and, Vietnam Vietnam National University, Ho Chi Minh City, Vietnam), Albert Luginov (University of Reading), and Muhammad Shahzad (University of Reading)*

## AI for Creative Visual Content Generation, Editing and Understanding

Geometry-Aware Texture Generation for 3D Head Modeling with Artist-driven Control ..... 6186

*Amin Fadaeinejad (Ubisoft LaForge, York University), Abdallah Dib (Ubisoft LaForge), Luiz Gustavo Hafemann (Ubisoft LaForge), Emeline Got (Ubisoft LaForge), Trevor Anderson (Ubisoft LaForge), Amaury Depierre (Ubisoft LaForge), Nikolaus F. Troje (York University), Marcus A. Brubaker (York University, Vector Institute), and Marc-André Carbonneau (Ubisoft LaForge)*

Temporal Consistent Semantic Video Color Transfer from Multiple References ..... 6197

*Aupendu Kar (Dolby Labs) and Guan-Ming Su (Dolby Labs)*

Semantic-Aware Local Image Editing with a Single Mask Operation .....	6206
<i>Dongchao Wen (IEIT SYSTEMS Co., Ltd.), Zijian Chen (IEIT SYSTEMS Co., Ltd.), Weihong Deng (IEIT SYSTEMS Co., Ltd.), Yujiang Tian (IEIT SYSTEMS Co., Ltd.), Hongzhi Shi (IEIT SYSTEMS Co., Ltd.), Yingjie Zhang (IEIT SYSTEMS Co., Ltd.), Xingchen Cui (IEIT SYSTEMS Co., Ltd.), Jian Zhao (IEIT SYSTEMS Co., Ltd.), Lingyan Liang (IEIT SYSTEMS Co., Ltd.), and Mei Wang (IEIT SYSTEMS Co., Ltd.)</i>	
OnlyFlow: Optical Flow based Motion Conditioning for Video Diffusion Models .....	6216
<i>Mathis Koroglu (Obvious Research &amp; ISIR - Sorbonne University, France), Hugo Caselles-Dupré (Obvious Research), Guillaume Jeanneret (ISIR - Sorbonne University, France), and Matthieu Cord (ISIR - Sorbonne University, France)</i>	
Training-free Color-Style Disentanglement for Constrained Text-to-Image Synthesis .....	6227
<i>Aishwarya Agarwal (Adobe Research, Bengaluru India), Srikrishna Karanam (Adobe Research, Bengaluru India), and Balaji Vasan Srinivasan (Adobe Research, Bengaluru India)</i>	
CLIPDraw++: Text-to-Sketch Synthesis with Simple Primitives .....	6237
<i>Nityanand Mathur (Smallest AI), Shyam Marjit (Indian Institute of Science), Abhra Chaudhuri (University of Exeter, University of Surrey), and Anjan Dutta (University Of Surrey)</i>	
Z-SASLM: Zero-Shot Style-Aligned SLI Blending Latent Manipulation .....	6247
<i>Alessio Borgi (Sapienza University of Rome), Luca Maiano (Sapienza University of Rome), and Irene Amerini (Sapienza University of Rome)</i>	
Generative AI for Film Creation: A Survey of Recent Advances .....	6257
<i>Ruihan Zhang (Google), Borou Yu (University of Santa Barbara, Media Arts and Technology), Jiajian Min (MYStudio), Yetong Xin (Harvard University), Zheng Wei (Hong Kong University of Science and Technology), Juncheng Nemo Shi (Reality Hack Inc.), Mingzhen Huang (SUNY Buffalo), Xianghao Kong (Hong Kong University of Science and Technology), Nix Liu Xin (Onceness), Shanshan Jiang (University of Southampton), Praagya Bahuguna (New York University), Mark Chan (New York University), Khushi Hora (New York University / Netflix ), Lijian Yang (Communication University of China), Yongqi Liang (Communication University of China), Runhe Bian (Harvard University), Yunlei Liu (University of Southern California), Isabela Campillo Valencia (Dodge College of Film and Media Arts), Patricia Morales Tredinick (Pratt Institute), Ilia Kozlov (Rubyspot), Sijia Jiang (Harvard University), Peiwen Huang (Harvard University), Na Chen (MIT), Xuanxuan Liu (Harvard University), and Anyi Rao (Hong Kong University of Science and Technology)</i>	
Defurnishing with X-Ray Vision: Joint Removal of Furniture from Panoramas and Mesh .....	6270
<i>Alan Dolhasz (Matterport), Chen Ma (Matterport), Dave Gausebeck (Matterport), Kevin Chen (Matterport), Gregor Miller (Matterport), Lucas Hayne (Matterport), Gunnar Hovden (Matterport), Azwad Sabik (Matterport), Olaf Brandt (Matterport), and Mira Slavcheva (Matterport)</i>	

