2025 IEEE/CVF Winter Conference on Applications of Computer Vision Workshops (WACVW 2025)

Tucson, Arizona, USA
28 February - 4 March 2025

Pages 1-805

IEEE Catalog Number: CFP25B39-POD
ISBN: 979-8-3315-3663-3

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:	CFP25B39-POD
ISBN (Print-On-Demand):	979-8-3315-3663-3
ISBN (Online):	979-8-3315-3662-6
ISSN:	2572-4398

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
Web: www.proceedings.com

2025 IEEE/CVF Winter Conference on Applications of Computer Vision Workshops (WACVW)

WACVW 2025

Table of Contents

Message from the Workshop Chairs ..... **v**

2nd CV4Smalls Workshop Computer Vision with Small Data: A Focus on Infants, Toddlers, and the Elderly (CV4Small)

Experimenting with Affective Computing Models in Video Interviews with Spanish-speaking Older Adults 1
Josep López Camuñas (Universitat Oberta de Catalunya, Spain), Cristina Bustos (Universitat Oberta de Catalunya, Spain), Yanjun Zhu (Northeastern University, USA), Raquel Ros (PAL Robotics, Spain), and Agata Lapedriza (Universitat Oberta de Catalunya, Spain)

Applying Computer Vision to Analyze Self-Injurious Behaviors in Children with Autism Spectrum Disorder 11
Ayda Eghbalian (University of Texas, USA), Md Mushfiqur Azam (University of Texas, USA), Katie Holloway (University of Texas, USA), Leslie Neely (University of Texas, USA), and Kevin Desai (University of Texas, USA)

Automatic Temporal Segmentation for Post-Stroke Rehabilitation: A Keypoint Detection and Temporal Segmentation Approach for Small Datasets 21
Jisoo Lee (Arizona State University, United States), Tamim Ahmed (University of Southern California, United States), Thanassis Rikakis (University of Southern California, United States), and Pavan Turaga (Arizona State University, United States)

SEER-ZSL: Semantic Encoder-Enhanced Representations for Generalized Zero-Shot Learning 30
William Heyden (Norwegian University of Life Sciences), Habib Ullah (Norwegian University of Life Sciences), Muhammad Salman Siddiqui (Norwegian University of Life Sciences), and Fadi Al Machot (Norwegian University of Life Sciences)

Classification of Infant Sleep–Wake States from Natural Overnight In-Crib Sleep Videos 42
Shayda Moezzi (Northeastern University, USA), Michael Wan (Northeastern University, USA), Sai Kumar Reddy Manne (Northeastern University, USA), Amal Mathew (Northeastern University, USA), Shaotong Zhu (Northeastern University, USA), Bishoy Galoaa (Northeastern University, USA), Elaheh Hatamimajoumerd (Northeastern University, USA), Emma Grace (University of Maine, USA), Cassandra B Rowan (University of Maine, USA), Emily Zimmerman (Northeastern University, USA), Briana J Taylor (Northeastern University, USA), Marie J Hayes (University of Maine, USA), and Sarah Ostadabbas (Northeastern University, USA)

UniMotion: Bridging 2D and 3D Representations for Human Motion Prediction 52
Yanjun Zhu (Northeastern University), Chen Bai (Xpeng Motors), Cheng Lu (Xpeng Motors), David Doermann (University at Buffalo), and Agata Lapedriza (Universitat Oberta de Catalunya)

CST: Character State Transformer for Object-Conditioned Human Motion Prediction 63
Kuan-Wei Tseng (Tokyo Institute of Technology), Rei Kawakami (Institute of Science Tokyo), Satoshi Ikehata (National Institute of Informatics /Institute of Science Tokyo), and Ikuro Sato (Institute of Science Tokyo / Denso IT Laboratory)

Advancing Multi-Person Tracking for Autism Behavior Analysis: Challenges, Opportunities, and Future Directions in Clinical Settings 73
Somaieh Amraee (Northeastern University, USA), Eva Blanco-Mallo (Northeastern University, USA), Deniz Erdogmus (Northeastern University, USA), Aston McCullough (Northeastern University, USA), Matthew Goodwin (Northeastern University, USA), and Sarah Ostadabbas (Northeastern University, USA)

Iris Recognition for Infants 83
Rasel Ahmed Bhuiyan (University of Notre Dame, IN, USA), Mateusz Trokielewicz (Warsaw University of Technology, Poland, and PayEye Poland), Piotr Maciejewicz (Medical University of Warsaw, Poland), Sherri Bucher (Indiana University Indianapolis, IN, USA), and Adam Czajka (University of Notre Dame, IN, USA)

ULTRRA: Unconstrained Large-scale Three-dimensional Reconstruction and Rendering across Altitudes (ULTRRA)

4th Workshop on Image/Video/Audio Quality in Computer Vision and Generative AI (ImageQuality)

Mahalanobis k-NN: A Statistical Lens for Robust Point-Cloud Registrations 93
Tejas Anvekar (University of Wyoming) and Shivanand Venkanna Sheshappanavar (University of Wyoming)

Improving Human Pose-Conditioned Generation: Fine-tuning ControlNet Models with Reinforcement Learning 103
Jeonghwan Lee (Yonsei University), Heywon Yun (Yonsei University), Jimin Kim (Ewha Womans University), and Homa Fashandi (LGE Toronto AI Lab)

PQD: Post-training Quantization for Efficient Diffusion Models 113
Jiaojiao Ye (University of Oxford), Zhen Wang (Zhejiang Lab), and Linnan Jiang (Zhejiang Lab)

High-Fidelity 4x Neural Reconstruction of Real-time Path Traced Images 120
Zhiqiang Lao (Futurewei Technology Inc), Yu Guo (Futurewei Technology Inc), Xiyun Song (Futurewei Technology Inc), Yubin Zhou (Futurewei Technology Inc), Zongfang Lin (Futurewei Technology Inc), Heather Yu (Futurewei Technology Inc), and Liang Peng (Futurewei Technology Inc)

LatentPS: Image Editing Using Latent Representations in Diffusion Models 130
Zilong Wu (The University of Tokyo, Japan), Hideki Murata (Dentsu Digital, Japan), Nayu Takahashi (Dentsu Digital, Japan), Qiyu Wu (The University of Tokyo, Japan), and Yoshimasa Tsuruoka (The University of Tokyo, Japan)

Lights, Camera, Matching: The Role of Image Illumination in Fair Face Recognition 140
Gabriella Pangelinan (Florida Institute of Technology), Grace Bezold (University of Notre Dame), Haiyu Wu (University of Notre Dame), Michael C King (Florida Institute of Technology), and Kevin W Bowyer (University of Notre Dame)

Quantifying Generative Stability: Mode Collapse Entropy Score for Mode Diversity Evaluation 150
Jens Duym (University of Antwerp), José Antonio Oramas Mogrovejo (University of Antwerp), and Ali Anwar (University of Antwerp)

Similarity Trajectories: Linking Sampling Process to Artifacts in Diffusion-Generated Images 160
Dennis Menn (University of Texas at Austin), Feng Liang (University of Texas at Austin), Hung-Yueh Chiang (University of Texas at Austin), and Diana Marculescu (University of Texas at Austin)

A Distortion Aware Image Quality Assessment Model 170
Ha Thu Nguyen (Norwegian University of Science and Technology), Katrien De Moor (Norwegian University of Science and Technology), Mohamed-Chaker Larabi (CNRS, Universite de Poitiers, XLIM, Poitiers), and Seyed Ali Amirshahi (Norwegian University of Science and Technology)

