2026 IEEE/CVF Winter Conference on Applications of Computer Vision Workshops (WACVW 2026)

Tucson, Arizona, USA
6-10 March 2026

Pages 1–612

IEEE Catalog Number: CFP26B39-POD
ISBN: 979-8-3315-9150-2

IEEE Advancing Technology for Humanity

Copyright © 2026, IEEE

All Rights Reserved

Copyright and Reprint Permissions:

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

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

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

IEEE Catalog Number:	CFP26B39-POD
ISBN (Print-On-Demand):	979-8-3315-9150-2
ISBN (Online):	979-8-3315-9149-6
ISSN:	2572-4398

Additional Copies of This Publication Are Available From:

Curran Associates, Inc
57 Morehouse Lane
Red Hook, NY 12571 USA

Phone: (845) 758-0400
Fax: (845) 758-2633
E-mail: curran@proceedings.com
Web: www.proceedings.com



TABLE OF CONTENTS

MedProbCLIP: Probabilistic Adaptation of Vision-Language Foundation Model for Reliable Radiograph-Report Retrieval 1
Ahmad Elallaf, Yu Zhang, Yuktha Priya Masupalli, Jeong Yang, Young Lee, Zechun Cao, Gongbo Liang

When AI Watches the Dose: Can Vision-Language Models Perform TB Medication Adherence Assessments Like Clinicians? 11
Yanqiao Wang, Patrick Kaggwa, Heeya Jolly, Suchendra M. Bhandarkar, Juliet Nabbuye Sekandi, Jin Sun

NEURO-GUARD: Neuro-Symbolic Generalization and Unbiased Adaptive Routing for Diagnostics - Explainable Medical AI 21
Midhat Urooj, Ayan Banerjee, Sandeep Gupta

Controllable Image Synthesis for Endoscopy: Leveraging Text and Spatial Guidance in Diffusion Models 30
Lu Xu, Anuja Vats, Marius Pedersen, Kiran Raja

Reamil: Reasoning- and Evidence-Aware Multiple Instance Learning for Whole-Slide Histopathology 40
Hyun Do Jung, Jungwon Choi, Hwiyoung Kim

K-Track: Kalman-Enhanced Tracking for Accelerating Deep Point Trackers on Edge Devices 46
Bishoy Galoaa, Pau Closas, Sarah Ostadabbas

PrismVAU: Prompt-Refined Inference System for Multimodal Video Anomaly Understanding 55
Iñaki Erregue, Kamal Nasrollahi, Sergio Escalera

TrafficRAG: Temporal Grounding for Traffic Violations via Retrieval-Augmented Generation 66
Soo Yong Kim, Joonyoung Kim, Jaewon Lee, Seong Wook Lee, Tae-San Eom, Jeonghun Chae

Detector-Augmented SAMURAI for Long-Duration Drone Tracking 75
Tamara R. Lenhard, Andreas Weinmann, Hichem Snoussi, Tobias Koch

Robust Thermal Image Object Detection Via Appearance-Guided Mixture of Experts 85
Andreas Aakerberg, Kamal Nasrollahi, Thomas B. Moeslund

SOVABench: A Vehicle Surveillance Action Retrieval Benchmark for Multimodal Large Language Models 95
Oriol Rabasseda, Zenjie Li, Kamal Nasrollahi, Sergio Escalera

Exploring Diffusion-generated Guidance for Thermal Image Super-resolution 105
Leo Thomas Ramos, Angel D. Sappa

Feature Alignment and Compositional Token for Human Pose Estimation 115
Po-Chi Hsu, Ming-Han Lee, Kun-Ru Wu, Yu-Chee Tseng

Training-Free Target Emphasis With Sam2 Pseudo-Masks for Robust Single Object Tracking 125
Byeongseong Lee, Jihong Min

SiamTrackCaps: A Siamese-Based Single Object Tracker Using Capsule Networks 132
Abdullah M. Algamdi, Victor Sanchez, Chang-Tsun Li

Robust Thermal Image Object Detection Challenge: Advancing Multi-Object Detection Performance Under Long-Term Thermal Drift 142
Anders Skaarup Johansen, Marco Parola, Andreas Aakerberg, Mario Giovanni Cosmimo Antonio Cimino Sergio Escalara, Thomas B. Moeslund, Kamal Nasrollahi

Robust Object Detection for Long-Term Thermal Surveillance via Explicit Background Modeling 152
Junwon Son, Jaehyun Cha, Joonha Park, Sihyung Kim, Siyoung Kim, Yoojoong Kim, Wonsuk Kim

Advancing Robust Infrared Object Detection: Practical Insights on Model Development and Evaluation 161
Daniel Stadler

Detecting Object Tracking Failure via Sequential Hypothesis Testing 171
Alejandro Monroy Muñoz, Rajeev Verma, Alexander Timans

FG-DINO: Inducing Foreground Approximation Into DINO for Low-Resolution Surveillance Videos 181
Malik Bouguila, Stefan Wolf, Jürgen Beyerer

Improving Thermal Object Detection Robustness via Zero-Shot Inpainting-Based Data Augmentation 193
Seonggyun Jeong, Wonil Lee, Hyunsik Oh

Seeing Isn't Believing: Context-Aware Adversarial Patch Synthesis via Conditional GAN 202
Roie Kazoom, Alon Goldberg, Hodaya Cohen, Ofer Hadar