PartStickers: Generating Parts of Objects for Rapid Prototyping .....	6281
<i>Mo Zhou (University of Colorado Boulder), Josh Myers-Dean (University of Colorado Boulder), and Danna Gurari (University of Colorado Boulder)</i>	
LAPIS: A novel dataset for personalized image aesthetic assessment .....	6292
<i>Anne-Sofie Maerten (dept. Brain and Cognition, KU Leuven), Li-Wei Chen (dept. Brain and Cognition, KU Leuven), Stefanie De Winter (dept. Art History, KU Leuven), Christophe Bossens (dept. Brain and Cognition, KU Leuven), and Johan Wagemans (dept. Brain and Cognition, KU Leuven)</i>	
WaveDIF: Wavelet sub-band based Deepfake Identification in Frequency Domain .....	6302
<i>Anurag Dutta (Department of Computer Science and Engineering, Indian Institute of Technology, Kharagpur, West Bengal, India), Arnab Kumar Das (Department of Information Technology, Indian Institute of Engineering Science and Technology, Shibpur, West Bengal, India), Ruchira Naskar (Department of Information Technology, Indian Institute of Engineering Science and Technology, Shibpur, West Bengal, India), and Rajat Subhra Chakraborty (Department of Computer Science and Engineering, Indian Institute of Technology, Kharagpur, West Bengal, India)</i>	
Progressive Autoregressive Video Diffusion Models .....	6312
<i>Desai Xie (Stony Brook University), Zhan Xu (Adobe Research), Yicong Hong (Adobe Research), Hao Tan (Adobe Research), Difan Liu (Adobe Research), Feng Liu (Adobe Research), Arie Kaufman (Stony Brook University), and Yang Zhou (Adobe Research)</i>	
STAM: Zero-Shot Style Transfer using Diffusion Model via Attention Modulation .....	6323
<i>Masud An Nur Islam Fahim (University of Vaasa), Nazmus Saqib (Jeju National University), and Jani Boutellier (University of Vaasa)</i>	
HopNet: Harmonizing Object Placement Network for Realistic Image Generation via Object Composition .....	6334
<i>Matthew Poska (The Pennsylvania State University), Sharon X Huang (The Pennsylvania State University), and Bin Hwang (Amazon Web Services)</i>	

## 2nd Workshop on Efficient and On-Device Generation (EDGE)

Scaling On-Device GPU Inference for Large Generative Models .....	6345
<i>Jiuqiang Tang (Google LLC), Raman Sorokin (Google LLC), Ekaterina Ignasheva (Meta Platforms, Inc.), Grant Jensen (Google LLC), Lin Chen (Google LLC), Juhyun Lee (Google LLC), Andrei Kulik (Google LLC), and Matthias Grundman (Google LLC)</i>	
ADAPTOR: Adaptive Token Reduction for Video Diffusion Transformers .....	6355
<i>Elia Peruzzo (University of Trento, Italy), Adil Karjauv (Qualcomm AI Research, US), Nicu Sebe (University of Trento, Italy), Amir Ghodrati (Qualcomm AI Research, US), and Amir Habibian (Qualcomm AI Research, US)</i>	
Latent Patched Efficient Diffusion Model For High Resolution Image Synthesis .....	6362
<i>Weiyun Jiang (Rice University, USA), Devendra Kumar Jangid (Samsung Research America, USA), Seok-Jun Lee (Samsung Research America, USA), and Hamid Rahim Sheikh (Samsung Research America, USA)</i>	

LLMPi: Optimizing LLMs for High-Throughput on Raspberry Pi .....	6369
<i>Mahsa Ardakani (University of South Carolina), Jinendra Malekar (University of South Carolina), and Ramtin Zand (University of South Carolina)</i>	
AdaVid: Adaptive Video-Language Pretraining .....	6379
<i>Chaitanya Patel (Stanford University), Juan Carlos Niebles (Stanford University), and Ehsan Adeli (Stanford University)</i>	
Geometric Consistency Refinement for Single Image Novel View Synthesis via Test-Time Adaptation of Diffusion Models .....	6389
<i>Josef Bengtson (Chalmers University of Technology, Sweden), David Nilsson (Chalmers University of Technology, Sweden), and Fredrik Kahl (Chalmers University of Technology, Sweden)</i>	