MambaTron: Efficient Cross-Modal Point Cloud Enhancement using Aggregate Selective State Space Modeling 180
Sai Tarun Inaganti (University of Minnesota) and Gennady Petrenko (Homothereum)

Diffusion Prism: Enhancing Diversity and Morphology Consistency in Mask-to-Image Diffusion... 191
Hao Wang (Clemson University, USA), Xiwen Chen (Clemson University, USA), Ashish Bastola (Clemson University, USA), Jiayou Qin (Stevens Institute of Technology, USA), and Abolfazl Razi (Clemson University, USA)

TE-NeRF: Triplane-Enhanced Neural Radiance Field for Artifact-Free Human Rendering 201
Sadia Mubashshira (University of Texas at SanAntonio) and Kevin Desai (University of Texas at San Antonio)

IP-FaceDiff: Identity-Preserving Facial Video Editing with Diffusion 211
Tharun Anand (IIT Madras), Aryan Garg (IIT Madras), and Kaushik Mitra (IIT Madras)

SST-EM: Advanced Metrics for Evaluating Semantic, Spatial and Temporal Aspects in Video Editing 222
Varun Biyyala (Yeshiva University), Bharat Chanderprakash Kathuria (Yeshiva University), Jialu Li (Yeshiva University), and Youshan Zhang (Yeshiva University)

Confident Pseudo-labeled Diffusion Augmentation for Canine Cardiomegaly Detection 232
Shiman Zhang (Yeshiva University), Lakshmikar Polamreddy (Yeshiva University), and Youshan Zhang (Yeshiva University)

Revealing Palimpsests with Latent Diffusion Models: A Generative Approach to Image Inpainting and Handwriting Reconstruction 242
Mahdi Jampour (Universität Hamburg, Hamburg, Germany)

Unsupervised Generative Approach for Anomaly Detection to Enhance the Quality of Unseen Medical Datasets 250
Zhemin Zhang (Arizona State University), Bhavika Patel (Mayo Clinic, AZ), Bhavik Patel (Mayo Clinic, AZ), and Imon Banerjee (Mayo Clinic, AZ)

Sparse Mixture-of-Experts for Non-Uniform Noise Reduction in MRI Images 260
Zeyun Deng (Purdue University) and Joseph Campbell (Purdue University)

HipyrNet: Hypernet-Guided Feature Pyramid network for mixed-exposure correction 269
Shaurya Rathore (Manipal Institute of Technology, MAHE), Aravind Shenoy (Manipal Institute of Technology, MAHE), Krish Didwania (Manipal Institute of Technology, MAHE), Aditya Kasliwal (Manipal Institute of Technology, MAHE), and Ujjwal Verma (Manipal Institute of Technology, MAHE)

DaBiT: Depth and Blur informed Transformer for Video Deblurring 278
Crispian Morris (University of Bristol), Nantheera Anantrasirichai (University of Bristol), Fan Zhang (University of Bristol), and David Bull (University of Bristol)

LS-GAN: Human Motion Synthesis with Latent-space GANs 289
Avinash Amballa (University of Massachusetts Amherst, USA), Gayathri Akkinapalli (University of Massachusetts Amherst, USA), and Vinitra Muralikrishnan (University of Massachusetts Amherst, USA)

Advancing Super-Resolution in Neural Radiance Fields via Variational Diffusion Strategies 299
Shrey Vishen (Monta Vista High School), Jatin Sarabu (Bellarmine College Preparatory), Saurav Kumar (UI Urbana-Champaign), Chinmay Bharathulwar (John P Stevens High School), Rithwick Lakshmanan (Pleasant Valley High School), and Vishnu Srinivas (Foothill High School)

V3SC – Video Surveillance Systems in Smart Cities: Integrating Traditional Cameras with High-Altitude and Advanced Monitoring Technologies (V3SC)

Invisibility Cloak: Hiding Anomalies in Videos via Adversarial Machine Learning Attacks 307
Hamza Karim (University of south Florida) and Yasin Yilmaz (University of South Florida)

AttriVision: Advancing Generalization in Pedestrian Attribute Recognition using CLIP 317
Mehran ADIBI SEDEH (Two-i, FRANCE), Assia BENBIHI (CIIRC, Czech Republic), Romain MARTIN (Two-i, FRANCE), Marianne CLAUSEL (université de Lorraine, FRANCE), and Cédric PRADALIER (GeorgiaTech Lorraine, FRANCE)

Video-to-Text Pedestrian Monitoring (VTPM): Leveraging Large Language Models for Privacy-Preserve Pedestrian Activity Monitoring at Intersections 329
Ahmed S Abdelrahman (University of Central Florida), Mohamed Abdel-Aty (University of Central Florida), and Dongdong Wang (University of Central Florida)

TakuNet: an Energy-Efficient CNN for Real-Time Inference on Embedded UAV systems in Emergency Response Scenarios 339
Daniel Rossi (University of Modena and Reggio Emilia, Italy), Guido Borghi (University of Modena and Reggio Emilia, Italy), and Roberto Vezzani (University of Modena and Reggio Emilia, Italy)

Gaze meets Computer Vision (GMCV): Bridging Human Attention and Machine Perception (GMCV)

Causal Representation-Based Domain Generalization on Gaze Estimation 349
Younghan Kim (Gachon University, Korea), Kangryun Moon (Sungkyunkwan University, Korea), Yongjun Park (Visualcamp, Korea), and Yonggyu Kim (Visualcamp, Korea)

Human Gaze Improves Vision Transformers by Token Masking 359
Jacqueline Kockwelp (University of Münster, Germany), Daniel Beckmann (University of Münster, Germany), and Benjamin Risse (University of Münster, Germany)

Addressing Age Bias in the Application of Appearance-Based Gaze-Tracking for Older Adults 369
Manuela Kunz (National Research Council Canada, Canada), Kathleen C. Fraser (National Research Council Canada, Canada), Bruce Wallace (Carleton University, Canada), Frank Knoefel (Elisabeth Bruyere Hospital, Canada), Rafik Goubran (Carleton University, Canada), Sina Shafiyan (University Ottawa, Canada), and Neil Thomas (Elisabeth Bruyere Hospital, Canada)

VISTA: A Visual and Textual Attention Dataset for Interpreting Multimodal Models 378
Harshit . (University of Utah, USA) and Tolga Tasdizen (University of Utah, USA)

On Segmenting Pupil Contours in Terms of Elliptical Fourier Series 386
Felix O'Mahony (Pupil Labs, Germany), Patrick Fürst (Pupil Labs, Germany), Michael Drews (Pupil Labs, Germany), and Kai Dierkes (Pupil Labs, Germany)

FlatTrack: Eye-tracking with ultra-thin lensless cameras 396
Purvam Jain (IIT Madras, India), Althaf M. Nazar (IIT Madras, India), Salman S. Khan (Rice University, USA), Kaushik Mitra (IIT Madras, India), and Praneeth Chakravarthula (UNC Chapel Hill, USA)

GeoCV: Computer Vision for Geospatial Image Analysis (GeoCV)

