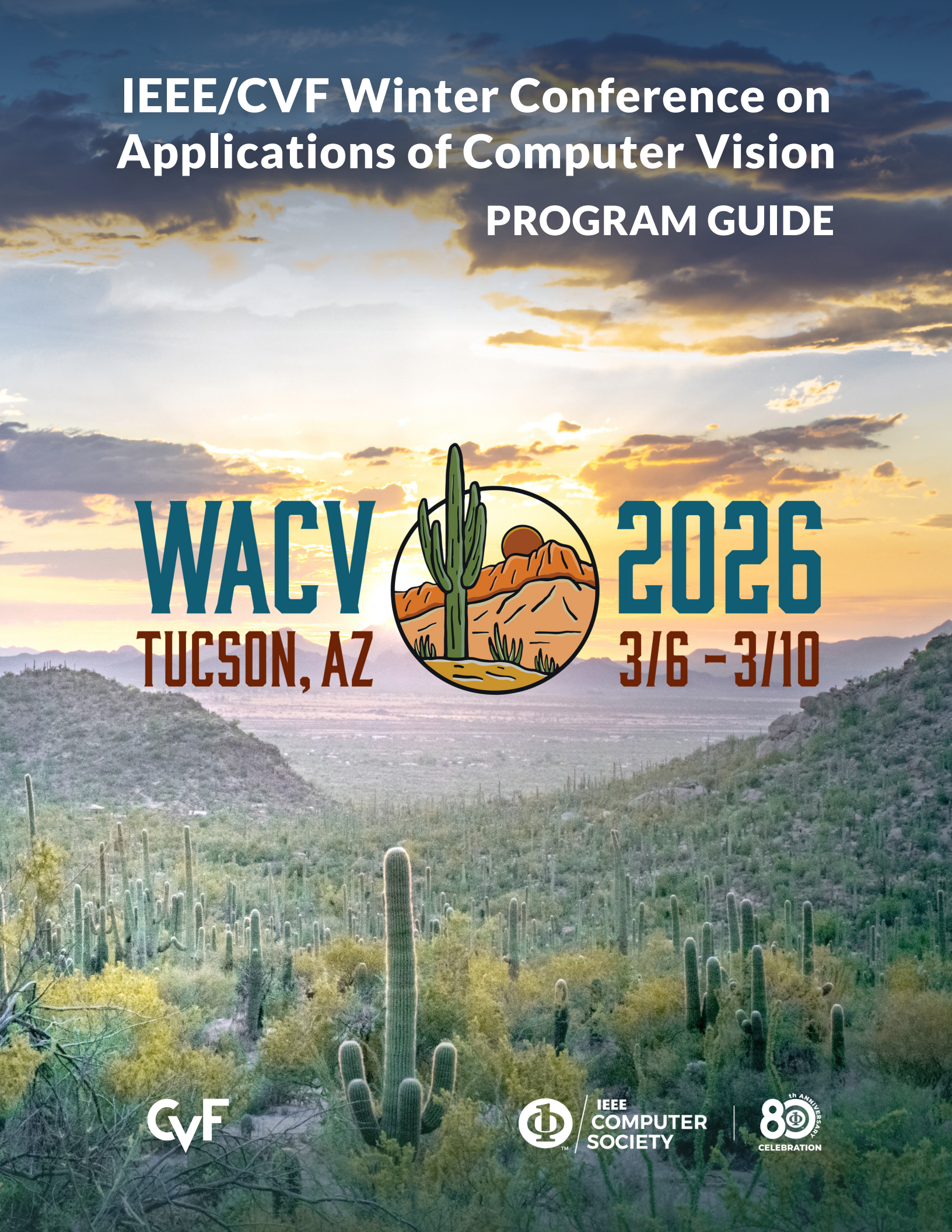
IEEE/CVF Winter Conference on Applications of Computer Vision

PROGRAM GUIDE



WACV 2026 At-A-Glance

Please check the workshop/tutorial websites for the full, up-to-date schedule and program.

Friday • March 6

REGISTRATION (AZ Ballroom Registration Desk) 7:30 - 17:00

POSTER PICKUP (Tucson Ballroom Registration Desk) | Poster sessions for all workshops (based on workshop schedule) will be in the AZ Ballroom Pre-Function space.

MORNING COFFEE BREAK 9:30 - 10:15 — Coffee breaks will be in the AZ Ballroom Pre-Function space. Breakfast and Lunch are not provided.

AFTERNOON COFFEE BREAK 15:00 - 15:45

- **T:** Machine Unlearning, Privacy, and AI Governance: Exploring Connections, Understanding Limitations, and Interrogating Policy Assumptions (AZ Ballroom Salon 10-11)
- **W:** Pixels to Patients: Bridging CV State-of-Art with Clinical Impact (P2P-CV) (AZ Ballroom Salon 1)
- **W:** Real-World Surveillance: Applications and Challenges, 6th (RWS) (AZ Ballroom Salon 2)
- **W:** International Workshop on Smart Waste Monitoring (WasteVision) (AZ Ballroom Salon 3-4)
- **W:** VisionDocs: 3rd Workshop on Computer Vision Systems for Document Analysis and Recognition (VisionDocs) (AZ Ballroom Salon 3-4)
- **W:** 4th Workshop on Computer Vision for Winter Sports (CV4WS) (AZ Ballroom Salon 5)
- **W:** The Second Workshop on Computer Vision for Geospatial Image Analysis (GeoCV) (AZ Ballroom Salon 5)
- **W:** Workshop on Large Foundation Models in Biology and Biomedicine (LFMBio) (AZ Ballroom Salon 6)
- **W:** HARVEST-Vision: International Workshop on Applications of CV and HPC in Agriculture (HARVEST-Vision) (AZ Ballroom Salon 7)
- **W:** EVGEN - Event-based Vision in the Era of Generative AI - Transforming Perception and Visual Innovation Summary (EVGEN) (AZ Ballroom Salon 7)
- **W:** Synthetic Realities and Data in Biometric Analysis and Security (SynRDinBAS) (AZ Ballroom Salon 8)
- **W:** SAFE 2026 – Synthetic & Adversarial ForEnsics (SAFE 2026) (AZ Ballroom Salon 9)

Saturday • March 7

REGISTRATION (AZ Ballroom Registration Desk)

POSTER PICKUP (Tucson Ballroom Registration Desk) | Poster sessions for all workshops (based on workshop schedule) will be in the AZ Ballroom Pre-Function space.

MORNING COFFEE BREAK 9:30 - 10:15 — Coffee breaks will be in the AZ Ballroom Pre-Function space. Breakfast and Lunch are not provided.

AFTERNOON COFFEE BREAK 15:00 - 15:45

- **T:** Beyond Vision: Multimodal Perspectives for Cross-View Geo-Localization (AZ Ballroom Salon 9)
- **W:** 5th Workshop on Image/Video/Audio Quality Assessment in Computer Vision, VLM and Diffusion Model (WVAQ) (AZ Ballroom Salon 1)
- **W:** LENS: Learning and Exploitation of Latent Space Geometries (LENS 2026) (AZ Ballroom Salon 2)
- **W:** WACV-2026 Workshop On Generative, Adversarial, Manipulation and Presentation Attacks In Biometrics (GAPBio) (AZ Ballroom Salon 3-4)
- **W:** WACV 2026 Workshop Proposal Scene Graph for Structured Intelligence (SG4SI) (AZ Ballroom Salon 3-4)
- **W:** 3rd Physical Retail AI Workshop (PRAW) (AZ Ballroom Salon 5)
- **W:** Workshop on Generative AI for Photography (GAIP) (AZ Ballroom Salon 5)
- **W:** Foundational Models Beyond the Visual Spectrum (FoMoV) (AZ Ballroom Salon 6)
- **W:** Large Language and Vision Models for Autonomous Driving (LLVM-AD) (AZ Ballroom Salon 6)
- **W:** Visual Art, Generative AI, and the Legal/Ethical Dilemma (VALED) (AZ Ballroom Salon 7)
- **W:** VReID-XFD: Video-based Human Recognition at Extreme Far Distances (VReID-XFD) (AZ Ballroom Salon 8)
- **W:** Robust and Generalized Lane Topology Understanding and HD Map Generation through CoT Design (TopoCoT) (AZ Ballroom Salon 8)
- **W:** 3rd Workshop on Computer Vision for Earth Observation Applications (CV4EO) (AZ Ballroom Salon 10-11)

Sunday • March 8

REGISTRATION (AZ Ballroom Registration Desk)

POSTER PICKUP (Tucson Ballroom Registration Desk)

EXHIBITS + DEMOS Tucson Ballroom 11:15 - 17:45

OPENING REMARKS & PAPER AWARDS
AZ Ballroom Salon 6 (Streamed to AZ Ballroom Salon 7)

KEYNOTE 1:
SPARSE VIEW SYNTHESIS
RAVI RAMAMOORTHI
AZ Ballroom Salon 6 (Streamed to AZ Ballroom Salon 7)

COURTESY BREAK 10:00-10:15

ORALS 1A: GENERATIVE MODELS I
AZ Ballroom Salon 6
10:15 - 11:15

ORALS 1B: 3D COMPUTER VISION I
AZ Ballroom Salon 7
10:15 - 11:15

POSTER SESSION 1
Tucson Ballroom & Pre-Function Space
11:15 - 13:00

LUNCH
Ania Terrace and Lawn
12:00 - 13:30

PANEL DISCUSSION
AZ Ballroom Salon 6
12:30 - 13:30

ORALS 2A: VISION + LANGUAGE AND OTHER MODALITIES I
AZ Ballroom Salon 6
13:45 - 14:45

ORALS 2B: BIOMETRICS, FACE, GESTURE, AND BODY POSE I
AZ Ballroom Salon 7
13:45 - 14:45

COURTESY BREAK 14:45 -15:00

ORALS 3A: LOW-LEVEL AND PHYSICS-BASED VISION
AZ Ballroom Salon 6
15:00 -16:00

ORALS 3B: IMAGE RECOGNITION AND UNDERSTANDING I
AZ Ballroom Salon 7
15:00 -16:00

POSTER SESSION 2
REFRESHMENTS
Tucson Ballroom & Pre-Function Space
16:00 - 17:45

Monday • March 9

REGISTRATION (AZ Ballroom Registration Desk)

POSTER PICKUP (Tucson Ballroom Registration Desk)

EXHIBITS + DEMOS Tucson Ballroom 10:45 - 18:00

KEYNOTE 2:
APPLICATIONS OF COMPUTER VISION IN HEALTHCARE: THE ROAD TO AUTONOMY
DORIN COMANICIU
AZ Ballroom Salon 6 (Streamed to AZ Ballroom Salon 7)

COURTESY BREAK 9:30 - 9:45

Orals 4A: Image Recognition and Understanding II
AZ Ballroom Salon 6
9:45 - 10:45

Orals 4B: Machine Learning I
AZ Ballroom Salon 7
9:45 - 10:45

POSTER SESSION 3
Tucson Ballroom & Pre-Function Space
10:45 - 12:30

LUNCH
Ania Terrace and Lawn
12:00 - 13:30

DOCTORAL CONSORTIUM
BY INVITATION ONLY
San Luis
12:00 - 14:00

Orals 5A: Generative Models II
AZ Ballroom Salon 6
13:30 - 14:30

Orals 5B: Remote Sensing and Sensors
AZ Ballroom Salon 7
13:30 - 14:30

COURTESY BREAK 14:30 - 14:45

ORALS 6A: 3D COMPUTER VISION II
AZ Ballroom Salon 6
14:45 -15:45

ORALS 6B: VIDEO RECOGNITION AND UNDERSTANDING I
AZ Ballroom Salon 7
14:45-15:45

PAMI TECHNICAL COMMITTEE MEETING
AZ Ballroom Salon 6
15:45 - 16:30

POSTER SESSION 4
RECEPTION
Tucson Ballroom & Pre-Function Space
16:30 - 18:15

Tuesday • March 10

REGISTRATION AND POSTER PICKUP (AZ Ballroom Registration Desk) No breakfast or lunch provided.

POSTER PICKUP (Tucson Ballroom Registration Desk)

EXHIBITS Tucson Ballroom 10:45 - 14:00

DEMOS Tucson Ballroom 10:45 - 17:30

KEYNOTE 3:
A SHORT HISTORY OF VIDEO UNDERSTANDING – PAST, PRESENT, AND FUTURE
HILDE KUHNE
AZ Ballroom Salon 6 (Streamed to AZ Ballroom Salon 7)

COURTESY BREAK 9:30 - 9:45

Orals 7A: Biometrics, Face, Gesture, and Body Pose II
AZ Ballroom Salon 6
9:45 - 10:45

Orals 7B: Vision + Language and Other Modalities II
AZ Ballroom Salon 7
9:45 - 10:45

POSTER SESSION 5
Tucson Ballroom
10:45 - 12:15

LUNCH
Ania Terrace and Lawn
12:00 - 13:30

ORALS 8A: BIOMEDICAL, HEALTHCARE, AND MEDICINE
AZ Ballroom Salon 6
13:30 - 14:30

ORALS 8B: VIDEO RECOGNITION AND UNDERSTANDING II
AZ Ballroom Salon 7
13:30 - 14:30

COURTESY BREAK 14:30 - 14:45

ORALS 9A: GENERATIVE MODELS III
AZ Ballroom Salon 6
14:45 -15:45

ORALS 9B: MACHINE LEARNING II
AZ Ballroom Salon 7
14:45 -15:45

POSTER SESSION 6
REFRESHMENTS
Tucson Ballroom & Pre-Function Space
15:45 - 17:30

GOLD SPONSORS

BRONZE SPONSORS

WELCOME MESSAGE WACV 2026

Welcome to Tucson, Arizona, the home of the Winter Conference on Applications of Computer Vision (WACV) for the second year! As the premier venue for advances in applications of computer vision, our meeting continues to grow along with the interest in AI and computer vision. We hope that you will enjoy the meeting and find time to enjoy Tucson.

WACV 2026 consists of a three-day main conference program and two days of workshops and tutorials. Following the WACV tradition, each day of the main conference includes a keynote presentation, followed by oral and poster presentations. There will also be ample opportunity to make new acquaintances and socialize with old colleagues during several social events that are part of the program of WACV this year.

We continue to attract a significant number of submissions, reaching a record 2,550 valid submissions–around 4% more than WACV 2025. To handle this number of submissions, we relied on 278 Area Chairs and 3,317 reviewers. We are especially grateful to the reviewers who were willing to conduct last-minute emergency reviews, ensuring that all papers received at least 3 reviews. The reviewing process was double-blind and followed best practices to avoid conflicts of interest. After receiving the reviews, the Area Chairs made decisions in groups of two, so that each paper was considered by multiple reviewers and multiple Area Chairs.

We continued the WACV tradition of a two-round submission process. By the Round 1 deadline in mid-July, we received 1,372 paper submissions. Of these, 84 were accepted directly after the first round of review, and an additional 736 were invited to be revised and resubmitted to Round 2, along with a rebuttal. Of the revised papers, 439 were accepted, whereas the rest were either rejected or withdrawn. The overall acceptance rate for Round 1 was therefore 38%. Authors could also submit new papers directly to Round 2, which had a deadline in mid-September. We received 1,178 new submissions in Round 2, of which 309 (26%) were accepted. Overall, across both rounds WACV received a total of 2,550 paper submissions and accepted a total of 832 papers, for an overall acceptance rate of 33%. These statistics illustrate how WACV's innovative two-round submission process allows authors to incorporate reviewer feedback to improve their papers (in Round 1), leading to a higher acceptance rate, as well as receiving quick feedback for Round 2.

The paper review process had two tracks: Applications and Algorithms. While papers in both tracks were expected to include original, novel, high-quality contributions, papers in the Applications track did not necessarily require algorithmic novelty. About 57% of papers were submitted to the Algorithms track, and 43% were submitted to the Applications track.

We gratefully acknowledge the contributions of our corporate supporters: Adobe, Encord, Kitware, Voxel51, and Edmund Optics Worldwide. We would also like to thank everyone involved in making WACV a success, including Tutorials Chairs Marta Gomez Barero and Vishal Patel, Workshop Chairs Scott M. McCloskey and Sarah Ostadabbas, Industry/Government Relations Chair Katarina Doctor, Ombudsperson Ryan Farrell, Doctoral Consortium Chairs Ehsan Azimi and Aparna Bharati, Demo and Exhibits Chairs Joel Brogan and Emanuela Marasco, Participation Broadening Chairs Soma Biswas and "YZ" Yezhou Yang, Publication Chair Srirangaraj Setlur, Publicity Chair Abby Stylianou, Finance Chair Raghavendra Ramachandra, Website Developer Lee Campbell, Technical Chair Yoshitomo Matsubara, and Awards Committee members Richard Souvenir (chair), Terry Boult, Yanxi Liu, Scott McCloskey, and Vishal Patel. We are grateful to our three keynote speakers, Ravi Ramamoorthi, Dorin Comaniciu, and Hilde Kühne. Special thanks go to our Conference Producer Nicole Finn, who has kept WACV running smoothly for many years. Finally, thanks to all of you for attending WACV and making it a success!

As we reflect on the many contributions that made WACV 2026 possible, we also wish to acknowledge that we began this journey with three General Chairs, including Margrit Betke, whose guidance helped shape the conference from the very start. Sadly, Margrit passed away during the organization effort. She was an exceptional scholar and mentor whose influential contributions to computer vision and its real-world applications will continue to inspire our community. We will miss her greatly, and we dedicate WACV 2026 in part to her memory.

Program Chairs:
Vlad Morariu, Brian Clipp, Naser Damer, Sharon X. Huang

General Chairs:
Vitomir Štruc, Rama Chellappa

WACV 2026 ORGANIZING COMMITTEE

General Chairs
Rama Chellappa (Johns Hopkins)
Vitomir Štruc (University of Ljubljana)

Program Chairs
Brian Clipp (Motional)
Naser Damer (TU Darmstadt/Fraunhofer IGD)
Sharon X. Huang (Penn State University)
Vlad Morariu (Adobe)

Website
Lee Campbell (WACV)

Tutorial Chairs
Marta Gomez Barero (Universität der Bundeswehr München)
Vishal Patel (Johns Hopkins University)

Workshop Chairs
Scott P. McCloskey (Kitware)
Sarah Ostadabbas (Northeastern University)

Industry/Government Relations Chair
Katarina Doctor (US Naval Research Laboratory)

Ombudsperson
Ryan Farrell (Brigham Young University)

Doctoral Consortium Chairs
Ehsan Azimi (University of Arizona)
Aparna Bharati (Lehigh University)

Demo/Exhibition Chairs
Joel Brogan (Oak Ridge National Laboratory)
Emanuela Marasco (Virginia Commonwealth University)

Participation Broadening Chairs
Soma Biswas (Indian Institute of Science)
"YZ" Yezhou Yang (Arizona State University)

Finance Chair
Raghavendra Ramachandra (Norwegian University of Science and Technology)

Publications Chair
Srirangaraj Setlur (State University of New York at Buffalo)

Publicity Chair
Abby Stylianou (Saint Louis University)

Technical Chair
Yoshitomo Matsubara (Yahoo!)

WACV 2026 AREA CHAIRS

Adam Czajka
Adam W Harley
Akihiro Sugimoto
Akshay Gadi Patil
Aleksei Tiulpin
Ales Leonardis
Alex Wong
Alimoor Reza
Amit Raj
Amit Roy-Chowdhury
Anh Tuan Tran
Anna Kukleva
Anthony Hoogs
Antonino Furnari
Aysegul Dundar
Behzad Bozorgtabar
Benjamin Busam
Benjamin Kimia
Benjamin Planche
Bing Su
Bingyi Kang
Bjorn Stenger
Bo Liu
Boqing Gong
Bo Wang
Bryan A. Plummer
Burhaneddin Yaman
Carlo Masone
Chaim Baskin
Chang Xu
Chaohui Wang
Chao Ma
Chen Chen
Chengcheng Tang
Cheng Deng
Chenyun Wu
Chih-Yuan Yang
Chris Thomas
Christopher Funk
Concetto Spampinato
Congcong Wen
C.V. Jawahar
Danna Gurari
Davide Moltisanti
David J. Crandall
Despoina Paschalidou
Di Huang
Dongdong Chen
Dong Gong
Du Tran
Dwarikanath Mahapatra
Ehsan Adeli
Emily Morgan Hand
Emre Akbas
Enrique Dunn
Fadi Boutros
Fatih Porikli
Federica Bogo
Feng Lu
Feng Yang
Fisher Yu
Francis Engelmann
Frederic Jurie
Fumio Okura
Georgios Pavlakos
Gianni Franchi
Gim Hee Lee
Gunhee Kim
Guorong Li
Gyeongsik Moon
Hao Huang
Haomiao Ni
Haoxiang Li
Heng Fan
Hengshuang Zhao
He Zhang
Hiroshi Kawasaki
Hoàng-Ân Lê
Hongbin Zha
Hongyang Li
Huan Fu
Hua Wei
Hyung Jin Chang
Hyungtae Lee
Hyunwoo J. Kim
Ilan Shimshoni
Ilke Demir
In Kyu Park
Javier Vazquez-Corral
Jiang Wang
Jiangxin Dong
Jiankang Deng
Jian Liang
Jianmin Bao
Jian Sun
Jianwen Xie
Jiasen Lu
Jiawen Chen
Jiaxing Huang
Jingjing Li
Jingkuan Song
Jingya Wang

WACV 2026 AREA CHAIRS

Jinshan Pan
Jinwei Ye
Jiuxiang Gu
John Collomosse
Jonathan Huang
Jonghyun Choi
Jongwon Choi
Jongwoo Lim
Joshua Harguess
Jungseock Joo
Junhwa Hur
Junseok Kwon
Junzhou Huang
Jyh-Jing Hwang
Kaifeng Chen
Kalle Åström
Kangfu Mei
Ko Nishino
Kuk-Jin Yoon
Kuldeep Purohit
Kushal Kafle
Kwang In Kim
Kwang Moo Yi
Laszlo A. Jeni
Le Lu
Le Wang
Liangyan Gui
Lianli Gao
Lin Gu
Lingxi Xie
Lorenzo Baraldi
Lu Qi
Maneesh Kumar Singh
Manmohan Chandraker
Marcella Cornia
Maria Vakalopoulou
Marius Leordeanu
Martin Weinmann
Matthew Joseph Leotta
Mayank Vatsa
Mei Han
Mengmeng Wang
Meng Tang
Mennatullah Siam
Miaomiao Liu
Michal Yarom
Ming-Hsuan Yang
Ming Li
Mingxing Tan
Mingyu Ding
Mohsen Ali
Mona Jalal
Mooi Choo Chuah
M. Saquib Sarfraz
Muhammad Haris Khan
Nalini K. Ratha
Nan Xue
Ngoc-Son Vu
Nianyi Li
Ning Xu
Norimichi Ukita
Ondrej Chum
Pavan K. Turaga
Philippos Mordohai
Philipp Terhörst
Pinar Yanardag
Pritam Sarkar
Qiang Qiu
Qianru Sun
Qi Bi
Qi Dai
Qi Mao
Radu Timofte
Raghavendra Ramachandra
Raja Giryes
Ran He
Raoul de Charette
Richa Singh
Roni Sengupta
Ruigang Yang
Ruiping Wang
Rui Yu
Ryan Farrell
Saeed Anwar
Salman Asif
Sandipan Banerjee
Sathyanarayanan N. Aakur
Scott McCloskey
Scott Workman
Shang-Hong Lai
Shayok Chakraborty
Shiliang Zhang
Sibei Yang
Simone Bianco
Siyu Zhu
SouYoung Jin
Srikrishna Karanam
Srinath Sridhar
Stephen Gould
Suha Kwak
Sunghyun Cho
Tae Hyun Kim
Tae-Kyun Kim
Taesup Moon
Takayuki Okatani
Tejas Gokhale
Tianhong Li
Vasileios Belagiannis
Venkatesh Babu Radhakrishnan
Vicente Ordonez
Viraj Shah
Wankou Yang
Waqas Sultani
Wei-Lun Chao
Wei Tang
Wei Yang
Wenbin Li
Wen Li
William Robson Schwartz
Wu Liu
Xiang Bai
Xiangtai Li
Xiangyu Xu
Xiaodan Liang
Xiaoqing Ye
Xilin Chen
Xilin CHEN
Xinggang Wang
Xin Ye
Xin Yu
Xinyu Zhang
Xi Peng
Xuequan Lu
Yadong MU
Yanbo Fan
Yang Liu
Yang Wang
Yan Huang
Yannick Hold-Geoffroy
Yanxi Liu
Yan Xu
Yiding Yang
Ying Tai
Yinjie Lei
Yin Li
Yogesh S Rawat
Yongkang Wong
Yosi Keller
Yuanjun Xiong
Yubei Chen
Yuchao Dai
Yu Cheng
Yu-Chiang Frank Wang
Yuexin Ma
Yuhui Yuan
Yujun Cai
Yuki M Asano
Yulan Guo
Yuliang Xiu
Yulun Zhang
Yuqian Zhou
Yusuke Matsui
Yusuke Sugano
Yu-Wing Tai
Yu-Xiong Wang
Zhanyu Ma
Zhaozheng Yin
Zhenfang Chen
Zhengzhong Tu
Zhen Lei
Zhenxing Niu
Zhipeng Zhang
Zhizhong Han
Zhongpai Gao
Zhouhui Lian
Zhuotao Tian

Abhinav Shrivastava
Aditya Nigam
Alejandro Fontan
Ana C Murillo
Andreas Maier
Andrei Bursuc
Andrew Markham
Aneesh Rangnekar
Anna Hilsmann
Balasubramanian Raman
Bernt Schiele
Carsten Steger
Chi Xie
Ching-Chun Huang
Christopher Wewer
Christos Sakaridis
Delin Qu
Dima Damen
Dung D. Le
Ender Konukoglu
Fabien Moutarde
Fabio Tosi
Gaetan Bahl
Gian Antonio Susto
Haidong Zhu
Hazel Doughty
He Zhao
Heming Du
Henghui Ding
Hirotaka Hachiya
Hiroto Honda
Hongbin Xu
Horst Possegger
Hyung-Jeong Yang
Ing Ren Tsang
J. Marius Zöllner
Jean Lahoud
Jia-Bin Huang
Juepeng Zheng
Jun-Jee Chao
Junmo Kim
Kai Han
Kailung Hua
Konstantinos Karantzalos
Kyungsu Kim
Lei Wang
Leonardo Iurada
Leonid Sigal
Liang Ding
Maks Ovsjanikov
Mathew Monfort
Matteo Matteucci
Matthew Stamm
Melissa Cote
Milad Khademi Nori
Min-Ling Zhang
Minkyu Choi
Mubarak Shah
Muhammad Abdullah Jamal
Nakul Agarwal
Nhat Chung
Nicolas Audebert
Nicolas Gonthier
Nicu Sebe
Nikolaos Zioulis
Peter Eisert
Qingdong He
Alex Colburn
Radu Grosu
Radu Marculescu
Reza Ghoddoosian
Ryota Yoshihashi
Sascha Hornauer
Seung-Wook Kim
Shaobo Wang
Xiaoyang Lyu
Shenghai Yuan
Simon Hadfield
Soufiane Belharbi
Soumik Sarkar
Sung Whan Yoon
Tatsuya Harada
Tianchen Deng
Tom Vercauteren
Tomasz Trzcinski
Tung Kieu
Vincent Lepetit
Wen-Huang Cheng
Xavier Giró-i-Nieto
Xingxing Zou
Xinqi Liu
Xu Tang
Yaxiong Chen
Yen-Yu Lin
Yezhou Yang
Yi Wang
Yiran Chen
Yizhou Wang
Yu-Lun Liu
Yuxiao Chen
Zequn Zhang
Zhan Ma
Zhenguo Li
Zhenzhen Liu
Zhiding Yu
Zhuowen Tu
Ziqiao Ma

Friday, March 6

Breakfast and Lunch are not provided.

7:30-17:00	**Registration**, AZ Ballroom Registration Desk
8:00-14:00	**Poster Pickup**, Tucson Ballroom Registration Desk
9:30-10:15	**Morning Coffee Break**, AZ Ballroom Pre-Function
15:00-15:45	**Afternoon Coffee Break**, AZ Ballroom Pre-Function
	Poster Sessions (based on workshop schedule), AZ Ballroom Pre-Function

WORKSHOPS

NOTE: **Use the QR code for each workshop's website for more information on that workshop (including the workshop's schedule). Here is the QR code to the WACV Workshops page.**

Workshop on Large Foundation Models in Biology and Biomedicine

Organizers: Suchendra M Bhandarkar, Shradha Agrawal, Kiran Raja
Time: Full Day
Location: Arizona Ballroom Salon 6
Summary: The rapid evolution of Large Foundation Models (LFMs) has transformed the landscape of biomedical research, clinical decision-making, and healthcare innovation. From decoding complex biological interactions to assisting in diagnosis and drug discovery, LFMs have demonstrated remarkable potential across a broad spectrum of biomedical applications. However, their adaptation to this highly specialized and sensitive domain presents unique challenges ranging from data scarcity and heterogeneity to issues of interpretability, fairness, and reproducibility. The Workshop on Large Foundation Models for Biology and Biomedicine (LFMBio 2026) aims to bring together researchers, practitioners, and industry experts to advance the science and practice of applying LFMs to biomedical problems. The workshop will serve as a forum for presenting original research, fostering interdisciplinary dialogue, and exploring cutting-edge innovations in model design, multimodal integration, trustworthiness, and ethical deployment. We invite contributions that span foundational model development, performance optimization, knowledge representation, real-world clinical applications, and the societal impact of these powerful technologies.

6th Real-World Surveillance: Applications and Challenges

Organizers: Andreas Specker, Mickael Cormier, Sergio Escalera Guerrero, Radu Ionescu, Fahad Shahbaz Khan, Kamal Nasrollahi
Time: Full Day
Location: Arizona Ballroom Salon 2
Summary: Real-World Surveillance: Applications and Challenges (RWS) is a full-day, 6th-edition WACV workshop that brings together researchers to close the gap between research and deployment in real-world surveillance settings.

SAFE 2026 – Synthetic & Adversarial ForEnsics

Organizers: Josue Martinez-Martinez, Ms. Danielle Sullivan, Dr. Giselle Zeno, Dr. Jonas Borgstrom, Dr. Pooya Khorrami, Dr. Sheila Alemany Blanco, Prof. David Harwath, Dr. Pedro Torres Carrasquillo
Time: Full Day
Location: Arizona Ballroom Salon 9
Summary: Modern machine learning systems must be robust against adversarial data to ensure reliable deployment. Simultaneously, there is a demand to develop authentication systems that can accurately identify synthetic content for the safety of its users. As generative AI and foundation models grow in complexity and become more accessible to the general public, malicious actors exploit model vulnerabilities to create realistic, malicious content, raising security and ethical concerns. This workshop bridges synthetic media forensics with adversarial robustness, focusing on provenance tracking, fingerprinting, and resilient detection methods considering real-time deployed systems requirements. It aims to foster collaboration across the AI, security, and policy domains, develop benchmarks and evaluation tools, and build a shared understanding of joint synthetic-adversarial threats to advance the deployment of trustworthy generative models and the verification of the authenticity of the media.

Pixels to Patients: Bridging CV State-of-Art with Clinical Impact

Organizers: Dr. Sahika Betul Yayli, Prof. Bradley J. Erickson
Time: Full Day
Location: Arizona Ballroom Salon 1
Summary: Pixels to Patients (P2P-CV) is a full-day workshop dedicated to bridging the gap between computer vision research and its safe, effective deployment in real-world clinical practice.

Medical imaging is among the most socially impactful domains of computer vision, yet a persistent gap remains between research breakthroughs and their safe, effective use in real-world clinics. WACV, with its strong tradition in applied computer vision, is the ideal venue to address this translation gap.

This workshop positions healthcare as a case study for broader CV challenges that arise whenever algorithms move from the lab to deployment. Topics such as domain generalization, fairness and bias mitigation, foundation models, trustworthy and explainable AI, and human–AI collaboration are central not only to medical imaging but also to autonomous driving, robotics, and other safety-critical applications.