## Domain Generalization: Evolution, Breakthroughs, and Future Horizons

FM-LoRA: Factorized Low-Rank Meta-Prompting for Continual Learning .....	6399
<i>Xiaobing Yu (Washington University, US), Jin Yang (Washington University in St. Louis), Xiao Wu (MBZUAI), Peijie Qiu (Washington University in St. Louis), and Xiaofeng Liu (Yale University)</i>	
T-SAM: Transductive Learning for Segment Anything Model .....	6409
<i>Rangel Daroya (University of Massachusetts Amherst), Deepak Chandran (Dolby Laboratories), Subhransu Maji (University of Massachusetts Amherst), and Andrea Fanelli (Dolby Laboratories)</i>	
Separating Shared and Domain-Specific LoRAs for Multi-Domain Learning .....	6419
<i>Yusaku Takama (Nagoya Institute of Technology), Ning Ding (Nagoya Institute of Technology), Tatsuya Yokota (Nagoya Institute of Technology), and Toru Tamaki (Nagoya Institute of Technology)</i>	
Domain Generalization for Semantic Segmentation: A Survey .....	6428
<i>Manuel Schwongerg (1TU Berlin, Mathematical Modeling of Industrial Life Cycles Group 2CARIAD SE) and Hanno Gottschalk (1TU Berlin, Mathematical Modeling of Industrial Life Cycles Group)</i>	
Improving Open-World Object Localization by Discovering Background .....	6440
<i>Ashish Singh (University of Massachusetts-Amherst), Michael Jones (MERL), Kuan-Chuan Peng (MERL), Anoop Cherian (MERL), Moitreyee Chatterjee (MERL), and Erik Learned-Miller (University of Massachusetts-Amherst)</i>	
Mixture-of-Shape-Experts (MoSE): End-to-End Shape Dictionary Framework to Prompt SAM for Generalizable Medical Segmentation .....	6450
<i>Jia Wei (Yale University), Xiaoqi Zhao (Yale University), Jonghye Woo (Massachusetts General Hospital and Harvard Medical School), Jinsong Ouyang (Yale University), Georges El Fakhri (Yale University), Qingyu Chen (Yale University), and Xiaofeng Liu (Yale University)</i>	
Domain Generalization through Attenuation of Domain-Specific Information .....	6461
<i>Reiji Saito (Meijo University) and Kazuhiro Hotta (Meijo University)</i>	

IMC: A Benchmark for Invariant Learning under Multiple Causes .....	6471
<i>Taero Kim (Yonsei University, Republic of Korea), Seonggyun Lee (Yonsei University, Republic of Korea), Joonseong Kang (Yonsei University, Republic of Korea), Youngjun Choi (Yonsei University, Republic of Korea), Wonsang Yun (Pohang University of Science and Technology (POSTECH), Republic of Korea), Nicole Hee-Yeon Kim (Korea Advanced Institute of Science &amp; Technology, Republic of Korea), Ziyu Chen (The Australian National University, Australia), Lexing Xie (The Australian National University, Australia), and Kyungwoo Song (Yonsei University, Republic of Korea)</i>	
Task-conditioned Ensemble of Expert Models for Continuous Learning .....	6481
<i>Renu Sharma (Michigan State University), Debasmita Pal (Michigan State University), and Arun Ross (Michigan State University)</i>	
Task-Level Contrastiveness for Cross-Domain Few-Shot Learning .....	6491
<i>Kristi Topollai (New York University) and Anna Choromanska (New York University)</i>	
MoPEFT: A Mixture-of-PEFTs for the Segment Anything Model .....	6502
<i>Rajat Sahay (Rochester Institute of Technology) and Andreas Savakis (Rochester Institute of Technology)</i>	
Confidence-calibrated covariate shift correction for few-shot classification in Vision-Language Models .....	6513
<i>Behraj Khan (Institute of Business Administration Karachi), Rizwan Qureshi (University of Central Florida), Nouman Muhammad Durrani (National University of Computer and Emerging Sciences, Karachi), and Tahir Qasim Syed (Institute of Business Administration Karachi)</i>	
ARDGen: Augmentation Regularization for Domain-Generalized Medical Report Generation .....	6526
<i>Syed Bilal Ahsan (NUCES FAST), Muhammad Ikhlas (BUIITEMS), Muhammad Muzamil Khan (ZHOBATS), Sana Ullah (ZHOBATS), and Muhammad Zaigham Zaheer (MBZUAI)</i>	
PiCaZo: Pixel-Aligned Contrastive Learning for Zero-Shot Domain Adaptation .....	6536
<i>Aniruddh Sikdar (Robert Bosch Centre for Cyber Physical Systems, Indian Institute of Science, Bengaluru, India), Arya Kishor (Indian Institute of Science, Bengaluru, India), Ishika Kadam (Birla Institute of Technology And Science, Pilani, India), and Suresh Sundaram (Department of Aerospace Engineering, Indian Institute of Science, Bengaluru, India)</i>	
Prototype-Based Continual Learning with Label-free Replay Buffer and Cluster Preservation Loss .....	6547
<i>Agil Aghasanli (Lancaster University), Yi Li (Lancaster University), and Plamen Angelov (Lancaster University)</i>	
VISTA-CLIP: Visual Incremental Self-Tuned Adaptation for Efficient Continual Panoptic Segmentation .....	6557
<i>Manjunath D (Indian Institute of Science), Shrikar Madhu (Indian Institute of Science), Aniruddh Sikdar (Indian Institute of Science), and Suresh Sundaram (Indian Institute of Science)</i>	
PhysNav-DG: A Novel Adaptive Framework for Robust VLM-Sensor Fusion in Navigation Applications .....	6566
<i>Trisanth Srinivasan (Cyrion Labs) and Santosh Patapati (Cyrion Labs)</i>	