CaLiSa-NeRF: Neural Radiance Field with Pinhole Camera Images, LiDAR point clouds and Satellite Imagery for Urban Scene Representation 405
Juyeop Han (MIT), Guilherme Cavalheiro (MIT), Josef Biberstein (MIT), Elham Alkabawi (KACST), Shahad Alqhatni (KACST), Fadwa Alaskar (KACST), Eman Bin Khunayn (KACST), and Sertac Karaman (MIT)

Advancing Open-Set Object Detection in Remote Sensing Using Multimodal Large Language Model 414
Nandini Saini (Indian Institute of Technology Jodhpur), Ashudeep Dubey (Indian Institute of Technology Jodhpur), Debasis Das (Indian Institute of Technology Jodhpur), and Chiranjoy Chattopadhyay (FLAME University Pune)

Adaptive Clustering for Efficient Phenotype Segmentation of UAV Hyperspectral Data 422
Ciem Cornelissen (UGhent), Sam Leroux (UGhent), and Pieter Simoens (UGhent)

A Zero-Shot Learning Approach for Ephemeral Gully Detection from Remote Sensing using Vision Language Models 432
Seyed Mohamad Ali Tousi (University of Missouri - Columbia, USA), Ramy Farag (University of Missouri - Columbia, USA), Jacket Demby's (University of Missouri - Columbia, USA), Gbenga Omotara (University of Missouri - Columbia, USA), John A. Lory (University of Missouri - Columbia, USA), and G. N. DeSouza (University of Missouri - Columbia, USA)

Temporal Resilience in Geo-Localization: Adapting to the Continuous Evolution of Urban and Rural Environments 442
Fabian Deuser (University of the Bundeswehr Munich), Wejdene Mansour (Technical University of Munich), Hao Li (Technical University of Munich), Konrad Habel (University of the Bundeswehr Munich), Martin Werner (Technical University of Munich), and Norbert Oswald (University of the Bundeswehr Munich)

PrecipFormer: Efficient Transformer for Precipitation Downscaling 452
Rohit Kumar (IISER Bhopal), Tanishq Sharma (IISER Bhopal), Vedanshi Vaghela (IISER Bhopal), Sanjeev K. Jha (IISER Bhopal), and Akshay Agarwal (IISER Bhopal)

CrossModalityDiffusion: Multi-Modal Novel View Synthesis with Unified Intermediate Representation 461
Alexander Berian (University of Arizona), Daniel Brignac (University of Arizona), Jhih Yang Wu (University of Arizona), Natnael Daba (University of Arizona), and Abhijit Mahalanobis (University of Arizona)

Enhancing Remote Sensing Representations Through Mixed-Modality Masked Autoencoding 470
Ori Linial (Google), George Leifman (Google), Yochai Blau (Google), Nadav Sherman (Google), Yotam Gigi (Google), Wojciech Sirko (Google), and Genady Beryozkin (Google)

Mapping Refugee Camps with AI: A Benchmark Dataset and Baseline Models for Humanitarian Applications 480
Amrita Gupta (Microsoft, United States of America), Anthony Ortiz (Microsoft, United States of America), Simone Fobi (Microsoft, United States of America), Duncan Kebut (Humanitarian OpenStreetMap Team, Kenya), Seema Iyer (USA for UNHCR, United States of America), Rahul Dodhia (Microsoft, United States of America), and Juan Lavista-Ferres (Microsoft, United States of America)

Leveraging Satellite Image Time Series for Accurate Extreme Event Detection 489
Heng Fang (KTH Royal Institute of Technology, Sweden) and Hossein Azizpour (KTH Royal Institute of Technology, Sweden)

MD-Glow: Multi-task Despeckling Glow for SAR Image Enhancement 499
Shunsuke Takao (University of Tsukuba)

Multiresolution Fusion and Classification of Hyperspectral and Panchromatic Remote Sensing Images 507
Martina Pastorino (University of Genoa), Gabriele Moser (University of Genoa), Sebastiano B. Serpico (University of Genoa), and Josiane Zerubia (Inria, Universite Cote d'Azur)

ProMM-RS: Exploring Probabilistic learning for Multi-Modal Remote Sensing Image Representations 517
Nicolas Houdré (Université Paris Cité, France), Diego Marcos (INRIA, France), Dino Ienco (INRAE, France), Laurent Wendling (Université Paris Cité, France), Camille Kurtz (Université Paris Cité, France), and Sylvain Lobry (Université Paris Cité, France)

Dfilled: Repurposing Edge-Enhancing Diffusion for Guided DSM Void Filling 526
Daniel Panangian (German Aerospace Center) and Ksenia Bittner (German Aerospace Center)

Enhancing Worldwide Image Geolocation by Ensembling Satellite-Based Ground-Level Attribute Predictors 535
Bianco Michael (Clarifai Inc.), David Eigen (Clarifai Inc.), and Michael Gormish (Clarifai Inc.)

Hyperspectral Pansharpening with Transformer-based Spectral Diffusion Priors 544
Hongcheng Jiang (University of Missouri Kansas City) and ZhiQiang Chen (University of Missouri Kansas City)

Direction guided Segmentation and Vectorisation of curbstones from high-resolution ortho-images 554
Mariya Jose (Leibniz University Hannover), Stefan Auer (German Aerospace Center), and Jiaojiao Tian (German Aerospace Center)

Layer Optimized Spatial Spectral Masked Autoencoder for Semantic Segmentation of Hyperspectral Imagery 562
Aaron Perez (University of Houston) and Saurabh Prasad (University of Houston)

Pre-training of Auto-generated Synthetic 3D Point Cloud Segmentation for Outdoor Scene 571
Takayuki Shinohara (National Institute of Advanced Industrial Science and Technology)

FuseForm: Multimodal Transformer for Semantic Segmentation 581
Justin McMillen (University of South Florida) and Yasin Yilmaz (University of South Florida)

The First Workshop on Out-of-Label Hazards in Autonomous Driving (COOOL)

Robustness to Perturbations in the Frequency Domain: Neural Network Verification and Certified Training 591
Harleen Hanspal (Imperial College London), Alessandro De Palma (INRIA, France), and Alessio Lomuscio (Imperial College London)

Model Weights Reflect a Continuous Space of Input Image Domains 601
Simen Cassiman (KU Leuven, Belgium), Marc Proesmans (KU Leuven, Belgium), Tinne Tuytelaars (KU Leuven, Belgium), and Luc Van Gool (em. KU Leuven, Belgium; INSAIT, Bulgaria)

Zero-shot hazard identification in Autonomous Driving: A Case Study on the COOOL Benchmark 611
Lukas Picek (INRIA / University of West Bohemia), Vojtech Cermak (Czech Technical University in Prague), and Marek Hanzl (University of West Bohemia)

Interpreting the Unexpected: A Multimodal Framework for Out-of-Label Hazard Detection and Explanation in Autonomous Driving 621
Mahdi Abbariki (Noshirvani University of Technology) and Maged Shoman (University of Tennessee)

SGNetPose+: Stepwise Goal-Driven Networks with Pose Information for Trajectory Prediction in Autonomous Driving 629
Akshat Ghiya (University of Colorado Colorado Springs, USA), Ali AlShami (University of Colorado Colorado Springs, USA), and Jugal Kalita (University of Colorado Colorado Springs, USA)

Open-World Hazard Detection and Captioning for Autonomous Driving with a Unified Multimodal Pipeline 638
Parisa Hatami (University of Tennessee at Chattanooga), Maged Shoman (University of Tennessee), and Mina Sartipi (University of Tennessee Chattanooga)