Towards Consistent and Efficient Decision-Based Attacks 212
Henning Duwe, Anna Münz, Holger H. Hoos

Forensic Detection of Generated MRI Imagery Using Autoregressive Modeling and Frequency Analysis 222
Arpan Mahara, Naphtali Rishe, Malek Adjouadi

Detecting Deepfake Talking Heads from Facial Biometric Anomalies 232
Justin D. Norman, Hany Farid

Read and Tell - Speech Dataset for Scripted and Spontaneous Scenarios 241
Ewelina Bartuzi-Trokielewicz, Alicja Martinek, Joanna Gajewska, Michał J. Ołowski, Donat Stankiewicz Adam Baran, Adrian Kordas, Michał Koźbiał, Elżbieta Gomulska, Jarosław Wójtowicz

Audio Deepfake Detectors vs. Real Fraud - the Fall of Benchmarks 251
Joanna Gajewska, Alicja Martinek, Ewelina Bartuzi Trokielewicz

Towards Facilitated Fairness Assessment of AI-Based Skin Lesion Classifiers Through GenAI-Based Image Synthesis 261
Ko Watanabe, Stanislav Frolov, Aya Hassan, David Dembinsky, Adriano Lucieri, Andreas Dengel

SGW-GAN: Sliced Gromov-Wasserstein Guided GANs for Retinal Fundus Image Enhancement 272
Yujian Xiong, Xuanzhao Dong, Wenhui Zhu, Xin Li, Oana Dumitrascu, Yalin Wang

nnMobileNet++: Towards Efficient Hybrid Networks for Retinal Image Analysis 282
Xin Li, Wenhui Zhu, Xuanzhao Dong, Hao Wang, Yujian Xiong, Oana Dumitrascu, Yalin Wang

Bridging Restoration and Diagnosis: A Comprehensive Benchmark for Retinal Fundus Enhancement 293
Xuanzhao Dong, Wenhui Zhu, Xiwen Chen, Hao Wang, Xin Li, Yujian Xiong, JiaJun Cheng, Zhipeng Wang Shao Tang, Oana M. Dumitrascu, Yalin Wang

What Matters in Reinforcement Learning Based Training for Medical Vision Language Model: an Empirical Study 303
Wenhui Zhu, Xuanzhao Dong, Xin Li, Peijie Qiu, Xiwen Chen, Walker Northrup, Abolfazl Razi, Aris Sotiras Yi Su, Zhipeng Wang, Shao Tang, Yalin Wang

VAOT: Vessel-Aware Optimal Transport for Retinal Fundus Enhancement 314
Xuanzhao Dong, Wenhui Zhu, Yujian Xiong, Xiwen Chen, Hao Wang, Xin Li, Jiajun Cheng, Zhipeng Wang Shao Tang, Oana Dumitrascu, Yalin Wang

LiteViT-CAM: A Lightweight and Intrinsically Explainable Framework for Realtime Clinical Video Analysis on Edge Devices 324
Gokaramaiah Thota, Nagaraju K, Sathya Babu Korra, Viswanath Pulabaigari

Basic: Bayesian Spiral Attention Classifier for Interpretable Medical Image Classification 333
Abhinav Sagar

Pmaf Loss: Probabilistic Margin-Aware Focal Loss for Robust Medical Image Classification 344
Abhinav Sagar

Balancing the Scales: Uncertainty-Aware Two-Stage RL for Multiview Child Malnutrition Screening 350
Misaal Khan, Mayank Vatsa, Kuldeep Singh, Richa Singh

TFFM: Topology-Aware Feature Fusion Module Via Latent Graph Reasoning for Retinal Vessel Segmentation 359
Iftekhar Ahmed, Shakib Absar, Aftar Ahmad Sami, Shadman Sakib, Debojyoti Biswas Seraj Al Mahmud Mostafa

GraDeCAR-Gradual Denoising by Contrastive Agreement-based Relabeling for Tackling Label Noise in Medical Imaging 369
Manish Gayen, Samiran Das, Shubhobrata Bhattacharya

EUS-FAT: Video-Based Multiple Instance Learning for Liver Fat Quantification from Endoscopic Ultrasound 379
Sahika Betul Yayli, Fatih Ibrahim Ozlugedik, Nicolo Pecco, Shuji Mitsuhashi, Vinay Chandrasekhara Bradley J Erickson

Bridging the Trust Gap: Interpretable AI for Clinical Decision Support in Arrhythmia Diagnostics 387
Nchofon Tagha Ghogomu, Damilare Emmanuel Olatunji, Carine Pierrette Mukamakuza

Improving Negation Understanding in Medical Vision-Language Models via Contrastive Fine-Tuning 395
Jasmine Vu, Shivanand Sheshappanavar

NeuroVLM: A Contrastive Vision-Language Model for Medical Reasoning in Alzheimer's Disease Diagnosis 405
Rakib Hossain Sajib, Md Baharul Islam, Md Kishor Morol, Marjia Sultana

OMNI-Dent: Towards an Accessible and Explainable AI Framework for Automated Dental Diagnosis 415
Leeje Jang, Yao-Yi Chiang, Angela M. Hastings, Patimaporn Pungchanchaikul, Martha B. Lucas Emily C. Schultz, Jeffrey P. Louie, Mohamed Estai, Wen-Chen Wang, Ryan H.L. Ip, Boyen Huang