4th Workshop on Computer Vision for Winter Sports

Organizers: Nicola Conci, Matteo Dunnhofer, Katja Ludwig, Rainer Lienhart, Hideki Koike, Christian Micheloni, Nicola Conci
Time: AM
Location: Arizona Ballroom Salon 5
Summary: The workshop invites paper submissions focusing on the analysis and interpretation of images and videos captured during winter sports and related summer activities such as mountain sports (e.g., mountaineering, downhill biking, climbing). Topics of interest for the workshop include machine learning solutions for video understanding, pose estimation of athletes, evaluation of athlete performance, forecasting performance, injury detection/prevention, crowd and spectator monitoring, augmented/virtual reality for fan engagement, and applications of computer vision/AI to various winter and mountain sports disciplines. Additionally, submissions can address challenges such as understanding images/videos in

harsh weather conditions, camera pose estimation in broadcast videos, trajectory reconstruction, winter scene reconstruction, snow/ice measurements, real-time processing algorithms, fusion of image/video data with other sensor data, and the creation of datasets and benchmarks.

HARVEST-Vision: International Workshop on Applications of CV and HPC in Agriculture

Organizers: Hari Subramoni, Scott Shearer
Time: AM
Location: Arizona Ballroom 7

Summary: By 2050, the global population is projected to reach 9.8 billion, with accelerating climate extremes and a halving of arable land per capita (vs. 1985), making food security and sustainability urgent.

HARVEST-Vision addresses this need by democratizing digital agriculture and translating end-to-end AI pipelines from the NSF ICICLE AI Institute community into accessible tools for:

- Small and large farms
- Agricultural research
- Staple and specialty crops
- Underrepresented communities

International Workshop on Smart Waste Monitoring (WasteVision)

Organizers: Antonio Greco, Bruno Vento, Carlo Sansone
Time: AM
Location: Arizona Ballroom 3-4

Summary: The International Workshop on Smart Waste Monitoring (WASTEVISION) aims to provide a dedicated forum for researchers, practitioners, and industry representatives to present and discuss innovative contributions in this emerging field. The workshop seeks to advance the state of the art in smart waste monitoring, illegal dumping detection, and environmental pollution surveillance, while promoting interdisciplinary collaboration and knowledge exchange.

EVGEN - Event-based Vision in the Era of Generative AI - Transforming Perception and Visual Innovation Summary

Organizers: Bharatesh Chakravarthi, Yezhou Yang, Cornelia M. Fermuller, Francisco Barranco, Federico Becattini, Aayush Atul Verma, Arpit Vaghela, Kaustav Chanda
Time: PM
Location: Arizona Ballroom 7

Summary: The 2nd Workshop on "Event-based Vision in the Era of Generative AI - Transforming Perception and Visual Innovation", held at WACV 2026, centers on the transformative impact of integrating event-based vision with generative AI. The workshop will explore how this synergy is reshaping visual perception, enabling advanced applications such as dynamic scene understanding, image and video generation, motion prediction, and conceptual reasoning for visual content. Additionally, topics such as multimodal fusion, gesture recognition, and applications in autonomous systems will be explored. By bringing together experts across event-based vision and AI, the workshop aims to highlight innovative approaches and inspire new research directions that push the boundaries of visual processing for enhanced perception. Our motivation is to initiate discussions in an emerging vertical that combines the strengths of event-based vision with the transformative capabilities of generative AI - a combination not fully explored previously.

VisionDocs: 3rd Workshop on Computer Vision Systems for Document Analysis and Recognition

Organizers: Axel De Nardin, Silvia Zottin, Silvia Cascianelli, Claudio Piciarelli, Gian Luca Foresti
Time: PM
Location: Arizona Ballroom 3-4

Summary: As digitalization accelerates, automated document analysis is becoming increasingly important in both industrial and cultural contexts. Modern documents are highly heterogeneous, varying in language, historical and geographical origin, visual appearance, writing styles, and layouts, which poses significant challenges for AI systems.

While recent advances in document analysis, particularly in computer vision, have achieved strong results, progress has largely focused on a narrow range of document types and tasks. Many areas remain under-explored, including low-resource languages, non-standard layouts, historical documents, generalization across diverse formats, and learning in low-data scenarios. Additionally, effectively integrating multimodal information such as text, visual, structural, and semantic cues remains an open challenge.

Document Analysis is inherently multidisciplinary, spanning diverse data sources, applications, and research communities. VisionDocs aims to foster cross-disciplinary collaboration and knowledge exchange to drive innovation and advance the field

The Second Workshop on Computer Vision for Geospatial Image Analysis

Organizers: Saurabh Prasad, Jocelyn Chanussot, Claudia Paris, Biplab Banerjee, Danfeng Hong
Time: PM
Location: Arizona Ballroom 5

Summary: GeoCV (the Second Workshop on Computer Vision for Geospatial Image Analysis) will convene researchers and practitioners working at the intersection of computer vision, machine learning, and geospatial imaging to address pressing societal challenges. The workshop will foster collaboration and cross-fertilization of ideas between academia and industry, highlighting both cutting-edge algorithms and high-impact applications. Topics will include geospatial foundation and large vision models; domain generalization and open-set/domain adaptation for out-of-distribution data; agentic AI for spatial reasoning; self-/weakly-/unsupervised learning; uncertainty quantification and explainable ML; multimodal perception (optical, SAR, LiDAR); and core tasks such as segmentation, detection, classification, and change/anomaly detection, with applications spanning ecological monitoring, precision agriculture, sustainable development, and disaster mapping. The program will feature peer-reviewed community contributions and keynote talks from leading researchers working in the field.

Synthetic Realities and Data in Biometric Analysis and Security

Organizers: Naser Damer; Vitomir Štruc, Fadi Boutros, Eduarda Caldeira, Laura Cassani, Marija Ivanovska, Vishal Patel, Ajita Rattani, Anderson Rocha, Matthew Stamm
Time: PM
Location: Arizona Ballroom Salon 8

Summary: Recent advancements in generative models like GANs, VAEs, and diffusion models have transformed data-driven tasks in computer vision and AI by enabling the creation of highly realistic synthetic data. These models address data scarcity challenges, offering versatile and ethical alternatives for training and testing machine learning algorithms. However, their realism also raises concerns, as synthetic data's indistinguishability from real data poses risks of misuse, manipulation, and potential harm when used unethically.

TUTORIALS

Machine Unlearning, Privacy, and AI Governance: Exploring Connections, Understanding Limitations, and Interrogating Policy Assumptions

Organizers: Kate Kaye
Time: Full Day
Location: Arizona Ballroom 10-11

Summary: Can computer vision and multimodal systems forget? Machine unlearning is often discussed in the context of privacy – particularly as a response to data removal requests in relation to Europe's General Data Protection Regulation's right to be forgotten. However, computer vision and computer science technologists rarely have the opportunity to engage directly with privacy and AI governance policy experts in Machine Unlearning (MU) discussions. This tutorial changes that! By bringing together researchers with MU technology expertise and others with privacy and AI governance policy expertise the tutorial aims to improve understanding between both groups. Expect presentations of cutting-edge MU approaches and active Q&A and discussion periods including about policy implications and limitations. Invited speakers include 'YZ' Yezhou Yang, associate professor, School of Computing and Augmented Intelligence at Arizona State University and Kairan Zhao, PhD candidate and teaching assistant in Machine Learning at the University of Warwick.

Saturday, March 7

Breakfast and Lunch are not provided.

8:00-17:00	**Registration**, AZ Ballroom Registration Desk
8:00-14:00	**Poster Pickup**, Tucson Ballroom Registration Desk
9:30-10:15	**Morning Coffee Break**, AZ Ballroom Pre-Function
15:00-15:45	**Afternoon Coffee Break**, AZ Ballroom Pre-Function
	Poster Sessions (based on workshop schedule), AZ Ballroom Pre-Function

WORKSHOPS

NOTE: **Use the QR code for each workshop's website for more information on that workshop (including the workshop's schedule). Here is the QR code to the WACV Workshops page.**

5th Workshop on Image/Video/Audio Quality Assessment in Computer Vision, VLM and Diffusion Model

Organizers: Joe Liu, Yarong Feng, Qipin Chen
Time: Full Day
Location: Arizona Ballroom Salon 1
Summary: Image, video, and audio quality critically impacts machine learning and computer vision systems, yet remains underexplored. Real-world applications—from autonomous vehicles to streaming services and generative AI—depend on robust quality assessment and enhancement techniques. While most visual learning systems assume high-quality inputs, artifacts from capture, compression, and transmission can severely degrade performance.

This workshop is particularly timely given the explosive growth of generative AI, which introduces new quality assessment challenges. Building on successful previous WACV workshops (2022-2025), we bring together industry and academic researchers to investigate quality impacts on visual learning tasks and develop innovative mitigation techniques. Topics include quality assessment methodologies, image/video enhancement, compression artifacts, multimodal applications, and evaluating generative AI outputs including hallucination effects in diffusion-based restoration.

LENS: Learning and Exploitation of Latent Space Geometries

Organizers: Anuj Srivastava, Sudeep Sarkar, Pavan Turaga
Time: Full Day
Location: Arizona Ballroom Salon 2
Summary: LENS brings together researchers studying the geometry of latent representations—their manifolds, Riemannian structures, intrinsic dimensions, and implications for model design and evaluation. We aim to bridge advances in geometric learning with practical computer vision applications, fostering dialogue between theory and deployments.

We welcome contributions that deepen our understanding of latent spaces (e.g., curvature, geodesics, topology), propose geometry-aware architectures and objectives, or demonstrate how latent geometry can improve robustness, generalization, fairness, privacy, and efficiency in real-world vision systems.

Visual Art, Generative AI, and the Legal/Ethical Dilemma

Organizers: Aparna Bharati, Mooi Choo Chuah, Qiuyu Tang
Time: PM
Location: Arizona Ballroom 7
Summary: Generative AI has transformed how visual art is created and circulated. Text-to-image generation systems such as Stable Diffusion, DALL·E, and Midjourney can instantly produce artworks inspired by centuries of human creativity. While these technologies democratize access to artistic tools, they also raise urgent questions about copyright, artistic integrity, and provenance. This workshop will bring together researchers, artists, legal scholars, and industry practitioners to critically examine the technical, legal, and societal challenges of visual art in the age of generative AI.

3rd Workshop on Computer Vision for Earth Observation (CV4EO) Applications

Organizers: Abby Stylianou, Philipe Dias, Dalton Lunga, Manil Maskey, Ronny Hänsch, Zhuo Zheng, Jiaqi Yang
Time: Full Day
Location: Arizona Ballroom 10-11
Summary: The 3rd Workshop on Computer Vision for Earth Observation (CV4EO) is conceived as a platform to foster application-oriented, multidisciplinary interactions between the computer vision community and experts from geoscience domains, across academia, non-profit organizations, Earth observation data providers, government agencies and other stakeholders leveraging AI, computer vision, and Earth observation technologies for decision-making in applications such as disaster response, national security, and environmental protection.

Foundational Models Beyond the Visual Spectrum

Organizers: Christopher Funk, Vishal M. Patel, Nathan Jacobs, Florence Yellin, Ritwik Gupta, Sean Hu
Time: AM
Location: Arizona Ballroom 6
Summary: The rapid rise of foundational models has transformed computer vision, but most progress has been confined to the visible spectrum. Many real-world applications in healthcare, maritime, biometrics, remote sensing, autonomous navigation, and defense rely on data modalities such as infrared, LIDAR, hyperspectral, depth, acoustic, event-cameras, RF, or radar, where foundational models remain underexplored. This workshop aims to bring together researchers working on extending and adapting foundational models beyond the visual spectrum, addressing challenges such as cross-modal pretraining, data scarcity, and domain adaptation. The motivation is to bridge the gap between visible-spectrum advances and broader multimodal sensing, which is both timely and relevant to the WACV community as it expands toward embodied AI and real-world deployment. The expected impact of the workshop is twofold: (i) to catalyze new research directions by highlighting the unique opportunities and challenges of non-visual modalities, and (ii) to foster collaborations across academia, industry, and government working in these critical areas. We anticipate outcomes including a clearer community roadmap, new benchmarks, and broader awareness of the importance of foundational models beyond the visual spectrum.

3rd Physical Retail AI Workshop (PRAW)

Organizers: Dr. Quanfu Fan, Dr. Shun Miao, Dr. Weijian Li, Dr. Sean Ma, Dr. David Woollard, Dr. Bruno Abbate, Dr. Davide Mazzini, Dr. Rocco Pietrini
Time: AM
Location: Arizona Ballroom Salon 5
Summary: In an evolving world in which consumers have a plurality of shopping methods available to them, physical ""brick and mortar"" stores continue to be the preferred means of shopping around the world. From groceries to clothing, customers continue to show strong demand for in person shopping.

Applications of vision-based Artificial Intelligence (AI) methods are increasingly present throughout society. Fueled by recent advances in Computer Vision, Deep Learning, web-scale training of vision and language models (""foundation models""), and edge compute, AI applications have expanded into a novel array of industries and products. In particular, the physical retail and grocery sectors have recently experienced an explosion of AI-enabled technologies, allowing for more efficient, effortless, and engaging experiences for shoppers, enabling the reduction of shrinkage for retailers, and providing insights on improving store efficiency, thereby reducing operational costs. Computer Vision applications are being deployed to numerous retail sectors, including small convenience stores, large grocery stores, fashion stores, and shopping carts, to name but a few."

VReID-XFD: Video-based Human Recognition at Extreme Far Distances

Organizers: Hugo Proença, Kailash A. Hambarde
Time: AM
Location: Arizona Ballroom Salon 8
Summary: VReID-XFD is a WACV 2026 workshop dedicated to video-based human recognition at extreme far distances, highlighting long-term re-identification with clothing changes, detection at-a-distance, and surveillance analytics.

The half-day programme pairs an archival paper track with the DetReIDX challenge, supplying baselines, hidden-label evaluation, and a forum for the community advancing robust UAV-driven recognition.

WACV-2026 Workshop On Generative, Adversarial, Manipulation and Presentation Attacks In Biometrics

Organizers: Kiran Raja, Naser Damer, Julian Fiérrez, Raghavendra Ramachandra
Time: AM
Location: Arizona Ballroom Salon 3-4
Summary: Recent advances in deep learning—particularly Generative Adversarial Networks (GANs), diffusion models, and Large Language Models (LLMs) have enabled the creation of highly realistic synthetic visual and textual content. While beneficial for applications such as entertainment and education, this realism poses serious threats to biometric security systems and the integrity of digital information. Image and video manipulation attacks, including morphing and DeepFakes, can bypass traditional biometric defenses, enable identity impersonation, and spread misinformation. LLM-generated text further escalates these risks through advanced phishing, social engineering, and disinformation campaigns. Modern attack strategies increasingly combine classical manipulation with adversarial machine learning techniques, prompting growing governmental investment in detection and mitigation research. This workshop, now in its seventh edition, focuses on recent advances in the creation, evaluation, impact, and mitigation of adversarial (soft and hard) attacks on biometric systems.

Workshop on Generative AI for Photography

Organizers: I-Sheng Fang, Daiqing Qi, Yu Yuan, Kuan-Chuan Peng, Jun-Cheng Chen
Time: PM
Location: Arizona Ballroom Salon 5
Summary: With the rapid breakthrough of Generative AI tools, it is timely and important to provide a platform to explore both the opportunities and challenges of these technologies. Our workshop on Generative AI for Photography brings together two rapidly evolving fields: cutting-edge AI and the creative art of visual storytelling. Generative AI tools are transforming how images can be created, enhanced, and manipulated, offering photographers new possibilities for creativity, efficiency, and accessibility. At the same time, they raise critical questions about authorship, authenticity, and ethical use. To tackle these challenges, researchers from both academia and industry must collaborate and make progress in relevant research and applied technologies. The organizing committee and keynote speakers of GAIP 2026 consist of experts from both academia and industry with rich experiences in designing and developing robust Generative AI tools and transferring them to real-world solutions. GAIP 2026 provides a focused venue to discuss and disseminate research related to Generative AI for photography, helping participants understand how AI can augment traditional practices, inspire new forms of expression, and shape the future of photography.

Robust and Generalized Lane Topology Understanding and HD Map Generation through CoT Design

Organizers: Dr. Zhen Li, Dr. Chao Zheng, Dr. Hongyang Li, Dr. Hao Zhao
Time: PM
Location: Arizona Ballroom Salon 8
Summary: The 1st WACV 2026 Workshop on Robust and Generalized Lane Topology Understanding and HD Map Generation through CoT Design (TopoCoT) seeks to provide a platform for industry experts and academics to brainstorm and exchange ideas about road understanding CoT and its derived outstanding works to advance autonomous driving. Through a half-day in-person event, the workshop will showcase regular and demo paper presentations and invited talks from famous researchers in academia and industry. Additionally, TopoCoT will launch two open-source real-world CoT lane topology reasoning datasets. The workshop will host a challenge based on this dataset to assess the capabilities of language and computer vision models in addressing HD map generation challenges.

WACV 2026 Workshop Proposal Scene Graph for Structured Intelligence

Organizers: Shengqiong Wu, Dennis Rotondi, Azade Farshad
Time: PM
Location: Arizona Ballroom Salon 3-4
Summary: Scene graphs provide a structured and interpretable representation of objects, attributes, and relationships in 2D, 3D, and even 4D scenes, serving as a vital bridge between raw visual data and high-level reasoning, which is critical for tasks such as visual reasoning, navigation, and embodied AI. With the rapid rise of multimodal foundation models, integrating scene graphs has become a timely and essential task, offering controllability, explainability, and stronger generalization across different domains and modalities.

This workshop will highlight the latest advances in scene graph generation, representation learning, and their applications in vision–language reasoning, multimodal generation, and robotics. We aim to establish new benchmarks, foster interdisciplinary collaboration, and

chart future directions toward the development of structured multimodal intelligence. By uniting researchers from computer vision, NLP, and robotics, the workshop will stimulate impactful discussions and accelerate progress toward trustworthy, general-purpose AI systems.

Large Language and Vision Models for Autonomous Driving

Organizers: Jiaru Zhang, SungYeon Park, Yifan Shen, Xiangbo Gao, Jiachen Li, Ziran Wang, Can Cui, Juntong Peng, Bohan Liu, Zhengzhong Tu, Ismini Lourentzou
Time: PM
Location: Arizona Ballroom 6
Summary: The 5th Workshop on Large Language and Vision Models for Autonomous Driving (LLVM-AD) at WACV 2026 aims to bring together professionals from academia and industry to explore the applications of large language and vision models in autonomous driving. We are particularly interested in bridging the gap between the rich image and language data found within the context of autonomous driving. Our primary areas of interest are: a) Traffic Scene Understanding enhanced by VLMs and b) Human-Autonomy Teaming driven by LLMs.

TUTORIAL

Beyond Vision: Multimodal Perspectives for Cross-View Geo-Localization

Organizers Chen Chen, Safwan Wshah, Xiaohan Zhang
Time: AM
Location: Arizona Ballroom 9
Summary: The increasing availability of geospatial data from heterogeneous modalities, including aerial and satellite imagery, ground-level views, and textual descriptions, has made cross-view geo-localization a critical research area with applications in autonomous navigation, urban monitoring, and augmented reality. Despite progress, challenges remain in handling extreme viewpoint variations, scaling across diverse domains, and integrating multimodal information. Recent developments in multimodal learning and Generative AI, such as Large Multimodal Models (LMMs), have introduced new paradigms for geo-localization. LMMs enable more generalized cross-view matching by incorporating language as an additional modality, supporting tasks such as text-based geo-localization, scene description, and multimodal reasoning. These capabilities not only improve performance but also expand the scope of cross-view geo-localization to broader multimodal applications. This tutorial will provide a comprehensive overview of these developments, highlighting the latest methodologies, datasets, and open research directions that are shaping the future of cross-view geo-localization.

Sunday, March 8

Reminder that all events are for registered attendees only.

8:00-17:00 **Registration,** AZ Ballroom Registration Desk
9:00-16:00 **Poster Pickup,** Tucson Ballroom Registration Desk
8:30-9:00 **Opening Remarks and Paper Awards,** AZ Ballroom 6 (streamed to 7)
9:00-10:00 **Keynote Talk 1,** AZ Ballroom 6 (streamed to 7)
Speaker: Ravi Ramamoorthi, **Title:** Sparse View Synthesis
10:00-10:15 **Courtesy Break**
10:15-11:15 **Oral Session 1A: Generative Models I,** AZ Ballroom 6
10:15-11:15 **Oral Session 1B: 3D Computer Vision I,** AZ Ballroom 7
11:15-17:45 **Exhibits + Demos,** Tucson Ballroom
11:15-13:00 **Poster Session 1,** Tucson Ballroom & Prefunction Space
12:00-13:30 **Lunch,** Ania Terrace and Lawn
12:30-13:30 **Panel Discussion,** AZ Ballroom 6
"Bridging the Gap Between Academic Benchmarks and Real-World Deployment in Computer Vision: The Path to Translation."
13:45-14:45 **Oral Session 2A: Vision+Language and Other Modalities I,** AZ Ballroom 6
13:45-14:45 **Oral Session 2B: Biometrics, Face, Gesture, and Body Pose I,** AZ Ballroom 7
14:45-15:00 **Courtesy Break**
15:00-16:00 **Oral Session 3A: Low-level and Physics-based Vision,** AZ Ballroom 6
15:00-16:00 **Oral Session 3B: Image Recognition and Understanding I,** AZ Ballroom 7
16:00-17:45 **Poster Session 2 + Refreshments,** Tucson Ballroom & Prefunction Space, *The poster session also includes demos and exhibitions.*

9:00-10:00 **Keynote Talk 1,** AZ Ballroom 6 (streamed to 7)
Speaker: Ravi Ramamoorthi, **Title:** Sparse View Synthesis

10:00-10:15 **Courtesy Break**

The papers in this oral session will also be presented as posters.
Format: 10 min. presentation, 2 min. questions

10:15-11:15 **Oral Session 1A: Generative Models I**
AZ Ballroom 6

1. DreamAnywhere: Object-Centric Panoramic 3D Scene Generation, *Edoardo A. Dominici, Jozef Hladký, Floor Verhoeven, Lukas Radl, Thomas Deixelberger, Stefan Ainetter, Philipp Drescher, Stefan Hauswiesner, Arno Coomans, Giacomo Nazzaro, Konstantinos Vardis, Markus Steinberger*
2. ViSTA: Visual Storytelling using Multi-modal Adapters for Text-to-Image Diffusion Models, *Sibo Dong, Ismail Shaheen, Maggie Shen, Rupayan Mallick, Sarah Adel Bargal*
3. Odo: Depth-Guided Diffusion for Identity-Preserving Body Reshaping, *Siddharth Khandelwal, Sridhar Kamath, Arjun Jain*
4. BiPO: Bidirectional Partial Occlusion Network for Text-to-Motion Synthesis, *Seong-Eun Hong, SooBin Lim, JuYeong Hwang, Minwook Chang, HyeongYeop Kang*
5. Reinforcement Learning-based Adaptive Control of Classifier-Free Guidance and Timestep Embeddings in Diffusion Models, *Haochen You, Baojing Liu, Hongyang He*

10:15-11:15 **Oral Session 1B: 3D Computer Vision I**
AZ Ballroom 7

1. TS-PCI: Point Cloud Frame Interpolation with Time-Aware Point Cloud Sampling and Self-Supervised Learning Strategy, *Kohei Matsuzaki, Keisuke Nonaka*
2. Enhanced Back-Projection of Vision Features for 3D Symmetry Detection, *Isaac Aguirre, Ivan Sipiran*
3. OracleGS: Grounding Generative Priors for Sparse-View Gaussian Splatting, *Atakan Topaloğlu, Kunyi Li, Michael Niemeyer, Nassir Navab, A. Murat Tekalp, Federico Tombari*
4. UnderWater SLAM with Laser-light sectioning method using ST-GAT, *Heyang Gao, Kazuto Ichimaru, Takafumi Iwaguchi, Hiroshi Kawasaki*
5. Leveraging Pretrained Representations for Cross-Modal Point Cloud Completion, *Kshitij Kale, Hrishikesh U, V sreenidhe, Shylaja S S*

11:15-17:45 **Exhibits + Demos,** Tucson Ballroom

1. A Personalized Language-Guided Video Summarization System Using Text Semantic Matching, *Tomoya Sugihara, The University of Tokyo*
2. NOx-Producing Hole Tracking in Quarry Blasts, *Xuesong Liu, Rochester Institute of Technology*
3. OceanAI: A Conversational Platform for Accurate, Transparent, Near-Real-Time Oceanographic Insights, *Bowen Chen, NCSU*
4. TalentProfiling: Assessing Athletic Potential via AlignmentBased Analysis of Fundamental Human Movements, *Chun-Min Yu, Academia Sinica*

11:15-13:00 **Poster Session 1,** Tucson Ballroom & Prefunction Space

1 DreamAnywhere: Object-Centric Panoramic 3D Scene Generation, *Edoardo A. Dominici, Jozef Hladký, Floor Verhoeven, Lukas Radl, Thomas Deixelberger, Stefan Ainetter, Philipp Drescher, Stefan Hauswiesner, Arno Coomans, Giacomo Nazzaro, Konstantinos Vardis, Markus Steinberger*
2 ViSTA: Visual Storytelling using Multi-modal Adapters for Text-to-Image Diffusion Models, *Sibo Dong, Ismail Shaheen, Maggie Shen, Rupayan Mallick, Sarah Adel Bargal*
3 Odo: Depth-Guided Diffusion for Identity-Preserving Body Reshaping, *Siddharth Khandelwal, Sridhar Kamath, Arjun Jain*
4 BiPO: Bidirectional Partial Occlusion Network for Text-to-Motion Synthesis, *Seong-Eun Hong, SooBin Lim, JuYeong Hwang, Minwook Chang, HyeongYeop Kang*
5 Reinforcement Learning-based Adaptive Control of Classifier-Free Guidance and Timestep Embeddings in Diffusion Models, *Haochen You, Baojing Liu, Hongyang He*
6 TS-PCI: Point Cloud Frame Interpolation with Time-Aware Point Cloud Sampling and Self-Supervised Learning Strategy, *Kohei Matsuzaki, Keisuke Nonaka*
7 Enhanced Back-Projection of Vision Features for 3D Symmetry Detection, *Isaac Aguirre, Ivan Sipiran*
8 OracleGS: Grounding Generative Priors for Sparse-View Gaussian Splatting, *Atakan Topaloğlu, Kunyi Li, Michael Niemeyer, Nassir Navab, A. Murat Tekalp, Federico Tombari*
9 UnderWater SLAM with Laser-light sectioning method using ST-GAT, *Heyang Gao, Kazuto Ichimaru, Takafumi Iwaguchi, Hiroshi Kawasaki*
10 Leveraging Pretrained Representations for Cross-Modal Point Cloud Completion, *Kshitij Kale, Hrishikesh U, V sreenidhe, Shylaja S S*
11 Referring Change Detection in Remote Sensing Imagery, *Yilmaz Korkmaz, Jay N. Paranjape, Celso M. de Melo, Vishal M. Patel*
12 Inpaint360GS: Efficient Object-Aware 3D Inpainting via Gaussian Splatting for 360° Scenes, *Shaoxiang Wang, Shihong Zhang, Christen Millerdurai, Rüdiger Westermann, Didier Stricker, Alain Pagani*
13 A Multi-Agent Diffusion Approach for MRI Anomaly Segmentation via Modality-Specific LoRA Specialization, *Wafa Al Ghallabi, Muhammad Zaigham Zaheer, Ritesh Thawkar, Omkar Thawakar, Salman Khan, Fahad Shahbaz Khan*
14 GenHSI: Controllable Generation of Human-Scene Interaction Videos, *Zekun Li, Rui Zhou, Rahul Sajnani, Xiaoyan Cong, Daniel Ritchie, Srinath Sridhar*
15 SymNet: A Multi-Task Network for Joint Radio Map Reconstruction and Transmitter Localization, *Lyuzhou Ye, Thanh Dat Le, Yan Huang*
16 LooC: Effective Low-Dimensional Codebook for Compositional Vector Quantization, *Jie Li, Kwan-Yee K. Wong, Kai Han*
17 End-to-End Fine-Tuning of 3D Texture Generation using Differentiable Rewards, *AmirHossein Zamani, Tianhao Xie, Amir G. Aghdam, Tiberiu Popa, Eugene Belilovsky*

18 IDEAL-M3D: Instance Diversity-Enriched Active Learning for Monocular 3D Detection, *Johannes Meier, Florian Günther, Riccardo Marin, Oussema Dhaouadi, Jacques Kaiser, Daniel Cremers*
19 FAE-Net: Fashion Attribute Editing via Disentangled Latent Conditioning in Diffusion Models, *P. Rajith Bhargav, Gaurab Bhattacharya, B S Vivek, Jayavardhana Gubbi*
20 DiT-VTON: Diffusion Transformer Framework for Unified Multi-Category Virtual Try-On and Virtual Try-All with Integrated Image Editing, *Qi Li, Shuwen Qiu, Kee Kiat Koo, Julien Han, Karim Bouyarmane*
21 MAESTRO: Masked AutoEncoders for Multimodal, Multitemporal, and Multispectral Earth Observation Data, *Antoine Labatie, Michael Vaccaro, Nina Lardiere, Anatol Garioud, Nicolas Gonthier*
22 TA-Prompting: Enhancing Video Large Language Models for Dense Video Captioning via Temporal Anchors, *Wei-Yuan Cheng, Kai-Po Chang, Chi-Pin Huang, Fu-En Yang, Yu-Chiang Frank Wang*
23 Towards Fast and Scalable Normal Integration using Continuous Components, *Francesco Milano, Jen Jen Chung, Lionel Ott, Roland Siegwart*
24 A Framework for Real-Time Surgical Phase Recognition with Application to Robot-Assisted Partial Nephrectomy, *Marco Mezzina, Tom Vercauteren, Tinne Tuytelaars, Matthew B. Blaschko*
25 A Woman with a Knife or A Knife with a Woman? Measuring Directional Bias Amplification in Image Captions, *Rahul Nair, Bhanu Tokas, Hannah Kerner*
26 Forget Less by Learning Together through Concept Consolidation, *Arjun Ramesh Kaushik, Naresh Kumar Devulapally, Vishnu Suresh Lokhande, Nalini Ratha, Venu Govindaraju*
27 Action Anticipation at a Glimpse: To What Extent Can Multimodal Cues Replace Video?, *Manuel Benavent-Lledo, Konstantinos Bacharidis, Victoria Manousaki, Konstantinos Papoutsakis, Antonis Argyros, José García-Rodríguez*
28 VISTA: A Vision and Intent-Aware Social Attention Framework for Multi-Agent Trajectory Prediction, *Stephane Da Silva Martins, Emanuel Aldea, Sylvie Le Hégarat-Mascle*
29 ChameleonTuner: Automatic ISP Color Tuning in Subjective Scenarios, *Zijie Tan, Yuxin Yue, Bahador Rashidi*
30 ART: Actor-Related Tubelet for Detecting Complex-shaped Action Tubes, *Jiaojiao Zhao*
31 Training-free Multi-view 4D Human Motion Reconstruction Virtual Reality System, *Yijie Li, Ce Zheng, Yijie He, Joel Julin, Ryosuke Ichikari, Satoki Ogiso, Satoshi Nakae, Akihiro Sato, Takeshi Kurata, Laszlo A. Jeni*
32 EmojiDiff: Advanced Facial Expression Control with High Identity Preservation in Portrait Generation, *Liangwei Jiang, Ruida Li, Zhifeng Zhang, Shuo Fang, Chenguang Ma*
33 OW-Rep: Open World Object Detection with Instance Representation Learning, *Sunoh Lee, Minsik Jeon, Jihong Min, Junwon Seo*
34 Cluster-Guided Adversarial Perturbations for Robust Contrastive Learning, *Seongyun Seo, Sungmin Han, Jeonghyun Lee, Sangkyun Lee*
35 A Universal Self-Attention Enhancement for Bridging Low-bit Quantization and Vision Transformers, *Jiahe Qian, Peisong Wang, Zhengyang Zhuge, Qinghao Hu, Jian Cheng*
36 Overcoming Fine-Grained Visual Challenges in Animal Re-Identification via Semantic Feature Alignment, *Yihao Wu, Di Zhao, Yuzhuo Li, Matthew Alajas, Alistair S. Glen, Jingfeng Zhang, Gillian Dobbie, Daniel Wilson, Yun Sing Koh*
37 M4U: Evaluating Multilingual Understanding and Reasoning for Large Multimodal Models, *Hongyu Wang, Jiayu Xu, Senwei Xie, Ruiping Wang, Jialin Li, Zhaojie Xie, Bin Zhang, Chuyan Xiong, Xilin Chen*
38 RAT4D: Rig and Animate Objects without Surface Templates in 4D, *Mosam Dabhi, Simon Lucey, Laszlo A. Jeni*
39 AFL-PRF: Adaptive Federated Learning for Low-Quality Data: Enhancing Performance, Robustness, and Fairness, *Pinrui Yu, Yiming Xie, Longtian Ye, Geng Yuan, Ningfang Mi, Xue Lin*
40 Eff-GRot: Efficient and Generalizable Rotation Estimation with Transformers, *Fanis Mathioulakis, Gorjan Radevski, Tinne Tuytelaars*
41 DreamMakeup: Face Makeup Customization using Latent Diffusion Models, *Geon Yeong Park, Inhwa Han, Serin Yang, Yeobin Hong, Seongmin Jeong, Heechan Jeon, Myeongjin Goh, Sung Won Yi, Jin Nam, Jong Chul Ye*
42 Morphing Through Time: Diffusion-Based Bridging of Temporal Gaps for Robust Alignment in Change Detection, *Seyedehanita Madani, Vishal M. Patel*