The 2nd Workshop on Artificial Intelligence for Multimedia Forensics and Disinformation Detection (AI4MFDD 2025)

DiffFake: Exposing Deepfakes using Differential Anomaly Detection 647
Sotirios Stamnas (University of Warwick, United Kingdom) and Victor Sanchez (University of Warwick, United Kingdom)

Improving the Perturbation-Based Explanation of Deepfake Detectors Through the Use of Adversarially-Generated Samples 658
Konstantinos Tsigos (Information Technologies Institute - CERTH), Evlampios Apostolidis (Information Technologies Institute - CERTH), and Vasileios Mezaris (Information Technologies Institute - CERTH)

IDTrust: Deep Identity Document Quality Detection with Bandpass Filtering 668
Musab Al-Ghadi (L3i Laboratory, La Rochelle University, France), Joris Voerman (L3i Laboratory, La Rochelle University, France), Mickael Coustaty (L3i Laboratory, La Rochelle University, France), Olivier Lessard (IMDS SAS, France), and Nicolas Sidere (L3i Laboratory, La Rochelle University, France)

HFMF: Hierarchical Fusion Meets Multi-Stream Models for Deepfake Detection 676
Anant Mehta (Texas A&M University), Bryant McArthur (Texas A&M University), Nagarjuna Kolloju (Texas A&M University), and Zhengzhong Tu (Texas A&M University)

GrDT: Towards Robust Deepfake Detection using Geometric Representation Distribution and Texture 686
Hongyang Xie (University of Warwick), Hongyang He (University of Warwick), Boyang Fu (University of Warwick), and Victor Sanchez (University of Warwick)

FoundPAD: Foundation Models Reloaded for Face Presentation Attack Detection 697
Guray Ozgur (Fraunhofer IGD, Germany), Eduarda Caldeira (Fraunhofer IGD, Germany), Tahar Chettaoui (Fraunhofer IGD, Germany), Fadi Boutros (Fraunhofer IGD, Germany), Raghavendra Ramachandra (Norwegian University of Science and Technology, Norway), and Naser Damer (TU Darmstadt and Fraunhofer IGD, Germany)

Disharmony: Forensics using Reverse Lighting Harmonization 708
Philip Wootaek Shin (Pennsylvania State University), Jack Sampson (Pennsylvania State University), Vijaykrishnan Narayanan (Pennsylvania State University), Andres Marquez (Pacific Northwest National Laboratory), and Mahantesh Halappanavar (Pacific Northwest National Laboratory)

FX-MAD: Frequency-domain Explainability and Explainability-driven Unsupervised Detection of Face Morphing Attacks 718
Marco Huber (Fraunhofer Institute for Computer Graphics Research IGD, Germany / Technische University, Germany), Pedro C. Neto (INESC TEC, Portugal / University of Porto, Portugal), Ana F. Sequeira (INESC TEC, Portugal / University of Porto, Portugal), and Naser Damer (Fraunhofer Institute for Computer Graphics Research IGD, Germany / Technische University, Germany)

Enhancing Synthetic Generated-Images Detection through Post-Hoc Calibration 729
Giovanna Maria Dimitri (University of Siena), Benedetta Tondi (University of Siena), and Mauro Barni (University of Siena)

Zero-Shot Warning Generation for Misinformative Multimodal Content 737
Giovanni Pio Delvecchio (University of Bologna, Italy), Huy Hong Nguyen (National Institute of Informatics, Japan), and Isao Echizen (National Institute of Informatics, Japan; The University of Tokyo, Japan)

Enhancing Ground-to-Aerial Image Matching for Visual Misinformation Detection Using Semantic Segmentation 747
Emanuele Mule (Sapienza University of Rome), Matteo Pannacci (Sapienza University of Rome), Ali Ghasemi Goudarzi (Sapienza University of Rome), Francesco Pro (Sapienza University of Rome), Lorenzo Papa (Sapienza University of Rome), Luca Maiano (Sapienza University of Rome), and Irene Amerini (Sapienza University of Rome)

Zero-training fraud detection in a large messaging platform? 756
Stephane Schwarz (University of Campinas, Brazil), Paulo Fonseca (Sinch AB, Sweden), and Anderson Rocha (University of Campinas, Brazil)

Real-World Surveillance: Applications and Challenges 5th (RWS)

SSTAR: Skeleton-based Spatio-Temporal Action Recognition for Intelligent Video Surveillance and Suicide Prevention in Metro Stations 765
Safwen Naimi (TÉLUQ), Wassim Bouachir (TÉLUQ), Guillaume-Alexandre Bilodeau (Polytechnique Montréal), and Brian Mishara (UQAM)

YOLO11-JDE: Fast and Accurate Multi-Object Tracking with Self-Supervised Re-ID 776
Iñaki Erregue (Universitat de Barcelona, Spain), Kamal Nasrollahi (Milestone Systems, Denmark), and Sergio Escalera (Universitat de Barcelona, Spain)

Evaluation of Spatio-Temporal Small Object Detection in Real-World Adverse Weather Conditions 786
Michel van Lier (TNO), Martin van Leeuwen (TNO), Bastian van Manen (TNO), Leo Kampmeijer (TNO), and Nicolas Boehrer (TNO)

Fall Detection: Leveraging Depth Information in Bayesian Networks 798
Marc Oliu (MILESTONE SYSTEMS A/S), Rohat Bozyil (MILESTONE SYSTEMS A/S), Mia Siemon (MILESTONE SYSTEMS A/S), Alejandro Martinez-Senent (MILESTONE SYSTEMS A/S), Sergio Escalera (Universitat de Barcelona), Thomas Moeslund (Aalborg Universitet), and Kamal Nasrollahi (Aalborg Universitet)

Interpreting Face Recognition Templates using Natural Language Descriptions 806
Anastasija Manojlovska (University of Ljubljana, Slovenia), Raghavendra Ramachandra (NTNU, Gjøvik, Norway), Georgios Spathoulas (NTNU, Gjøvik, Norway), Vitomir Štruc (University of Ljubljana, Slovenia), and Klemen Grm (University of Ljubljana, Slovenia)

MCTR: Multi Camera Tracking Transformer 816
Alexandru Niculescu-Mizil (NEC Laboratories America), Deep Patel (NEC Laboratories America), and Iain Melvin (NEC Labs America)

Joint Audio-Visual Idling Vehicle Detection with Streamlined Input Dependencies 827
Xiwen Li (University of Utah), Rehman Mohammed (University of Utah), Tristalee Mangin (University of Utah), Surojit Saha (University of Utah), Kerry Kelly (Uof U), Ross Whitaker (Scientific Computing and Imaging Institute, University of Utah), and Tolga Tasdizen (University of Utah, USA)

Simultaneous Multi-Object Multi-Camera Trajectory Forecasting (SMO-MCTF) 837
Amey Noolkar (University of Glasgow, UK) and Victor Sanchez (University of Warwick, UK)

B-FPGM: Lightweight Face Detection via Bayesian-Optimized Soft FPGM Pruning 844
Nikolaos Kaparinos (CERTH-ITI) and Vasileios Mezaris (CERTH-ITI)

EVGEN: "Event-based Vision in the Era of Generative AI" – Transforming Perception and Visual Innovation (EVGEN)

BeSplat: Gaussian Splatting from a Single Blurry Image and Event Stream 854
Gopi Raju Matta (IIT Madras), Trisha Reddypalli (IIIT RGUKT RK Valley), and Kaushik Mitra (IIT Madras)

Simple Transformer with Single Leaky Neuron for Event Vision 865
Himanshu Kumar (NA) and Aniket Konkar (The George Washington University)

Combined Physics and Event Camera Simulator for Slip Detection 872
Thilo Reinold (TU Berlin), Suman Ghosh (TU Berlin), and Guillermo Gallego (TU Berlin)

Continuous Histogram for Event-based Vision Camera Systems 881
JongHun Park (ATUS Inc.) and MunHo Hong (ATUS Inc.)