TranSegNet: A Transformer Guided Framework for Enhanced Segmentation of Infarct Core and Affected ASPECTS Regions for Ischemic Stroke Assessment 425
Anagha Harindranath, Sony M. S., Sumi Suresh M. S., Vivek Nambiar, Vivek Menon

LLM Augmented Intervenable Multimodal Adaptor for Post-operative Complication Prediction in Lung Cancer Surgery 434
Shubham Pandey, Bhavin Jawade, Srirangaraj Setlur, Venu Govindaraju, Kenneth Seastedt

From Downhill to Slalom: A Multi-Period Probabilistic Framework for Adaptive Snow Management in Alpine Ski Racing 444
Zichao Li

Reference-based Positive Ski Coaching using Foot Pressure and MLLMs 452
Toshihiro Hirano, Hitoshi Yoshihara, Chen-Chieh Liao, Erwin Wu, Hideki Koike

eSkiTB: A Synthetic Event-Based Dataset for Tracking Skiers 463
Krishna Vinod, Joseph Raj Vishal, Kaustav Chanda, Prithvi Jai Ramesh, Yezhou Yang Bharatesh Chakravarthi

Nordic Skiing Dataset (NSD): A Pose Estimation Dataset for Performance Feedback in Cross Country Skiing 473
Emil Lundin, Britta Ayano Wilde, Karl Löwenmark, Jérémy Angulo, Ashish Kumar Singh, Petter Kyösti Andreas Almqvist, Joakim Abrahamsson, Hans-Christer Holmberg, Hamam Mokayed

On the Evaluation of Multimodal Large Language Models for Agricultural Image Classification Across Diverse Tasks 482
Anindya Bijoy Das, Shibbir Ahmed, Shahnewaz Karim Sakib

Supporting Ultra-High-Resolution Digital Agriculture Tasks with Fully Synthetic Curriculum Learning 491
Jacob Hatef, Quentin Anthony, Nawras Alnaasan, Dhabaleswar Panda

Cott-ADNet: Lightweight Real-Time Cotton Boll and Flower Detection Under Field Conditions 500
Rui-Feng Wang, Mingrui Xu, Matthew C. Bauer, Iago Beffart Schardong, Xiaowen Ma, Peng W. Chee Kangning Cui

FUME: Fused Unified Multi-Gas Emission Network for Livestock Rumen Acidosis Detection 510
Taminul Islam, Toqi Tahamid Sarker, Mohamed Embaby, Khaled R Ahmed, Amer AbuGhazaleh

A Multi-View Photometric Stereo Pipeline for Specular 3D Fruit Reconstruction 520
Ariel Zuniga-Santana, Gabriele Facciolo, Shohei Nobuhara, Rodrigo Verschae

Advancing Precision Livestock Farming: Robust Country Chicken Detection via FeatherNet Fusion-YOLO and Hensense 530
Maneesh Reddy Bhavanam, Pavan Suraj Kanna, Tarun Sai Reddy Tamma, Chaitanya Alapati

Illegal Waste Dumping Detection 539
Thierry Bouwmans, Antonio Greco, Sébastien Piérard, Andrea Vincenzo Ricciardi, Carlo Sansone Marc Van Droogenbroeck, Bruno Vento

“ScatSpotter” - A Dog Poop Detection Dataset 549
Jonathan Crall

Vision-Language Temporal Analysis for Illegal Waste Dumping Detection in Surveillance Videos 559
Magzhan Kairanbay

RoLID-11K: A Dashcam Dataset for Small-Object Roadside Litter Detection 568
Tao Wu, Qing Xu, Xiangjian He, Oakleigh Weekes, James Brown, Wenting Duan

SortWaste: A Densely Annotated Dataset for Object Detection in Industrial Waste Sorting 576
Sara Inácio, Hugo Proença, João C. Neves

TrashDet: Iterative Neural Architecture Search for Efficient Waste Detection 585
Tony Tran, Bin Hu

Garberus: Three-Headed Video Classifier to Guard Against Illegal Waste Dumps 594
Vincenzo M. Scarrica, Alessio Pierluigi Placitelli, Antonino Staiano

X3D-based Illegal Waste Dumping Detection with Temporal Localization 603
Sajed Almorsy, Marwan Torki

K-Fold Ensemble of Hiera and VideoMAE for Illegal Waste Dumping Detection 613
Seongrok Kim, Donghun Ryu, Donghwan Hwang, Sungwoo Byeon, Minhyeok Lee, Wonsuk Kim

A Lightweight Temporal Detection Framework for Illegal Waste Dumping in Real Surveillance Footage 621
Lorenzo Putzu, Rita Delussu, Mauro Fadda

Hybrid Temporal-Spatial Novelty Detection for Illegal Waste Dumping in Surveillance Video 628
Andrei Daniel Zbuce, Lukas Patrick Laza, Andrei Ionut Ticărat, Elisa Valentina Moisi Daniela Elena Popescu

Boundary-Sensitive Start-Time Estimation with Onset-centric Temporal Detection for Illegal Waste Dumping in Surveillance Video 637
Mudasir Hussain Bhat, Daw-Tung Lin

UEOF: A Benchmark Dataset for Underwater Event-Based Optical Flow 645
Nick Truong, Pritam P. Karmokar, William J. Beksi

EUMD: Event-Based Unrolled Motion Deblurring Via Learnable Prior Optimization 656
Haichuan Zhang, Dongdong Fu, Jonathan Miller, Anustup Choudhury, Chul Lee, Vishal Monga