43 AdaptViG: Adaptive Vision GNN with Exponential Decay Gating, *Mustafa Munir, Md Mostafijur Rahman, Radu Marculescu*
44 MooTrack360: A Novel Fisheye Camera Dataset for Robust Multi Diary Cow Detection and Tracking, *Rasmus Gjerlund K. Christiansen, Toan Van Nguyen, Lasse Rose Malskær, Leon Bodenhagen, Dirk Kraft*
45 MDUNet: Multimodal Decoding UNet for Passive Occluder-Aided Non-line-of-sight 3D Imaging, *Fadlullah Raji, John Murray-Bruce*
46 Interleaved Vision-and-Language Generation via Generative Voken, *Kaizhi Zheng, Xuehai He, Xin Eric Wang*
47 Root Completion from Intraoral Scans of Tooth Crowns using Diffusion with Patch Perturbation, *Yohan Jang, In-Seok Song, Seung Jun Baek*
48 Systematic Analysis of the Unintentional CSAM-Generation-Potential of Text-to-Image Models, *Nicolas Göller, Martin Steinebach*
49 SGPMIL: Sparse Gaussian Process Multiple Instance Learning, *Andreas Lolos, Stergios Christodoulidis, Maria Vakalopoulou, Jose Dolz, Aris L. Moustakas*
50 Understanding Human-Like Biases in VLMs via Subjective Face Analytics, *Chaitanya Roygaga, Aparna Bharati*
51 M-ErasureBench: A Comprehensive Multimodal Evaluation Benchmark for Concept Erasure in Diffusion Models, *Ju Hsuan Weng, Jia-Wei Liao, Cheng-Fu Chou, Jun-Cheng Chen*
52 Revisiting Layer Normalization for Point Cloud Test Time Adaptation, *Moslem Yazdanpanah, Ali Bahri, Mehrdad Noori, Sahar Dastani, Samuel Barbeau, David Osowiechi, Gustavo Adolfo Vargas Hakim, Ismail Ben Ayed, Christian Desrosiers*
53 Learning to Animate Images from A Few Videos to Portray Delicate Human Actions, *Haoxin Li, Yingchen Yu, Qilong Wu, Hanwang Zhang, Song Bai, Boyang Li*
54 One-Shot Fine-Grained Re-Identification of Paint Marked Honey Bees using Vision Foundation Models, *Luke Meyers, Josué A. Rodríguez-Cordero, Rémi Mégret*
55 From Street to Orbit: Training-Free Cross-View Retrieval via Location Semantics and LLM Guidance, *Jeongho Min, Dongyoung Kim, Jaehyup Lee*
56 Mitigating Backdoor Attacks via Trigger Reconstruction and Model Hardening, *Guanhong Tao, Siyuan Cheng, Guangyu Shen, Yingqi Liu, Shengwei An, Zhuo Zhang, Zhenting Wang, Hanxi Guo, Xiangyu Zhang*
57 Network-agnostic Distortion-robust Projections for Wide-angle Image Understanding, *Akshaya Athwale, Ola Ahmad, Jean-Francois Lalonde*
58 4D-Animal: Freely Reconstructing Animatable 3D Animals from Videos, *Shanshan Zhong, Jiawei Peng, Zehan Zheng, Zhongzhan Huang, Wufei Ma, Guofeng Zhang, Qihao Liu, Alan Yuille, Jieneng Chen*
59 Text Slider: Efficient and Plug-and-Play Continuous Concept Control for Image/Video Synthesis via LoRA Adapters, *Pin-Yen Chiu, I-Sheng Fang, Jun-Cheng Chen*
60 Cluster-based Pseudo-labeling for Semi-Supervised LiDAR Semantic Segmentation, *Qingju Guo, Shuang Li, Jing Geng, Binhui Xie, Jiawei Shan, Wei Li*
61 Human Pose Aggregation for Multi-View Temporal Video Alignment, *Fabien Delattre, Tsung-Wei Huang, Guan-Ming Su, Erik Learned-Miller*
62 Marshaled Learning: Bridging Large Neural Networks with Memory-Constrained Trusted Execution Environments in Federated Learning, *Shiwei Ding, Xiaoyong Yuan, Zhenlin Wang, Lan Emily Zhang, Giuseppe Ateniese*
63 Feature-Disentangling RGB-NIR Fusion Network for Remote Driver Physiological Measurement, *Tayssir Bouraffa, Ziyuan Wang, Daniel Strüber*
64 Prompt-OT: An Optimal Transport Regularization Paradigm for Knowledge Preservation in Vision-Language Model Adaptation, *Xiwen Chen, Wenhui Zhu, Peijie Qiu, Hao Wang, Huayu Li, Haiyu Wu, Xuanzhao Dong, Aristeidis Sotiras, Yalin Wang, Abolfazl Razi*
65 Zero-Shot Video Deraining with Video Diffusion Models, *Tuomas Varanka, Juan Luis Gonzalez, Hyeongwoo Kim, Pablo Garrido, Xu Yao*
66 Joint Modeling of Corruption-Driven and Information-Limited Uncertainty for Robust 3D Gaussian Splatting, *Zeji Hui, Amirali Khodadadian Gostar, WeiQin Chuah, Alireza Bab-Hadiashar, Ruwan Tennakoon*
67 Data-Driven Lipschitz Continuity: A Cost-Effective Approach to Improve Adversarial Robustness, *Erh-Chung Chen, Pin-Yu Chen, I-Hsin Chung, Che-Rung Lee*
68 CADE: Continual Weakly-supervised Video Anomaly Detection with Ensembles, *Satoshi Hashimoto, Tatsuya Konishi, Tomoya Kaichi, Kazunori Matsumoto, Mori Kurokawa*
69 PaRaChute: Pathology-Radiology Cross-Modal Fusion for Missing-Modality-Robust Survival Prediction, *Pietro Caforio, Isabella Poles, Marco D. Santambrogio*
70 How to Design and Train Your Implicit Neural Representation for Video Compression, *Matthew Gwilliam, Roy Zhang, Namitha Padmanabhan, Hongyang Du, Abhinav Shrivastava*
71 Can We Challenge Open-Vocabulary Object Detectors with Generated Content in Street Scenes?, *Annika Mütze, Sadia Ilyas, Christian Dörpelkus, Matthias Rottmann*
72 Conjuring Positive Pairs for Efficient Unification of Representation Learning and Image Synthesis, *Imanol G. Estepa, Jesús M. Rodríguez-de-Vera, Ignacio Sarasúa, Bhalaji Nagarajan, Petia Radeva*
73 DICE: Discrete Inversion Enabling Controllable Editing for Masked Generative Models, *Xiaoxiao He, Quan Dao, Ligong Han, Song Wen, Minhao Bai, Di Liu, Han Zhang, Felix Juefei-Xu, Chaowei Tan, Bo Liu, Martin Renqiang Min, Kang Li, Faez Ahmed, Akash Srivastava, Hongdong Li, Junzhou Huang, Dimitris N. Metaxas*
74 Deepfake Detection that Generalizes Across Benchmarks, *Andrii Yermakov, Jan Cech, Jiri Matas, Mario Fritz*
75 Unified Video Anomaly Detection Model for Detecting Different Anomaly Types, *Kijung Lee, Youngwan Jo, Sunghyun Ahn, Sanghyun Park*
76 SCORP: Scene-Consistent Object Refinement via Proxy Generation and Tuning, *Ziwei Chen, Ziling Liu, Zitong Huang, Mingqi Gao, Feng Zheng*
77 CountingDINO: A Training-free Pipeline for Class-Agnostic Counting using Unsupervised Backbones, *Giacomo Pacini, Lorenzo Bianchi, Luca Ciampi, Nicola Messina, Giuseppe Amato, Fabrizio Falchi*
78 ViGG: Robust RGB-D Point Cloud Registration using Visual-Geometric Mutual Guidance, *Congjia Chen, Shen Yan, Yufu Qu*
79 Gaussian Representations for Video, *Sachin Shah, Anustup Choudhury, Guan-Ming Su, Jaclyn Pytlarz, Christopher A. Metzler, Trisha Mittal*
80 MSRTrack: LLM-Powered Object Tracking with Motion and Semantic Reasoning, *Tong Shen, Di Wang, José M. F. Moura*
81 Enabling High-Quality In-the-Wild Imaging from Severely Aberrated Metalens Bursts, *Debabrata Mandal, Zhihan Peng, Yujie Wang, Praneeth Chakravarthula*
82 Boosting Medical Vision-Language Pretraining via Momentum Self-Distillation under Limited Computing Resources, *Phuc Pham, Nhu Pham, Ngoc Quoc Ly*
83 Beyond Realism: Learning the Art of Expressive Composition with StickerNet, *Haoming Lu, David Kocharian, Humphrey Shi*
84 Analysis of Text Accuracy and Visual Alignment in Vision-Language Models for Artistic Text Generation, *Fatima Alderazi, Motaz Alfarraj*
85 MEDAL: multi-modal MEta-space Distillation and ALignment for Visual Compatibility Learning, *Dween Rabius Sanny, Vinay Kumar Verma, Prateek Sircar, Deepak Gupta*
86 PS3: Part Level Instance Segmentation in 3D, *Hong-Xuan Yen, Chiamin Chen, Yanqing Wang, Yu-Lun Liu, Min Sun*
87 Saliency-Guided DETR for Moment Retrieval and Highlight Detection, *Aleksandr Gordeev, Vladimir Dokholyan, Irina Tolstykh, Maksim Kuprashevich*
88 Accelerated Dose Generation in Gamma Knife Radiosurgery Using a Wavelet Diffusion Model for Sparse Representation, *Sangyoon Lee, Shubuendu Mishra, Yoichi Watanabe*

89 DermEVAL: A Dermatologist-Reviewed Benchmark for Multimodal Large Language Models, *Hongjin Zhao, Weihao Li, Zhenyue Qin, Ge-Peng Ji, Yang Liu, Tom Gedeon, Nick Barnes*
90 SynPlay: Large-Scale Synthetic Human Data with Real-World Diversity for Aerial-View Perception, *Jinsub Yim, Hyungtae Lee, Sungmin Eum, Yi-Ting Shen, Yan Zhang, Heesung Kwon, Shuvra S. Bhattacharyya*
91 Do Generative Video Models Understand Physical Principles?, *Saman Motamed, Laura Culp, Kevin Swersky, Priyank Jaini, Robert Geirhos*
92 ZonUI-3B: Competitive GUI Grounding with a 3B VLM Trained on a Single Consumer GPU, *ZongHan Hsieh, ShengJing Yang, Tzer-Jen Wei*
93 No MoCap Needed: Post-Training Motion Diffusion Models with Reinforcement Learning using Only Textual Prompts, *Macaluso Girolamo, Mandelli Lorenzo, Mirko Bicchierai, Stefano Berretti, Andrew D. Bagdanov*
94 MIX-based Foreground and Background Patch Augmentation Guided by Physics and Material Properties for X-ray Detection, *Xintong Liu, Dongliang Chang, Yujun Tong, Zhanyu Ma*
95 Reverse Personalization, *Han-Wei Kung, Tuomas Varanka, Nicu Sebe*
96 Beyond the Encoder: Joint Encoder-Decoder Contrastive Pre-Training Improves Dense Prediction, *Sébastien Quetin, Tapotosh Ghosh, Farhad Maleki*
97 HyperPose: Hyper-pose Embeddings for 3D-Aware Generative Models with Self-Supervised Disentangling of Pose and Scene, *Mijeong Kim, Namgi Kim, Bohyung Han*
98 Learning Mask-Aware Offsets: Two-branch Deformable Attention Networks for Inpainting with Masked Region Avoidance, *Hyeongseok Oh, Joonki Paik*
99 Salience-SGG: Enhancing Unbiased Scene Graph Generation with Iterative Salience Estimation, *Runfeng Qu, Ole Hall, Pia K Bideau, Julie Ouerfelli-Ethier, Martin Rolfs, Klaus Obermayer, Olaf Hellwich*
100 SmokeBench: Evaluating Multimodal Large Language Models for Wildfire Smoke Detection, *Tianye Qi, Weihao Li, Nick Barnes*
101 Visibility guided Self-Supervised Occlusion-Resilient Human Pose Estimation, *Arindam Dutta, Sarosij Bose, Rohit Kundu, Calvin-Khang Ta, Saketh Bachu, Konstantinos Karydis, Amit Roy-Chowdhury*
102 Diverse Sketch Colorization with Content-Enhanced Style Representation and Recolorization Distillation, *Shuangming Mao, Haixiang Zhu*
103 GateFusion: Hierarchical Gated Cross-Modal Fusion for Active Speaker Detection, *Yu Wang, Juhyung Ha, Frangil M. Ramirez, Yuchen Wang, David J. Crandall*
104 MR-Pruner: Training-free Multi-resolution Visual Token Pruning for Multi-modal Large Language Models, *Seunghoon Han, Hyewon Lee, Soyoung Park, Jong-Ryul Lee, Sungsu Lim*
105 SpikeRain: Towards Energy-Efficient Single Image Deraining with Spiking Neural Networks, *Md Tanvir Islam, Inzamamul Alam, Sambit Bakshi, Khan Muhammad, Javier Del Ser, Sangtae Ahn*
106 MBTI: Metric-Based Textual Inversion for Fine-Grained Image Generation, *Byungkwan Chae, Youngjae Choi, Heewon Kim*
107 ImageNet-sES: A First Systematic Study of Sensor–Environment Simulation Anchored by Real Recaptures, *Ji-yoon Kim, Eunsu Baek, Hyung-Sin Kim*
108 MomentMix Augmentation with Length-Aware DETR for Temporally Robust Moment Retrieval, *Seojeong Park, Jiho Choi, Kyungjune Baek, Hyunjung Shim*
109 From Lightweight CNNs to SpikeNets: Benchmarking Accuracy–Energy Tradeoffs with Pruned Spiking SqueezeNet, *Radib Bin Kabir, Tawsif Tashwar Dipto, Mehedi Ahamed, Sabbir Ahmed, Md Hasanul Kabir*
110 AFRAgent : An Adaptive Feature Renormalization Based High Resolution Aware GUI agent, *Neeraj Anand, Rishabh Jain, Sohan Patnaik, Mausoom Sarkar, Balaji Krishnamurthy*
111 BanglaProtha: Evaluating Vision Language Models in Underrepresented Long-tail Cultural Contexts, *Md Fahim, Md Sakib Ul Rahman, Akm Moshiur Rahman, Md Farhan Ishmam, Md Tasmim Rahman, Fariha Tanjim Shifat, Fabiha Haider, Md Farhad Alam Bhuiyan*
112 CommonForms: A Large, Diverse Dataset for Form Field Detection, *Joe Barrow*
113 Beyond Real Weights: Hypercomplex Representations for Stable Quantization, *Jawad Ibn Ahad, Maisha Rahman, Amrijit Biswas, Muhammad Rafsan Kabir, Robin Krambroeckers, Sifat Momen, Nabeel Mohammed, Shafin Rahman*
114 mEOL: Training-Free Instruction-Guided Multimodal Embedder for Vector Graphics and Image Retrieval, *Kyeong Seon Kim, Baek Seong-Eun, Lee Jung-Mok, Tae-Hyun Oh*
115 Contrastive Integrated Gradients: A Feature Attribution-Based Method for Explaining Whole Slide Image Classification, *Anh Mai Vu, Tuan L. Vo, Ngoc Lam Quang Bui, Nam N. B. Le, Akash Awasthi, Huy Q. Vo, Thanh-Huy Nguyen, Zhu Han, Chandra Mohan, Hien Van Nguyen*
116 PSA-MIL: A Probabilistic Spatial Attention-Based Multiple Instance Learning for Whole Slide Image Classification, *Sharon Peled, Yosef E. Maruvka, Moti Freiman*
117 Fully Unsupervised Self-debiasing of Text-to-Image Diffusion Models, *Korada Sri Vardhana, Shrikrishna Lolla, Soma Biswas*
118 Model-free Domain Adaptation for Concealed Multimodal Large-Language Models, *Yu Mitsuzumi, Akisato Kimura, Hisashi Kashima*
119 CAAC: Confidence-Aware Attention Calibration to Reduce Hallucinations in Large Vision-Language Models, *Mehrdad Fazli, Bowen Wei, Ahmet Sari, Ziwei Zhu*
120 LiDAR-DHMT: LiDAR-Adaptive Dual Hierarchical Mask Transformer for Robust Freespace Detection and Semantic Segmentation, *Siyu Chen, Ting Han, Changshe Zhang, Xin Luo, Huan Chen, Meiliu Wu, Guorong Cai, Jinhe Su*
121 Mixed Diffusion for 3D Indoor Scene Synthesis, *Siyi Hu, Diego Martín Arroyo, Stephanie Debats, Fabian Manhardt, Luca Carlone, Federico Tombari*
122 PiSA: A Self-Augmented Data Engine and Training Strategy for 3D Understanding with Large Models, *Zilu Guo, Hongbin Lin, Zhihao Yuan, Chaoda Zheng, Pengshuo Qiu, Dongzhi Jiang, Renrui Zhang, Chun-Mei Feng, Zhen Li*
123 MARS: a Multimodal Alignment and Ranking System for Few-Shot Segmentation, *Nico Catalano, Stefano Samele, Paolo Pertino, Matteo Matteucci*
124 SilverLining: Data-First Mitigation of Spatial and Spectral Shortcuts Without Introducing New Confounders, *Balagopal Unnikrishnan, Michael Brudno, Chris McIntosh*
125 Distilling Diversity and Control in Diffusion Models, *Rohit Gandikota, David Bau*
126 Narrating For You: Prompt-guided Audio-visual Narrating Face Generation Employing Multi-entangled Latent Space, *Aashish Chandra K, Aashutosh A V, Abhijit Das*
127 Explaining the Unseen: Multimodal Vision-Language Reasoning for Situational Awareness in Underground Mining Disasters, *Mizanur Rahman Jewel, Mohamed Elmahallawy, Sanjay Madria, Samuel Frimpong*
128 FALCONEye: Finding Answers and Localizing Content in ONE-hour-long videos with multi-modal LLMs, *Carlos Plou, Cesar Borja, Ruben Martinez-Cantin, Ana C. Murillo*
129 Autocorrelation-based Fiducial Markers for Traceability, *Ismail Bencheikh, Max Dunitz, Marie d'Autume, Marc Pic, Enric Meinhardt-Llopis, Gabriele Facciolo, Pablo Musé*
130 GrounDiff: Diffusion-Based Ground Surface Generation from Digital Surface Models, *Oussema Dhaouadi, Johannes Meier, Jacques Kaiser, Daniel Cremers*
131 QuadraNet V2: Efficient and Sustainable Training of High-Order Neural Networks with Quadratic Adaptation, *Chenhui Xu, Fuxun Yu, Jinjun Xiong, Xiang Chen*
132 AutoSew: A Geometric Approach to Stitching Prediction with Graph Neural Networks, *Pablo Ríos-Navarro, Elena Garces, Jorge Lopez-Moreno*
133 SaccadeX: Directed Acyclic Graph-based Semi-Supervised Learning of Continuous Ocular Dynamics from Sparse Neuromorphic Streams, *Nuwan Bandara, Thivya Kandappu, Archan Misra*

134 Federated Model Synchronization for Diagnostic Redefinition through a Novel Selective Parameter Unlearning, *Mayank Kumar Kundalwal, Mamta Mamta, Deepak Mishra, Asif Ekbal*
135 A Fast, Simple, and Flexible Scale Informative Feature Transform Module for Arbitrary Scale Image Super-Resolution, *Aupendu Kar, Prabir Kumar Biswas*

12:00-13:30 **Lunch,** Ania Terrace and Lawn

12:30-13:30 **Panel Discussion,** AZ Ballroom 6
"Bridging the Gap Between Academic Benchmarks and Real-World Deployment in Computer Vision: The Path to Translation."

13:45-14:45 **Oral Session 2A: Vision+Language and Other Modalities I,** AZ Ballroom 6
1. MageBench: Bridging Large Multimodal Models to Agents, *Miaosen Zhang, Qi Dai, Yifan Yang, Jianmin Bao, Dongdong Chen, Kai Qiu, Chong Luo, Xin Geng, Baining Guo*
2. You May Speak Freely: Improving the Fine-Grained Visual Recognition Capabilities of Multimodal Large Language Models with Answer Extraction, *Logan Lawrence, Oindrila Saha, Megan Wei, Chen Sun, Subhransu Maji, Grant Van Horn*
3. InteracTalker: Prompt-Based Human-Object Interaction with Co-Speech Gesture Generation, *Sreehari Rajan, Kunal Bhosikar, Charu Sharma*
4. ITSELF: Attention Guided Fine-Grained Alignment for Vision-Language Retrieval, *Tien-Huy Nguyen, Huu-Loc Tran, Thanh Duc Ngo*
5. MarineEval: Assessing the Marine Intelligence of Vision-Language Models, *Wong Yuk Kwan, Tuan-An To, Jipeng Zhang, Zheng Ziqiang, Sai-Kit Yeung*

13:45-14:45 **Oral Session 2B: Biometrics, Face, Gesture, and Body Pose I,** AZ Ballroom 7
1. Identity Verification from Human Scent using Channel Representation of 2D Gas Chromatography-Mass Spectrometry Data, *Radim Spetlik, Jan Hlavsa, Jana Čechová, Petra Pojmanová, Jiri Matas, Štěpán Urban*
2. milliMamba: Specular-Aware Human Pose Estimation via Dual mmWave Radar with Multi-Frame Mamba Fusion, *Niraj Prakash Kini, Shiau-Rung Tsai, Guan-Hsun Lin, Wen-Hsiao Peng, Ching-Wen Ma, Jenq-Neng Hwang*
3. OpenCowID: Zero-Shot Visual Identification of Dairy Cows, *Omkar Prabhune, Younghyun Kim*
4. QCFace: Image Quality Control for boosting Face Representation & Recognition, *Duc-Phuong Doan-Ngo, Thanh-Dang Diep, Thanh Nguyen-Duc, Thanh-Sach LE, Nam Thoai*
5. MMHOI: Modeling Complex 3D Multi-Human Multi-Object Interactions, *Kaen Kogashi, Anoop Cherian, Meng-Yu Jennifer Kuo*

14:45-15:00 **Courtesy Break**

15:00-16:00 **Oral Session 3A: Low-level and Physics-based Vision,** AZ Ballroom 6
1. BrightRate: Quality Assessment for User-Generated HDR Videos, *Shreshth Saini, Bowen Chen, Yilin Wang, Neil Birkbeck, Balu Adsumilli, Alan C. Bovik*
2. Reviving Unsupervised Optical Flow: Concept Reevaluation, Multi-Scale Advances and Full Open-Source Release, *Azin Jahedi, Marc Rivinius, Noah Berenguel Senn, Andres Bruhn*
3. UniCoRN: Latent Diffusion-based Unified Controllable Image Restoration Network across Multiple Degradations, *Debabrata Mandal, Soumitri Chattopadhyay, Guansen Tong, Praneeth Chakravarthula*
4. DRWKV: Focusing on Object Edges for Low-Light Image Enhancement, *Xuecheng Bai, Yuxiang Wang, Boyu Hu, Qinyuan Jie, Chuanzhi Xu, Kechen Li, Hongru Xiao, Vera Chung*

15:00-16:00 **Oral Session 3B: Image Recognition and Understanding I,** AZ Ballroom 7
1. Layout Anything: One Transformer for Universal Room Layout Estimation, *Md Sohag Mia, Muhammad Abdullah Adnan*
2. BOP-Distrib: Revisiting 6D Pose Estimation Benchmarks for Better Evaluation under Visual Ambiguities, *Boris Meden, Asma Brazi, Fabrice Mayran de Chamisso, Steve Bourgeois, Vincent Lepetit*
3. Cosine Similarity is Almost All You Need (for Prototypical-Part Models), *Luke Moffett, Frank Willard, Maximillian Machado, Emmanuel Mokel, Jon Donnelly, Zhicheng Guo, Adam Costarino, Julia Yang, Giyoung Kim, Alina Jade Barnett, Cynthia Rudin*
4. ORCA: Object Recognition and Comprehension for Archiving Marine Species, *Yuk-Kwan Wong, Haixin Liang, Zeyu Ma, Yiwei Chen, Ziqiang Zheng, Rinaldi Gotama, Pascal Sebastian, Lauren D. Sparks, Sai-Kit Yeung*

16:00-17:45 **Poster Session 2 + Refreshments,** Tucson Ballroom & Prefunction Space, *The poster session also includes demos and exhibitions.*
1 MageBench: Bridging Large Multimodal Models to Agents, *Miaosen Zhang, Qi Dai, Yifan Yang, Jianmin Bao, Dongdong Chen, Kai Qiu, Chong Luo, Xin Geng, Baining Guo*
2 You May Speak Freely: Improving the Fine-Grained Visual Recognition Capabilities of Multimodal Large Language Models with Answer Extraction, *Logan Lawrence, Oindrila Saha, Megan Wei, Chen Sun, Subhransu Maji, Grant Van Horn*
3 InteracTalker: Prompt-Based Human-Object Interaction with Co-Speech Gesture Generation, *Sreehari Rajan, Kunal Bhosikar, Charu Sharma*
4 ITSELF: Attention Guided Fine-Grained Alignment for Vision-Language Retrieval, *Tien-Huy Nguyen, Huu-Loc Tran, Thanh Duc Ngo*
5 MarineEval: Assessing the Marine Intelligence of Vision-Language Models, *Wong Yuk Kwan, Tuan-An To, Jipeng Zhang, Zheng Ziqiang, Sai-Kit Yeung*
6 Identity Verification from Human Scent using Channel Representation of 2D Gas Chromatography-Mass Spectrometry Data, *Radim Spetlik, Jan Hlavsa, Jana Čechová, Petra Pojmanová, Jiri Matas, Štěpán Urban*
7 milliMamba: Specular-Aware Human Pose Estimation via Dual mmWave Radar with Multi-Frame Mamba Fusion, *Niraj Prakash Kini, Shiau-Rung Tsai, Guan-Hsun Lin, Wen-Hsiao Peng, Ching-Wen Ma, Jenq-Neng Hwang*
8 OpenCowID: Zero-Shot Visual Identification of Dairy Cows, *Omkar Prabhune, Younghyun Kim*
9 QCFace: Image Quality Control for boosting Face Representation & Recognition, *Duc-Phuong Doan-Ngo, Thanh-Dang Diep, Thanh Nguyen-Duc, Thanh-Sach LE, Nam Thoai*
10 MMHOI: Modeling Complex 3D Multi-Human Multi-Object Interactions, *Kaen Kogashi, Anoop Cherian, Meng-Yu Jennifer Kuo*
11 BrightRate: Quality Assessment for User-Generated HDR Videos, *Shreshth Saini, Bowen Chen, Yilin Wang, Neil Birkbeck, Balu Adsumilli, Alan C. Bovik*
12 Reviving Unsupervised Optical Flow: Concept Reevaluation, Multi-Scale Advances and Full Open-Source Release, *Azin Jahedi, Marc Rivinius, Noah Berenguel Senn, Andres Bruhn*
13 UniCoRN: Latent Diffusion-based Unified Controllable Image Restoration Network across Multiple Degradations, *Debabrata Mandal, Soumitri Chattopadhyay, Guansen Tong, Praneeth Chakravarthula*
14 DRWKV: Focusing on Object Edges for Low-Light Image Enhancement, *Xuecheng Bai, Yuxiang Wang, Boyu Hu, Qinyuan Jie, Chuanzhi Xu, Kechen Li, Hongru Xiao, Vera Chung*
15 Layout Anything: One Transformer for Universal Room Layout Estimation, *Md Sohag Mia, Muhammad Abdullah Adnan*
16 BOP-Distrib: Revisiting 6D Pose Estimation Benchmarks for Better Evaluation under Visual Ambiguities, *Boris Meden, Asma Brazi, Fabrice Mayran de Chamisso, Steve Bourgeois, Vincent Lepetit*