Ethical development and evaluation of multi-modal AI in healthcare (EMMA)

3rdWorkshop on Large Language and Vision Models for Autonomous Driving (LLVMAD)

Language-Driven Active Learning for Diverse Open-Set 3D Object Detection 890
Ross Greer (University of California, Merced), Bjørk Antoniussen (Aalborg Universitet), Andreas Møgelmose (Aalborg Universitet), and Mohan Trivedi (University of California San Diego)

SenseRAG: Constructing Environmental Knowledge Bases with Proactive Querying for LLM-Based Autonomous Driving 899
Xuewen Luo (Monash University), Chenxi Liu (University of Utah), Fan Ding (Monash University), Fengze Yang (University of Utah), Yang Zhou (Texas A&M University), Junnyong Loo (Monash University), and Hwa Hui Tew (Monash University)

Glimpse of MCQ based VQA in Road & Traffic Scenarios 907
Ambarish Parthasarathy (Indian Institute of Technology), Athira Krishnan R (Indian Institute of Technology), and Sumukha BG (Indian Institute of Technology)

OpenEMMA: Open-Source Multimodal Model for End-to-End Autonomous Driving 911
Shuo Xing (Texas A&M University), Chengyuan Qian (Texas A&M University), Yuping Wang (University of Michigan), Hongyuan Hua (University of Toronto), Kexin Tian (Texas A&M University), Yang Zhou (Texas A&M University), and Zhengzhong Tu (Texas A&M University)

Position: Prospective of Autonomous Driving - Multimodal LLMs, World Models, Embodied Intelligence, AI Alignment, and Mamba 920
Yunsheng Ma (Purdue University), Wenqian Ye (University of Virginia), Can Cui (Purdue University), Haiming Zhang (CUHK–SZ), Shuo Xing (Texas A&M University), Fucai Ke (Monash University), Jinhong Wang (Zhejiang University), Chenglin Miao (Iowa State University), Jintai Chen (HKUST-GZ), Hamid Rezatofighi (Monash University), Zhen Li (CUHK–SZ), Guangtao Zheng (University of Virginia), Chao Zheng (Tencent), Tianjiao He (University of Toronto), Manmohan Chandraker (University of California San Diego), Burhaneddin Yaman (Bosch Center for Artificial Intelligence), Xin Ye (Bosch Center for Artificial Intelligence), Hang Zhao (Tsinghua University), and Xu Cao (University of Illinois Urbana-Champaign)

Evaluating Multimodal Vision-Language Model Prompting Strategies for Visual Question Answering in Road Scene Understanding 937
Aryan Keskar (University of California, Merced), Srinivasa Perisetla (University of California, Merced), and Ross Greer (University of California, Merced)

Scenario Understanding of Traffic Scenes Through Large Visual Language Models 947
Esteban Rivera (Technische Universität München (TUM), Germany), Jannik Lübberstedt (Technische Universität München (TUM), Germany), Nico Uhlemann (Technische Universität München (TUM), Germany), and Markus Lienkamp (Technische Universität München (TUM), Germany)

Enhancing Weakly-Supervised Object Detection on Static Images through (Hallucinated) Motion 956
Cagri Gungor (University of Pittsburgh) and Adriana Kovashka (University of Pittsburgh)

Query3D: LLM-Powered Open-Vocabulary Scene Segmentation with Language Embedded 3D Gaussians 961
Amirhosein Chahe (Drexel University) and Lifeng Zhou (Drexel University)

ScVLM: Enhancing Vision-Language Model for Safety-Critical Event Understanding 971
Liang Shi (Virginia Polytechnic Institute and State University, USA), Boyu Jiang (Virginia Polytechnic Institute and State University, USA), Tong Zeng (Virginia Polytechnic Institute and State University, USA), and Feng Guo (Virginia Polytechnic Institute and State University, USA)

VLMine: Long-Tail Data Mining with Vision Language Models 982
Mao Ye (The University of Texas at Austin), Greg P. Meyer (Cruise), Zaiwei Zhang (Cruise), Dennis Park (Cruise), Siva Karthik Mustikovela (Cruise), Yuning Chai (Cruise), and Eric Wolff (Cruise)

Workshop on Automated Spatial and Temporal Anomaly Detection (ASTAD)

Detecting Contextual Anomalies by Discovering Consistent Spatial Regions 993
Zhengye Yang (Rensselaer Polytechnic Institute) and Richard J Radke (Rensselaer Polytechnic Institute)

ComplexVAD: Detecting Interaction Anomalies in Video 1003
Furkan Mumcu (University of South Florida, USA), Michael Jones (Mitsubishi Electric Research Laboratories, USA), Yasin Yilmaz (University of South Florida, USA), and Anoop Cherian (Mitsubishi Electric Research Laboratories, USA)

ML-JET: A Benchmark Dataset for Classifying Heavy Ion Collisions 1013
Haydar Mehryar (Wayne State University, USA), Chengzhi Mao (Columbia University, USA), and Loren Schwiebert (Wayne State University, USA)

AnoFPDM: Anomaly Detection with Forward Process of Diffusion Models for Brain MRI 1023
Yiming Che (Arizona State University, United States), Fazle Rafsani (Arizona State University, United States), Jay Shah (Arizona State University, United States), Md Mahfuzur Rahman Siddiquee (Arizona State University, United States), and Teresa Wu (Arizona State University, United States)

Exploring Pose-Based Anomaly Detection for Retail Security: A Real-World Shoplifting Dataset and Benchmark 1033
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), Shanle Yao (University of North Carolina at Charlotte), and Hamed Tabkhi (University of North Carolina at Charlotte)

PCAD: A Real-World Dataset for 6D Pose Industrial Anomaly Detection 1042
Robert Maack (Institute for Technologies and Management of Digital Transformation (TMDT), University of Wuppertal), Lars Thun (Institute for Technologies and Management of Digital Transformation (TMDT), University of Wuppertal), Thomas Liang (University of Wisconsin–Madison), Hasan Tercan (Institute for Technologies and Management of Digital Transformation (TMDT), University of Wuppertal), and Tobias Meisen (Institute for Technologies and Management of Digital Transformation (TMDT), University of Wuppertal)

The International Workshop on Human-Autonomous Vehicle Interaction (HAVI)

AAT-DA: Accident Anticipation Transformer with Driver Attention 1052
Yuto Kumamoto (Nagoya University), Kento Ohtani (Nagoya University), Daiki Suzuki (DENSO CORPORATION), Minori Yamataka (DENSO CORPORATION), and Kazuya Takeda (Nagoya University)

Snapshot: Towards Application-centered Models for Pedestrian Trajectory Prediction in Urban Traffic Environments 1062
Nico Uhlemann (Technische Universität München (TUM), Germany), Yipeng Zhou (Technische Universität München (TUM), Germany), Tobias Simeon Mohr (Technische Universität München (TUM), Germany), and Markus Lienkamp (Technische Universität München (TUM), Germany)