PhysEye: A Physics-Informed Framework for Unsupervised 3D Eye Tracking with Event Cameras 666
Hicham Sekkati, Jean-Francois Lapointe

SynSacc: A Blender-to-V2E Pipeline for Synthetic Neuromorphic Eye-Movement Data and Sim-to-Real Spiking Model Training 675
Khadija Iddrisu, Waseem Shariff, Suzanne Little, Noel O'Connor

Gram-Schmidt Feature Reduction for Disentangling Style from Degradation in Historical Writer Identification 685
Kaveh Safavigerdini, Bahram Yaghooti, Amir Erfan Z. S. Abadi, Bruno Sinopoli, Kannappan Palaniappan

Pseudo Contrastive Learning for Diagram Comprehension in Multimodal Models 692
Hiroshi Sasaki

An Empirical Study of Siamese Vision Transformers for Scribe Re-Identification 701
Alessio Fagioli, Carmine Fabbri, Luigi Cinque, Emanuela Colombi, Gian Luca Foresti

Adaptive Multi-Scale Feature Fusion for Paragraph-Level Handwritten Text Recognition via Spatial Attention 710
Baha Edine Harrath, Mohammed Hamdan, Mohamed Cheriet

Savior: Sample-Efficient Adaptation of Vision-Language Models for OCR Representation 719
Akshata A Bhat, Sharath Naganna, Saiful Haq, Prashant Khatri, Krishna Chaitanya Reddy Tamataam Neha Arun, Niyati Chhaya, Pushpak Bhattacharyya

Cross-Lingual Transfer for Complex Scripts: A Benchmark on End-to-End Khmer Scene Text Spotting 729
Vannkinh Nom, Saly Keo, Souhail Bakkali, Muhammad Muzzamil Luqman, Mickaël Coustaty Jean-Marc Ogier

SynthForm: Towards a DLA-Free E2E Form Understanding Model 739
Andre Fu, Egil Karlsen

Unsupervised Spatially Aware Gaussian Mixture Model via Implicit Deep Priors 749
Herbert Rakotonirina, Théophile Lohier, Julien Baptiste

Multi-Receptive Field Ensemble with Cross-Entropy Masking for Class Imbalance in Remote Sensing Change Detection 759
Humza Naveed, Xina Zeng, Mitch Bryson, Nagita Mehr Seresht

GLACIA: Instance-Aware Positional Reasoning for Glacial Lake Segmentation via Multimodal Large Language Model 767
Lalit Maurya, Saurabh Kaushik, Beth Tellman

RAVEN: A Rapid Agentic Vision Framework for Emergency Response in Vulnerable Settlements 777
Arkadip Maitra, Suvajit Patra, Sumantro Mukherjee

Agentic Ai in Remote Sensing: Foundations, Taxonomy, and Emerging Systems 786
Niloufar Alipour Talemi, Julia Boone, Fatemeh Afghah

K-Vehicles: A Remote Sensing Dataset for Vehicle Detection in Aerial Imagery 800
Leo Thomas Ramos, Edmundo Casas, Cristian Romero, Francklin Rivas-Echeverría

Gradient-Based Active Learning for Geospatial Semantic Segmentation with Large Vision Models 810
Ryan Faulkenberry, Saurabh Prasad

Type-Aware Ranking of Urban Similarity From Aerial Imagery 821
Idan Kligvasser, Yotam Intrator, George Leifman, Yuval Desheh, Aviad Barzilai, Niv Efron, Ehud Rivlin

Bridging the Domain Gap in Agricultural Vision: Parameter-Efficient VLM Adaptation via Expert Descriptions 830
Deeksha Aggarwal, Yash Mittal, Uttam Kumar

Satellite-Based Detection of Looted Archaeological Sites Using Machine Learning 840
Girmaw Abebe Tadesse, Titien Bartette, Andrew Hassanali, Allen Kim, Jonathan Chemla, Andrew Zolli Yves Ubelmann, Caleb Robinson, Inbal Becker-Reshef, Juan Lavista Ferres

Foundation Models for Phenotyping Segmentation of Wheat Stripe Rust Resistance from UAV Hyperspectral Images 849
Jie Deng, Tailai Chen, Chenyu Li, Danfeng Hong

Point-to-Dense Supervision Framework for SAM Based Remote Sensing Segmentation with Prototype Refinement and Structure Aware, Confidence Guided Topology Learning 859
Anisha Chakravorty, Shounak Chakraborty

The King(dom) is Naked: Lightweight Machine Learning for Hyperspectral Bare Soil Detection 868
Agata M. Wijata, Bogdan Ruszczak, Lukasz Tulczyjew, Nicolas Longépé, Jakub Nalepa

Contrast Then Confidence (C^{2}): Contrastive Pretraining for Uncertainty-Aware Out-of-Distribution Detection in Satellite Imagery 877
Rawhatur Rabbi, Fabliha Afaf Sarwar, Md Farhadul Islam, Tashik Ahmed, Sidu Rudro Md. Mahfujur Rahman, Meem Arafat Manab, Jannatun Noor Mukta

On-The-Fly OVD Adaptation with Flame: Few-Shot Localization Via Active Marginal-Samples Exploration 886
Yehonathan Refael, Amit Aides, Aviad Barzilai, George Leifman, Vered Silverman, Bolous Jaber Genady Beryozkin