17 Cosine Similarity is Almost All You Need (for Prototypical-Part Models), *Luke Moffett, Frank Willard, Maximillian Machado, Emmanuel Mokel, Jon Donnelly, Zhicheng Guo, Adam Costarino, Julia Yang, Giyoung Kim, Alina Jade Barnett, Cynthia Rudin*
18 ORCA: Object Recognition and Comprehension for Archiving Marine Species, *Yuk-Kwan Wong, Haixin Liang, Zeyu Ma, Yiwei Chen, Ziqiang Zheng, Rinaldi Gotama, Pascal Sebastian, Lauren D. Sparks, Sai-Kit Yeung*
19 Multimodal Medical Image Binding via Shared Text Embeddings, *Yunhao Liu, Suyang Xi, Shiqi Liu, Hong Ding, Chicheng Jin, Chong Zhong, Junjun He, Catherine C. Liu, Yiqing Shen*
20 PHYSPLAT: a Framework for Photorealistic Hybrid Simulation of Real and Synthetic Elements using 3D Gaussian Splatting, *Mario Alfonso-Arsuaga, Henar Dominguez-Elvira, Jorge Casas-Guerrero, Andrea Castiella-Aguirrezabala, Lorenzo Costábile Domínguez, Jorge García-González, Maria Naranjo-Almeida, Marc Comino-Trinidad, Jorge Lopez-Moreno*
21 ArchitectHead: Continuous Level of Detail Control for 3D Gaussian Head Avatars, *Peizhi Yan, Rabab Ward, Qiang Tang, Shan Du*
22 Uncertainty-Aware Subset Selection for Robust Visual Explainability under Distribution Shifts, *Madhav Gupta, Vishak Prasad, Ganesh Ramakrishnan*
23 The Perceptual Observatory Characterizing Robustness and Grounding in MLLMs, *Tejas Anvekar, Fenil Bardoliya, Pavan K. Turaga, Chitta Baral, Vivek Gupta*
24 FAST-EQA: Efficient Embodied Question Answering with Global and Local Region Relevancy, *Haochen Zhang, Nirav Savaliya, Faizan Siddiqui, Enna Sachdeva*
25 RapidMV: Leveraging Spatio-Angular Latent Space for Efficient and Consistent Text-to-Multi-View Synthesis, *Seungwook Kim, Yichun Shi, Kejie Li, Minsu Cho, Peng Wang*
26 Diffusion-Based Authentication of Copy Detection Patterns: A Multimodal Framework with Printer Signature Conditioning, *Bolutife Atoki, Iuliia Tkachenko, Bertrand Kerautret, Carlos Crispim-Junior*
27 AD2: Analysis and Detection of Adversarial Threats in Visual Perception for End-to-End Autonomous Driving Systems, *Ishan Sahu, Somnath Hazra, Somak Aditya, Soumyajit Dey*
28 SSplain: Sparse and Smooth Explainer for Retinopathy of Prematurity Classification, *Elifnur Sunger, Tales Imbiriba, Peter Campbell, Deniz Erdogmus, Stratis Ioannidis, Jennifer Dy*
29 Tables Guide Vision: Learning to See the Heart through Tabular Data, *Marta Hasny, Maxime Di Folco, Keno Bressem, Julia Schnabel*
30 SAFER-AiD: Saccade-Assisted Foveal-peripheral vision Enhanced Reconstruction for Adversarial Defense, *Jiayang Liu, Daniel Ts'o, Yiming Bu, Qinru Qiu*
31 Enhancing Object Detection Training via Joint Image-Annotation Generation, *Roy Uziel, Oded Bialer*
32 Descrip3D: Enhancing Large Language Model-based 3D Scene Understanding with Object-Level Text Descriptions, *Jintang Xue, Ganning Zhao, Jie-En Yao, Hong-En Chen, Yue Hu, Meida Chen, Suya You, C.-C. Jay Kuo*
33 From SAM to DINOv2: Towards Distilling Foundation Models to Lightweight Baselines for Generalized Polyp Segmentation, *Shivanshu Agnihotri, Snehashis Majhi, Deepak Ranjan Nayak, Debesh Jha*
34 Not Like Transformers: Drop the Beat Representation for Dance Generation with Mamba-Based Diffusion Model, *Sangjune Park, Inhyeok Choi, Donghyeon Soon, Youngwoo Jeon, Kyungdon Joo*
35 BoxSplitGen: A Generative Model for 3D Part Bounding Boxes in Varying Granularity, *Juil Koo, Wei-Tung Lin, Chanho Park, Chanhyeok Park, Minhyuk Sung*
36 3D Gaussian Point Encoders, *Jim James, Benjamin Wilson, Simon Lucey, James Hays*
37 HumanGuideNet: Adapter-Based Alignment of Deep Neural Networks with Human Similarity Judgments, *Xufu Liu, Yifan Yang, Zhengxin Zhang*
38 Low-Rank Expert Merging for Multi-Source Domain Adaptation in Person Re-Identification, *Taha Mustapha Nehdi, Nairouz Mrabah, Atif Belal, Marco Pedersoli, Eric Granger*
39 MEGA-PCC: A Mamba-based Efficient Approach for Joint Geometry and Attribute Point Cloud Compression, *Kai-Hsiang Hsieh, Monyneath Yim, Wen-Hsiao Peng, Jui-Chiu Chiang*
40 HyPCA-Net: Advancing Multimodal Fusion in Medical Image Analysis, *Joy Dhar, Manish Kumar Pandey, Debashis Das Chakladar, Maryam Haghighat, Azadeh Alavi, Sajib Mistry, Nayyar Zaidi*
41 CycleSL: Server-Client Cyclical Update Driven Scalable Split Learning, *Mengdi Wang, Efe Bozkir, Enkelejda Kasneci*
42 3DSceneEditor: Controllable 3D Scene Editing with Gaussian Splatting, *Ziyang Yan, Yihua Shao, Minwen Liao, Siyu Chen, Nan Wang, Muyuan Lin, Jenq-Neng Hwang, Hao Zhao, Fabio Remondino, Lei Li*
43 Segmentation-Aware Latent Diffusion for Satellite Image Super-Resolution: Enabling Smallholder Farm Boundary Delineation, *Aditi Agarwal, Anjali Jain, Nikita Saxena, Ishan Deshpande, Michal Kazmierski, Abigail Annkah, Nadav Sherman, Karthikeyan Shanmugam, Alok Talekar, Vaibhav Rajan*
44 mmWEAVER: Environment-Specific mmWave Signal Synthesis from a Photo and Activity Description, *Mahathir Monjur, Shahriar Nirjon*
45 Detection-Driven Object Count Optimization for Text-to-Image Diffusion Models, *Oz Zafar, Yuval Cohen, Lior Wolf, Idan Schwartz*
46 Roadside Monocular 3D Detection Prompted by 2D Detection, *Yechi Ma, Yanan Li, Wei Hua, Shu Kong*
47 UniCalib: Targetless LiDAR-camera Calibration via Probabilistic Flow on Unified Depth Representations, *Shu Han, Xubo Zhu, Ji Wu, Ximeng Cai, Wen Yang, Huai Yu, Gui-Song Xia*
48 Color Bind: Exploring Color Perception in Text-to-Image Models, *Shay Shomer-Chai, Wenxuan Peng, Bharath Hariharan, Hadar Averbuch-Elor*
49 ASC: Learning Augmentation Severity-Consistent Representations Improves Generalization via Augmentation Search, *Amirhossein Alamdar, Hossein Jafarinia, Mahdi Nouri, Mohammad Hossein Rohban*
50 Detecting Out-of-Distribution Objects through Class-Conditioned Inpainting, *Quang-Huy Nguyen, Jin Peng Zhou, Zhenzhen Liu, Khanh-Huyen Bui, Kilian Q. Weinberger, Wei-Lun Chao, Dung D. Le*
51 Snapmoji: Instant Generation of Animatable Dual-Stylized Avatars, *Eric Ming Chen, Di Liu, Sizhuo Ma, Michael Vasilkovsky, Bing Zhou, Qiang Gao, Wenzhou Wang, Jiahao Luo, Dimitris N. Metaxas, Vincent Sitzmann, Jian Wang*
52 Seeing is Believing (and Predicting): Context-Aware Multi-Human Behavior Prediction with Vision Language Models, *Utsav Panchal, Yuchen Liu, Luigi Palmieri, Ilche Georgievski, Marco Aiello*
53 False Alarm Rectification for Early Smoke Segmentation, *Hongjin Zhao, Weihao Li, Ge-Peng Ji, Nick Barnes*
54 Interaction-via-Actions: Cattle Interaction Detection with Joint Learning of Action-Interaction Latent Space, *Ren Nakagawa, Yang Yang, Risa Shinoda, Hiroaki Santo, Kenji Oyama, Fumio Okura, Takenao Ohkawa*
55 Semi-Supervised Hierarchical Open-Set Classification, *Erik Wallin, Fredrik Kahl, Lars Hammarstrand*
56 ClusterMine: Robust Label-Free Visual Out-Of Distribution Detection via Concept Mining from Text Corpora, *Nikolas Adaloglou, Diana Petrusheva, Mohamed Asker, Felix Michels, Markus Kollmann*
57 CURIO: Curvature-Aligned and Efficient OCR for Low-Resource Historical Manuscripts, *Sai Madhusudan Gunda, Tathagata Ghosh, Simran Singh Sandral, Ravi Kiran Sarvadevabhatla*
58 DoTA: Latent Distribution Conditioned Data Attribution for Diffusion Models, *Ninad Joshi, Vivek Srivastava, Shirish Karande*
59 Learnable Query-Enhanced Pose Transformation, *Yi-Zhen Wang, Hong-Han Shuai*
60 CAMP-VQA: Caption-Embedded Multimodal Perception for No-Reference Quality Assessment of Compressed Video, *Xinyi Wang, Angeliki Katsenou, Junxiao Shen, David Bull*

61 Evaluating Text-to-Image and Text-to-Video Synthesis with a Conditional Frechet Distance, *Jaywon Koo, Jefferson Hernandez, Moayed Haji-Ali, Ziyan Yang, Vicente Ordonez*
62 From Bands to Depth: Understanding Bathymetry Decisions on Sentinel-2, *Satyaki Roy Chowdhury, Aswathnarayan Radhakrishnan, Hari Subramoni*
63 Pyramidal Spectrum: Frequency-based Hierarchically Vector Quantized VAE for Videos, *Tushar Prakash, Onkar Susladkar, Sparsh Mittal, Inderjit S. Dhillon*
64 PRISM-CAFO: Prior-conditioned Remote-sensing Infrastructure Segmentation and Mapping for CAFOs, *Oishee Bintey Hoque, Nibir Chandra Mandal, Kyle Luong, Mandy L Wilson, Samarth Swarup, Madhav Marathe, Abhijin Adiga*
65 Dressing the Imagination: A Dataset for AI-Powered Translation of Text into Fashion Outfits and A Novel NeRA Adapter for Enhanced Feature Adaptation, *Gayatri Deshmukh, Somsubhra De, Chirag Sehgal, Jishu Sen Gupta, Sparsh Mittal*
66 EllipssianNet: Image-guided Sampling of 2D Gaussians for Gaussian Splatting, *MyoungGon Kim, JeongHyeon Ahn, Seohyeon Park, Hyemi Kim, Seunghyun Park, Jung Ho Hwang, JungHyun Han*
67 Zero-Shot Coreset Selection via Iterative Subspace Sampling, *Brent A. Griffin, Jacob Marks, Jason J. Corso*
68 MorphXAI: An Explainable Framework for Morphological Analysis of Parasites in Blood Smear Images, *Aqsa Yousaf, Sint Sint Win, Megan Coffee, Habeeb Olufowobi*
69 WWE-UIE: A Wavelet & White Balance Efficient Network for Underwater Image Enhancement, *Ching-Heng Cheng, Jen-Wei Lee, Chia-Ming Lee, Chih-Chung Hsu*
70 SPAR-Det: Segmentation-guided and Prior-Aided Routing for Small Object Detection, *Seungchan Kwon, Gyuil Lim, Youngjoon Han*
71 GeoHSAF: Geometric Hippocampus Shape Analysis Framework for Longitudinal Alzheimer's Disease Classification, *Mubarak Olaoluwa, Heni Loukil, Arafet Sbei, Hassen Drira*
72 BAFIS: Dataset + Framework to Assess Occupational Bias and Human Preference in Modern Text-to-image Models, *Thomas Klassert, Adrian Ulges, Biying Fu*
73 Imitating the Functionality of Image-to-Image Models Using a Single Example, *Nurit Spingarn, Tomer Michaeli*
74 Beyond the Highlights: Video Retrieval with Salient and Surrounding Contexts, *Jaehun Bang, Moon Ye-Bin, Tae-Hyun Oh, Kyungdon Joo*
75 SCORE: Soft Label Compression-Centric Dataset Condensation via Coding Rate Optimization, *Bowen Yuan, Yuxia Fu, Zijian Wang, Zi Huang, Yadan Luo*
76 Sketch2Stitch: GANs for Abstract Sketch-Based Dress Synthesis, *Faizan Farooq Khan, Eslam Abdelrahman Bakr, Davide Morelli, Marcella Cornia, Rita Cucchiara, Mohamed Elhoseiny*
77 AnyBald: Toward Realistic Diffusion-Based Hair Removal In-The-Wild, *Yongjun Choi, Seungoh Han, Soomin Kim, Sumin Son, Mohsen Rohani, Edgar Maucourant, Dongbo Min, Kyungdon Joo*
78 Understanding Generative AI Capabilities in Everyday Image Editing Tasks, *Brandon Collins, Mohammad Reza Taesiri, Logan Bolton, Viet Dac Lai, Franck Dernoncourt, Trung Bui, Anh Totti Nguyen*
79 Reconstructing Realistic and Relightable Eyes, *Wesley Khademi, Jogendra Kundu, Yatong An, Alexander Fix, David Colmenares*
80 1LoRA: Summation Compression for Very Low-Rank Adaptation, *Alessio Quercia, Zhuo Cao, Arya Bangun, Richard D. Paul, Abigail Morrison, Ira Assent, Hanno Scharr*
81 Illuminating Darkness: Learning to Enhance Low-light Images In-the-Wild, *S. M. A. Sharif, Abdur Rehman, Zain Ul Abidin, Fayaz Ali Dharejo, Radu Timofte, Rizwan Ali Naqvi*
82 DOODLE: Diffusion-based Out-of-Distribution Learning for Open-set LiDAR Semantic Segmentation, *Changgyoon Oh, Hyeonseong Kim, Daehyun We, Jongoh Jeong, Yujeong Chae, Kuk-Jin Yoon*
83 RobustFormer: Noise-Robust Pre-training for Images and Videos, *Ashish Bastola, Nishant Luitel, Hao Wang, Danda Pani Paudel, Roshni Poudel, Abolfazl Razi*
84 CVP: Central-Peripheral Vision-Inspired Multimodal Model for Spatial Reasoning, *Zeyuan Chen, Xiang Zhang, Haiyang Xu, Jianwen Xie, Zhuowen Tu*
85 Trajectory Tactics: When Transformers Learn Exploration to Generate Online Signature, *Anurag Pandey, Aditya Nigam, Arnav Bhavsar, Ashutosh Sharma, Basu Verma, Divya Acharya, Mohd Amir*
86 BrandFusion: Aligning Image Generation with Brand Styles, *Parul Gupta, Varun Khurana, Yaman Kumar Singla, Balaji Krishnamurthy, Abhinav Dhall*
87 Intra-Class Probabilistic Embeddings for Uncertainty Estimation in Vision-Language Models, *Zhenxiang Lin, Maryam Haghighat, Will Browne, Dimity Miller*
88 FARF-Net: Frequency-guided Adaptive Receptive Field Network for Edge-enhanced Polyp Segmentation, *Xue Li, Aiwen Jiang, Hongqian Yu, Yang Xiao*
89 Discrete Facial Encoding: A Framework for Data-driven Facial Display Discovery, *Minh Tran, Maksim Siniukov, Zhangyu Jin, Mohammad Soleymani*
90 Knowledge to Sight: Reasoning over Visual Attributes via Knowledge Decomposition for Abnormality Grounding, *Jun Li, Che Liu, Wenjia Bai, Mingxuan Liu, Rossella Arcucci, Cosmin I. Bercea, Julia A. Schnabel*
91 DenseBEV: Transforming BEV Grid Cells into 3D Objects, *Marius Dähling, Sebastian Krebs, J. Marius Zöllner*
92 Human Knowledge Integrated Multi-modal Learning for Single Source Domain Generalization, *Ayan Banerjee, Kuntal Thakur, Sandeep Gupta*
93 GraspDiffusion: Synthesizing Realistic Whole-body Hand-Object Interaction, *Patrick Kwon, Chen Chen, Hanbyul Joo*
94 ScoliGaitX: A Deep Multi-Modal Fusion Network for Scoliosis Assessment via Gait Video Analysis, *Kaushik Vishwakarma, Aditya Nigam*
95 Non-Contact Blood Pressure Estimation from Face Videos via Physiology-Aware Contrastive Learning, *JaeHyuk Son, Young-Seok Choi*
96 Ordinal-Aware Multimodal Engagement Recognition for Collaborative Learning, *Nha Tran, Dat Ly, Phi Ta, Hung Nguyen, Hien D. Nguyen*
97 DynaGSLAM: Real-Time Gaussian-Splatting SLAM for Online Rendering, Tracking, Motion Predictions of Moving Objects in Dynamic Scenes, *Runfa Blark Li, Mahdi Shaghaghi, Keito Suzuki, Xinshuang Liu, Varun Moparthi, Bang Du, Walker Curtis, Martin Renschler, Ki Myung Brian Lee, Nikolay Atanasov, Truong Nguyen*
98 Test-Time Adaptation through Semantically-guided Feature Decomposition for Few-shot Chest X-ray Diagnosis, *Jayant Mahawar, Angshuman Paul*
99 FlowMorph: Revealing an Optimizable Flow Latent Space for Controlled Image Morphing, *Yan Zheng, Yi Yang, Lanqing Guo, Zhangyang Wang*
100 PVeRA: Probabilistic Vector-Based Random Matrix Adaptation, *Leo Fillioux, Enzo Ferrante, Paul-Henry Cournède, Maria Vakalopoulou, Stergios Christodoulidis*
101 Harnessing Object Grounding for Time-Sensitive Video Understanding, *Tz-Ying Wu, Sharath Nittur Sridhar, Subarna Tripathi*
102 TaxonRL: Reinforcement Learning with Intermediate Rewards for Interpretable Fine-Grained Visual Reasoning, *Siddhant Gole, Akash Pal, Ankit Jha, Subhasis Chaudhuri, Biplab Banerjee*
103 Revisiting Retentive Networks for Fast Range-View 3D LiDAR Semantic Segmentation, *Simone Mosco, Daniel Fusaro, Wanmeng Li, Alberto Pretto*
104 Zero-shot Hierarchical Plant Segmentation via Foundation Segmentation Models and Text-to-image Attention, *Junhao Xing, Ryohei Miyakawa, Yang Yang, Xinpeng Liu, Risa Shinoda, Hiroaki Santo, Yosuke Toda, Fumio Okura*
105 Moiré Zero: An Efficient and High-Performance Neural Architecture for Moiré Removal, *Seungryong Lee, Woojeong Baek, Younghyun Kim, Eunwoo Kim, Haru Moon, Donggon Yoo, Eunbyung Park*

106 LENVIZ: A High-Resolution Low-Exposure Night Vision Benchmark Dataset, *Manjushree Aithal, Rosaura G. Vidal Mata, Manikandtan Kartha, Gong Chen, Eashan Adhikarla, Lucas N. Kirsten, Zhicheng Fu, Nikhil A. Madhusudhana, Joe Nasti*
107 RobustGait: Robustness Analysis for Appearance Based Gait Recognition, *Reeshoon Sayera, Akash Kumar, Sirshapan Mitra, Prudvi Kamtam, Yogesh S Rawat*
108 A-V Representation Learning via Audio Shift Prediction for Multimodal Deepfake Detection and Temporal Localization, *Ashutosh Anshul, Eng Siong Chng, Deepu Rajan*
109 CraftSVG: Multi-Object Text-to-SVG Synthesis via Layout Guided Diffusion, *Ayan Banerjee, Nityanand Mathur, Josep Llados, Umapada Pal, Anjan Dutta*
110 Shift-Equivariant Complex-Valued Convolutional Neural Networks, *Quentin Gabot, Teck-Yian Lim, Jeremy Fix, Joana Frontera-Pons, Chengfang Ren, Jean-Philippe Ovarlez*
111 ObjectMeshDeform : Towards Recovering Precise 3D Geometry of Real Objects via Image-guided Mesh Deformation of 3D Generative Priors, *Siddharth Katageri, Sanjana Sinha, Sourav Ghosh, Soumyadip Maity, Brojeshwar Bhowmick*
112 Memory-Augmented Representation for Efficient Event-based Visuomotor Policy Learning with Adaptive Perception and Control, *Uday Kamal, Saibal Mukhopadhyay*
113 PADM: A Physics-aware Diffusion Model for Attenuation Correction, *Trung Kien Pham, Hoang Minh Vu, Anh Duc Chu, Dac Thai Nguyen, Trung Thanh Nguyen, Thao Nguyen Truong, Mai Hong Son, Thanh Trung Nguyen, Phi Le Nguyen*
114 NerVast: Compression-Efficient Scaling of Implicit Neural Video Representations via Scene-based Parameter-sharing, *Yunheon Lee, Jaehong Kim, Juncheol Ye, Dongsu Han*
115 CineVerse: Consistent Keyframe Synthesis for Cinematic Scene Composition, *Quynh Phung, Long Mai, Fabian David Caba Heilbron, Feng Liu, Jia-Bin Huang, Cusuh Ham*
116 MMCM: Multimodality-aware Metric using Clustering-based Modes for Probabilistic Human Motion Prediction, *Kyotaro Tokoro, Hiromu Taketsugu, Norimichi Ukita*
117 Face-LLaVA: Facial Expression and Attribute Understanding through Instruction Tuning, *Ashutosh Chaubey, Xulang Guan, Mohammad Soleymani*
118 ConsensusXAI: A Framework to Examine Class-wise Agreement in Medical Imaging, *Abbas Haider, David Wright, Ruth Hogg, Hui Wang, Tunde Peto, Richard Gault*
119 Real-Time Tracking of Flexible Markers in Low-Contrast Fluoroscopy Using a Deep Neural Network Trained Solely on Synthetic Data, *Tomoki Uchiyama, Yukinobu Sakata, Ryusuke Hirai, Hitoshi Ishikawa, Shinichiro Mori*
120 OpenLVLM-MIA: A Controlled Benchmark Revealing the Limits of Membership Inference Attacks on Large Vision-Language Models, *Ryoto Miyamoto, Xin Fan, Fuyuko Kido, Tsuneo Matsumoto, Hayato Yamana*
121 DMAT: An End-to-End Framework for Joint Atmospheric Turbulence Mitigation and Object Detection, *Paul Hill, Zhiming Liu, Alin Achim, David Bull, Nantheera Anantrasirichai*
122 Divide and Refine: Enhancing Multimodal Representation and Explainability for Emotion Recognition in Conversation, *Anh-Tuan Mai, Cam-Van Thi Nguyen, Duc-Trong Le*
123 iMotion-LLM: Instruction-Conditioned Trajectory Generation, *Abdulwahab Felemban, Nussair Hroub, Jian Ding, Eslam Abdelrahman, Xiaoqian Shen, Abduallah Mohamed, Mohamed Elhoseiny*
124 SasMamba: A Lightweight Structure-Aware Stride State Space Model for 3D Human Pose Estimation, *Hu Cui, Wenqiang Hua, Renjing Huang, Shurui Jia, Tessai Hayama*
125 SUGAR: A Sweeter Spot for Generative Unlearning of Many Identities, *Dung Thuy Nguyen, Quang Nguyen, Preston K. Robinette, Eli Jiang, Taylor T. Johnson, Kevin Leach*
126 Mitigating the Modality Gap: Few-Shot Out-of-Distribution Detection with Multi-modal Prototypes and Image Bias Estimation, *Yimu Wang, Evelien Riddell, Adrian Chow, Sean Sedwards, Krzysztof Czarnecki*
127 Matching Semantically Similar Non-Identical Objects, *Yusuke Marumo, Kazuhiko Kawamoto, Satomi Tanaka, Shigenobu Hirano, Hiroshi Kera*
128 UI-Styler: Ultrasound Image Style Transfer with Class-Aware Prompts for Cross-Device Diagnosis Using a Frozen Black-Box Inference Network, *Nhat-Tuong Do-Tran, Ngoc-Hoang-Lam Le, Ching-Chun Huang*
129 Tables Decoded: DELTA for Structure, TARQA for Understanding, *Jahanvi Rajput, Dhruv Kudale, Saikiran Kasturi, Utkarsh Verma, Ganesh Ramakrishnan*
130 What Happens When: Learning Temporal Orders of Events in Videos, *Daechul Ahn, Yura Choi, Hyeonbeom Choi, Seongwon Cho, San Kim, Jonghyun Choi*
131 UNO: Unifying One-stage Video Scene Graph Generation via Object-Centric Visual Representation Learning, *Huy Le, Nhat Chung, Tung Kieu, Jingkang Yang, Ngan Le*
132 Personalized Image Privacy Advisors via Federated Daisy-Chaining, *Sourasekhar Banerjee, Vengateswaran Subramaniam, Debaditya Roy, Vigneshwaran Subbaraju, Monowar Bhuyan*
133 DiRe: Diversity-promoting Regularization for Dataset Condensation, *Saumyaranjan Mohanty, Aravind Reddy, Konda Reddy Mopuri*
134 Vision-informed Semantic Text Alignment for Open-set Recognition in Remote Sensing, *Maximilian von Klinski, Maximilian Schall*
135 Anatomy-VLM: A Fine-grained Vision-Language Model for Medical Interpretation, *Difei Gu, Yunhe Gao, Mu Zhou, Dimitris Metaxas*
136 Temporal Object Captioning for Street Scene Videos from LiDAR Tracks, *Vignesh Gopinathan, Urs Zimmermann, Michael Arnold, Matthias Rottmann*
137 STARS: Self-supervised Tuning for 3D Action Recognition in Skeleton Sequences, *Soroush Mehraban, Mohammad Javad Rajabi, Andrea Iaboni, Babak Taati*
138 RAVU: Retrieval Augmented Video Understanding with Compositional Reasoning over Graph, *Sameer Malik, Ayush Singh, Moyuru Yamada, Dishank Aggarwal*
139 Conditional Text-to-Image Generation with Reference Guidance, *Taewook Kim, Ze Wang, Zhengyuan Yang, Jiang Wang, Lijuan Wang, Zicheng Liu, Qiang Qiu*
140 Improved Wildfire Spread Prediction with Time-Series Data and the WSTS+ Benchmark, *Saad Lahrichi, Jake Bova, Jesse Johnson, Jordan Malof*
141 From Few-Shot to Zero-Shot Pallet Load Recognition: A Deployed Embedding-Based Vision System for Industrial Logistics, *Juan Jesús Losada del Olmo, Emilio Pardo Ballesteros, Pedro E. López-de-Teruel, Alberto Ruiz*
142 Graph-Based Spectral Attention with Multi-Spectral Images for Illuminant Estimation, *Dong-Hoon Kang, Seung-Yeop Baek, Jong-Ok Kim*

Monday, March 9

Reminder that all events are for registered attendees only.

8:00-17:00 **Registration,** AZ Ballroom Registration Desk
9:00-16:00 **Poster Pickup ,** Tucson Ballroom Registration Desk
8:30-9:30 **Keynote Talk 2,** AZ Ballroom 6 (streamed to 7)
Speaker: Dorin Comaniciu, **Title:** Applications of Computer Vision in Healthcare: The Road to Autonomy
9:30-9:45 **Courtesy Break,** AZ Ballroom Pre-Function
9:45-10:45 **Oral Session 4A: Image Recognition and Understanding II,** AZ Ballroom 6
9:45-10:45 **Oral Session 4B: Machine Learning I,** AZ Ballroom 7
10:45-18:00 **Exhibits + Demos,** Tucson Ballroom
10:45-12:30 **Poster Session 3,** Tucson Ballroom & Prefunction Space
12:00-14:00 **Doctoral Consortium Event,** San Luis, *The Doctoral consortium is by invitation only.*
12:00-13:30 **Lunch,** Ania Terrace and Lawn
13:30-14:30 **Oral Session 5A: Generative Models II,** AZ Ballroom 6
13:30-14:30 **Oral Session 5B: Remote Sensing and Sensors,** AZ Ballroom 7
14:30-14:45 **Courtesy Break**
14:45-15:45 **Oral Session 6A: 3D Computer Vision II,** AZ Ballroom 6
14:45-15:45 **Oral Session 6B: Video Recognition and Understanding I,** AZ Ballroom 7
15:45-16:30 **PAMI-TC meeting,** AZ Ballroom Salon 6
16:30-18:15 **Poster Session 4 + Reception,** Tucson Ballroom & Prefunction Space

8:30-9:30 **Keynote Talk 2,** AZ Ballroom 6 (streamed to 7)
Speaker: Dorin Comaniciu, **Title:** Applications of Computer Vision in Healthcare: The Road to Autonomy

9:30 - 9:45 **Courtesy Break**

The papers in this oral session will also be presented as posters.
Format: 10 min. presentation, 2 min. questions

9:45-10:45 **Oral Session 4A: Image Recognition and Understanding II,** AZ Ballroom 6

1. Fast Vision Mamba: Pooling Spatial Dimensions for Accelerated Processing, *Saarthak Kapse, Robin M. Betz, Srinivasan Sivanandan*
2. Extreme Amodal Face Detection, *Changlin Song, Yunzhong Hou, Michael Randall Barnes, Rahul Shome, Dylan Campbell*
3. ENCORE : A Neural Collapse Perspective on Out-of-Distribution Detection in Deep Neural Networks, *A. Q. M. Sazzad Sayyed, Nathaniel D. Bastian, Francesco Restuccia*
4. Performance of Conformal Prediction in Capturing Aleatoric Uncertainty, *Misgina Tsighe Hagos, Claes Lundström*

9:45-10:45 **Oral Session 4B: Machine Learning I,** AZ Ballroom 7

1. Scalpel: Fine-Grained Alignment of Attention Activation Manifolds via Mixture Gaussian Bridges to Mitigate Multimodal Hallucination, *Ziqiang Shi, Rujie Liu, Shanshan Yu, Satoshi Munakata, Koichi Shirahata*
2. Unified Alignment Protocol: Making Sense of the Unlabeled Data in New Domains, *Sabbir Ahmed, Mamshad Nayeem Rizve, Abdullah Al Arafat, Jacqueline Tiffany Liu, Rahim Hossain, Mohaiminul Al Nahian, Adnan Siraj Rakin*
3. Feedback Alignment Meets Low-Rank Manifolds: A Structured Recipe for Local Learning, *Arani Roy, Marco P. Apolinario, Shristi Das Biswas, Kaushik Roy*
4. Learning from Unknown for Open-Set Test-Time Adaptation, *Taki Hasan Rafi, Amit Agarwal, Hitesh L. Patel, Dong-Kyu Chae*
5. Streaming Real-Time Trajectory Prediction Using Endpoint-Aware Modeling, *Alexander Prutsch, David Schinagl, Horst Possegger*

10:45-18:00 **Exhibits + Demos,** Tucson Ballroom

1. GorillaWatch: A Platform for Large-Scale In-the-Wild Gorilla Re-Identification and Population Monitoring, *Maximilian Schall, Hasso Plattner Institute*
2. Demo on setup of xLAM the Large Action Model, *Pradnya Desai, Salesforce*
3. Hybrid Rendering for Multimodal Autonomous Driving: Merging Neural and Physics-Based Simulation, *Máté Tóth, aiMotive*
4. Live Demonstration: Field Work AI Agent Using Unsupervised Discovery of Long-Term Spatiotemporal Periodic Workflows, *Takuto Sato, Fujitsu*

10:45-12:30 **Poster Session 3,** Tucson Ballroom & Prefunction Space

1 Fast Vision Mamba: Pooling Spatial Dimensions for Accelerated Processing, *Saarthak Kapse, Robin M. Betz, Srinivasan Sivanandan*
2 Extreme Amodal Face Detection, *Changlin Song, Yunzhong Hou, Michael Randall Barnes, Rahul Shome, Dylan Campbell*
3 ENCORE : A Neural Collapse Perspective on Out-of-Distribution Detection in Deep Neural Networks, *A. Q. M. Sazzad Sayyed, Nathaniel D. Bastian, Francesco Restuccia*
4 Performance of Conformal Prediction in Capturing Aleatoric Uncertainty, *Misgina Tsighe Hagos, Claes Lundström*
5 Scalpel: Fine-Grained Alignment of Attention Activation Manifolds via Mixture Gaussian Bridges to Mitigate Multimodal Hallucination, *Ziqiang Shi, Rujie Liu, Shanshan Yu, Satoshi Munakata, Koichi Shirahata*
6 Unified Alignment Protocol: Making Sense of the Unlabeled Data in New Domains, *Sabbir Ahmed, Mamshad Nayeem Rizve, Abdullah Al Arafat, Jacqueline Tiffany Liu, Rahim Hossain, Mohaiminul Al Nahian, Adnan Siraj Rakin*
7 Feedback Alignment Meets Low-Rank Manifolds: A Structured Recipe for Local Learning, *Arani Roy, Marco P. Apolinario, Shristi Das Biswas, Kaushik Roy*
8 Learning from Unknown for Open-Set Test-Time Adaptation, *Taki Hasan Rafi, Amit Agarwal, Hitesh L. Patel, Dong-Kyu Chae*
9 StreetView-Waste: A Multi-Task Dataset for Urban Waste Management, *Diogo J. Paulo, João Martins, Hugo Proença, João C. Neves*
10 AnyAnomaly: Zero-Shot Customizable Video Anomaly Detection with LVLM, *Sunghyun Ahn, Youngwan Jo, Kijung Lee, Sein Kwon, Inpyo Hong, Sanghyun Park*
11 SkelSplat: Robust Multi-view 3D Human Pose Estimation with Differentiable Gaussian Rendering, *Laura Bragagnolo, Leonardo Barcellona, Stefano Ghidoni*
12 Evaluating the Capability of Video Question Generation for Expert Knowledge Elicitation, *Huaying Zhang, Atsushi Hashimoto, Tosho Hirasawa*
13 Dragonite: Single-Step Drag-based Image Editing with Geometric-Semantic Guidance, *Meng-Ting Jhong, Tai-Ming Huang, Shang-Fu Chen, Wen-Huang Cheng, Kai-Lung Hua*
14 Augmenting with NeRFs: Fast Relocalization on Densified Datasets, *Michael Tomadakis, Rebecca Borissova, Yuxuan Zhang, Sanjeev Koppal*
15 GASP: Unifying Geometric and Semantic Self-Supervised Pre-training for Autonomous Driving, *William Ljungbergh, Adam Lilja, Adam Tonderski, Arvid Laveno Ling, Carl Lindström, Willem Verbeke, Junsheng Fu, Christoffer Petersson, Lars Hammarstrand, Michael Felsberg*
16 Towards Reliable Test-Time Adaptation: Style Invariance as a Correctness Likelihood, *Gilhyun Nam, Taewon Kim, Joonhyun Jeong, Eunho Yang*
17 TalkingPose: Efficient Face and Gesture Animation with Feedback-guided Diffusion Model, *Alireza Javanmardi, Pragati Jaiswal, Tewodros Amberbir Habtegebrial, Christen Millerdurai, Shaoxiang Wang, Alain Pagani, Didier Stricker*
18 Large Sign Language Models: Toward 3D American Sign Language Translation, *Sen Zhang, Xiaoxiao He, Di Liu, Zhaoyang Xia, Mingyu Zhao, Chaowei Tan, Vivian Li, Bo Liu, Dimitris N. Metaxas, Mubbasir Kapadia*
19 Crafting Descriptive Information for a Zero-shot Method to Improve Knowledge-Based Visual Question Answering Performance, *Mohammad Mahdi Moradi, Sudhir Mudur*