"What's Happening"- A Human-centered Multimodal Interpreter Explaining the Actions of Autonomous Vehicles 1073
Xuewen Luo (Monash University), Fan Ding (Monash University), Rishikesh Panda (BITS Pilani K K Birla Goa Campus), Ruiqi Chen (University of Washington), Junnyong Loo (Monash University), and Shuyun Zhang (King's College London)

Deep Learning-based rPPG Models towards Automotive Applications: A Benchmark Study 1081
Tayssir Bouraffa (Chalmers University of Technology Department of Computer Science and Engineering Gothenburg, Sweden), Dimitrios Koutsakis (Chalmers University of Technology Gothenburg, Sweden), and Salvija Zelvyte (Chalmers University of Technology Gothenburg, Sweden)

2nd Workshop on Computer Vision for Earth Observation Applications (CV4EO)

Location generalizability of image-based air quality models 1091
Eleanor Byler (Pacific Northwest National Laboratory), Christian Svinth (Pacific Northwest National Laboratory), and Kirsten Chojnicki (Pacific Northwest National Laboratory)

FineAir: Finest-grained Airplanes in High-resolution Satellite Images 1101
Murat Osswald (University of the Bundeswehr Munich, Germany), Louis Niederloehner (University of the Bundeswehr Munich, Germany), Sascha Koejer (University of the Bundeswehr Munich, Germany), Tobias Ziedorn (University of the Bundeswehr Munich, Germany), Valerio Gulli (European Space Imaging, Germany), Michael Mommert (Stuttgart University of Applied Sciences, Germany), and Helmut Mayer (University of the Bundeswehr Munich, Germany)

Model Compression Meets Resolution Scaling for Efficient Remote Sensing Classification 1110
Tushar Shinde (IIT Madras Zanzibar)

Adaptive Structure-Aware Connectivity-Preserving Loss for Improved Road Segmentation in Remote Sensing Images 1120
Sara Shojaei (University of Missouri-Columbia, USA), Trevor Bohl (University of Missouri-Columbia, USA), Kannappan Palaniappan (University of Missouri-Columbia, USA), and Filiz Bunyak (University of Missouri-Columbia, USA)

Interactive Rotated Object Detection for Novel Class Detection in Remotely Sensed Imagery 1129
Marvin Burges (TU Wien, Austria), Philipe Ambrozio Dias (Oak Ridge National Laboratory, USA), Dalton Lunga (Oak Ridge National Laboratory, USA), Carson Woody (Oak Ridge National Laboratory, USA), and Sarah Walters (Oak Ridge National Laboratory, USA)

EarthView: A Large Scale Remote Sensing Dataset for Self-Supervision 1138
Diego Velazquez (Computer Vision Center, ServiceNow Research), Pau Rodriguez (Apple Research), Sergio Alonso (Satellogic), Josep M. Gonfaus (Satellogic), Jordi Gonzalez (Computer Vision Center, Universitat Autonoma de Barcelona), Gerardo Richarte (Satellogic), Javier Marin (Satellogic Solutions), Yoshua Bengio (Mila - Quebec AI Institute), and Alexandre Lacoste (ServiceNow Research)

Semantic Neural Radiance Fields for Multi-Date Satellite Data 1148
Valentin Wagner (Fraunhofer IOSB), Sebastian Bullinger (Fraunhofer IOSB), Christoph Bodensteiner (Fraunhofer IOSB), and Michael Arens (Fraunhofer IOSB)

3rd Workshop on Computer Vision for Winter Sports (CV4WS)

SkipClick: Combining Quick Responses and Low-Level Features for Interactive Segmentation in Winter Sports Contexts 1157
Robin Schön (University of Augsburg, Germany), Julian Lorenz (University of Augsburg, Germany), Daniel Kienzle (University of Augsburg, Germany), and Rainer Lienhart (University of Augsburg, Germany)

Gate-Shift-Pose: Enhancing Action Recognition in Sports with Skeleton Information 1167
Edoardo Bianchi (Free University of Bozen-Bolzano) and Oswald Lanz (Free University of Bozen-Bolzano)

Towards long-term player tracking with graph hierarchies and domain-specific features 1175
Maria Koshkina (York University) and James H. Elder (York University)

CLaP - Contrast, Label, Predict: a quest for cheaper labeling in 3D human pose estimation 1186
Davide Cavicchini (University of Trento, Italy), Alessia Pivotto (University of Trento, Italy), Sofia Lorengo (University of Trento, Italy), Andrea Rosani (Free University of Bozen-Bolzano, Italy), and Nicola Garau (University of Trento, Italy)

VisionDocs: 1st Workshop on Computer Vision Systems for Document Analysis and Recognition (VisionDocs)

Mixed-Precision is All You Need for Efficient Document Image Classification 1195
Tushar Shinde (IIT Madras Zanzibar) and Shivam Bhardwaj (IIT Madras Zanzibar)

Improving the Identification of Layers in 3D Images of Ancient Papyrus using Artificial Neural Networks 1204
Nicolas Klenert (Zuse Institute Berlin, Germany), Finn Schwoerer (Zuse Institute Berlin, Germany), Noushin Hajarolasvadi (Zuse Institute Berlin, Germany), Siloé Bournez (Zuse Institute Berlin, Germany), Tobias Arlt (Helmholtz-Zentrum Berlin für Materialien und Energie, Germany), Heinz-Eberhard Mahnke (Helmholtz-Zentrum Berlin für Materialien und Energie, Germany), Verena Lepper (Ägyptisches Museum und Papyrussammlung, Germany), and Daniel Baum (Zuse Institute Berlin, Germany)

DocSum: Domain-Adaptive Pre-training for Document Abstractive Summarization 1213
Phan Phuong Mai Chau (University of Science and Technology of Hanoi), Souhail Bakkali (L3i-lab, La Rochelle Université), and Antoine Doucet (L3i-lab, La Rochelle Université)

Multi-Modal Large Language Model driven Augmented Reality Situated Visualization: the Case of Wine Recognition 1223
Vincenzo Armandi (University of Bologna), Andrea Loretti (University of Bologna), Lorenzo Stacchio (University of Macerata), Pasquale Cascarano (University of Bologna), and Gustavo Marfia (University of Bologna)

Low-Rank Adaptation vs. Fine-Tuning for Handwritten Text Recognition 1233
Lukas Hüttner (Pattern Recognition Lab, Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany), Martin Mayr (Pattern Recognition Lab, Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany), Thomas Gorges (Pattern Recognition Lab, Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany), Fei Wu (Pattern Recognition Lab, Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany), Mathias Seuret (Pattern Recognition Lab, Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany), Andreas Maier (Pattern Recognition Lab, Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany), and Vincent Christlein (Pattern Recognition Lab, Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany)

RAPTOR: Refined Approach for Product Table Object Recognition 1243
Eliott Thomas (Yooz, France), Mickael Coustaty (La Rochelle Université, France), Aurélie Joseph (Yooz, France), Gaspar Delion (Yooz, France), Elodie Carel (Yooz, France), Vincent Poulain D'andecy (Yooz, France), and Jean Marc Ogier (La Rochelle Université, France)

A Comparative Analysis of OCR Models on Diverse Datasets: Insights from Memes and Hiertext Dataset 1253
Iknoor Singh (University of Sheffield, UK), Miguel Colom (ENS Paris-Saclay, France), and Kalina Bontcheva (University of Sheffield, UK)