Segment Anything but Farms: Comparing Segmentation Paradigms for Rural UAV Captured Ultra-High-Resolution Imagery 895
Snehalraj Chugh, Dharmendra Singh Chaudhary, Subash Sigdel, Shubham Thapa, Lalit BC Nishan Ghimire, Bipendra Basnyat, Nirmalya Roy

GenEava: Generating Cartoon Avatars With Fine-Grained Facial Expressions From Realistic Diffusion-Based Faces 904
Hao Yu, Rupayan Mallick, Margrit Betke, Sarah Adel Bargal

Towards Inclusive Biometrics: Synthetic Generation of Vitiligo Faces and Their Impact on Face Image Quality 915
A. Dörsch, A. Laguna Liang, C. Rathgeb, C. Busch

Ignoring the Decoy: Exposing and Tackling Forensic Distractions in Image Forgery Localization Using Masked Convolutions 924
Xander Staelens, Peter Lambert, Glenn Van Wallendael, Hannes Mareen

The Safe Image Authenticity Challenge: Detecting and Localizing Partial and Fully Synthetic Manipulations 933
Tai Nguyen, Matthew Stamm, Jill Crisman, Jesse Hostetler, Laura Cassano, Michael Davinroy Peter Bautista

Faces of Deception - a Multi-Level Analysis of Lip-Sync-Based Fake Ads Involving Politicians and Their Impact on Content Semantics and Recognition Systems 946
Ewelina Bartuzi-Trokielewicz, Katarzyna Roszczewska, Alicja Martinek

Orthogonal Projection Steering for Bias Mitigation and ICAO Compliance in Diffusion Transformers 956
Raghavendra Mudgalgundurao, Aryan Khurana, Patrick Schuch, Raghavendra Ramachandra, Kiran Raja

SynCAF: Synthetic Counterfactual Activations for Fairness in Facial Cognition Signals 966
Kushbu Agrawal, Nalin Rajput, Raghavendra Ramachandra, Sumarga Kumar Sah Tyagi Deepak Kumar Jain, Rahee Walambe, Ketan Kotecha

Mitigating Longitudinal Performance Degradation in Child Face Recognition Using Synthetic Data 975
Afzal Hossain, Stephanie Schuckers

HiFi-Deblur: High-Frequency Intense Image Deblurring with Frequency-Decoupled U-Net and Discrete Wavelet Transform 984
Jeewoo Lim, Cheolhee Yu, Jin Tae Kwak

Transforming Video Subjective Testing with Training, Engagement, and Real-Time Feedback 994
Kumar Rahul, Sriram Sethuraman, Andrew Segall, Yixu Chen

Diffuse4D: Completing Nerf-Stereo Depth via Diffusion-Driven Restoration in Dynamic Scenes 1003
Yue Hu, Daniel Wang, Xiyun Song, Zhiqiang Lao, Yubin Zhou, Meida Chen, Andrew Feng, Heather Yu

Efficient Deep Demosaicing with Spatially Downsampled Isotropic Networks 1012
Cory Fan, Wenchao Zhang

STEC: A Spatio-Temporal Entropy Coverage Metric for Evaluating Sampled Video Frames 1021
Shih-Yao Lin

VIBEFACE - Video and Image Biometric Dataset for Evaluation of Faces 1028
Ewelina Bartuzi-Trokielewicz, Michał Koźbiał, Adam Baran, Adrian Kordas, Jarosław Wójtowicz
Elżbieta Gomulska, Michał Banach, Alicja Martinek, Michał J. Ołowski, Joanna Gajewska

We Still See Broken Limbs: Towards Anatomical Realism in GenAI Via Human Preference Learning 1038
Stephanie Käs, Erik Schönherr, Bastian Leibe

Cost Savings from Automatic Quality Assessment of Generated Images 1048
Xavier Giró-i-Nieto, Nefeli Andreou, Anqi Liang, Manel Baradad, Francesc Moreno-Noguer
Aleix Martinez

Handsurge: Localized Neural Surgery for Diffusion-Generated Hand Deformity Restoration 1058
Zhiqiang Lao, Dizhi Ma, Heather Yu

REMinD: Balancing Robust Concept Unlearning and Image Quality in Diffusion Models 1068
An Huang, Junggab Son, Zuobin Xiong

SPoRC-VIST: A Benchmark for Evaluating Generative Natural Narrative in Vision-Language Models 1078
Yunlin Zeng

Reason Then Ground: Multilingual Text/Logo Grounding on Movie Posters 1087
Ganesh Samarth, Yarong Feng, Solale Tabarestani, Caren Chen, Xiaonan Zhao

Enhancement as Augmentation: Improving Detection in Highly Degraded Underwater Images Through Mixed-Domain Training 1096
Ashraf Saleem, Alex Hromada, Ali Awad, Amna Mazen, Evan Lucas, Timothy C. Havens

From Filters to VLMs: Benchmarking Defogging Methods Through Object Detection and Segmentation Performance 1106
Ardalan Aryashad, Parsa Razmara, Amin Mahjoub, Seyedarmin Azizi, Mahdi Salmani, Arad Firouzkouhi

Image-Specific Adaptation of Transformer Encoders for Compute-Efficient Segmentation 1116
Manyi Yao, Abhishek Aich, Yumin Suh, Amit Roy-Chowdhury, Christian Shelton, Manmohan Chandraker