20 Learning Action Hierarchies via Hybrid Geometric Diffusion, *Arjun Ramesh Kaushik, Nalini K. Ratha, Venu Govindaraju*
21 Learning Group Actions In Disentangled Latent Image Representations, *Farhana Hossain Swarnali, Miaomiao Zhang, Tonmoy Hossain*
22 GAITGen: Disentangled Motion-Pathology Impaired Gait Generative Model -- Bringing Motion Generation to the Clinical Domain, *Vida Adeli, Soroush Mehraban, Majid Mirmehdi, Alan L. Whone, Benjamin Filtjens, Amirhossein Dadashzadeh, Alfonso Fasano, Andrea Iaboni, Babak Taati*
23 Gradient-Free Classifier Guidance for Diffusion Model Sampling, *Rahul Shenoy, Zhihong Pan, Kaushik Balakrishnan, Qisen Cheng, Yongmoon Jeon, Heejune Yang, Jaewon Kim*
24 Mitigating Object and Action Hallucinations in Multimodal LLMs via Self-Augmented Contrastive Alignment, *Kai-Po Chang, Wei-Yuan Cheng, Chi-Pin Huang, Fu-En Yang, Yu-Chiang Frank Wang*
25 Guided Model Merging for Hybrid Data Learning: Leveraging Centralized Data to Refine Decentralized Models, *Junyi Zhu, Ruicong Yao, Taha Ceritli, Savas Ozkan, Matthew B. Blaschko, Eunchung Noh, Jeongwon Min, Cho Jung Min, Mete Ozay*
26 Fused Similarity Measure Based Alignment with Dual-Scale Adaptive Selection for Weakly Supervised Video Anomaly Detection, *Yue-Gao Lu, Hong-Jie Xing, Chun-Guo Li*
27 PointNet4D: A Lightweight 4D Point Cloud Video Backbone for Online and Offline Perception in Robotic Applications, *Yunze Liu, Zifan Wang, Peiran Wu, Jiayang Ao*
28 Cross-Modal Event Encoder: Bridging Image–Text Knowledge to Event Streams, *Sungheon Jeong, Hanning Chen, Sanggeon Yun, Suhyeon Cho, Wenjun Huang, Xiangjian Liu, Mohsen Imani*
29 SAIL: Self-supervised Learning of Lighting-Invariant Representations from Real Images with Latent Diffusion, *Hala Djeghim, Céline Loscos, Désiré Sidibé*
30 PSDiffusion: Harmonized Multi-Layer Image Generation via Layout and Appearance Alignment, *Dingbang Huang, Wenbo Li, Yifei Zhao, Xinyu Pan, Yanhong Zeng, Bo Dai*
31 NERVE: Neighbourhood & Entropy-Guided Random-Walk for Training Free Open-Vocabulary Segmentation, *Kunal Mahatha, Jose Dolz, Christian Desrosiers*
32 CLIP's Visual Embedding Projector is a Few-shot Cornucopia, *Mohammad Fahes, Tuan-Hung Vu, Andrei Bursuc, Patrick Perez, Raoul De Charette*
33 RegionAligner: Bridging Ego-Exo Views for Object Correspondence via Unified Text-Visual Learning, *Yuhao Su, Ehsan Elhamifar*
34 Towards Streaming LiDAR Object Detection with Point Clouds as Egocentric Sequences, *Mellon M. Zhang, Glen Chou, Saibal Mukhopadhyay*
35 Show Me: Unifying Instructional Image and Video Generation with Diffusion Models, *Yujiang Pu, Zhanbo Huang, Vishnu Boddeti, Yu Kong*
36 VOCAL: Visual Odometry via ContrAstive Learning, *Chi-Yao Huang, Zeel Shaileshkumar Bhatt, Yezhou Yang*
37 Pose-Diverse Multi-View Virtual Try-on from a Single Frontal Image via Diffusion Transformer, *Seonghee Han, Minchang Chung, Gyeongsu Cho, Kyungdon Joo, Taehwan Kim*
38 SimForce: Force and Surface Electromyography from Full Body Video with Graph Neural Nets, *Esha Dasgupta, Boeun Kim, Sang Hoon Yeo, Hyung Jin Chang*
39 SCATR: Mitigating New Instance Suppression in LiDAR-based Tracking-by-Attention via Second Chance Assignment and Track Query Dropout, *Brian Cheong, Letian Wang, Sandro Papais, Steven L. Waslander*
40 AirLock+: Scaling UAV-to-Satellite Image Registration for Target Geolocalization and Geospatial Augmented Reality, *Zhiyun Deng, Austin Case, Luis Sentis*
41 Blur2Sharp: Human Novel Pose and View Synthesis with Generative Prior Refinement, *Chia-Hern Lai, I-Hsuan Lo, Yen-Ku Yeh, Thanh-Nguyen Truong, Ching-Chun Huang*
42 Power of Boundary and Reflection: Semantic Transparent Object Segmentation using Pyramid Vision Transformer with Transparent Cues, *Tuan-Anh Vu, Hai Nguyen-Truong, Ziqiang Zheng, Binh-Son Hua, Qing Guo, Ivor W. Tsang, Sai-Kit Yeung*
43 HEART-PFL: Stable Personalized Federated Learning under Heterogeneity with Hierarchical Directional Alignment and Adversarial Knowledge Transfer, *Minjun Kim, Minje Kim*
44 Detecting Social Engagement of Elderly From Lifelog Image-streams to Identify Effective Cues for Autobiographic Recall, *Vengateswaran Subramaniam, Vigneshwaran Subbaraju, Debaditya Roy, Pramath Krishna, Thivya Kandappu, Qianli Xu*
45 HistoMILKD: A Multiple Instance Learning based Multi-Teacher Knowledge Distillation Framework for Whole Slide Image Classification, *Mayur Mallya, Ali Khajegili Mirabadi, Hossein Farahani, Ali Bashashati*
46 Stroke Modeling Enables Vectorized Character Generation with Large Vectorized Glyph Model, *Xinyue Zhang, Haolong Li, Jiawei Ma, Chen Ye*
47 CasTex: Cascaded Text-to-Texture Synthesis via Explicit Texture Maps and Physically-Based Shading, *Mishan Aliev, Dmitry Baranchuk, Kirill Struminsky*
48 DATTA: Domain-Adversarial Test-Time Adaptation for Cross-Domain WiFi-Based Human Activity Recognition, *Julian Strohmayer, Rafael Sterzinger, Matthias Wödlinger, Martin Kampel*
49 MixER: From Cross-Modal to Mixed-Modal Visible-Infrared Re-Identification, *Mahdi Alehdaghi, Rajarshi Bhattacharya, Dai Yannick, Pourya Shamsolmoali, Rafael M. O. Cruz, Eric Granger*
50 LighthouseGS: Indoor Structure-aware 3D Gaussian Splatting for Panorama-Style Mobile Captures, *Seungoh Han, Jaehoon Jang, Hyunsu Kim, Jaeheung Surh, Junhyung Kwak, Hyowon Ha, Kyungdon Joo*
51 FLARES: Fast and Accurate LiDAR Multi-Range Semantic Segmentation, *Bin Yang, Alexandru Paul Condurache*
52 Language Integration in Fine-Tuning Multimodal Large Language Models for Image-Based Regression, *Roy H. Jennings, Genady Paikin, Roy Shaul, Evgeny Soloveichik*
53 MANTA: Physics-Informed Generalized Underwater Object Tracking, *Suhas Srinath, Hemang Jamadagni, Aditya Chandrasekar, Prathosh A P*
54 KD360-VoxelBEV: LiDAR and 360-degree Camera Cross Modality Knowledge Distillation for Bird's-Eye-View Segmentation, *Wenke E, Yixin Sun, Jiaxu Liu, Hubert P. H. Shum, Amir Atapour-Abarghouei, Toby P. Breckon*
55 MAFM^3: Modular Adaptation of Foundation Models for Multi-Modal Medical AI, *Mohammad Areeb Qazi, Munachiso S Nwadike, Ibrahim Almakky, Mohammad Yaqub, Numan Saeed*
56 SPOC: Spatially-Progressing Object State Change Segmentation in Video, *Priyanka Mandikal, Tushar Nagarajan, Alex Stoken, Zihui Xue, Kristen Grauman*
57 Self-Supervised Visual Prompting for Cross-Domain Road Damage Detection, *Xi Xiao, Zhuxuanzi Wang, Mingqiao Mo, Chen Liu, Chenrui Ma, Yanshu Li, Smita Krishnaswamy, Xiao Wang, Tianyang Wang*
58 Data-Driven Loss Functions for Inference-Time Optimization in Text-to-Image, *Sapir Esther Yiflach, Yuval Atzmon, Gal Chechik*
59 Multi-Modal Soccer Scene Analysis with Masked Pre-Training, *Marc Peral, Guillem Capellera, Luis Ferraz, Antonio Rubio, Antonio Agudo*
60 Visual Detector Compression via Location-Aware Discriminant Analysis, *Qizhen Lan, Jung Im Choi, Qing Tian*
61 Latent Uncertainty-Aware Multi-View SDF Scan Completion, *Faezeh Zakeri, Lukas Ruppert, Raphael Braun, Hendrik P.A. Lensch*
62 Comp4D: Compositional 4D Scene Generation, *Hanwen Liang, Dejia Xu, Neel P. Bhatt, Hezhen Hu, Hanxue Liang, Konstantinos N. Plataniotis*
63 SENCA-st: Integrating Spatial Transcriptomics and Histopathology with Cross Attention Shared Encoder for Region Identification in Cancer Pathology, *Shanaka Liyanaarachchi, Chathurya Wijethunga, Shihab Aaqil Ahamed, Akthas Absar, Ranga Rodrigo*

64 Towards High-Fidelity, Identity-Preserving Real-Time Makeup Transfer: Decoupling Style Generation, *Lydia Chau, Zhi Yu, Ruowei Jiang*
65 Feature Inversion as a Lens on Vision Encoders, *Eduard Allakhverdov, Dmitrii Tarasov, Elizaveta Goncharova, Andrey Kuznetsov*
66 MuSACo: Multimodal Subject-Specific Selection and Adaptation for Expression Recognition with Co-Training, *Muhammad Osama Zeeshan, Natacha Gillet, Alessandro Lameiras Koerich, Marco Pedersoli, Francois Bremond, Eric Granger*
67 SCALEX: Scalable Concept and Latent Exploration for Diffusion Models, *E. Zhixuan Zeng, Yuhao Chen, Alexander Wong*
68 DCSHARP: 3D Gaussian Splatting with Direction Cosine Spherical Harmonics and Shape-Aware Pruning, *Ahmed Hasssan, Jian Meng, Yuanbo Xiangli, Jae-sun Seo*
69 DOTGraph: CLIP-Driven Feature Disentanglement and Optimal Transport based Graph Learning for Few-Shot Segmentation, *Shreya Biswas, Zhaozheng Yin*
70 Delta-LLaVA: Base-then-Specialize Alignment for Token-Efficient Vision-Language Models, *Mohamad Zamini, Diksha Shukla*
71 Sketch-guided Cage-based 3D Gaussian Splatting Deformation, *Tianhao Xie, Noam Aigerman, Eugene Belilovsky, Tiberiu Popa*
72 Autoregressive Styled Text Image Generation, but Make it Reliable, *Carmine Zaccagnino, Fabio Quattrini, Vittorio Pippi, Silvia Cascianelli, Alessio Tonioni, Rita Cucchiara*
73 Universal Neural Architecture Space: Covering ConvNets, Transformers and Everything in Between, *Ondrej Tybl, Lukas Neumann*
74 Sea-CLIP: Mining Semantic-Aware Representations for Few-Shot Anomaly Detection with CLIP, *Xiao Guo, Zhimin Chen, Carlos D. Castillo, Hongcheng Wang, Xiaoming Liu*
75 CLIP-IT: CLIP-based Pairing of Histology Images with Privileged Textual Information, *Banafsheh Karimian, Giulia Avanzato, Soufiane Belharbi, Alexis Guichemerre, Luke McCaffrey, Mohammadhadi Shateri, Eric Granger*
76 LightGazeNet: A Lightweight GNN-based Architecture for Gaze Estimation, *Heena Patel, Anirban Chowdhury, Pooja Jigar Choksy, Samiksha Pradeep Pachade, Ajinkya Puar*
77 Codebook Knowledge with Mamba-Transformer For Low-Light Image Enhancement, *Runhua Deng, Aiwen Jiang, Qiuhai Yan, Long Peng*
78 TiCLS: Tightly Coupled Language Text Spotter, *Leeje Jang, Yijun Lin, Yao-Yi Chiang, Jerod Weinman*
79 Perception-Inspired Color Space Design for Photo White Balance Editing, *Yang Cheng, Ziteng Cui, Lin Gu, Shenghan Su, Zenghui Zhang*
80 CLUE: Bringing Machine Unlearning to Mobile Devices, *A.Q.M. Sazzad Sayyed, Nathaniel D. Bastian, Michael De Lucia, Ananthram Swami, Francesco Restuccia*
81 FairScene: Learning Class-Disentangled 2D/3D Representations for Semantic Scene Completion, *Dian Jia, Pei Yu, Wei Tang*
82 Frequency Is What You Need: Considering Word Frequency When Text Masking Benefits Vision-Language Model Pre-training, *Mingliang Liang, Martha Larson*
83 Event-based Graph Representation with Spatial and Motion Vectors for Asynchronous Object Detection, *Aayush Atul Verma, Arpitsinh Vaghela, Bharatesh Chakravarthi, Kaustav Chanda, Yezhou Yang*
84 ART-ASyn: Anatomy-aware Realistic Texture-based Anomaly Synthesis Framework for Chest X-Rays, *Qinyi Cao, Jianan Fan, Weidong Cai*
85 High-Rate Mixout: Revisiting Mixout for Robust Domain Generalization, *Masih Aminbeidokhti, Heitor Rapela Medeiros, Srikanth Muralidharan, Eric Granger, Marco Pedersoli*
86 MuseDance: A Diffusion-based Music-Driven Image Animation System, *Zhikang Dong, Weituo Hao, Ju-Chiang Wang, Peng Zhang, Pawel Polak*
87 A Little More Like This: Text-to-Image Retrieval with Vision-Language Models Using Relevance Feedback, *Bulat Khaertdinov, Mirela Popa, Nava Tintarev*
88 Bi-ICE: An Inner Interpretable Framework for Image Classification via Bi-directional Interactions between Concept and Input Embeddings, *Jinyung Hong, Yearim Kim, Keun Hee Park, Sangyu Han, Nojun Kwak, Theodore P. Pavlic*
89 DTMIR-Pro: Domain Translation with Prompt-based Latent-Space Generalization for Multi-Weather Image Restoration, *Ashutosh Kulkarni, Prashant W. Patil, Santosh Kumar Vipparthi, Subrahmanyam Murala, Balasubramanian Raman*
90 ObjectCore – Efficient Few-shot Logical Anomaly Detection using Object Representations, *Matic Fučka, Vitjan Zavrtanik, Danijel Skocaj*
91 A Novel Metric for Detecting Memorization in Generative Models for Brain MRI Synthesis, *Antonio Scardace, Lemuel Puglisi, Francesco Guarnera, Sebastiano Battiato, Daniele Ravi*
92 Unsupervised Modular Adaptive Region Growing and RegionMix Classification for Wind Turbine Segmentation, *Raül Pérez-Gonzalo, Riccardo Magro, Andreas Espersen, Antonio Agudo*
93 Gen-AFFECT: Generation of Avatar Fine-grained Facial Expressions with Consistent identiTy, *Hao Yu, Rupayan Mallick, Margrit Betke, Sarah Adel Bargal*
94 FlowEO: Generative Unsupervised Domain Adaptation for Earth Observation, *Georges Le Bellier, Nicolas Audebert*
95 Efficient Vision Transformers via Token Merging with Head-wise Attention Correction, *Yuki Ichikawa, Masato Motomura, Thiem Van Chu, Daichi Fujiki*
96 2S-CEDiff: A Two-Stage Diffusion Framework for Generating High-Fidelity Contrast-Enhanced CT Images from Non-Contrast Scans, *Yibang Wu, Tzung-Dau Wang, Shang-Hong Lai*
97 JOCA: Task-Driven Joint Optimisation of Camera Hardware and Adaptive Camera Control Algorithms, *Chengyang Yan, Mitch Bryson, Donald G. Dansereau*
98 CAPE: A CLIP-Aware Pointing Ensemble of Complementary Heatmap Cues for Embodied Reference Understanding, *Fevziye Irem Eyiokur, Dogucan Yaman, Hazım Kemal Ekenel, Alexander Waibel*
99 Online Episodic Memory Visual Query Localization with Egocentric Streaming Object Memory, *Zaira Manigrasso, Matteo Dunnhofer, Antonino Furnari, Moritz Nottebaum, Antonio Finocchiaro, Davide Marana, Rosario Forte, Giovanni Maria Farinella, Christian Micheloni*
100 Mobile-Oriented Video Diffusion: Enabling Text-to-Video Generation on Mobile Devices Without Retraining, Compression, or Pruning, *Bosung Kim, Kyuhwan Lee, Isu Jeong, Jungmin Cheon, Yeojin Lee, Seulki Lee*
101 WorkZone3D: A Multimodal Dataset for 3D Work Zone Perception in Autonomous Driving, *Shounak Sural, Nishad Sahu, Ragunathan Rajkumar*
102 Unconditional Priors Matter! Improving Conditional Generation of Fine-Tuned Diffusion Models, *Prin Phunyaphibarn, Phillip Y. Lee, Jaihoon Kim, Minhyuk Sung*
103 Training-free Detection of Text-to-video Generations via Over-coherence, *Jonathan Brokman, Oren Rachmil, Omer Hofman, Roy Betser, Amit Giloni, Roman Vainshtein, Hisashi Kojima*
104 ReBrain: Brain MRI Reconstruction from Sparse CT Slice via Retrieval-Augmented Diffusion, *Junming Liu, Yifei Sun, Weihua Cheng, Yujin Kang, Yirong Chen, Ding Wang, Guosun Zeng*
105 Efficient Text-Guided Convolutional Adapter for the Diffusion Model, *Aryan Das, Koushik Biswas, Swalpa Kumar Roy, Badri Narayana Patro, Vinay Kumar Verma*
106 FLoMo-Net: A Novel Task-Adaptive Mixture of Experts Routing Framework with Frequency and Uncertainty Correction for Medical Image Segmentation, *Md Rayhan Ahmed, Patricia Lasserre*
107 From Detection to Anticipation: Online Understanding of Struggles across Various Tasks and Activities, *Shijia Feng, Michael Wray, Walterio W. Mayol-Cuevas*
108 CaRS: A Causal Intervention Segmentation Framework and Benchmark Dataset for Autonomous Driving under Transitional Weather Conditions, *Kondapally Madhavi, K Naveen Kumar, C Krishna Mohan, Sobhan Babu*
109 Domain Generalizing DINO for Visual Regression via Latent Distractor Subspace Consistency, *Nikhil Reddy, Chetan Arora, Mahsa Baktashmotlagh*

110 WALDO: Where Unseen Model-based 6D Pose Estimation Meets Occlusion, *Sajjad Pakdamansavoji, Yintao Ma, Amir Rasouli, Tongtong Cao*
111 Decoupling Shape and Texture in SAM-2 via Controlled Texture Replacement, *Inbal Cohen, Boaz Meivar, Peihan Tu, Shai Avidan, Gal Oren*
112 HumanBench: Two Heads, No Legs, But Mostly Human, the State of Generative Capabilities in T2I Models, *Anubhooti Jain, Mayank Vatsa, Richa Singh*
113 SOAF: Scene Occlusion-aware Neural Acoustic Field, *Huiyu Gao, Jiahao Ma, David Ahmedt-Aristizabal, Chuong Nguyen, Miaomiao Liu*
114 FedSCAI: Leveraging Server and Client Alignment for Unsupervised Federated Source-Free Domain Adaptation, *M. Yashwanth, Sampath Koti, Arunabh Singh, Shyam Marjit, Anirban Chakraborty*
115 Structure-Aware Feature Rectification with Region Adjacency Graphs for Training-free Open-Vocabulary Semantic Segmentation, *Qiming Huang, Hao Ai, Jianbo Jiao*
116 One Model, Many Behaviors: Training-Induced Effects on Out-of-Distribution Detection, *Gerhard Krumpl, Henning Avenhaus, Horst Possegger*
117 TalkingHeadBench: A Multi-Modal Benchmark & Analysis of Talking-Head DeepFake Detection, *Xinqi Xiong, Prakrut Patel, Qingyuan Fan, Amisha Wadhwa, Sarathy Selvam, Xiao Guo, Luchao Qi, Xiaoming Liu, Roni Sengupta*
118 Rethinking Real Image Editing: Unleashing Diverse Editing Operators via Multi-Objective Optimization, *Siyuan Wang, Xi Yang, Zihao Zhou, Huiru Shao, Rui Zhang, Qiufeng Wang, Guangliang Cheng, Kaizhu Huang*
119 Decomposition Sampling for Efficient Region Annotations in Active Learning, *Jingna Qiu, Frauke Wilm, Mathias Öttl, Jonas Utz, Maja Schlereth, Moritz Schillinger, Marc Aubreville, Katharina Breininger*
120 DNA: Dual-branch Network with Adaptation for Open-Set Online Handwriting Generation, *Tsai-Ling Huang, Nhat-Tuong Do-Tran, Ngoc-Hoang-Lam Le, Hong-Han Shuai, Ching-Chun Huang*
121 STEG-AIW: Spatio-Temporal Gating and Adaptive-Timestep Inference for Efficient Spiking Neural Networks, *Gulfam Ahmed Saju, Anton Spirkin, Felipe Marcelino, Yuchou Chang*
122 Enhancing Visual Planning with Auxiliary Tasks and Multi-token Prediction, *Ce Zhang, Yale Song, Ruta Desai, Michael Louis Iuzzolino, Joseph Tighe, Gedas Bertasius, Satwik Kottur*
123 Self-Supervised Compression and Artifact Correction for Streaming Underwater Imaging Sonar, *Rongsheng Qian, Chi Xu, Xiaoqiang Ma, Hao Fang, Yili Jin, William I. Atlas, Jiangchuan Liu*
124 Perceptually Guided 3DGS Streaming and Rendering for Mixed Reality, *Yunxiang Zhang, Sai Harsha Mupparaju, Kenneth Chen, Jenna Kang, Xinyu Zhang, Maito Omori, Kazuyuki Arimatsu, Qi Sun*
125 Unlocking Vision-Language Models for Video Anomaly Detection via Fine-Grained Prompting, *Shu Zou, Xinyu Tian, Lukas Wesemann, Fabian Waschkowski, Zhaoyuan Yang, Jing Zhang*
126 Unsupervised Memorability Modeling from Tip-of-the-Tongue Retrieval Queries, *Sree Bhattacharyya, Yaman K. Singla, Sudhir Yarram, Somesh Singh, Harini SI, James Z. Wang*
127 Learning Compact Video Representations for Efficient Long-form Video Understanding in Large Multimodal Models, *Yuxiao Chen, Jue Wang, Zhikang Zhang, Jingru Yi, Xu Zhang, Yang Zou, Zhaowei Cai, Jianbo Yuan, Xinyu Li, Hao Yang, Davide Modolo*
128 Guided Texture Segmentation via Coordinate-Aware Class-Ratio Mapping, *Bishal Ranjan Swain, Kyung Joo Cheoi, Jaepil Ko*
129 Empowering Source-Free Domain Adaptation via MLLM-Guided Reliability-Based Curriculum Learning, *Dongjie Chen, Kartik Patwari, Xiaoguang Zhu, Zhengfeng Lai, Sen-Ching Samson Cheung, Chen-Nee Chuah*
130 ExDDV: A New Dataset for Explainable Deepfake Detection in Video, *Vlad Hondru, Eduard Hogea, Darian M. Onchis, Radu Tudor Ionescu*
131 AGENet: Adaptive Edge-aware Geodesic Distance Learning for Few-Shot Medical Image Segmentation, *Ziyuan Gao*
132 KMOPS: Keypoint-Driven Method for Multi-Object Pose and Metric Size Estimation from Stereo Images, *Ying-Kun Wu, Yi Shen, Tzuhsuan Huang, I-Sheng Fang, Jun-Cheng Chen*
133 OPFormer: Object Pose Estimation Leveraging Foundation Model with Geometric Encoding, *Artem Moroz, Vít Zeman, Martin Mikšík, Elizaveta Isianova, Miroslav David, Pavel Burget, Varun Burde*
134 Streaming Real-Time Trajectory Prediction Using Endpoint-Aware Modeling, *Alexander Prutsch, David Schinagl, Horst Possegger*
135 Hybrid State Representation for Video Procedure Planning, *Woo Suk Choi, Youwon Jang, Minsu Lee, Byoung-Tak Zhang*

12:00-14:00 **Doctoral Consortium Event,** San Luis,
The Doctoral consortium is by invitation only.

12:00-13:30 **Lunch,** Ania Terrace and Lawn

13:30-14:30 **Oral Session 5A: Generative Models II,** AZ Ballroom 6
1. CONSTANT: Towards High-Quality One-Shot Handwriting Generation with Patch Contrastive Enhancement and Style-Aware Quantization, *Anh-Duy Le, Van-Linh Pham, Thanh-Nam Vo, Xuan Toan Mai, Tuan-Anh Tran*
2. DCText: Scheduled Attention Masking for Visual Text Generation via Divide-and-Conquer Strategy, *Jaewoo Song, Jooyoung Choi, Kanghyun Baek, Sangyub Lee, Daemin Park, Sungroh Yoon*
3. VFace: A Training-Free Approach for Diffusion-Based Video Face Swapping, *Sanoojan Baliah, Yohan Abeysinghe, Rusiru Thushara, Khan Muhammad, Abhinav Dhall, Karthik Nandakumar, Muhammad Haris Khan*
4. VividAnimator: An End-to-End Audio and Pose-driven Half-Body Human Animation Framework, *Donglin Huang, Tianhang Liu, Junming Huang, Xiaoda Yang, Yongyuan Li, Chi Wang, Weiwei Xu*
5. Fine-grained Defocus Blur Control for Generative Image Models, *Ayush Shrivastava, Connelly Barnes, Xuaner Zhang, Lingzhi Zhang, Andrew Owens, Sohrab Amirghodsi, Eli Shechtman*

13:30-14:30 **Oral Session 5B: Remote Sensing and Sensors,** AZ Ballroom 7
1. CalibBEV: LiDAR-Camera Calibration via BEV Alignment, *Filippo D'Addeo, Lorenzo Cipelli, Adriano Cardace, Emanuele Ghelfi, Andrea Zinelli, Massimo Bertozzi*
2. X-JEPA: A Novel Joint Learning Cross-Modal Predictive Alignment Framework for Remote Sensing Image Retrieval, *Shabnam Choudhury, Yash Salunkhe, Vaibhav Rajan, Subhasis Chaudhuri, Biplab Banerjee*
3. SSMRadNet : A Sample-wise State-Space Framework for Efficient and Ultra-Light Radar Segmentation and Object Detection, *Anuvab Sen, Mir Sayeed Mohammad, Saibal Mukhopadhyay*
4. Rank-based Geographical Regularization: Revisiting Contrastive Self-Supervised Learning for Multispectral Remote Sensing Imagery, *Tom Burgert, Leonard Hackel, Paolo Rota, Begüm Demir*

14:30-14:45 **Courtesy Break**

14:45-15:45 **Oral Session 6A: 3D Computer Vision II,** AZ Ballroom 6
1. OMeGa: Joint Optimization of Explicit Meshes and Gaussian Splats for Robust Scene-Level Surface Reconstruction, *Yuhang Cao, Haojun Yan, Danya Yao*
2. Confidence Through Parallel Attention for Depth and Uncertainty Estimation in Dynamic Environments, *Onkar Susladkar, Rohit Pawar, Chirag Sehgal, Samaksh Ujjawal, Sparsh Mittal*
3. BiNAR: A Bi-Modal Framework for Non-Aligned RGB-IR 3D Reconstruction via Gaussian Splatting, *Zhongwen Wang, Han Ling, Weihao Zhang, Yinghui Sun, Quansen Sun*
4. Spec-Gloss Surfels and Normal-Diffuse Priors for Relightable Glossy Objects, *Georgios Kouros, Minye Wu, Tinne Tuytelaars*

5. Occlusion Boundary and Depth: Mutual Enhancement via Multi-Task Learning, *Lintao Xu, Yinghao Wang, Chaohui Wang*

14:45-15:45 Oral Session 6B: Video Recognition and Understanding I, AZ Ballroom 7

1. Ego-EXTRA: video-language Egocentric Dataset for EXpert-TRAinee assistance, *Francesco Ragusa, Michele Mazzamuto, Rosario Forte, Irene D'Ambra, James Fort, Jakob Engel, Antonino Furnari, Giovanni Maria Farinella*
2. Similarity-aware Probabilistic Embeddings Modeling for Video-Text Retrieval, *Yuliang Huang, Pengxu Wei, Zhicheng Dong, Liang Lin*
3. PromptGAR: Flexible Promptive Group Activity Recognition, *Zhangyu Jin, Andrew Feng, Ankur Chemburkar, Celso M De Melo*
4. Spacewalk-18: A Benchmark for Multimodal and Long-form Procedural Video Understanding in Novel Domains, *Zitian Tang, Rohan Myer Krishnan, Zhiqiu Yu, Chen Sun*
5. Broadcast2Pitch: Game State Reconstruction from Unconstrained Soccer Videos, *Yin May Oo, Yewon Hwang, Muhammad Amrulloh Robbani, Vanyi Chao, Ankhzaya Jamsrandorj, Hoang Quoc Nguyen, Kyung-Ryoul Mun, Jinwook Kim*

15:45-16:30 PAMI-TC meeting, AZ Ballroom Salon 6

16:30-18:15 Poster Session 4 + Reception, Tucson Ballroom & Prefunction Space

1 CONSTANT: Towards High-Quality One-Shot Handwriting Generation with Patch Contrastive Enhancement and Style-Aware Quantization, *Anh-Duy Le, Van-Linh Pham, Thanh-Nam Vo, Xuan Toan Mai, Tuan-Anh Tran*

2 DCText: Scheduled Attention Masking for Visual Text Generation via Divide-and-Conquer Strategy, *Jaewoo Song, Jooyoung Choi, Kanghyun Baek, Sangyub Lee, Daemin Park, Sungroh Yoon*

3 VFace: A Training-Free Approach for Diffusion-Based Video Face Swapping, *Sanoojan Baliah, Yohan Abeysinghe, Rusiru Thushara, Khan Muhammad, Abhinav Dhall, Karthik Nandakumar, Muhammad Haris Khan*