# Catch UAVs that Want to Watch You: Detection and Tracking of Unmanned Aerial Vehicle (UAV) in the Wild and the 4th Anti-UAV Workshop & Challenge

Strong Baseline: Multi-UAV Tracking via YOLOv12 with BoT-SORT-ReID .....	6575
<i>Yu-Hsi Chen (The University of Melbourne)</i>	
PPTracker: Tracking UAV Swarms with Prior Prompt .....	6585
<i>Haolin Qin Qin (Beijing Institute of Technology), Tianhao Li (Beijing Institute of Technology), Tingfa Xu (Beijing Institute of Technology), Jingxuan Xu (Beijing Institute of Technology), Yuqiang Fang (Space Engineering University), and Jianan Li (Beijing Institute of Technology)</i>	
The Power of Augmentations in IR Object Detection .....	6593
<i>İhsan Emre Üstün (Aselsan, Turkey) and Cevahir Çığla (Aselsan, Turkey)</i>	
Dist-Tracker: A Small Object-aware Detector and Tracker for UAV Tracking .....	6603
<i>Wenzhen Wang (Xi'an Jiaotong University), Jing Fu (Xi'an Jiaotong University), Jiayi Song (Xi'an Jiaotong University), Kaiyu Li (Xi'an Jiaotong University), Hui Qiao (China Telecom Shaanxi Branch), Jiang Liu (China Telecom Artificial Intelligence Technology Co., Ltd), Hao Sun (China Telecom Artificial Intelligence Technology Co., Ltd), and Xiangyong Cao (Xi'an Jiaotong University)</i>	
Enhancing Few-Shot Class-Incremental Learning via Frozen Feature Augmentation .....	6612
<i>Shimou Ling (University of Electronic Science and Technology of China, China), Shengkai Gan (University of Electronic Science and Technology of China, China), Caoxin Wang (University of Electronic Science and Technology of China, China), Lili Pan (University of Electronic Science and Technology of China, China), and Hongliang Li (University of Electronic Science and Technology of China, China)</i>	
StrongSiamTracker: A Siamese Tracker with Dynamic Global Detection for Robust Anti-UAV Tracking .....	6621
<i>Xiaolong Cui (Nanjing University of Science and Technology), Liu Wan (Nanjing University of Science and Technology), Lingqi Kong (Nanjing University of Science and Technology), Jimin Li (Nanjing University of Science and Technology), Chaojie Zhang (Nanjing University of Science and Technology), Ruohan Zhao (Nanjing University of Science and Technology), Panlong Wu (Nanjing University of Science and Technology), and Shan He (Nanjing University of Science and Technology)</i>	
A Simple Detector with Frame Dynamics is a Strong Tracker .....	6632
<i>Chenxu Peng (Nankai University), Chenxu Wang (Nankai University), Minrui Zou (Nankai University), Danyang Li (Nankai University), Zhengpeng Yang (Nankai University), Yimian Dai (Nankai University), Ming-Ming Cheng (Nankai University), and Xiang Li (Nankai University)</i>	

DLST: Dual-Template Co-Evolution Learning for Robust Long-Term Drone Tracking in Dynamic Environments .....	6643
<i>Jiahao Zhang (Beijing University of Technology), Yixin Wei (Beijing University of Technology), Jinli Zhang (Beijing University of Technology), Zongli Jiang (Beijing University of Technology), Peiwen Yu (Fudan University), Yufei Ma (Beijing University of Technology), and Runan Jin (Beijing University of Technology)</i>	
Detection and Localization of Drones and UAVs Using Sound and Vision .....	6652
<i>Erik Tegler (Lund University), Max Modig (Lund University), Per Skarin (Saab Kockums), Kalle Åström (Lund University), Magnus Oskarsson (Lund University), and Gabrielle Flood (Lund University)</i>	
Securing the Skies: A Comprehensive Survey on Anti-UAV Methods, Benchmarking, and Future Directions .....	6661
<i>Yifei Dong (University of Washington), Fengyi Wu (University of Washington), Sanjian Zhang (University of Washington), Guangyu Chen (University of Washington), Yuzhi Hu (University of Washington), Masumi Yano (University of Washington), Jingdong Sun (Carnegie Mellon University), Siyu Huang (Clemson University), Feng Liu (Drexel University), Qi Dai (Microsoft Research), and Zhi-Qi Cheng (University of Washington)</i>	

## Author Index