Vision Language Models Learn to Assess Images with Specialists 1126
Quyet V. Do, Seunghyun Yoon, Ruiyi Zhang, Thiloshon Nagarajah, Trung Bui, Viet Dac Lai

Quality-Driven and Diversity-Aware Sample Expansion for Robust Marine Obstacle Segmentation 1136
Miaohua Zhang, Mohammad Ali Armin, Xuesong Li, Sisi Liang, Lars Petersson, Changming Sun
David Ahmedt-Aristizabal, Zeeshan Hayder

Seeing in the Dark: Synthesizing Underexposure for More Robust Underwater Image Augmentation 1146
Ali Awad, Evan Lucas, Ashraf Saleem, Shadi Moradi, Anthony Geglio, Angus Galloway, Timothy Havens
Peter Esselman

YOLO-OSA: A ShuffleAttention-Enhanced YOLO Model for FOD Detection with Comprehensive Benchmarking on MS COCO 1154
Soumen Biswas, Abhishek Mitra, Ananth Ganesh, Kingshuk Banerjee

CARLA-Haze: A Synthetic Benchmark for Outdoor Image Dehazing 1164
Henry O. Velesaca, Leo Thomas Ramos, Angel D. Sappa

When Probe and Gallery Are Low Quality: Decreasing Accuracy and Increasing Demographic Disparities in 1:N Identification 1174
Gabriella Pangelinan, Michael C. King, Kevin W. Bowyer

Fast 2DGS: Efficient Image Representation with Deep Gaussian Prior 1184
Hao Wang, Ashish Bastola, Chaoyi Zhou, Wenhui Zhu, Xiwen Chen, Xuanzhao Dong, Siyu Huang Abolfazl Razi

VideoForge: Efficient Domain Adaptation for Video Generation Through Quality-Driven Rewards and Enhanced LoRA 1194
Xiaogang Wang, Bill Cai, Girish Dilip Patil, Yash Shah

ViTNT-FIQA: Training-Free Face Image Quality Assessment With Vision Transformers 1203
Guray Ozgur, Eduarda Caldeira, Tahar Chettaoui, Jan Niklas Kolf, Marco Huber, Naser Damer Fadi Boutros

JetBench: Quality-Aware Benchmarking of Vision Models for Jet Parameter Classification in Heavy-Ion Physics 1214
Haydar Mehryar, Loren Schwiebert

Device-Robust Spectral Grading and Origin Detection from UV-VIS-NIR Images: Towards Practical Gemstone Quality Assessment 1224
Mahule Roy, Srikanth Thudumu, Jason Fisher, Rajesh Vasa, Kon Mouzakis

Subtle Motion Blur Detection and Segmentation from Static Image Artworks 1233
Ganesh Samarth, Sibendu Paul, Solale Tabarestani, Caren Chen

Can You Find the Difference? Visually Identical Image Detection 1243
Haiyun Henry Jin, Yarong Feng, Jiaojiao Zhao, Minmin Shen, Caren Chen

Geo-Dmae: Geometric Deep Multi-Autoencoders for Monitoring Heterogeneous Normal Aging in Brain Subcortex 1250
Yuanyao Tan, Anuj Srivastava, Chao Huang

Latent Space Manifold Geometry for Robust Image Classification 1259
Shenyuan Liang, Benjamin Beaudett, Anuj Srivastava, Mauricio Pamplona Segundo, Sudeep Sarkar

Learning Domain Agnostic Latent Embeddings of 3D Faces for Zero-Shot Animal Expression Transfer 1269
Yue Wang, Lawrence Amadi, Xiang Gao, Yazheng Chen, Yuanpeng Liu, Ning Lu, Xianfeng David Gu

Face Identity Unlearning for Retrieval via Embedding Dispersion 1277
Mikhail Zakharov

Attention as Geometric Transformation: Revisiting Feature Distillation for Semantic Segmentation 1287
Amir M. Mansourian, Arya Jalali, Rozhan Ahmadi, Shohreh Kasaei

ARGS: Advanced Regularization on Aligning Gaussians Over the Surface 1298
Jeong Uk Lee, Sung Hee Choi

Stabilizing Intrinsic Explanations: A Geometric Perspective on Concept Leakage 1307
Keun Hee Park, Jinyung Hong, Theodore P. Pavlic, Pavan Turaga

Improving Shape Bias in Learnable Geometric Moment Representations 1317
Sangmin Jung, Anirudh Rayas, Reza Rahimi Azghan, Hassan Ghasemzadeh, Yezhou Yang, Pavan Turaga

Hyperbolic Embeddings Improve Narrative Quality in Retrieval Augmented Generation Models 1326
Mayukh Debnath, Ankita Shukla

OrthoMDA: Curvature-Preserving Manifold Visualization for Regression Networks via Gram-Schmidt Orthogonalization 1334
Kaveh Safavigerdini, Bahram Yaghooti, Bruno Sinopoli, Kannappan Palaniappan

EzBlender: Efficient 3D Editing with Plan-and-ReAct Agent 1343
Hao Wang, Wenhui Zhu, Shao Tang, Zhipeng Wang, Xuanzhao Dong, Xin Li, Xiwen Chen, Ashish Bastola Xinhao Huang, Yalin Wang, Abolfazl Razi

Tone Matters: The Impact of Linguistic Tone on Hallucination in VLMs 1353
Weihao Hong, Zhiyuan Jiang, Bingyu Shen, Xinlei Guan, Yangyi Feng, Meng Xu, Boyang Li