4 VividAnimator: An End-to-End Audio and Pose-driven Half-Body Human Animation Framework, *Donglin Huang, Tianhang Liu, Junming Huang, Xiaoda Yang, Yongyuan Li, Chi Wang, Weiwei Xu*

5 Fine-grained Defocus Blur Control for Generative Image Models, *Ayush Shrivastava, Connelly Barnes, Xuaner Zhang, Lingzhi Zhang, Andrew Owens, Sohrab Amirghodsi, Eli Shechtman*

6 CalibBEV: LiDAR-Camera Calibration via BEV Alignment, *Filippo D'Addeo, Lorenzo Cipelli, Adriano Cardace, Emanuele Ghelfi, Andrea Zinelli, Massimo Bertozzi*

7 X-JEPA: A Novel Joint Learning Cross-Modal Predictive Alignment Framework for Remote Sensing Image Retrieval, *Shabnam Choudhury, Yash Salunkhe, Vaibhav Rajan, Subhasis Chaudhuri, Biplab Banerjee*

8 SSMRadNet : A Sample-wise State-Space Framework for Efficient and Ultra-Light Radar Segmentation and Object Detection, *Anuvab Sen, Mir Sayeed Mohammad, Saibal Mukhopadhyay*

9 Rank-based Geographical Regularization: Revisiting Contrastive Self-Supervised Learning for Multispectral Remote Sensing Imagery, *Tom Burgert, Leonard Hackel, Paolo Rota, Begüm Demir*

10 OMeGa: Joint Optimization of Explicit Meshes and Gaussian Splats for Robust Scene-Level Surface Reconstruction, *Yuhang Cao, Haojun Yan, Danya Yao*

11 Confidence Through Parallel Attention for Depth and Uncertainty Estimation in Dynamic Environments, *Onkar Susladkar, Rohit Pawar, Chirag Sehgal, Samaksh Ujjawal, Sparsh Mittal*

12 BiNAR: A Bi-Modal Framework for Non-Aligned RGB-IR 3D Reconstruction via Gaussian Splatting, *Zhongwen Wang, Han Ling, Weihao Zhang, Yinghui Sun, Quansen Sun*

13 Spec-Gloss Surfels and Normal-Diffuse Priors for Relightable Glossy Objects, *Georgios Kouros, Minye Wu, Tinne Tuytelaars*

14 Occlusion Boundary and Depth: Mutual Enhancement via Multi-Task Learning, *Lintao Xu, Yinghao Wang, Chaohui Wang*

15 Ego-EXTRA: video-language Egocentric Dataset for EXpert-TRAinee assistance, *Francesco Ragusa, Michele Mazzamuto, Rosario Forte, Irene D'Ambra, James Fort, Jakob Engel, Antonino Furnari, Giovanni Maria Farinella*

16 Similarity-aware Probabilistic Embeddings Modeling for Video-Text Retrieval, *Yuliang Huang, Pengxu Wei, Zhicheng Dong, Liang Lin*

17 PromptGAR: Flexible Promptive Group Activity Recognition, *Zhangyu Jin, Andrew Feng, Ankur Chemburkar, Celso M De Melo*

18 Spacewalk-18: A Benchmark for Multimodal and Long-form Procedural Video Understanding in Novel Domains, *Zitian Tang, Rohan Myer Krishnan, Zhiqiu Yu, Chen Sun*

19 Broadcast2Pitch: Game State Reconstruction from Unconstrained Soccer Videos, *Yin May Oo, Yewon Hwang, Muhammad Amrulloh Robbani, Vanyi Chao, Ankhzaya Jamsrandorj, Hoang Quoc Nguyen, Kyung-Ryoul Mun, Jinwook Kim*

20 VLMs Guided Interpretable Decision Making in Autonomous Driving, *Xin Hu, Taotao Jing, Renran Tian, Zhengming Ding*

21 DARB-Splatting: Generalizing Splatting with Decaying Anisotropic Radial Basis Functions, *Hashiru Pramuditha, Vinasirajan Viruthshaan, Vishagar Arunan, Saeedha Nazar, Ranga Rodrigo, Sameera Ramasinghe, Simon Lucey*

22 Surgical Gaussian Surfels: Highly Accurate Real-time Surgical Scene Rendering using Gaussian Surfels, *Idris O. Sunmola, Zhenjun Zhao, Samuel Schmidgall, Paul Maria Scheikl, Yumeng Wang, Viet Pham, Axel Krieger*

23 Gated Temporal Fusion Transformers for Robust Multi-Object Tracking, *Jinho Kim, Kuk-Jin Yoon*

24 SIAM: Synchronous Interaction Attention for Human Mesh Recovery, *Niaz Ahmad, Saif Ullah, Youngmoon Lee, Guanghui Wang*

25 SynchroRaMa : Lip-Synchronized and Emotion-Aware Talking Face Generation via Multi-Modal Emotion Embedding, *Phyo Thet Yee, Dimitrios Kollias, Sudeepta Mishra, Abhinav Dhall*

26 Are All Marine Species Created Equal? Performance Disparities in Underwater Object Detection, *Melanie Wille, Tobias Fischer, Scarlett Raine*

27 LVM-Lite: Training Large Vision Models with Efficient Sequential Modeling, *Xianhang Li, Hongru Zhu, Sucheng Ren, Linjie Yang, Peng Wang, Heng Wang, Xiaohui Shen, Qing Liu, Cihang Xie*

28 HiGlassRM: Learning to Remove High-prescription Glasses via Synthetic Dataset Generation, *Sebin Lee, Heewon Kim*

29 Transformer-Based Inpainting for Real-Time 3D Streaming in Sparse Multi-Camera Setups, *Leif Van Holland, Domenic Zingsheim, Mana Takhsha, Hannah Dröge, Patrick Stotko, Markus Plack, Reinhard Klein*

30 SeaClips: A Video Dataset for Maritime Object Detection, *Franziska Denk, Christian Rankl, Shaban Almouahed, David Moser, Robert Sablatnig*

31 UniDiff: Parameter-Efficient Adaptation of Diffusion Models for Land Cover Classification with Multi-Modal Remotely Sensed Imagery and Sparse Annotations, *Yuzhen Hu, Saurabh Prasad*

32 CropAT: Leveraging Diffusion-Generated Target-Like Cropped Objects for Pseudo-Label Refinement in Domain-Adaptive Object Detection, *Chen-Che Huang, Tzuhsuan Huang, Jun-Cheng Chen*

33 Beyond Faces: A Multimodal Person Clustering for Unconstrained Environments, *Sahngmin Yoo, Sangwon Lee, Seongin Jo*

34 Eye-for-an-eye: Appearance Transfer with Dense Semantic Correspondence in Diffusion Models, *Sooyeon Go, Kyungmook Choi, Minjung Shin, Youngjung Uh*

35 Towards Photorealistic Style Transfer with Multimodal Guidance and Robustness to Content Images in Arbitrary Styles, *Ruikai Zhou, Yating Liu, Yi Xu*

36 Hierarchical Adaptive networks with Task vectors for Test-Time Adaptation, *Sameer Ambekar, Marta Hasny, Laura Alexandra Daza, Daniel M. Lang, Julia Schnabel*

37 Automated Pore Detection from In-Situ FDM 3D Printing Video: A Comparative Evaluation of Modern Segmentation Models, *Abdullah Al Ahad Khan, Md Shariful Islam, Lin Li, Lai Jiang, Noushin Ghaffari*

38 Safe Vision-Language Models via Unsafe Weights Manipulation, *Moreno D'incà, Elia Peruzzo, Xingqian Xu, Humphrey Shi, Nicu Sebe, Massimiliano Mancini*
39 SOPHY: Generating Simulation-Ready Objects with Physical Materials, *Junyi Cao, Evangelos Kalogerakis*
40 Generalization of Real World Video Deblurring By Image-to-Image Translation, *Kassymzhomart Aitbek, Seungjoon Yang*
41 A Dataset and Framework for Learning State-invariant Object Representations, *Rohan Sarkar, Avinash Kak*
42 Better Safe Than Sorry? Overreaction Problem of Vision Language Models in Visual Emergency Recognition, *Dasol Choi, Seunghyun Lee, Youngsook Song*
43 Optimizing LVLMs with On-Policy Data for Effective Hallucination Mitigation, *Chengzhi Yu, Yifan Xu, Yifan Chen, Wenyi Zhang*
44 Countering Multi-modal Representation Collapse through Rank-targeted Fusion, *Seulgi Kim, Kiran Kokilepersaud, Mohit Prabhushankar, Ghassan AlRegib*
45 Learning Unified Spatio-temporal Representations for Efficient Compressed Video Understanding, *Shristi Das Biswas, Efstathia Soufleri, Arani Roy, Kaushik Roy*
46 More Than Memory Savings: Zeroth-Order Optimization Mitigates Forgetting in Continual Learning, *Wanhao Yu, Zheng Wang, Shuteng Niu, Sen Lin, Li Yang*
47 Global Focal and Radial Distortion Averaging from Radial Fundamental Matrices for Robust Self-Calibration, *Sergei Solonets, Daniil Sinitsyn, Daniel Cremers*
48 UCDSC: Open Set UnCertainty aware Deep Simplex Classifier for Medical Image Datasets, *Arnav Aditya, Nitin Kumar, Saurabh Shigwan*
49 DODA: Adapting Object Detectors to Dynamic Agricultural Environments in Real-Time with Diffusion, *Shuai Xiang, Pieter M. Blok, James Burridge, Haozhou Wang, Wei Guo*
50 Structured Context Learning for Generic Event Boundary Detection, *Xin Gu, Congcong Li, Xinyao Wang, Dexiang Hong, Heng Fan, Libo Zhang, Longyin Wen, Tiejian Luo*
51 Sun-E: Dataset and Benchmark for Event-Based Sun Sensing, *Sydney Dolan, Alessandro Golkar*
52 FCC: Fully Connected Correlation for One-Shot Segmentation, *Seonghyeon Moon, Haein Kong, Muhammad Haris Khan, Mubbasir Kapadia, Yuewei Lin*
53 MoRe: Monocular Geometry Refinement via Graph Optimization for Cross-View Consistency, *Dongki Jung, Jaehoon Choi, Yonghan Lee, Sungmin Eum, Heesung Kwon, Dinesh Manocha*
54 ProSkill: Segment-Level Skill Assessment in Procedural Videos, *Michele Mazzamuto, Daniele Di Mauro, Gianpiero Francesca, Giovanni Maria Farinella, Antonino Furnari*
55 MergeSlide: Continual Model Merging and Task-to-Class Prompt-Aligned Inference for Lifelong Learning on Whole Slide Images, *Doanh C. Bui, Ba Hung Ngo, Hoai Luan Pham, Khang Nguyen, Maï K. Nguyen, Yasuhiko Nakashima*
56 CoL2A: Convolution-free Local Linear Attention for SpatioTemporal Event Processing, *Yusuke Sekikawa, Itsumi Araki, Jun Nagata, Andreu Girbau*
57 Dronaquatics: Real-time Swimming Analytics Using Drone Captured Imagery, *Thu Tran, Harold Abraham Joseph, Kichang Lee, Kenny Tsu Wei Choo, Dong Ma, Shaohui Foong, Thivya Kandappu, Jeonggil Ko, Rajesh Balan*
58 Histopath-C: Towards Realistic Domain Shifts for Histopathology Vision-Language Adaptation, *Mehrdad Noori, Gustavo A. Vargas Hakim, David Osowiechi, Fereshteh Shakeri, Ali Bahri, Moslem Yazdanpanah, Sahar Dastani, Ismail Ben Ayed, Christian Desrosiers*
59 Robust Multimodal Emotion Recognition from Incomplete Modalities via Query-Based Unimodal and Cross-Modal Learning, *Ryo Miyoshi, Mayu Otani, Yuki Okafuji*
60 WiSE-OD: Benchmarking Robustness in Infrared Object Detection, *Heitor R. Medeiros, Atif Belal, Masih Aminbeidokhti, Eric Granger, Marco Pedersoli*
61 Patch-wise Retrieval: A Bag of Practical Techniques for Instance-level Matching, *Wonseok Choi, Sohwi Lim, Nam Hyeon-Woo, Moon Ye-Bin, Dong-Ju Jeong, Jinyoung Hwang, Tae-Hyun Oh*
62 Gaussian Swaying: Surface-Based Framework for Aerodynamic Simulation with 3D Gaussians, *Hongru Yan, Xiang Zhang, Zeyuan Chen, Fangyin Wei, Zhuowen Tu*
63 Restora-Flow: Mask-Guided Image Restoration with Flow Matching, *Arnela Hadzic, Franz Thaler, Lea Bogensperger, Simon Johannes Joham, Martin Urschler*
64 PrevMatch: Revisiting and Maximizing Temporal Knowledge in Semi-Supervised Semantic Segmentation, *Wooseok Shin, Hyun Joon Park, Jin Sob Kim, Juan Yun, Se Hong Park, Sung Won Han*
65 One-shot Portrait Stylizaiton via Geometric Alignment, *Xinrui Wang, Zilin Guo, Zhuoru Li, Jinze Yu, Heng Zhang, Yusuke Iwasawa, Yutaka Matsuo, Jiaxian Guo*
66 Zero-Shot Table Extraction in Business Documents: A Unified Benchmark with Error Taxonomy and Ecological Analysis, *Eliott Thomas, Mickael Coustaty, Aurélie Joseph, Tri-Cong Pham, Gaspar Deloin, Elodie Carel, Vincent Poulain D'andecy, Jean-Marc Ogier*
67 SegMango: Early Deep Mango Yield Prediction based on Flower Segmentation and Weather Data, *Janaksinh Ven, Charu Sharma, Azeemuddin Syed*
68 Geo3DVQA: Evaluating Vision-Language Models for 3D Geospatial Reasoning from Aerial Imagery, *Mai Tsujimoto, Junjue Wang, Weihao Xuan, Naoto Yokoya*
69 PoseGaussian: Pose-Driven Novel View Synthesis for Robust 3D Human Reconstruction, *Ju Shen, Chen Chen, Tam V. Nguyen, Vijayan K. Asari*
70 Timestamp Query Transformer for Temporal Action Segmentation, *Tieqiao Wang, Sinisa Todorovic*
71 QC-SF: Improving Computer Vision for Airborne LiDAR Point Clouds of Boreal Forests with Quebec Simulated Forest Dataset, *Olivier Stocker, Reza Mahmoudi Kouhi, Omid Reisi Gahrouei, Thierry Badard, Eric Guilbert*
72 RemEdit: Efficient Diffusion Editing with Riemannian Geometry, *Eashan Adhikarla, Brian D. Davison*
73 GrowTAS: Progressive Expansion from Small to Large Subnets for Efficient ViT Architecture Search, *Hyunju Lee, Youngmin Oh, Jeimin Jeon, Donghyeon Baek, Bumsub Ham*
74 Subspace-Guided Knowledge Distillation for Efficient Model Transfer, *Zeeshan Hayder, Ali Cheraghian, Lars Petersson, Mehrtash Harandi*
75 TimeRefine: Temporal Grounding with Time Refining Video LLM, *Xizi Wang, Feng Cheng, Ziyang Wang, Huiyu Wang, Md Mohaiminul Islam, Lorenzo Torresani, Mohit Bansal, Gedas Bertasius, David Crandall*
76 Curve Skeletonization in Continuous domain for Meshes and Point Clouds, *Jai Bardhan, Ramya Hebbalaguppe, Aravind Udupa*
77 Gene-DML: Dual-Pathway Multi-Level Discrimination for Gene Expression Prediction from Histopathology Images, *Yaxuan Song, Jianan Fan, Hang Chang, Weidong Cai*
78 Color Preserving CMOS-SPAD Fusion for Multi-Frame HDR, *Aleksi Suonsivu, Lauri Salmela, Lassi Helin, Leevi Uosukainen, Giacomo Boracchi*
79 Unsupervised Segmentation by Diffusing, Walking and Cutting, *Daniela Ivanova, Marco Aversa, Paul Henderson, John H. Williamson*
80 Learning Spatio-temporal Feature Representations for Video-based Gaze Estimation, *Alexandre Personnic, Mihai Bace*
81 ImageChain: Advancing Sequential Image-to-Text Reasoning in Multimodal Large Language Models, *Danae Sanchez Villegas, Ingo Ziegler, Desmond Elliott*
82 Edge-Aware Image Manipulation via Diffusion Models with a Novel Structure-Preservation Loss, *Minsu Gong, Nuri Ryu, Jungseul Ok, Sunghyun Cho*
83 3D Superquadric Splatting, *Daniel MacSwayne, Ales Leonardis, Jianbo Jiao*

84 Controllable Long-term Motion Generation with Extended Joint Targets, *Eunjong Lee, Eunhee Kim, Sanghoon Hong, Eunho Jung, Jihoon Kim*
85 ST-Think: How Multimodal Large Language Models Reason About 4D Worlds from Ego-Centric Videos, *Peiran Wu, Yunze Liu, Miao Liu, Junxiao Shen*
86 RPT-SR: Regional Prior attention Transformer for infrared image Super-Resolution, *Youngwan Jin, Incheol Park, Sanghyeop Yeo, Hyeongjin Ju, Yagiz Nalcakan, Shiho Kim*
87 LASOR: Towards Clinically Transparent and Explainable Ophthalmic Report Generation via Lesion-Aware Segmentation, *Jian Park, Hyunseon Won, JeeEun Kim, Joon Seo Hwang, Jeong Mo Han, Ji In Park, Daniel Duck-Jin Hwang, Jinyoung Han*
88 Quantifying the Limits of Segmentation Foundation Models: Modeling Challenges in Segmenting Tree-Like and Low-Contrast Objects, *Yixin Zhang, Nicholas Konz, Kevin Kramer, Maciej A Mazurowski*
89 Lorentz Entailment Cone for Semantic Segmentation, *Zahid Hasan, Masud Ahmed, Nirmalya Roy*
90 WarpRF: Multi-View Consistency for Training-Free Uncertainty Quantification and Applications in Radiance Fields, *Sadra Safadoust, Fabio Tosi, Fatma Güney, Matteo Poggi*
91 GAEA: A Geolocation Aware Conversational Assistant, *Ron Campos, Ashmal Vayani, Parth Parag Kulkarni, Rohit Gupta, Aizan Zafar, Aritra Dutta, Mubarak Shah*
92 WSSSP-Net: Weakly Supervised Semantic Segmentation Plugin Network for Face Anti-Spoofing, *Krzysztof Galus, Piotr Syga, Piotr Kawa*
93 CONCORD: Concept-Informed Diffusion for Dataset Distillation, *Jianyang Gu, Haonan Wang, Ruoxi Jia, Saeed Vahidian, Vyacheslav Kungurtsev, Wei Jiang, Yiran Chen*
94 Improving Out-of-Distribution Detection Using Segmented Images and Cross-View Attention Fusion, *Alexander Politowicz, Sahisnu Mazumder, Bing Liu*
95 An improved architecture for part-based animal re-identification through semantic segmentation distillation, *Eugênio Dias Ribeiro Neto, Marc Chaumont, Gérard Subsol, Michel De Garine-Wichatitsky, Hélène Guis*
96 FB-4D: Spatial-Temporal Coherent Dynamic 3D Content Generation with Feature Banks, *Jinwei Li, Huan-ang Gao, Wenyi Li, Haohan Chi, Chenyu Liu, Chenxi Du, Yiqian Liu, Mingju Gao, Zongzheng Zhang, Guiyu Zhang, Jingwei Zhao, Hongyang Li, Yao Yao, Li Yi, Yikai Wang, Hao Zhao*
97 Hestia: Voxel-Face-Aware Hierarchical Next-Best-View Acquisition for Efficient 3D Reconstruction, *Cheng-You Lu, Zhuoli Zhuang, Nguyen Thanh Trung Le, Da Xiao, Yu-Cheng Chang, Thomas Do, Srinath Sridhar, Chin-Teng Lin*
98 R3: Reconstruction, Raw, and Rain: Deraining Directly in the Bayer Domain, *Nate Rothschild, Moshe Kimhi, Avi Mendelson, Chaim Baskin*
99 An Efficient Multi-Rater Setup Towards Personalized and Diversified Medical Image Segmentation, *Sajed Almorsy, Ayman Khalafallah, Marwan Torki,*
100 HiMix : Hierarchical Visual-Textual Mixing Network for Lesion Segmentation, *Soojin Hwang, Jaeyoon Sim, Won Hwa Kim*
101 FSP-DETR: Few-Shot Prototypical Parasitic Ova Detection, *Shubham Trehan, Udhav Ramachandran, Akash Rao, Ruth Scimeca, Sathyanarayanan N. Aakur*
102 DF-Mamba: Deformable State Space Modeling for 3D Hand Pose Estimation in Interactions, *Yifan Zhou, Takehiko Ohkawa, Guwenxiao Zhou, Kanoko Goto, Takumi Hirose, Yusuke Sekikawa, Nakamasa Inoue*
103 Context-Preserving Dermoscopic Editing: Mask-Guided Lesion-Aware Diffusion for Attribute Modification, *Tao Sun, Yun Jiang, Yarong Jin, Zequn Zhang, Huanting Guo*
104 How I Met Your Bias: Investigating Bias Amplification in Diffusion Models, *Nathan Roos, Ani Gjergji, Ekaterina Iakovleva, Vito Paolo Pastore, Enzo Tartaglione*
105 BREEN: Bridge Data-Efficient Encoder-Free Multimodal Learning with Learnable Queries, *Tianle Li, Yongming Rao, Winston Hu, Yu Cheng*
106 ProtoGMVAE: A Variational Auto-Encoder with True Gaussian Mixture Prior for Prototypical-based Self-Explainability, *Martin Blanchard, Christophe Ducottet, Damien Muselet, Olivier Delézay*
107 AEON: Adaptive Embedding Optimized Noise for Robust Watermarking in Diffusion Models, *Muhammad Shahid Muneer, Simon S. Woo*
108 Revisiting Vision–Language Foundations for No-Reference Image Quality Assessment, *Ankit Yadav, Ta Duc Huy, Lingqiao Liu*
109 Semi-supervised Domain Adaptation via Mutual Alignment through Joint Error, *Dexuan Zhang, Thomas Westfechtel, Tatsuya Harada*
110 Unified Control for Inference-Time Guidance of Denoising Diffusion Models, *Maurya Goyal, Anuj Singh, Hadi Jamali-Rad*
111 Learning Subglacial Bed Topography from Sparse Radar with Physics-Guided Residuals, *Bayu Adhi Tama, Jianwu Wang, Vandana Janeja, Mostafa Cham*
112 4D Multimodal Co-attention Fusion Network with Latent Contrastive Alignment for Alzheimer's Diagnosis, *Yuxiang Wei, Yanteng Zhang, Xi Xiao, Tianyang Wang, Xiao Wang, Vince D. Calhoun*
113 One-Cycle Structured Pruning via Stability-Driven Subnetwork Search, *Deepak Ghimire, Dayoung Kil, Seonghwan Jeong, Jaesik Park, Seong-heum Kim*
114 PredMapNet: Future and Historical Reasoning for Consistent Online HD Vectorized Map Construction, *Bo Lang, Nirav Savaliya, Zhihao Zheng, Jinglun Feng, Zheng-Hang Yeh, Mooi Choo Chuah*
115 Multi-view Stereo with Multiple Projectors for Oneshot Entire Shape Scan based on Neural SDF and DSSS Demultiplexing, *Kota Nishihara, Ryo Furukawa, Ryusuke Sagawa, Hiroshi Kawasaki*
116 FreeCond: Free Lunch in the Input Conditions of Text-Guided Inpainting, *Teng-Fang Hsiao, Bo-Kai Ruan, Sung-Lin Tsai, Yi-Lun Wu, Hong-Han Shuai*
117 ControlEvents: Controllable Synthesis of Event Camera Data with Foundational Prior from Image Diffusion Models, *Yixuan Hu, Yuxuan Xue, Simon Klenk, Daniel Cremers, Gerard Pons-Moll*
118 DPBridge: Latent Diffusion Bridge for Dense Prediction, *Haorui Ji, Taojun Lin, Hongdong Li*
119 PatchEAD: Unifying Industrial Visual Prompting Frameworks for Patch-Exclusive Anomaly Detection, *Po-Han Huang, Jeng-Lin Li, Po-Hsuan Huang, Ming-Ching Chang, Wei-Chao Chen*
120 SurfDist: Interpretable Three-Dimensional Instance Segmentation Using Curved Surface Patches, *Jackson Borchardt, Saul Kato*
121 CRISP: Cylindrical Rendering for In-Stream Point Clouds, *Hyungwoo Kang, Seonyoung Jang, Yeojun Yoon, Byung Tae Oh*
122 INRetouch: Context Aware Implicit Neural Representation for Photography Retouching, *Omar Elezabi, Marcos V. Conde, Zongwei Wu, Radu Timofte*
123 Line Art Colorization with Offset Prior-based Diffusion Model, *Xuan Zhu, Miao Cao, Fang-Lue Zhang, Yu-Kun Lai, Paul L Rosin*
124 Food Image Generation on Multi-Noun Categories, *Xinyue Pan, Yuhao Chen, Jiangpeng He, Fengqing Zhu*
125 RobuMTL: Enhancing Multi-Task Learning Robustness Against Weather Conditions, *Tasneem Shaffee, Sherief Reda*
126 EndoPBR: Photorealistic Synthetic Data for Surgical 3D Vision via Physically-based Rendering, *John J. Han, Jie Ying Wu*
127 Inpainting of Sparse Depth Maps from Monocular Depth-from-Focus on Pixel Processor Arrays, *Maciej Lewandowski, Piotr Dudek*
128 DMS2F-HAD: A Dual-branch Mamba-based Spatial–Spectral Fusion Network for Hyperspectral Anomaly Detection, *Aayushma Pant, Lakpa Tamang, Tsz-Kwan Lee, Sunil Aryal*
129 F-ViTA: Foundation Model Guided Visible to Infrared Translation, *Jay Nitin Paranjape, Celso M De Melo, Vishal M. Patel*

130 KFS-Bench: Comprehensive Evaluation of Key Frame Sampling in Long Video Understanding, *Zongyao Li, Kengo Ishida, Satoshi Yamazaki, Xiaotong Ji, Jianquan Liu*
131 FujiView: Multimodal Late-Fusion for Predicting Scenic Visibility, *Bryceton Bible, Nehal Hasnaeen, Hairong Qi*
132 Improving Animal Pose Estimation through Species Similarity Measures and Rigorous Label Definition, *Medhashree Parhy, Shaan Chanchani, Claire Kim, Josh Mansky, Parth Thakre, Zian Pan, Haoyu Chen, Amy R. Reibman*
133 Grounding Descriptions in Images informs Zero-Shot Visual Recognition, *Shaunak Halbe, Junjiao Tian, Joseph K J, James Seale Smith, Katherine Stevo, Vineeth N. Balasubramanian, Zsolt Kira*
134 Semi-supervised Key-Point Estimation for Echocardiography Video, *Seok-Hwan Oh, Hyeon-Jik Lee, Guil Jung, Myeong-Gee Kim, Young-Min Kim, Hyuksool Kwon, Hyeon-Min Bae*
135 Anatomically-guided Masked Autoencoder Pre-training for Aneurysm Detection, *Alberto M. Ceballos Arroyo, Jisoo Kim, Chu-Hsuan Lin, Lei Qin, Geoffrey S. Young, Huaizu Jiang*
136 Style-Friendly SNR Sampler for Style-Driven Generation, *Jooyoung Choi, Chaehun Shin, Yeongtak Oh, Heeseung Kim, Jungbeom Lee, Sungroh Yoon*
137 Lose Your Self (LoYS): An Adversarial Entropy-based Unsupervised Approach for Model Debiasing, *Vito Paolo Pastore, Massimiliano Ciranni, Vittorio Murino*
138 MagicDrive3D: Controllable 3D Generation for Any-View Rendering in Street Scenes, *Ruiyuan Gao, Kai Chen, Zhihao Li, Lanqing Hong, Zhenguo Li, Qiang Xu*
139 Grounding Degradations in Natural Language for All-In-One Video Restoration, *Muhammad Kamran Janjua, Amirhosein Ghasemabadi, Kunlin Zhang, Mohammad Salameh, Chao Gao, Di Niu*
140 ControlVP: Interactive Geometric Refinement of AI-Generated Images with Consistent Vanishing Points, *Ryota Okumura, Kaede Shiohara, Toshihiko Yamasaki*
141 Enhancing Vision Language Corruption Robustness using Cross-Distribution & Prompted Denoisers, *Sameer Shafayet Latif, Sadab Shiper, K. M. Rahiduzzaman Kiran, Md Farhan Ishmam, Md Azam Hossain, Abu Raihan Mostofa Kamal, Md Hamjajul Ashmafee*
142 BlendCLIP: Bridging Synthetic and Real Domains for Zero-Shot 3D Object Classification with Multimodal Pretraining, *Ajinkya Khoche, Gergő László Nagy, Maciej Wozniak, Thomas Gustafsson, Patric Jensfelt*
143 Towards Egocentric 3D Hand Pose Estimation in Unseen Domains, *Wiktor Mucha, Michael Wray, Martin Kampel*
144 Direct Visual Grounding by Directing Attention of Visual Tokens, *Parsa Esmaeilkhani, Longin Jan Latecki*
145 CaFlow: Enhancing Long-Term Action Quality Assessment with Causal Counterfactual Flow, *Ruisheng Han, Kanglei Zhou, Shuang Chen, Amir Atapour-Abarghouei, Hubert P. H. Shum*

Tuesday, March 10

Reminder that all events are for registered attendees only.