Offline Signature Verification in the Banking Domain 1264
Valentina Arrigoni (UniCredit, Italy)

2nd Physical Retail AI Workshop (PRAW)

WTPose: Waterfall Transformer for Multi-person Pose Estimation 1274
Navin Ranjan (Rochester Institute of Technology), Bruno Artacho (Rochester Institute of Technology), and Andreas Savakis (Rochester Institute of Technology)

What Matters when Building Vision Language Models for Product Image Analysis? 1282
Ameni Trabelsi (Amazon, USA), Maria Zontak (Amazon, USA), Yiming Qian (Amazon, Canada), Brian Jackson (Amazon, USA), Suleiman Khan (Amazon, USA), and Umit Batur (Amazon, USA)

Vision-Language Models (VLMs) in Open Traffic Scene Understanding

Synthetic Realities and Data in Biometric Analysis and Security (SynRDinBAS)

SIGN-GAIL: Rewarding Online Signature Generation for Digital Imitation 1292
Anurag Pandey (Ph.D Research Scholar), Arnav Bhavsar Vinayak (Associate Professor , IIT Mandi), Aditya Nigam (Associate Professor IIT Mandi), Divya Acharya (Senior Technical Lead), Balaji Rao K (Senior Technical Lead at HCLTech), and Basu Verma (Technical Lead at HCLTech)

Exploring Correlated Facial Attributes in Text-to-Image Models: Unintended Consequences in Synthetic Face Generation 1302
Sander De Coninck (IDLab, Ghent University - imec), Sam Leroux (IDLab, Ghent University - imec), and Pieter Simoens (IDLab, Ghent University - imec)

On the generalisation capability of local surface frames in detecting diffusion-based facial images 1312
Andrea Ciamarra (CNIT, National Interuniversity Consortium for Telecommunications, Italy), Roberto Caldelli (CNIT, National Interuniversity Consortium for Telecommunications and Universitas Mercatorum, Italy), and Alberto Del Bimbo (University of Florence, Italy)

Generating Realistic Forehead-Creases for User Verification via Conditioned Piecewise Polynomial Curves 1322
Abhishek Tandon (Indian Institute of Technology Mandi, India), Geetanjali Sharma (Indian Institute of Technology Mandi, India), Gaurav Jaswal (Technology Innovation Hub - Indian Institute of Technology, Mandi), Aditya Nigam (Indian Institute of Technology Mandi, India), and Raghavendra Ramachandra (Norwegian University of Science and Technology (NTNU), Norway)

Face-swapping based Data Augmentation for ID Document and Selfie Face Verification 1331
Elmokhtar Mohamed Moussa (IDnow AI & ML Center of Excellence, France), Ioannis Sarridis (Centre for Research and Technology Hellas, Greece), Emmanouil Krasanakis (Centre for Research and Technology Hellas, Greece), Nathan Ramoly (IDnow AI & ML Center of Excellence, France), Symeon Papadopoulos (Centre for Research and Technology Hellas, Greece), Ahmad-Montaser Awal (IDnow AI & ML Center of Excellence, France), and Lara Younes (IDnow AI & ML Center of Excellence, France)

Unified Face Matching and Physical-Digital Spoofing Attack Detection 1339
Arun Kunwar (University of North Texas, USA) and Ajita Rattani (University of North Texas, USA)

Latent Diffusion Shield - Mitigating Malicious Use of Diffusion Models through Latent Space Adversarial Perturbations 1350
Huy Phan (Amazon, Rutgers University), Boshi Huang (Amazon), Ayush Jaiswal (Amazon), Ekraam Sabir (Amazon), Prateek Singhal (Amazon), and Bo Yuan (Rutgers University)

Parallel Prints: Generating Realistic Cancelable Fingerprint Templates 1359
Rishabh shukla (Indian Institute of Technology Jammu), Harkeerat Kaur (Indian Institute of Technology Jammu), and Isao Echizen (National Institute of Informatics, Japan)

TextureCrop: Enhancing Synthetic Image Detection through Texture-based Cropping 1369
Despina Konstantinidou (Center for Research and Technology Hellas), Christos Koutlis (Center for Research and Technology Hellas), and Symeon Papadopoulos (Center for Research and Technology Hellas)

Digi2Real: Bridging the Realism Gap in Synthetic Data Face Recognition via Foundation Models 1379
Anjith George (Idiap Research Institute) and Sebastien Marcel (Idiap Research Institute)

2nd Workshop on the Applications of Computational Imaging (WACI)

Fashionability-Enhancing Outfit Image Editing with Conditional Diffusion Models 1389
Qice Qin (Waseda University, Japan), Yuki Hirakawa (ZOZO Research, Japan), Ryotaro Shimizu (ZOZO Research, Japan), Takuya Furusawa (ZOZO Research, Japan), and Edgar Simo-Serra (Waseda University, Japan)

Segment Anything in Light Fields for Real-Time Applications via Constrained Prompting 1400
Nikolai Goncharov (University of Sydney) and Donald Dansereau (University of Sydney)

M-GAID: A Real-World Dataset for Ghosting Artifact Detection and Removal in Mobile Imaging. 1407
Muneeb A Khan (Sangmyung University, South Korea), Hyungsub Kim (Sangmyung University, South Korea), Jiho Eum (Sangmyung University, South Korea), Yihyun Myung (Sangmyung University, South Korea), Yujin Choi (Sangmyung University, South Korea), and Heemin Park (Sangmyung University, South Korea)

MaCVi 2025: 3rd Workshop on Maritime Computer Vision (MaCVi)

Automatic Fish Age Prediction using Deep Machine Learning: Combining Otolith Image, NIR Spectra and Metadata Features 1417
Aotian Zheng (University of Washington), Jenq-Neng Hwang (University of Washington), Yudong Liu (University of Washington), Qiancheng Li (University of Washington), Beverly Barnett (Southeast Fisheries Science Center), Farron Wallace (Southeast Fisheries Science Center), Irina Benson (Alaska Fisheries Science Center), and Thomas Helser (Alaska Fisheries Science Center)

Marine Event Vision: Harnessing Event Cameras For Robust Object Detection In Marine Scenarios 1425
Nenyi Kweku Nkensen Dadson (Louisiana State University, United States of America) and Corina Barbalata (Louisiana State University, United States of America)

Underwater Image Enhancement and Object Detection: Are Poor Object Detection Results On Enhanced Images Due to Missing Human Labels? 1435
Evan Lucas (Michigan Technological University), Ali Awad (Michigan Technological University), Anthony Geglio (Michigan Technological University), Shadi Moradi (Michigan Technological University), Ashraf Saleem (Michigan Technological University), Timothy Havens (Michigan Technological University), Angus Galloway (Engineering Technologies Canada Ltd.), and Sidike Paheding (Fairfield University)

Camera-based Intruder Detection and Monitoring of Ship Crew Work Hours 1441
Md Mahmuddun Nabi Murad (University of South Florida, USA), Bora San Turgut (Istanbul University, Turkey), Awwab Ahmed (University of South Florida, USA), Gokhan Camliyurt (Korea Maritime and Ocean University, South Korea), and Yasin Yilmaz (University of South Florida, USA)