Under-Canopy Terrain Reconstruction in Dense Forests Using RGB Imaging and Neural 3D Reconstruction 1363
Refael Sheffer, Chen Pinchover, Haim Zisman, Dror Ozeri, Roee Litman

DIS2:Dis entanglement Meets Dis tillation with Classwise Attention for Robust Remote Sensing Segmentation Under Missing Modalities 1373
Nhi Kieu, Kien Nguyen, Arnold Wiliem, Clinton Fookes, Sridha Sridharan

FLASH-SAR: Fast Learning Self-Supervised Hierarchical Architecture for SAR 1382
Sai Shruti Prakhya, Uttam Kumar

Neighborhood Feature Pooling for Remote Sensing Image Classification 1393
Fahimeh Orvati Nia, Amirmohammad Mohammadi, Salim Al Kharsa, Pragati Naikare Zigfried Hampel-Arias, Joshua Peeples

When Less is More: Evaluating Structural Pruning in Geospatial Foundation Models 1403
Amina Said, Julia Dietlmeier, Margaret McCaul, Noel E. O'Connor

Subimage Overlap Prediction: Task-Aligned Self-Supervised Pretraining for Semantic Segmentation in Remote Sensing Imagery 1414
Lakshay Sharma, Alex Marin

SSeg: Active Sparse Point-Label Augmentation for Semantic Segmentation 1424
Cesar Borja, Carlos Plou, Ruben Martinez-Cantin, Ana C. Murillo

Prithvi-Complimentary Adaptive Fusion Encoder (CAFE): Unlocking Full-Potential for Flood Inundation Mapping 1435
Saurabh Kaushik, Lalit Maurya, Beth Tellman

SPEAR: Self-Supervised Sample Efficient Pixel-Level Multi-Modal Spectral Fusion for Earth Observation Applications 1445
Rajiv Ranjan, Udaiveer Singh, Anjali Aggarwal, Shashank Tamaskar, Dharmendra Saraswat

Accuracy Versus Efficiency in Model Selection for Remote Sensing Scientific Workflows 1455
Diellza Sherifi Iseni, Jonathan Bader, Francisco Mena, Odej Kao

Application of AlphaEarth GeoFoundation Model for High-Accuracy Crop Yield Estimation 1465
Jichao Fang, Mingda Wu, Zhou Zhang, Wei Luo

Steering Instruction-Tuned Vision-Language Models for Multi-Label Landcover Classification 1475
Weijia Wang, Jichao Fang, Wei Luo

Enhancing Remote Sensing Change Detection Via Masked Edge Reconstruction 1484
Shuaiyu Chen, Fu Wang, Tianjin Huang, Xiang Li, Siyang Song, Guangliang Cheng, Chunbo Luo, Zeyu Fu

FIRE-VLM: A Vision-Language-Driven Reinforcement Learning Framework for UAV Wildfire Tracking in a Physics-Grounded Fire Digital Twin 1493
Chris Webb, Mobin Habibpour, Mayamin Hamid Raha, Ali Reza Tavakkoli, Janice Coen, Fatemeh Afghah

A Novel Reliability Aware Patch Selection Based Few Shot Learning for Target Recognition and Open-Set Identification in Ill Posed SAR Images 1503
Anisha Chakravorty, Abhishek Vimukt, Shounak Chakraborty, Abhishek Soni

Multi-Label Classification in Remote Sensing: Leveraging High-Resolution Patches for Low-Resolution Tasks 1512
Shreya Pandey, Pragna Echuri, Vishnu Meher Vemulapalli, Shounak Chakraborty

Spectra without Spectra: Learning Spectral Reconstruction from Real-World RGB Data with Sparse Supervision 1521
Shira Werman, Genady Paikin, Roy Shaul, Evgeny Soloveichik

Few-Shot LoRA Adaptation of a Flow-Matching Foundation Model for Cross-Spectral Object Detection 1531
Maxim Clouser, Kia Khezeli, John Kalantari

RangeSAM: On the Potential of Visual Foundation Models for Range-View Represented LiDAR Segmentation 1540
Paul Julius Kühn, Duc Anh Nguyen, Arjan Kuijper, Saptarshi Neil Sinha

Rethinking Semantics for Complex-Scene Thermal Image Generation 1549
Tayeba Qazi, Ayush Maheshwari, Prerana Mukherjee, Brejesh Lall

Addressing Data Scarcity in Materials Science Research with Deep Generative Models 1556
Yilmaz Korkmaz, Deepti Hegde, Kerri-Lee Chintersingh, Milad Alemohammad, Velat Kilic
Michael R. Flickinger, Amee L. Polk, Megan Bohkoor, Cole Peters, Rami O. Knio, Todd C. Hufnagel
Mark A. Foster, Timothy P. Weihs, Vishal M. Patel

Zero-Shot Product Attribute Labeling with Vision-Language Models: A Three-Tier Evaluation Framework 1565
Shubham Shukla, Kunal Sonalkar

CHURI: Contour-Gated, Resource-Intelligent Retail Tracking System 1574
Riya Waghmare

From Pixels to Purchase: Building and Evaluating a Taxonomy-Decoupled Visual Search Engine for Home Goods E-Commerce 1582
Cheng Lyu, Jingyue Zhang, Ryan Maunu, Mengwei Li, Vinny DeGenova, Yuanli Pei