8:00-14:00 **Registration,** AZ Ballroom Registration Desk
9:00-14:00 **Poster Pickup**, Tucson Ballroom Registration Desk
8:30-9:30 **Keynote Talk 3**, AZ Ballroom 6 (streamed to 7)
Speaker: Hilde Kuhne, **Title:** A Short History of Video Understanding - Past, Present, and Future
9:30-9:45 **Courtesy Break**, AZ Ballroom Prefunction
9:45-10:45 **Oral Session 7A: Biometrics, Face, Gesture, and Body Pose II**, AZ Ballroom 6
9:45-10:45 **Oral Session 7B: Vision+Language and Other Modalities II**, AZ Ballroom 7
10:45-14:00 **Exhibits**, Tucson Ballroom
10:45-17:30 **Demos**, Tucson Ballroom
10:45-12:15 **Poster Session 5**, Tucson Ballroom & Prefunction Space
12:00-13:30 **Lunch**, Ania Terrace and Lawn
13:30-14:30 **Oral Session 8A: Biomedical, Healthcare, and Medicine**, AZ Ballroom 6
13:30-14:30 **Oral Session 8B: Video Recognition and Understanding II**, AZ Ballroom 7
14:30-14:45 **Courtesy Break**
14:45-15:45 **Oral Session 9A: Generative Models III,** AZ Ballroom 6
14:45-15:45 **Oral Session 9B: Machine Learning II**, AZ Ballroom 7
15:45-17:30 **Poster Session 6 + Refreshments**, Tucson Ballroom & Prefunction Space

8:30-9:30 **Keynote Talk 3**, AZ Ballroom 6 (streamed to 7)
Speaker: Hilde Kuhne, **Title:** A Short History of Video Understanding - Past, Present, and Future

9:30 - 9:45 **Courtesy Break**

The papers in this oral session will also be presented as posters.
Format: 10 min. presentation, 2 min. questions

9:45-10:45 **Oral Session 7A: Biometrics, Face, Gesture, and Body Pose II**, AZ Ballroom 6

1. Motion-Aware Graph Fusion Network for 3D Human Pose Estimation, *Yen Pham, Xiaohui Yuan, Chengyuan Zhuang*
2. UniGaze: Towards Universal Gaze Estimation via Large-scale Pre-Training, *Jiawei Qin, Xucong Zhang, Yusuke Sugano*
3. Unsupervised Discovery of Long-Term Spatiotemporal Periodic Workflows in Human Activities, *Fan Yang, Quanting Xie, Atsunori Moteki, Shoichi Masui, Shan Jiang, Kanji Uchino, Yonatan Bisk, Graham Neubig*
4. VAST-ReID: A Low-Light Benchmark Dataset for Person Re-Identification with Visual and Attribute-Rich Semantic Tracking, *Hammad Khan, Rakesh Kumar Giri, Kamalakar Vijay Thakare, Heeseung Choi, Hyungjoo Jung, Debi Prosad Dogra, Ig-Jae Kim*
5. DexAvatar: 3D Sign Language Reconstruction with Hand and Body Pose Priors, *Kaustubh Kundu, Hrishav Bakul Barua, Lucy Robertson-Bell, Zhixi Cai, Kalin Stefanov*

9:45-10:45 **Oral Session 7B: Vision+Language and Other Modalities II**, AZ Ballroom 7

1. DREAM: Dynamic Prompts and GuidedMix for Efficient Continual Adaptation of Visual-Language Models, *Evelyn Chee, Mong Li Lee, Wynne Hsu*
2. brat: Aligned Multi-View Embeddings for Brain MRI Analysis, *Maxime Kayser, Maksim Gridnev, Wanting Wang, Max Bain, Aneesh Rangnekar, Avijit Chatterjee, Aleksandr Petrov, Harini Veeraraghavan, Nathaniel C. Swinburne*
3. Towards Fine-Grained Adaptation of CLIP via a Self-Trained Alignment Score, *Eman Ali, Sathira Silva, Chetan Arora, Muhammad Haris Khan*
4. Advancing Multimodal LLMs by Large-Scale 3D Visual Instruction Dataset Generation, *Liu He, Xiao Zeng, Yizhi Song, Albert Y. C. Chen, Lu Xia, Shashwat Verma, Sankalp Dayal, Min Sun, Cheng-Hao Kuo, Daniel Aliaga*
5. CLIP-UP: CLIP-Based Unanswerable Problem Detection for Visual Question Answering, *Ben Vardi, Oron Nir, Ariel Shamir*

10:45-14:00 **Exhibits**, Tucson Ballroom

10:45-17:30 **Demos**, Tucson Ballroom

1. LayoutWeaver: Interactive Image-Space Correction for Spatial Text Datasets, *Prashant Khatri, Hyprbots Systems*
2. GroundEye: Visually Grounded Evidence Retriever for Enterprise Document Agents, *Krishna Chaitanya Reddy Tamataam, Hyprbots Systems*
3. When Ink meets Pixel: Decoding Human Intent in Financial Documents with DocIntent-OCR, *Akshata Bhat, Hyprbots Systems*
4. Will My Model Generalize? An Interactive System for Quantifying Domain Shift, *Kuntal Thakur, ASU*

10:45-12:15 **Poster Session 5**, Tucson Ballroom & Prefunction Space

1 Motion-Aware Graph Fusion Network for 3D Human Pose Estimation, *Yen Pham, Xiaohui Yuan, Chengyuan Zhuang*
2 UniGaze: Towards Universal Gaze Estimation via Large-scale Pre-Training, *Jiawei Qin, Xucong Zhang, Yusuke Sugano*
3 Unsupervised Discovery of Long-Term Spatiotemporal Periodic Workflows in Human Activities, *Fan Yang, Quanting Xie, Atsunori Moteki, Shoichi Masui, Shan Jiang, Kanji Uchino, Yonatan Bisk, Graham Neubig*
4 VAST-ReID: A Low-Light Benchmark Dataset for Person Re-Identification with Visual and Attribute-Rich Semantic Tracking, *Hammad Khan, Rakesh Kumar Giri, Kamalakar Vijay Thakare, Heeseung Choi, Hyungjoo Jung, Debi Prosad Dogra, Ig-Jae Kim*
5 DexAvatar: 3D Sign Language Reconstruction with Hand and Body Pose Priors, *Kaustubh Kundu, Hrishav Bakul Barua, Lucy Robertson-Bell, Zhixi Cai, Kalin Stefanov*
6 DREAM: Dynamic Prompts and GuidedMix for Efficient Continual Adaptation of Visual-Language Models, *Evelyn Chee, Mong Li Lee, Wynne Hsu*
7 brat: Aligned Multi-View Embeddings for Brain MRI Analysis, *Maxime Kayser, Maksim Gridnev, Wanting Wang, Max Bain, Aneesh Rangnekar, Avijit Chatterjee, Aleksandr Petrov, Harini Veeraraghavan, Nathaniel C. Swinburne*
8 Towards Fine-Grained Adaptation of CLIP via a Self-Trained Alignment Score, *Eman Ali, Sathira Silva, Chetan Arora, Muhammad Haris Khan*
9 Advancing Multimodal LLMs by Large-Scale 3D Visual Instruction Dataset Generation, *Liu He, Xiao Zeng, Yizhi Song, Albert Y. C. Chen, Lu Xia, Shashwat Verma, Sankalp Dayal, Min Sun, Cheng-Hao Kuo, Daniel Aliaga*
10 CLIP-UP: CLIP-Based Unanswerable Problem Detection for Visual Question Answering, *Ben Vardi, Oron Nir, Ariel Shamir*
11 Isolating the Role of Temporal Information in Video Saliency: A Controlled Experimental Analysis, *Peter El-Jiz, Matthias Kuemmerer, Matthias Tangemann, Matthias Bethge, Andreas Bartels, Michael Mario Bannert*
12 Diffusion-Based Action Recognition Generalizes to Untrained Domains, *Rogério Guimarães, Frank Xiao, Pietro Perona, Markus Marks*
13 CORA: Consistency-Guided Semi-Supervised Framework for Reasoning Segmentation, *Prantik Howlader, Hoang Nguyen-Canh, Srijan Das, Jingyi Xu, Hieu Le, Dimitris Samaras*
14 Hierarchical Instance Tracking to Balance Privacy Preservation with Accessible Information, *Neelima Prasad, Jarek Tyler Reynolds, Neel Karsanbhai, Tanusree Sharma, Lotus Zhang, Abigale Stangl, Yang Wang, Leah Findlater, Danna Gurari*
15 Causality-Driven Audits of Model Robustness, *Nathan Drenkow, William Paul, Chris Ribaudo, Mathias Unberath*
16 BAFLE-DCT: Bypassing Adversarial Filters via Frequency-Selective Embedding in the DCT Domain, *Thilina Mendis, Farah Kandah, Sathyanarayanan N. Aakur*
17 Logit-Adjusted Test-Time Adaptation under Partial Class Imbalance,

Thilina Weerasinghe, Ruwan Tennakoon, WeiQin Chuah, Alireza Bab-Hadiashar

18 Test Time Adaptation Using Adaptive Quantile Recalibration, *Paria Mehrbod, Pedro Vianna, Geraldin Nanfack, Guy Wolf, Eugene Belilovsky*

19 OSEG: Improving Diffusion sampling through Orthogonal Smoothed Energy Guidance, *Masud An Nur Islam Fahim, Nazmus Saqib, Joon-Min Gil*

20 Hymavi : A Hybrid Mamba-Attention Network in Multi-View Framework for Volumetric Medical Image Segmentation, *Sy Dat Tran, Jin Kyu Gahm*

21 RoadBench: A Vision-Language Foundation Model and Benchmark for Road Damage Understanding, *Xi Xiao, Yunbei Zhang, Janet Wang, Lin Zhao, Yuxiang Wei, Hengjia Li, Yanshu Li, Xiao Wang, Swalpa Kumar Roy, Hao Xu, Tianyang Wang*

22 IMKD: Intensity-Aware Multi-Level Knowledge Distillation for Camera-Radar Fusion, *Shashank Mishra, Karan Patil, Didier Stricker, Jason Rambach*

23 LogicCBMs: Logic-Enhanced Concept-Based Learning, *Deepika SN Vemuri, Gautham Bellamkonda, Aditya Pola, Vineeth N Balasubramanian*

24 Understanding the Visual Projection Space of Multimodal LLMs, *Sungheon Jeong, Yoojeong Song, Hyungjoon Kim*

25 NoHumansRequired: Autonomous High-Quality Image Editing Triplet Mining, *Maksim Kuprashevich, Grigorii Alekseenko, Irina Tolstykh, Georgii Fedorov, Bulat Suleimanov, Vladimir Dokholyan, Aleksandr Gordeev*

26 SSMT-Net: A Semi-Supervised Multitask Transformer-Based Network for Thyroid Nodule Segmentation in Ultrasound Images, *Muhammad Umar Farooq, Abd Ur Rehman, Azka Rehman, Muhammad Usman, Dong-Kyu Chae*

27 Gaussian Splatting Map Registration with Orthographic Bird's-Eye-View Renderings, *H. Leblond, G. Simon, R. Martins, C. Demonceaux, M.-O. Berger*

28 MUSE: Model-based Uncertainty-aware Similarity Estimation for zero-shot 2D Object Detection and Segmentation, *Sungmin Cho, Sungbum Park, Insoo Oh*

29 MVAT: Multi-View Aware Teacher for Weakly Supervised 3D Object Detection, *Saad Lahlali, Alexandre Fournier-Montgieux, Nicolas Granger, Hervé Le Borgne, Quoc-Cuong Pham*

30 ODEt(ODEl): Shortcutting the Time and the Length in Diffusion and Flow Models for Faster Sampling, *Denis Gudovskiy, Wenzhao Zheng, Tomoyuki Okuno, Yohei Nakata, Kurt Keutzer*

31 TM-Adapter: Temporal Merge Adapter for Efficient Global Temporal Modeling, *Woo Joo Hahm, Seungwoo Jang, Hyeon Tak Kim, Daeun Lee, Kwangsu Kim*

32 Diagnose Like A REAL Pathologist: An Uncertainty-Focused Approach for Trustworthy Multi-Resolution Multiple Instance Learning, *Sungrae Hong, Sol Lee, Jisu Shin, Jiwon Jeong, Mun Yong Yi*

33 Align Video Diffusion Model with Online Video-Centric Preference Optimization, *Jiacheng Zhang, Jie Wu, Weifeng Chen, Yatai Ji, Weilin Huang, Xuefeng Xiao, Kai Han*

34 SceneProp: Combining Neural Network and Markov Random Field for Scene-Graph Grounding, *Keita Otani, Tatsuya Harada*

35 GHOST: Getting to the Bottom of Hallucinations with A Multi-round Consistency Benchmark, *Vibashan VS, Nadine Chang, Jenny Schmalfuss, Vishal M. Patel, Zhiding Yu, Jose M. Alvarez*

36 GeneVA: A Dataset of Human Annotations for Generative Text to Video Artifacts, *Jenna Kang, Maria Beatriz Silva, Patsorn Sangkloy, Kenneth Chen, Niall L. Williams, Qi Sun*

37 Zero-Shot Domain Generalisation via Prompt-Driven Feature Refinement, *Tingrui Qiao, Di Zhao, Caroline Walker, Chris Cunningham, Yun Sing Koh*

38 Can Image Splicing and Copy-Move Forgery Be Detected by the Same Model? Forensim: An Attention-Based State-Space Approach, *Soumyaroop Nandi, Prem Natarajan*

39 Distilling Offline Action Detection Models into Real-Time Streaming Models, *Deep Patel, Yasunori Babazaki, Yasuto Nagase, Iain Melvin, Martin Renqiang Min*

40 AuthGuard: Generalizable Deepfake Detection via Language Guidance, *Guangyu Shen, Zhihua Li, Xiang Xu, Tianchen Zhao, Zheng Zhang, Dongsheng An, Zhuowen Tu, Yifan Xing, Qin Zhang*

41 GroupPortrait: Multi-ID Portrait Generation with High Identity Preservation and Fine-Grained Control, *Meijia Huang, Ruida Li, Bing Ma, Liangwei Jiang, Shuo Fang, Chenguang Ma*

42 GFT-GCN: Privacy-Preserving 3D Face Mesh Recognition with Spectral Diffusion, *Hichem Felouat, Hanrui Wang, Isao Echizen*

43 Denoise, Divide, Distill, and Predict (D3P): Towards Forecasting Long-horizon Real-world Anomaly from Normalcy, *Quentin Mérilleau, Snehashis Majhi, Antitza Dantcheva, Quan Kong, Lorenzo Garattoni, Gianpiero Francesca, Francois Bremond*

44 See, Record, Do: Automated Generation of UI Workflows from Tutorial Videos, *Adam Beauchaine, Craig A. Shue*

45 Video and Language Alignment in 2D Systems for 3D Multi-object Scenes with Multi-Information Derivative-Free Control, *Jason Armitage, Rico Sennrich*

46 FocalComm: Hard Instance-Aware Multi-Agent Perception, *Dereje Shenkut, Vijayakumar Bhagavatula*

47 SFMNet: Sparse Focal Modulation for 3D Object Detection, *Oren Shrout, Ayellet Tal*

48 MoSCo: Real-time and Efficient Text-to-Motion Synthesis via Delta Training, *Zhiyuan Zhang, Lingqiao Liu*

49 VLMDiff: Leveraging Vision-Language Models for Multi-Class Anomaly Detection with Diffusion, *Samet Hicsonmez, Abd El Rahman Shabayek, Djamila Aouada*

50 Predicting Task fMRI Contrasts from Resting-State fMRI Using Sparse 3D Convolutions, *Ivan Sviridov, Maria Boyko, Maksim Sharaev*

51 SceneEdited: A City-Scale Benchmark for 3D HD Map Updating via Image-Guided Change Detection, *Chun-Jung Lin, Tat-Jun Chin, Sourav Garg, Feras Dayoub*

52 Overcoming Small Data Limitations in Video-Based Infant Respiration Estimation, *Liyang Song, Hardik Bishnoi, Sai Kumar Reddy Manne, Sarah Ostadabbas, Briana J. Taylor, Michael Wan*

53 Non-Aligned Reference Image Quality Assessment for Novel View Synthesis, *Abhijay Ghildyal, Rajesh Sureddi, Nabajeet Barman, Saman Zadtootaghaj, Alan C Bovik*

54 Pointmap-Conditioned Diffusion for Consistent Novel View Synthesis, *Thang-Anh-Quan Nguyen, Laurent Caraffa, Jean-Philippe Tarel, Roland Brémond*

55 TED-4DGS: Temporally Activated and Embedding-based Deformation for 4DGS Compression, *Cheng-Yuan Ho, He-Bi Yang, Jui-Chiu Chiang, Yu-Lun Liu, Wen-Hsiao Peng*

56 QUOTA: Quantifying Objects with Text-to-Image Models for Any Domain, *Wenfang Sun, Yingjun Du, Gaowen Liu, Yefeng Zheng, Cees G. M. Snoek*

57 Single-step Diffusion for Image Compression at Ultra-Low Bitrates, *Chanung Park, Joo Chan Lee, Jong Hwan Ko*

58 Virtually Unrolling the Herculaneum Papyri by Diffeomorphic Spiral Fitting, *Paul Henderson*

59 Diffusion Noise Optimization for Synthetic VLM Training, *Ren Ohkubo, Rintaro Yanagi, Hirokatsu Kataoka, Yutaka Satoh*

60 Histogram Assisted Quality Aware Generative Model for Resolution Invariant NIR Image Colorization, *Abhinav Attri, Rajeev Ranjan Dwivedi, Samiran Das, Vinod Kumar Kurmi*

61 Polymorph: Energy-Efficient Multi-Label Classification for Video Streams on Embedded Devices, *Saeid Ghafouri, Mohsen Fayyaz, Xiangchen Li, Deepu John, Bo Ji, Dimitrios S. Nikolopoulos, Hans Vandierendonck*

62 A Unified Diffusion-Based Framework for Multi-Agent Trajectory Prediction Integrating Structured Multi-Modal Representations, *Chenxi Yang, Suyang Xi, Hong Ding, Yiqing Shen, Yunhao Liu*

63 ChartQA-X: Generating Explanations for Visual Chart Reasoning, *Shamanthak Hegde, Pooyan Fazli, Hasti Seifi*

64 Distribution Highlighted Reference-based Label Distribution Learning for Facial Age Estimation, *Satoshi Suzuki, Shin'ya Yamaguchi, Shoichiro Takeda, Takuhiro Kaneko, Shota Orihashi, Ryo Masumura*

65 T2VWorldBench: A Benchmark for Evaluating World Knowledge in Text-to-Video Generation, *Yubin Chen, Xuyang Guo, Zhenmei Shi, Zhao Song, Jiahao Zhang*
66 UniTabBank: A Large Scale Multi-Lingual, Multi-Layout, Multi-Type, Multi-Format Dataset for Table Detection, *Ajoy Mondal, Saumya Mundra, Avijit Dasgupta, C.V. Jawahar*
67 R-MMA: Enhancing Vision-Language Models with Recurrent Adapters for Few-Shot and Cross-Domain Generalization, *Md Fahim, Md Farhan Ishmam, Mir Sazzat Hossain, M Ashraful Amin, Amin Ahsan Ali, AKM Mahbubur Rahman*
69 Training-Free Few-Shot Segmentation via Vision-Language Guided Prompting, *Euihyun Yoon, Taejin Park, Jaekoo Lee*
70 High-Level Semantics and Low-Level Features Fusion for Multi-Scale Object Detection in Dynamic Construction Environments, *Mahdi Bonyani, Maryam Soleymani, Chao Wang*
71 Mean-Shift Distillation for Diffusion Mode Seeking, *Vikas Thamizharasan, Nikitas Chatzis, Iliyan Georgiev, Matthew Fisher, Evangelos Kalogerakis, Difan Liu, Nanxuan Zhao, Michal Lukáč*
72 TacticalCalib: End-to-End 6-DoF Camera Pose Regression for Tactical Camera Calibration, *Liang Fan, Xiaoqian Liu, Zhi Chen, Lingkai Yang*
73 F-INR: Functional Tensor Decomposition for Implicit Neural Representations, *Sai Karthikeya Vemuri, Tim Büchner, Joachim Denzler*
74 V2XScene: Multi-View Consistent 3D Scene Simulation for Collaborative Perception, *Yanfei Li, Yuan Zeng, Yi Gong*
75 WiSAR3D - Aerial LiDAR Dataset for 3D Object Detection, *Oren Shrout, Ori Nizan, Yizhak Ben-Shabat, Ayellet Tal*
76 A Deep Network for Object Detection on Inland Waters, *Dennis Griesser, Bastian Goldluecke, Matthias O. Franz, Georg Umlauf*
77 CoreCaption: Core Caption based Text-to-Video Retrieval, *Junkyu Jang*
78 ZebraPose: Zebra Detection and Pose Estimation using only Synthetic Data, *Elia Bonetto, Aamir Ahmad*
79 FAIR-SIGHT: Fairness Assurance in Image Recognition via Simultaneous Conformal Thresholding and Dynamic Output Repair, *Arya Fayyazi, Mehdi Kamal, Massoud Pedram*
80 GDoFS: Gaussian DoF Separation for Plausible 3D Geometry in Sparse-View 3DGS, *Yongsung Kim, Jooyoung Choi, Sungroh Yoon*
81 Learning Beyond Labels: Self-Supervised Handwritten Text Recognition, *Shree Mitra, Ajoy Mondal, C.V. Jawahar*
82 Guiding What Not to Generate: Automated Negative Prompting for Text-Image Alignment, *Sangha Park, Eunji Kim, Yeongtak Oh, Jooyoung Choi, Sungroh Yoon*
83 Neural Geometry Image-Based Representations with Optimal Transport (OT), *Xiang Gao, Yuanpeng Liu, Xinmu Wang, Jiazhi Li, Minghao Guo, Yu Guo, Xiyun Song, Heather Yu, Zhiqiang Lao, Xianfeng David Gu*
84 RealDroneVision: Dataset and Architecture Advancements for Small-Object Drone Detection, *Arun Kumar Sivapuram, Pranav R T Peddinti, Harish Puppala, Komuravelli Prashanth, Jaladi Sri Harsha, Rama Krishna Sai Gorthi*
85 PerVL-Bench: Benchmarking Multimodal Personalization for Large Vision–Language Models, *Minsung Kim*
86 TRACE: Confounder-free Adversarial Fine-tuning for Robust Object Detection, *Wonho Lee, Jisu Lee, Hyunsik Na, Sohee Park, Daeseon Choi*
87 Zero-LEAD: Source-Free Universal Domain Adaptation for Abdominal Multi-Organ Segmentation, *Ahmed El-Sayed, Marwan Torki*
88 Correcting and Quantifying Systematic Errors in 3D Box Annotations for Autonomous Driving, *Alexandre Justo Miro, Ludvig af Klinteberg, Bogdan Timus, Aron Asefaw, Ajinkya Khoche, Thomas Gustafsson, Sina Sharif Mansouri, Masoud Daneshtalab*
89 FastHMR: Accelerating Human Mesh Recovery via Token and Layer Merging with Diffusion Decoding, *Soroush Mehraban, Andrea Iaboni, Babak Taati*
90 Point2Pose: A Generative Framework for 3D Human Pose Estimation with Multi-View Point Cloud Dataset, *Hyunsoo Lee, Daeum Jeon, Hyeokjae Oh*
91 UniVid: Unifying Vision Tasks with Pre-trained Video Generation Models, *Lan Chen, Yuchao Gu, Qi Mao*
92 Optimal Transport for Rectified Flow Image Editing: Unifying Inversion-Based and Direct Methods, *Marian Lupaşcu, Mihai-Sorin Stupariu*
93 Synthesizing Compositional Videos from Text Description, *Prajwal Singh, Kuldeep Kulkarni, Shanmuganathan Raman, Harsh Rangwani*
94 S2O: Static to Openable Enhancement for Articulated 3D Objects, *Denys Iliash, Hanxiao Jiang, Yiming Zhang, Manolis Savva, Angel X. Chang*
95 HodgeFormer: Transformers for Learnable Operators on Triangular Meshes through Data-Driven Hodge Matrices, *Akis Nousias, Stavros Nousias*
96 Exploring Automated Recognition of Instructional Activity and Discourse from Multimodal Classroom Data, *Ivo Bueno, Ruikun Hou, Babette Bühler, Tim Fütterer, James Drimalla, Jonathan K. Foster, Peter Youngs, Peter Gerjets, Ulrich Trautwein, Enkelejda Kasneci*
97 From Prompt to Production: Automating Brand-Safe Marketing Imagery with Text-to-Image Models, *Parmida Atighehchian, Henry Wang, Andrei Kapustin, Boris Lerner, Tiancheng Jiang, Taylor Jensen, Negin Sokhandan*
98 Equivariant Sampling for Improving Diffusion Model-based Image Restoration, *Chenxu Wu, Qingpeng Kong, Peiang Zhao, Wendi Yang, Wenxin Ma, Fenghe Tang, Zihang Jiang, S.Kevin Zhou*
99 PoseAdapt: Sustainable Human Pose Estimation via Continual Learning Benchmarks and Toolkit, *Muhammad Saif Ullah Khan, Didier Stricker*
100 SAVeD: Learning to Denoise Low-SNR Video for Improved Downstream Performance, *Suzanne Stathatos, Michael Hobley, Pietro Perona, Markus Marks*
101 ATM: Enhanced Alignment for Text-to-Motion Generation, *Ke Han, Yueming Lyu, Weichen Yu, Nicu Sebe*
102 Memoire: Learning User Personas from Gallery Tags for Personalized Photo Curation, *Praful Mathur, Mohsin Iftekhar, Aman Sharma, Sarvesh Tiwari, Meghali Deka, Sathish Cherukuri, K Roopa Sheshadri, Rakesh Valusa*
103 CSGaussian: Progressive Rate-Distortion Compression and Segmentation for 3D Gaussian Splatting, *Yu-Jen Tseng, Chia-Hao Kao, Jing-Zhong Chen, Alessandro Gnutti, Shao-Yuan Lo, Yen-Yu Lin, Wen-Hsiao Peng*
104 Test-Time Adaptation for Video Highlight Detection Using Meta-Auxiliary Learning and Cross-Modality Hallucinations, *Zahidul Islam, Sujoy Paul, Mrigank Rochan*
105 SD-CSFL: A Synthetic Data-Driven Conformity Scoring Framework for Robust Federated Learning, *Ebtisaam Alharbi, Abdulrahman Kerim, Leandro Soriano Marcolino, Qiang Ni*
106 Alignment and Distillation: A Robust Framework for Multimodal Domain Generalizable Human Action Recognition, *Hyeonbin Ji, Juyeob Lee, Eunil Park*
107 Procedure Learning via Regularized Gromov-Wasserstein Optimal Transport, *Syed Ahmed Mahmood, Ali Shah Ali, Umer Ahmed, Fawad Javed Fateh, M. Zeeshan Zia, Quoc-Huy Tran*
108 Clear Sights on Site: A Spatial-Adaptive Channel Network for Deblurring Construction Site Images, *Mahdi Bonyani, Maryam Soleymani, Chao Wang*
109 SegMo: Segment-aligned Text to 3D Human Motion Generation, *Bowen Dang, Lin Wu, Xiaohang Yang, Zheng Yuan, Zhixiang Chen*
110 From Darkness to Detail: Frequency-Aware SSMs for Low-Light Vision, *Eashan Adhikarla, Kai Zhang, Gong Chen, John Nicholson, Brian D. Davison*
111 Multimodal Adversarial Defense for Vision-Language Models by Leveraging One-To-Many Relationships, *Futa Waseda, Antonio Tejero-de-Pablos, Isao Echizen*
112 ForestSplats: Deformable Transient Field for Gaussian Splatting in the Wild, *Wongi Park, Myeongseok Nam, Siwon Kim, Sangwoo Jo, Soomok Lee*
113 Graph Query Networks for Object Detection with Automotive Radar, *Loveneet Saini, Hasan Tercan, Tobias Meisen*

114 FlowCLAS: Enhancing Normalizing Flow-Based Anomaly Segmentation Via Contrastive Learning, *Chang Won Lee, Selina Leveugle, Paul Grouchy, Chris Langley, Svetlana Stolpner, Jonathan Kelly, Steven L. Waslander*
115 ScoreNet: Netting Lightweight Quality Scores for Better Visual Assessment with Large Multi-Modality Models, *Bahador Rashidi, Kiarash Aghakasiri, Shupei Zhang, Amirmohsen Sattarifard, Yue Zhang, Chao Gao*
116 VectorSynth: Fine-Grained Satellite Image Synthesis with Structured Semantics, *Daniel Cher, Brian Wei, Srikumar Sastry, Nathan Jacobs*
117 Digital Forensic AI You Can Explain: A Case Study on Video Source Camera Identification, *Maryna Veksler, Kemal Akkaya, Selcuk Uluagac*
118 Modeling and Learning Multiple Hypotheses for Monocular 3D Object Detection, *Hyeonjeong Park, Peixi Xiong, Pei Yu, Wei Tang*
119 SGD-Mix: Enhancing Domain-Specific Image Classification with Label-Preserving Data Augmentation, *Yixuan Dong, Fang-Yi Su, Jung-Hsien Chiang*
120 DreamCatcher: Efficient Multi-Concept Customization via Representation Finetuning, *Jungwon Lee, Changhun Lee, Eunhyeok Park*
121 Enhancing Monocular 3D Hand Reconstruction with Learned Texture Priors, *Giorgos Karvounas, Nikolaos Kyriazis, Iason Oikonomidis, Georgios Pavlakos, Antonis A. Argyros*
122 PALMS+: Modular Image-Based Floor Plan Localization Leveraging Depth Foundation Model, *Yunqian Cheng, Roberto Manduchi, Benjamin John Princen*
123 IPCD: Intrinsic Point-Cloud Decomposition, *Shogo Sato, Takuhiro Kaneko, Shoichiro Takeda, Tomoyasu Shimada, Kazuhiko Murasaki, Taiga Yoshida, Ryuichi Tanida, Akisato Kimura*
124 Multimodal Graph Representation Learning over Arbitrary Sets of Modalities, *Santosh Patapati, Trisanth Srinivasan*
125 FNOPT: Resolution-Agnostic, Self-Supervised Cloth Simulation using Meta-Optimization with Fourier Neural Operators, *Ruochen Chen, Thuy Tran, Shaifali Parashar*
126 D2Mamba: Dual Domain Guided Informed Search in State Space Model for Underwater Image Enhancement, *Alik Pramanick, Soumajit Roy, Arijit Sur*
127 Image-Guided Semantic Pseudo-LiDAR Point Generation for 3D Object Detection, *Minseung Lee, Seokha Moon, Seung Joon Lee, Reza Mahjourian, Jinkyu Kim*
128 HABIT: Human Action Benchmark for Interactive Traffic in CARLA, *Mohan Ramesh, Mark Azer, Fabian Flohr*
129 EVTP-IVS: Effective Visual Token Pruning For Unifying Instruction Visual Segmentation In Multi-Modal Large Language Models, *Wenhui Zhu, Xiwen Chen, Zhipeng Wang, Shao Tang, Sayan Ghosh, Xuanzhao Dong, Rajat Koner, Yalin Wang*
130 AuViRe: Audio-visual Speech Representation Reconstruction for Deepfake Temporal Localization, *Christos Koutlis, Symeon Papadopoulos*
131 Exploring the Boundaries of Diffusion Models for Offline Writer Identification with Sparse and Intra-Variable Data, *Aritra Dey, Chandranath Adak, Kumari Priya, Soumi Chattopadhyay, Sukalpa Chanda*
132 CLoCKDistill: Consistent Location and Context aware Knowledge Distillation for DETRs, *Qizhen Lan, Qing Tian*
133 MaxInfo: A Training-Free Key-Frame Selection Method Using Maximum Volume for Enhanced Video Understanding, *Pengyi Li, Irina Abdullaeva, Alexander Gambashidze, Andrey Kuznetsov, Ivan Oseledets*

13:30-14:30 Oral Session 8A: Biomedical, Healthcare, and Medicine, AZ Ballroom 6

1. Cycle-Consistent Multi-Graph Matching for Self-Supervised Annotation of C. Elegans, *Sebastian Stricker, Christoph Karg, Lisa Hutschenreiter, Dagmar Kainmueller, Bogdan Savchynskyy*
2. Automated Suturing Skill Assessment in Robot-assisted Surgery from Endoscopic Videos using Clinically-guided Evaluation Criteria, *Atharva Sunil Deo, Ujjwal Pasupulety, Nicholas Matsumoto, Jay Moran, Cherine Yang, Jeanine Kim, Rafal Dariusz Kocielnik, Aurash Naser-Tavakolian, Andrew Hung*
3. Deep Image Decomposition for Medical Imaging Anonymization and Curation, *Yael Elkin, Gal Ben-Arie, Tammy Riklin-Raviv*
4. Intraoperative 2D/3D Registration via Spherical Similarity Learning and Differentiable Levenberg-Marquardt Optimization, *Minheng Chen, Youyong Kong*
5. ACuRE: Accurate Continuity-Regularized SpO2 Estimation Using Liquid Time-Constant Networks, *Shahzad Ahmad, Divya Mishra, Sania Bano, Sukalpa Chanda, Yogesh Singh Rawat*

13:30-14:30 Oral Session 8B: Video Recognition and Understanding II, AZ Ballroom 7

1. CAST: Evaluating Multi-Object Trackers with Context-Aware Switch and Transfer Scores, *Jin Bai, Gregory D. Hager*
2. Advancing Player Identification and Tracking with Global ID Fusion (GIF), *Karol Wojtulewicz, Minxing Liu, Niklas Carlsson*
3. Distilling What and Why: Enhancing Driver Intention Prediction with MLLMs, *Sainithin Artham, Avijit Dasgupta, Shankar Gangisetty, C.V. Jawahar*
4. LASER: Lip Landmark Assisted Speaker Detection for Robustness, *Le Thien Phuc Nguyen, Zhuoran Yu, Yong Jae Lee*
5. VADER: Towards Causal Video Anomaly Understanding with Relation-Aware Large Language Models, *Ying Cheng, Yu-Ho Lin, Min-Hung Chen, Fu-En Yang, Shang-Hong Lai*