Vessel registration number detection and recognition system 1450
Matej Fabijanic (University of Zagreb Faculty of Electrical Engineering and Computing), Fausto Ferreira (University of Zagreb Faculty of Electrical Engineering and Computing), Maja Magdalenic (University of Zagreb Faculty of Electrical Engineering and Computing), Juraj Obradovic (University of Zagreb Faculty of Electrical Engineering and Computing), Nadir Kapetanovic (University of Zagreb Faculty of Electrical Engineering and Computing), and Nikola Miskovic (University of Zagreb Faculty of Electrical Engineering and Computing)

3rd Workshop on Maritime Computer Vision (MaCVi) 2025: Challenge Results 1457
Benjamin Kiefer (Universität Tübingen), Lojze Zust (University of Ljubljana), Matej Kristan (University of Ljubljana), Janez Pers (University of Ljubljana), Matija Tersek (Luxonis), Uma Mudenagudi (KLE Technological University), Chaitra Desai (KLE Technological University), Arnold Wiliem (Shield AI), Marten Kreis (Universität Tübingen), Nikhil Akalwadi (KLE Technological University), Zhiqiang Zhong (Nanjing University of Science and Technology), Zhe Zhang (Nanjing University of Science and Technology), Sujie Liu (Nanjing University of Science and Technology), Xuran Chen (Nanjing University of Science and Technology), Yang Yang (Nanjing University of Science and Technology), Matej Fabijanic, Fausto Ferreira (University of Zagreb), Seongju Lee (Gwangju Institute of Science & Technology), Shanliang Yao (University of Liverpool), Himanshu Kumar (Invertis University), Aurelius Marcus (Beijing University of Posts and Telecommunications), Gregor Novak (Univerzi v Ljubljani), Yuan Feng (Dalian Maritime University), Annie Cheng (ACVLab, National Cheng Kung University), Thien Nguyen (University of Information Technology), Jannik Sheikh (Fraunhofer), Josip Saric (University of Ljubljana), Zhuoxiao Li (Institute of Automation), Yutang Lu (N/A), Yipeng Lin (N/A), Xiang Yang (N/A), Ching-Heng Cheng (N/A), Ali Awad (N/A), Jon Muhovič (University of Ljubljana), Yitong Quan (University of Tuebingen), Junseok Lee (Gwangju Institute of Science and Technology (GIST)), Kyoobin Lee (Gwangju Institute of Science and Technology (GIST)), Runwei Guan (University of Liverpool), Xiaoyu Huang (University of Liverpool), Yi Ni (Xi'an Jiaotong-Liverpool University), Tzu-Yu Lin (National Cheng Kung University), Chia-Ming Lee (National Cheng Kung University), Chih-Chung Hsu (National Cheng Kung University), Andreas Michel (Fraunhofer), Wolfgang Gross (Fraunhofer), Nan Jiang (Nanjing University), Fei Feng (Beijing University of Posts and Telecommunications), Evan Lucas (Michigan Technological University), Ashraf Saleem (Michigan Technological University), Yu-Fan Lin (National Cheng Kung University), and Martin Weinmann (Karlsruhe Institute of Technology)

A Framework for Imbalanced SAR Ship Classification: Curriculum Learning, Weighted Loss Functions, and a Novel Evaluation Metric 1485
Ch Muhammad Awais (University of Pisa & ISTI-CNR, Italy), Marco Reggiannini (National Biodiversity Future Center & ISTI-CNR, Italy), and Davide Moroni (ISTI-CNR, Italy)

MTReD: 3D Reconstruction Dataset for Fly-over Videos of Maritime Domain 1494
Rui Yi Yong (Shield AI), Samuel Picosson (Shield AI), and Arnold Wiliem (Shield AI)

In-domain self-supervised learning for plankton image classification on a budget 1503
Massimiliano Ciranni (University of Genoa, Italy), Ani Gjergji (University of Genoa, Italy), Andrea Maracani (Istituto Italiano di Tecnologia, Italy), Vittorio Murino (Istituto Italiano di Tecnologia, Italy), and Vito Paolo Pastore (University of Genoa, Italy)

AutoFish: Dataset and Benchmark for Fine-grained Analysis of Fish 1513
Stefan Hein Bengtson (Aalborg University, Denmark), Daniel Lehotský (Aalborg University, Denmark), Vasiliki Ismiroglou (Aalborg University, Denmark), Niels Madsen (Aalborg University, Denmark), Thomas B. Moeslund (Aalborg University, Denmark), and Malte Pedersen (Aalborg University, Denmark)

Navigating Coreset Selection and Model Compression for Efficient Maritime Image Classification 1523
Tushar Shinde (IIT Madras Zanzibar), Avinash Kumar Sharma (IIT Madras Zanzibar), Shivam Bhardwaj (IIT Madras Zanzibar), and Ahmed Silima Vuai (IIT Madras Zanzibar)

FalconEye: Efficient Centroid-Based Object Detection for Maritime High Altitude UAV Images on Embedded Devices 1532
Rohan Sawahn (Hasso Plattner Institute and SearchWing e.V.) and Maximilian Schall (Hasso Plattner Institute)

Workshop on Manipulation, Generative, Adversarial, and Presentation Attack in Biometrics (MAPA)

Extracting Local Information from Global Representations for Interpretable Deepfake Detection 1544
Elahe Soltandoost (University of Padova, Italy), Richard Plesh (Clarkson University, USA), Stephanie Schuckers (University of North Carolina Charlotte, USA), Peter Peer (University of Ljubljana, Slovenia), and Vitomir Štruc (University of Ljubljana, Slovenia)

Transferable Adversarial Attacks on Audio Deepfake Detection 1555
Muhammad Umar Farooq (University of Michigan-Flint, USA), Awais Khan (University of Michigan-Flint, USA), Kutub Uddin (University of Michigan-Flint, USA), and Khalid Mahmood Malik (University of Michigan-Flint, USA)

MADation: Face Morphing Attack Detection with Foundation Models 1565
Eduarda Caldeira (Fraunhofer IGD, Germany), Guray Ozgur (Fraunhofer IGD, Germany), Tahar Chettaoui (Fraunhofer IGD, Germany), Marija Ivanovska (University of Ljubljana, Slovenia), Peter Peer (University of Ljubljana, Slovenia), Fadi Boutros (Fraunhofer IGD, Germany), Vitomir Struc (University of Ljubljana, Slovenia), and Naser Damer (Fraunhofer IGD and TU Darmstadt, Germany)

Wavelet-Driven Generalizable Framework for Deepfake Face Forgery Detection 1576
Lalith Bharadwaj Baru (International Institute of Information Technology Hyderabad, India), Rohit Boddeda (International Institute of Information Technology Hyderabad, India), Shilhora Akshay Patel (Indian Institute of Technology Hyderabad, India), and Sai Mohan Gajapaka (Michigan State University, USA)

Metric for Evaluating Performance of Reference-Free Demorphing Methods 1585
Nitish Shukla (Michigan State University) and Arun Ross (MSU)

Face detection and recognition under real-world scenarios - dealing with deepfake incidents and malicious data distortions 1592
Ewelina Bartuzi-Trokielewicz (NASK - National Research Institute), Alicja Martinek (NASK - National Research Institute, AGH University of Kraków), and Adrian Kordas (NASK - National Research Institute)

Exploring ChatGPT for Face Presentation Attack Detection in Zero and Few-Shot in-Context Learning 1602
Alain Komaty (Idiap Research Institute), Hatef Otroshi Shahreza (Idiap Research Institute), Anjith George (Idiap Research Institute), and Sébastien Marcel (Idiap Research Institute)

Author Index