Enhancing Aerial–Ground Video Person Re-Identification via DFGS-Guided CLIP Sampling and Inference-Time Uncertainty-Aware Fusion 1591
Thi Ngoc Ha Nguyen, Tien-Dung Mai

SAS-VPReID: A Scale-Adaptive Framework with Shape Priors for Video-Based Person Re-IDentification at Extreme Far Distances 1599
Qiwei Yang, Pingping Zhang, Yuhao Wang, Zijing Gong

S3-CLIP: Video Super Resolution for Person-ReID 1609
Tamás Endrei, György Cserey

EAGLE-ReID: Strategic Alignment and Delta Consistency for Extreme Far-Distance Aerial-Ground Re-Identification 1618
Cheng-Jun Kang, Jin-Hui Jiang, Yu-Fan Lin, Chih-Chung Hsu

VReID-XFD: Video-Based Person Re-Identification at Extreme Far Distance Challenge Results 1625
Kailash A. Hambarde, Hugo Proença, Md Rashidunnabi, Pranita Samale, Qiwei Yang, Pingping Zhang Zijing Gong, Yuhao Wang, Xi Zhang, Ruoshui Qu, Qiaoyun He, Yuhang Zhang, Thi Ngoc Ha Nguyen Tien-Dung Mai, Cheng-Jun Kang, Yu-Fan Lin, Jin-Hui Jiang, Chih-Chung Hsu, Tamás Endrei György Cserey, Ashwat Rajbhandari

Moving Masks: a Preliminary Study on Face Presentation Attack On-The-Move 1635
Raghavendra Ramachandra, Narayan Vetrekar, Krishna Patel, Marissa Ataide, Sushma Venkatesh Rajendra Gad

Robustness of Presentation Attack Detection in Remote Identity Validation Scenarios 1646
John J. Howard, Richard O. Plesh, Yevgeniy B. Sirotin, Jerry L. Tipton, Arun R. Vemury

When Humans Judge Irises: Pupil Size Normalization as an Aid and Synthetic Irises as a Challenge 1654
Mahsa Mitcheff, Adam Czajka

StructFormer: Structure-Consistent Face De-Identification Under Strong Privacy Constraints* 1663
Haini Zhu, Deepak K. Jain, Xudong Zhao, Muyu Li, Vitomir Štruc, Sumarga K. Sah Tyagi

SIRR-LMM: Single-Image Reflection Removal via Large Multimodal Model 1674
Yu Guo, Zhiqiang Lao, Xiyun Song, Yubin Zhou, Heather Yu

OV-SGT: Open Vocabulary Semantic Graph Transformer for Scene Graph Generation 1685
George Vanica, Adrian G. Bors

Probabilistic Scene Graph Prompting: Uncertainty-Aware Structured Reasoning in Multimodal LLMs 1695
Zichao Li

A Structured Analysis and Taxonomy of Scene Graph Representations for Group Activity Understanding 1703
Narthana Sivalingam, Santhirarajah Sivasthigan, Thamayanthi Mahendranathan, Roshan Godaliyadda Vijitha Herath, Parakrama Ekanayake

LOST-3DSG: Lightweight Open-Vocabulary 3D Scene Graphs with Semantic Tracking in Dynamic Environments 1712
Sara Micol Ferraina, Michele Brienza, Francesco Argenziano, Emanuele Musumeci, Vincenzo Suriani Domenico D. Bloisi, Daniele Nardi

Lightweight Multi-Scale Fusion for Real-Time Autonomous Driving Segmentation 1721
Muhammad Nadeem, HyeYoung Lee

FROST-Drive: Scalable and Efficient End-to-End Driving with a Frozen Vision Encoder 1729
Zeyu Dong, Yimin Zhu, Yu Wu, Yu Sun

Efficient Visual Question Answering Pipeline for Autonomous Driving via Scene Region Compression 1739
Yuliang Cai, Dongqiangzi Ye, Zitian Chen, Chongruo Wu

Benchmarking Vision-Language Models for Traffic Scene Understanding in Inclement Winter Weather: The AWDB Benchmark 1747
Ali Awad, Tagore Rao Kosireddy, Krishna Mokkapati, Nibendra Bajracharya, Ashraf Saleem, Vinh Nguyen Evan Lucas

Role of Language-Guidance in Knowledge Distillation for Semantic Segmentation Under Limited Field-Of-View Autonomous Driving 1761
Yeon Woo Shin, Keun Hee Park, Pavan Turaga, Eun Som Jeon

Less is More: Agentic Prompt Design for Safe VLM Action Selection 1772
Elahe Yahyapour, Chengbo Ai

Trust-Guided Multimodal LLM Integration with Reinforcement Learning for Autonomous Driving 1780
Sairam Chennaka, Jaswanth Nidamanuri

VLA4CoDrive: Vision-Language-Action Dataset for Cooperative Autonomous Driving 1789
Sayed Pedram Haeri Boroujeni, Abolfazl Razi

2COOOL: An Evaluation Benchmark for Generating Incident Reports on Out-of-Distribution Hazards in Autonomous Driving 1800
Ali K. AlShami, Ryan Rabinowitz, Terrance E. Boult, Jugal Kalita

GATEPose: A Graph Attention Transformer Enhanced with Pose and Orientation Angles for Pedestrian Crossing Intention Prediction 1810
Ali K. AlShami, Terrance Boult, Jugal Kalita