14:30-14:45 Courtesy Break

14:45-15:45 Oral Session 9A: Generative Models III, AZ Ballroom 6

1. SCAdapter: Content-Style Disentanglement for Diffusion Style Transfer, *Luan Thanh Trinh*
2. T2LF: LLM-Guided Multimodal Diffusion for Text-to-Light Field Synthesis, *Soyoung Yoon, Namhyuk Ahn, In Kyu Park*
3. VideoSketcher: A Training-Free Approach for Coherent Video Sketch Transfer, *Huining Li, Bangzhen Liu, Rui Yang, Yang Zhou, Chenshu Xu, Xufang Pang, Shengfeng He*
4. Zero-Shot Audio-Visual Editing via Cross-Modal Delta Denoising, *Yan-Bo Lin, Kevin Lin, Zhengyuan Yang, Linjie Li, Jianfeng Wang, Chung-Ching Lin, Xiaofei Wang, Gedas Bertasius, Lijuan Wang*
5. SceneEval: Evaluating Semantic Coherence in Text-Conditioned 3D Indoor Scene Synthesis, *Hou In Ivan Tam, Hou In Derek Pun, Austin T. Wang, Angel X. Chang, Manolis Savva*

14:45-15:45 Oral Session 9B: Machine Learning II, AZ Ballroom 7

1. IPTQ-ViT: Post-Training Quantization of Non-linear Functions for Integer-only Vision Transformers, *Gihwan Kim, Jemin Lee, Hyungshin Kim*
2. MM-TS: Multi-Modal Temperature and Margin Schedules for Contrastive Learning with Long-Tail Data, *Siarhei Sheludzko, Dhimitrios Duka, Bernt Schiele, Hilde Kuehne, Anna Kukleva*
3. Boosting Unsupervised Video Instance Segmentation with Automatic Quality-Guided Self-Training, *Kaixuan Lu, Mehmet Onurcan Kaya, Dim P. Papadopoulos*
4. Locally Explaining Prediction Behavior via Gradual Interventions and Measuring Property Gradients, *Niklas Penzel, Joachim Denzler*

15:45-17:30 Poster Session 6 + Refreshments, Tucson Ballroom & Prefunction Space

1 Cycle-Consistent Multi-Graph Matching for Self-Supervised Annotation of C. Elegans, *Sebastian Stricker, Christoph Karg, Lisa Hutschenreiter, Dagmar Kainmueller, Bogdan Savchynskyy*
2 Automated Suturing Skill Assessment in Robot-assisted Surgery from Endoscopic Videos using Clinically-guided Evaluation Criteria, *Atharva Sunil Deo, Ujjwal Pasupulety, Nicholas Matsumoto, Jay Moran, Cherine Yang, Jeanine Kim, Rafal Dariusz Kocielnik, Aurash Naser-Tavakolian, Andrew Hung*
3 Deep Image Decomposition for Medical Imaging Anonymization and Curation, *Yael Elkin, Gal Ben-Arie, Tammy Riklin-Raviv*

4 Intraoperative 2D/3D Registration via Spherical Similarity Learning and Differentiable Levenberg-Marquardt Optimization, *Minheng Chen, Youyong Kong*
5 ACuRE: Accurate Continuity-Regularized SpO2 Estimation Using Liquid Time-Constant Networks, *Shahzad Ahmad, Divya Mishra, Sania Bano, Sukalpa Chanda, Yogesh Singh Rawat*
6 CAST: Evaluating Multi-Object Trackers with Context-Aware Switch and Transfer Scores, *Jin Bai, Gregory D. Hager*
7 Advancing Player Identification and Tracking with Global ID Fusion (GIF), *Karol Wojtulewicz, Minxing Liu, Niklas Carlsson*
8 Distilling What and Why: Enhancing Driver Intention Prediction with MLLMs, *Sainithin Artham, Avijit Dasgupta, Shankar Gangisetty, C.V. Jawahar*
9 LASER: Lip Landmark Assisted Speaker Detection for Robustness, *Le Thien Phuc Nguyen, Zhuoran Yu, Yong Jae Lee*
10 VADER: Towards Causal Video Anomaly Understanding with Relation-Aware Large Language Models, *Ying Cheng, Yu-Ho Lin, Min-Hung Chen, Fu-En Yang, Shang-Hong Lai*
11 SCAdapter: Content-Style Disentanglement for Diffusion Style Transfer, *Luan Thanh Trinh*
12 T2LF: LLM-Guided Multimodal Diffusion for Text-to-Light Field Synthesis, *Soyoung Yoon, Namhyuk Ahn, In Kyu Park*
13 VideoSketcher: A Training-Free Approach for Coherent Video Sketch Transfer, *Huining Li, Bangzhen Liu, Rui Yang, Yang Zhou, Chenshu Xu, Xufang Pang, Shengfeng He*
14 Zero-Shot Audio-Visual Editing via Cross-Modal Delta Denoising, *Yan-Bo Lin, Kevin Lin, Zhengyuan Yang, Linjie Li, Jianfeng Wang, Chung-Ching Lin, Xiaofei Wang, Gedas Bertasius, Lijuan Wang*
15 SceneEval: Evaluating Semantic Coherence in Text-Conditioned 3D Indoor Scene Synthesis, *Hou In Ivan Tam, Hou In Derek Pun, Austin T. Wang, Angel X. Chang, Manolis Savva*
16 IPTQ-ViT: Post-Training Quantization of Non-linear Functions for Integer-only Vision Transformers, *Gihwan Kim, Jemin Lee, Hyungshin Kim*
17 MM-TS: Multi-Modal Temperature and Margin Schedules for Contrastive Learning with Long-Tail Data, *Siarhei Sheludzko, Dhimitrios Duka, Bernt Schiele, Hilde Kuehne, Anna Kukleva*
18 Boosting Unsupervised Video Instance Segmentation with Automatic Quality-Guided Self-Training, *Kaixuan Lu, Mehmet Onurcan Kaya, Dim P. Papadopoulos*
19 Locally Explaining Prediction Behavior via Gradual Interventions and Measuring Property Gradients, *Niklas Penzel, Joachim Denzler*
20 Beyond Paired Data: Self-Supervised UAV Geo-Localization from Reference Imagery Alone, *Tristan Amadei, Enric Meinhardt-Llopis, Benedicte Bascle, Corentin Abgrall, Gabriele Facciolo*
21 Where is the Watermark? Interpretable Watermark Detection at the Block Level, *Maria Bulychev, Neil G Marchant, Benjamin I. P. Rubinstein*
22 PointSt3R: Point Tracking through 3D Ground Correspondence, *Rhodri Guerrier, Adam W. Harley, Dima Damen*
23 Uplifting Table Tennis: A Robust, Real-World Application for 3D Trajectory and Spin Estimation, *Daniel Kienzle, Katja Ludwig, Julian Lorenz, Shin'Ichi Satoh, Rainer Lienhart*
24 FairVLM: Enhancing Fairness and Prompt Sensitivity in Vision Language Models for Medical Image Segmentation, *Md Motiur Rahman, Saeka Rahman, Smriti Bhatt, Miad Faezipour*
25 SHaSaM: Submodular Hard Sample Mining for Fair Facial Attribute Recognition, *Anay Majee, Rishabh Iyer*
26 DM3Net: Dual-Camera Super-Resolution via Domain Modulation and Multi-scale Matching, *Cong Guan, Jiacheng Ying, Osamu Yoshie, Yuya Ieiri*
27 SuperRivolution: Fine-Scale Rivers from Coarse Temporal Satellite Imagery, *Rangel Daroya, Subhransu Maji*
28 Adversarial Pseudo-replay for Exemplar-free Class-incremental Learning, *Hiroto Honda*
29 DiffRegCD: Integrated Registration and Change Detection with Diffusion Features, *Seyedehanita Madani, Rama Chellappa, Vishal M. Patel*
30 HDR Reconstruction Boosting with Training-Free and Exposure-Consistent Diffusion, *Yo-Tin Lin, Su-Kai Chen, Hou-Ning Hu, Yen-Yu Lin, Yu-Lun Liu*
31 Optimizing against Infeasible Inclusions from Data for Semantic Segmentation through Morphology, *Shamik Basu, Luc Van Gool, Christos Sakaridis*
32 3D Cell Oversegmentation Correction via Geo-Wasserstein Divergence, *Peter Chen, Bryan Chang, Olivia A Creasey, Julie Beth Sneddon, Zev J Gartner, Yining Liu*
33 TopoRec: Point Cloud Recognition Using Topological Data Analysis, *Anirban Ghosh, Iliya Kulbaka, Ian Dahlin, Ayan Dutta*
34 MapleGrasp: Mask-guided Feature Pooling for Language-driven Efficient Robotic Grasping, *Vineet Bhat, Naman Patel, Prashanth Krishnamurthy, Ramesh Karri, Farshad Khorrami*
35 Remote Sensing Forestry Similarity Convolution, *Shikuan Wang, Yuangong Chen, Jianzhou Gong, Lingyi Meng, Mengquan Wu, Longxing Liu, Haiwei Yuan, Mingbin Guo*
36 RampWatch: An In-the-Wild Dataset and Text-Guided Detection Framework for Recreational Vessels, *Malik Muhammad Asim, Claire B. Smallwood, Abdullah Tariq, Johnny Lo, Syed Zulqarnain Gilani*
37 Enhancing Reverse Distillation with Core Exemplar Learning for Unified Multi-Class Anomaly Detection, *Heechul Lim, Min-Soo Kim, Hyun-Boo Lee, Suk-Ju Kang, Kang-Wook Chon, Haeyun Lee*
38 Leveraging Sparsity for Privacy in Collaborative Inference, *Maximilian Andreas Hoefler, Karsten Mueller, Wojciech Samek*
39 Improvise, Adapt, Overcome — Telescopic Adapters for Efficient Fine-tuning of Vision Language Models in Medical Imaging, *Ujjwal Mishra, Vinita Shukla, Praful Hambarde, Amit Shukla*
40 SVD-Det: A Lightweight Framework for Video Forgery Detection Using Semantic and Visual Defect Cues, *Tsung-Shan Yang, Tianyu Zhang, Feng Qian, Bing Yan, C.-C. Jay Kuo*
41 Joint Optimization of Camera Model and Deep Neural Network for Image Recognition, *Youta Noboru, Yuko Ozasa, Masayuki Tanaka*
42 Training-free Conditional Image Embedding Framework Leveraging Large Vision Language Models, *Masayuki Kawarada, Kosuke Yamada, Antonio Tejero-de-Pablos, Naoto Inoue*
43 ReFineVQA: Iterative Refinement of Video Description via Feedback Generation for Video Question Answering, *Jeongwan Shin, Chan Hur, Seongmin Cho, Jaeho Choi, Hyeyoung Park*
44 MIST: Multilingual Incidental Dataset for Scene Text Detection, *Saumya Mundra, Ajoy Mondal, C.V. Jawahar*
45 Multi-Grained Text-Guided Image Fusion for Multi-Exposure and Multi-Focus Scenarios, *Mingwei Tang, Jiahao Nie, Guang Yang, Ziqing Cui, Jie Li*
46 MemeTAG: Keyword-Driven Meme Classification through Tag Embedding Reconstruction, *Akshit Sharma, Prashant W Patil*
47 Leveraging Semantic Attribute Binding for Free-Lunch Color Control in Diffusion Models, *Héctor Laria, Alexandra Gomez-Villa, Jiang Qin, Muhammad Atif Butt, Bogdan Raducanu, Javier Vazquez-Corral, Joost van de Weijer, Kai Wang*
48 MedPEFT-CL: Dual-Phase Parameter-Efficient Continual Learning with Medical Semantic Adapter and Bidirectional Memory Consolidation, *Ziyuan Gao, Philippe Morel*
49 Splatter Layout: Geometry-embedded 3D Reconstruction via Surface Unfolding, *Bryan Heryanto, Tackgeun You, Chanwoo Kim, Hwasup Lim*
50 HOLO: Holistic Lightweight Optimization for Scene Understanding with Auto-Annotation and Multimodal Learning, *Xiaoyun Hu, Xiaohan Yan, Nan Wang, Gang Wei, Zhicheng Wang*
51 PDV: Prompt Directional Vectors for Zero-shot Composed Image Retrieval, *Osman Tursun, Sinan Kalkan, Simon Denman, Clinton Fookes*
52 GRAPE (Gaussian Rendering for Accelerated Pixel Enhancement) Brings Fast and Lightweight Arbitrary Super-Resolution, *Jung In Jang, Kyong Hwan Jin*
53 Fetal and Neonatal Cortical Surface Reconstruction with Anatomical Normal-guidance and Perceptual Enhancements, *Jiyang Lee, Woori Bae, U-Geun Ji, Hanyeol Yang, Jong-Min Lee*

54 View-aware Cross-modal Distillation for Multi-view Action Recognition, *Trung Thanh Nguyen, Yasutomo Kawanishi, Vijay John, Takahiro Komamizu, Ichiro Ide*
55 NRGMark: Localized Watermarking for Energy Transparency in Images, *Shruti Agarwal, Élie Michel, Vishal Asnani, Tania Mathern, John Collomosse*
56 Test-Time Consistency in Vision Language Models, *Shih-Han Chou, Shivam Chandhok, James J. Little, Leonid Sigal*
57 Revisiting an Old Perspective Projection for Monocular 3D Morphable Models Regression, *Toby Chong, Ryota Nakajima*
58 General and Domain-Specific Zero-shot Detection of Generated Images via Conditional Likelihood, *Roy Betser, Omer Hofman, Roman Vainshtein, Guy Gilboa*
59 Co-STAR: Collaborative Curriculum Self-Training with Adaptive Regularization for Source-Free Video Domain Adaptation, *Amirhossein Dadashzadeh, Parsa Esmati, Majid Mirmehdi*
60 Being Positive about Negative Queries: Exclusion Aware Multimodal Retrieval using Disentangled Representations, *Prachi Jha, Sumit Bhatia, Srikanta Bedathur*
61 DualRes: Production-ready Dynamic Object Detection, *Jibril El Hassani, Thomas Verelst*
62 Semantic Map Guided Bird's-Eye View Learning for Online HD Map Construction, *Huantao Ren, Hesham M. Eraqi, ABM Musa, Mohamed Moustafa*
63 ISALux: Illumination and Semantics-Aware Transformer Employing Mixture of Experts for Low Light Image Enhancement, *Raul Balmez, Alexandru Brateanu, Ciprian Orhei, Codruta O. Ancuti, Cosmin Ancuti*
64 FastPose-ViT: A Vision Transformer for Real-Time Spacecraft Pose Estimation, *Pierre Ancey, Andrew Price, Saqib Javed, Mathieu Salzmann*
65 QuEENet: Quantum-Enhanced Expressive Network for Image Classification, *Shashank Bayal, Rushikesh Govind Dawane, Komal Komal, Santosh Kumar Vipparthi, Subrahmanyam Murala*
66 SOLAR: Switchable Output Layer for Accuracy and Robustness in Once-for-All Training, *Shaharyar Ahmed Khan Tareen, Lei Fan, Xiaojing Yuan, Qin Lin, Bin Hu*
67 FG-TRACER: Tracing Information Flow in Multimodal Large Language Models in Free-Form Generation, *Alessia Saporita, Vittorio Pipoli, Federico Bolelli, Lorenzo Baraldi, Andrea Acquaviva, Elisa Ficarra*
68 Patch Your Matcher: Correspondence-Aware Image-to-Image Translation Unlocks Cross-Modal Matching via Single-Modality Priors, *Anton Frolov, Volker Rodehorst*
69 Diversity Preserving Coresets for Image Quality Assessment, *Arpita Nema, Hanwei Zhu, Xi Zhang, Weisi Lin*
70 SAVE: Sparse Autoencoder-Driven Visual Information Enhancement for Mitigating Object Hallucination, *Sangha Park, Seungryong Yoo, Jisoo Mok, Sungroh Yoon*
71 NavMapFusion: Diffusion-based Fusion of Navigation Maps for Online Vectorized HD Map Construction, *Thomas Monninger, Zihan Zhang, Steffen Staab, Sihao Ding*
72 GFT: Graph Feature Tuning for Efficient Point Cloud Analysis, *Manish Dhakal, Venkat R Dasari, Rajshekhar Sunderraman, Yi Ding*
73 QAL : A Loss for Recall–Precision Balance in 3D Reconstruction, *Pranay Meshram, Yash Turkar, Kartikeya Singh, Praveen Raj Masilamani, Charuvahan Adhivarahan, Karthik Dantu*
74 Meta-YOLO: Metadata-Guided Real-Time Object Detector in Aerial Imagery, *Deukryeol Yoon, Seonghak Kim, Young Hwa Sung, Jinho Jung*
75 Q-Former Autoencoder: A Modern Framework for Medical Anomaly Detection, *Francesco Dalmonte, Emirhan Bayar, Emre Akbas, Mariana-Iuliana Georgescu*
76 AusSmoke meets MultiNatSmoke: a fully-labelled diverse smoke segmentation dataset, *Weihao Li, Hongjin Zhao, Gao Zhu, Ge-Peng Ji, Nicholas Wilson, Marta Yebra, Nick Barnes*
77 NAPP: Noise-Adaptive Prototype Perturbation for Few-Shot Learning, *Ilhwan Kim, Sangwoo Yun, Dongheon Lee, Seongsu Kim, Joonki Paik*
78 GaussianHeadTalk: Wobble-Free 3D Talking Heads with Audio Driven Gaussian Splatting, *Madhav Agarwal, Mingtian Zhang, Laura Sevilla-Lara, Steven McDonagh*
79 Splannequin: Freezing Monocular Mannequin-Challenge Footage with Dual-Detection Splatting, *Hao-Jen Chien, Yi-Chuan Huang, Chung-Ho Wu, Wei-Lun Chao, Yu-Lun Liu*
80 Optimization-Free Style Transfer for 3D Gaussian Splats, *Raphael Du Sablon, David Hart*
81 PEaRL: Pathway-Enhanced Representation Learning for Gene and Pathway Expression Prediction from Histology, *Sejuti Majumder, Saarthak Kapse, Moinak Bhattacharya, Xuan Xu, Alisa Yurovsky, Prateek Prasanna*
82 Flood-LDM: Generalizable Latent Diffusion Models for rapid and accurate zero-shot High-Resolution Flood Mapping, *Sun Han Neo, Sachith Seneviratne, Viraj Vidura Herath Herath Mudiyanselage, Abhishek Saha, Sanka Rasnayaka, Lucy Amanda Marshall*
83 LangPose: Language-Aligned Motion for Robust 3D Human Pose Estimation, *Longyun Liao, Rong Zheng*
84 SphereEdit: Spherical Semantic Editing in Diffusion Models, *Salamata Konate, Hassan Hamidi, Frank Rudzicz, Elham Dolatabadi, Laleh Seyyed-Kalantari*
85 FedEFC: Federated Learning Using Enhanced Forward Correction Against Noisy Labels, *Seunghun Yu, Jin-Hyun Ahn, Joonhyuk Kang*
86 Photo Dating by Facial Age Aggregation, *Jakub Paplhám, Vojtech Franc*
87 Scalable Video Action Anticipation with Cross Linear Attentive Memory, *Zeyun Zhong, Manuel Martin, David Schneider, David J. Lerch, Chengzhi Wu, Frederik Diederichs, Juergen Gall, Jürgen Beyerer*
88 DUDA: Distilled Unsupervised Domain Adaptation for Lightweight Semantic Segmentation, *Beomseok Kang, Niluthpol Chowdhury Mithun, Abhinav Rajvanshi, Han-Pang Chiu, Supun Samarasekera*
89 Generalizing Sports Feedback Generation by Watching Competitions and Reading Books: A Rock Climbing Case Study, *Arushi Rai, Adriana Kovashka*
90 START: Spatial and Textual Learning for Chart Understanding, *Zhuoming Liu, Xiaofeng Gao, Feiyang Niu, Qiaozi Gao, Liu Liu, Robinson Piramuthu*
91 VRAgent: Self-Refining Agent for Zero-Shot Multimodal Video Retrieval, *Ketul Shah, Pankaj Nathani, Rama Chellappa, Fabian Caba Heilbron*
92 MapVerse: A Benchmark for Geospatial Question Answering on Diverse Real-World Maps, *Sharat Bhat, Harshita Khandelwal, Tushar Kataria, Vivek Gupta*
93 IMPACT: Interpretable Most Important Person Analysis and Classification using Transformer-based Models, *Akshat Rampuria, Kamakshya Prasad Nayak, Kamalakar Vijay Thakare, Tushar Joshi, Aditya Dhananjay Singh, Haesol Park, Heeseung Choi, Debi Prosad Dogra, Ig-Jae Kim*
94 SurgXBench: Explainable Vision-Language Model Benchmark for Surgery, *Jiajun Cheng, Xianwu Zhao, Sainan Liu, Xiaofan Yu, Ravi Prakash, Patrick J. Codd, Jonathan Elliott Katz, Shan Lin*
95 SeqFeedNet: Sequential Feature Feedback Network for Background Subtraction, *Yu-Shun Huang, Jing-Ming Guo, Yi-Xiang Yang*
96 Robust Scene Coordinate Regression via Geometrically-Consistent Global Descriptors, *Son Tung Nguyen, Alejandro Fontan, Michael Milford, Tobias Fischer*
97 VitaGlyph: Vitalizing Artistic Typography with Flexible Dual-branch Diffusion Models, *Kailai Feng, Yabo Zhang, Haodong Yu, Zhilong Ji, Jinfeng Bai, Hongzhi Zhang, Wangmeng Zuo*
98 AortaDiff: A Unified Multitask Diffusion Framework for Contrast-Free AAA Imaging, *Yuxuan Ou, Ning Bi, Jiazhen Pan, Boliang Yu, Jiancheng Yang, Usama Zidan, Regent Lee, Vicente Grau*
99 DirectDrag: High-Fidelity, Mask-Free, Prompt-Free Drag-based Image Editing via Readout-Guided Feature Alignment, *Sheng-Hao Liao, Shang-Fu Chen, Tai-Ming Huang, Wen-Huang Cheng, Kai-Lung Hua*

100 Bridging the Domain Gap in Small Multimodal Models: A Dual-level Alignment Perspective, *Aveen Dayal, Peketi Divya, Nidhi Tiwari, Linga Reddy Cenkeramaddi, C Krishna Mohan, Abhinav Kumar*
101 Crash2DocAI: Automated Integration of Post-Crash Car Part Images into Technical Reports, *Vaclav Divis, Marek Hrúz, Jessica Giovagnola, Khalil Ben Chikha*
102 Conversational Image Generation: Towards Multi-Round Personalized Generation with Multi-Modal Language Models, *Haochen Zhang, Animesh Sinha, Felix Juefei-Xu, Haoyu Ma, Kunpeng Li, Zhipeng Fan, Xiaoliang Dai, Tingbo Hou, Peizhao Zhang, Zecheng He*
103 CSF-Net: Context-Semantic Fusion Network for Large Mask Inpainting, *Chae-Yeon Heo, Yeong-Jun Cho*
104 SceneShine: Illumination-aware Human Scene Gaussian Re-Splatting from Mobile Device Video, *Xuqian Ren, Wenjia Wang, Mai Ngoc Nguyen, Juho Kannala, Esa Rahtu*
105 See, Think, Learn: A Self-Taught Multimodal Reasoner, *Sourabh Sharma, Sonam Gupta, Sadbhawna Sadbhawna*
106 SpecGen: Neural Spectral BRDF Generation via Spectral-Spatial Tri-plane Aggregation, *Zhenyu Jin, Wenjie Li, Zhanyu Ma, Heng Guo*
107 Pretraining Helps When Capacity Allows: Evidence from Ultra-Small ConvNets, *Srikanth Muralidharan, Heitor R. Medeiros, Masih Aminbeidokhti, Eric Granger, Marco Pedersoli*
108 Mem-MLP: Real-Time 3D Human Motion Generation from Sparse Inputs, *Sinan Mutlu, Georgios F. Angelis, Savas Ozkan, Paul Wisbey, Anastasios Drosou, Mete Ozay*
109 UltraClean: A Simple Framework to Train Robust Neural Networks against Backdoor Attacks, *Bingyin Zhao, Yingjie Lao*
110 GorillaWatch: An Automated System for In-the-Wild Gorilla Re-Identification and Population Monitoring, *Maximilian Schall, Felix Leonard Knöfel, Noah Elias König, Jan Jonas Kubeler, Maximilian von Klinski, Joan Wilhelm Linnemann, Xiaoshi Liu, Iven Jelle Schlegelmilch, Ole Woyciniuk, Alexandra Schild, Dante Wasmuht, Magdalena Bermejo Espinet, German Illera Basas, Gerard de Melo*
111 Reciprocal Teaching: Dynamic Multi-Model Teacher-Student Learning for Multiple Noisy Annotations, *Wenjie Ai, Cuong C. Nguyen, Gustavo Carneiro, Adrian Hilton*
112 DuPLUS: Dual-Prompt Vision-Language Model for Universal Medical Image Segmentation and Prognosis, *Numan Saeed, Tausifa Jan Saleem, Fadillah Maani, Muhammad Ridzuan, Hu Wang, Mohammad Yaqub*
113 TriaGS: Differentiable Triangulation-Guided Geometric Consistency for 3D Gaussian Splatting, *Quan Tran, Tuan Dang*
114 SDT-6D: Fully Sparse Depth-Transformer for Staged End-to-End 6D Pose Estimation in Industrial Multi-View Bin Picking, *Nico Leuze, Maximilian Hoh, Alfred Schoettl, Samed Doğan, Nicolas R.-Peña*
115 Generalized Category Discovery for LiDAR Semantic Segmentation, *Minseok Kim, Jiyong Boo, Kuk-Jin Yoon*
116 ICONIC-444: A 3.1-Million-Image Dataset for OOD Detection Research, *Gerhard Krumpl, Henning Avenhaus, Horst Possegger*
117 Any Detector Can Detect Anything, *Thomas E. Huang, Siyuan Li, Martin Danelljan, Henghui Ding, Luc Van Gool, Fisher Yu*
118 Towards Unconstrained Cross-View Pose Estimation, *Alexander Wollam, Kyle Ashley, Maxim Shugaev, Oliver Arend, Ilya Semenov, Hadis Dashtestani, Sumved Ravi, Nathan Jacobs*
119 Disentangle and Regularize: Sign Language Production with Articulator-Based Disentanglement and Channel-Aware Regularization, *Sümeyye Meryem Taşyürek, Tuğçe Kızıltepe, Hacer Yalim Keles*
120 SmoothDiffusion-VE: Real-time Generative Video Editing Using Adaptive Feature Cache, *Mustafa Munir, Sophia Zalewski, Shiqiu Liu, David Tarjan, Sushmitha Belede, Anjul Patney, Radu Marculescu*
121 SafeguardGS: 3D Gaussian Primitive Pruning While Avoiding Catastrophic Scene Destruction, *Yongjae Lee, Zhaoliang Zhang, Deliang Fan*
122 Uncertainty-Aware Vision-Language Segmentation for Medical Imaging, *Aryan Das, Tanishq Rachamalla, Koushik Biswas, Swalpa Kumar Roy, Vinay Kumar Verma*
123 Stabilizing Direct Training of Spiking Neural Networks: Membrane Potential Initialization and Threshold-robust Surrogate Gradient, *Hyunho Kook, Byeongho Yu, Jeong Min Oh, Eunhyeok Park*
124 DocWaveDiff: A Predict-and-Refine approach for Document Image Enhancement with Wavelet U-Nets and Diffusion models, *Matteo Marulli, Marco Bertini*
125 Chain-of-Look Spatial Reasoning for Dense Surgical Instrument Counting, *Rishikesh Bhyri, Brian R Quaranto, Junsong Yuan, Peter C W Kim, Nan Xi*
126 Rethinking Latent Variable in Learned Image Compression, *Fangzhou Yi, Zhicheng Gong, Hui Zeng*
127 AugMapNet: Improving Spatial Latent Structure via BEV Grid Augmentation for Enhanced Vectorized Online HD Map Construction, *Thomas Monninger, Md Zafar Anwar, Stanislaw Antol, Steffen Staab, Sihao Ding*
128 From Cognitive Priors to Instance Semantics: A Unified Framework for Multi-task Affective Computing, *Guanyu Hu, Dimitrios Kollias, Xinyu Yang*
129 FuLLaMa: Training-free Diffusion-based Object Removal with Context Preservation, *Ilke Demir, Umur Aybars Ciftci*
130 STRinGS: Selective Text Refinement in Gaussian Splatting, *Abhinav Raundhal, Gaurav Behera, P J Narayanan, Ravi Kiran Sarvadevabhatla, Makarand Tapaswi*
131 VIZOR: Viewpoint-Invariant Zero-Shot Scene Graph Generation for 3D Scene Reasoning, *Vivek Madhavaram, Vartika Sengar, Arkadipta De, Charu Sharma*
132 Relevance-aware Multi-context Contrastive Decoding for Retrieval-augmented Visual Question Answering, *Jongha Kim, Byungoh Ko, Jeehye Na, Jinsung Yoon, Hyunwoo J. Kim*
133 CanKD: Cross-Attention-based Non-local Operation for Feature-based Knowledge Distillation, *Shizhe Sun, Wataru Ohyama*
134 FlyPose: Towards Robust Human Pose Estimation From Aerial Views, *Hassaan Farooq, Marvin Brenner, Peter Stütz*
135 MedROV: Towards Real-Time Open-Vocabulary Detection Across Diverse Medical Imaging Modalities, *Tooba Tehreem Sheikh, Jean Lahoud, Rao Muhammad Anwer, Fahad Shahbaz Khan, Salman Khan, Hisham Cholakkal*
136 Exploiting Label-Independent Regularization from Spatial Patterns for Whole Slide Image Analysis, *Weiyi Wu, Xinwen Xu, Chongyang Gao, Xingjian Diao, Siting Li, Jiang Gui*
137 SVS-GAN for Semantic Synthesis of Traffic Videos for Autonomous Driving, *Khaled M. Seyam, Julian Wiederer, Markus Braun, Bin Yang*
138 Integrating Multi-scale and Multi-filtration Topological Features for Medical Image Classification, *Pengfei Gu, Huimin Li, Haoteng Tang, Dongkuan Xu, Erik Enriquez, DongChul Kim, Bin Fu, Danny Z. Chen*
139 NeuroBridge: Few-Shot Cross-Modal Neuron Re-identification via Dual-Channel Deep Metric Learning, *Wenwei Li, Mingwei Liao, Lingyi Cai, Anan Li*
140 Sketch3R: Rapid and Realistic 3D VR Sketch Creation to Shape Retrieval, *Mritunjoy Halder, Shivam Ashok Shukla, Lokender Tiwari, Raghav Mittal, Brojeshwar Bhowmick*
141 PhyEduVideo: A Benchmark for Evaluating Text-to-Video Models for Physics Education, *Megha Mariam K.M, Aditya Arun, Zakaria Laskar, C.V. Jawahar*
142 Dual-Domain Multimodal Hyperbolic Fusion for Cardiopulmonary Disease Diagnosis in Emergency Care, *Ke Nan, Maggie Samaan, Benjamin Burns, Xia Ning, Yuchi Han, Yuan Xue*

NOTES


To Starr Pass Golf Club, Tennis Courts and Catalina Barbeque Co. & Sports Bar (1 mile)
WACV 2026
TUCSON, AZ
3/6 - 3/10
Kehg-Nena Path
JW MARRIOTT
TUCSON STARR PASS
To Front Desk
Spa Lap Pool
Ania Lawn
Ania Terrace
Starbucks
(Level 3)
Resort Marketplace
Aki Bridge
Building 6
Executive Boardroom 2
Executive Boardroom 1
ATM
Hashani Spa
La Estrella
A
San Luis 2 1
San Ignacio
San Xavier
San Pedro 2 1
Passaggio
B
3
To Front Desk
Building 7 (Level 3)
(Level 2)
Spa Parking
Self Parking
Self Parking Below
Tucson Ballroom
J I H G F E D C B A
Private Balcony
Tucson Registration Desk
Arizona Ballroom
12 11 10 9 8 7 6 5 4 3 2 1
Arizona Registration Desk
Lower Level
Starr Circle Foyer
Starr Circle (Level 3)
Service Drive (Lower Level)
A 24 hour Self Serve Business Center
B Elevator to Self Parking/ Conference Center
1 Guest Room Elevators
Handicapped Accessible
Room Location Key
First Number - Floor
Second Number - Building
Last 2 Numbers - Room Number
S
E
W
N