

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

**Tucson, Arizona, USA
6-10 March 2026**

Pages 1–666

IEEE Catalog Number: CFP26082-POD
ISBN: 979-8-3315-5512-2



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:	CFP26082-POD
ISBN (Print-On-Demand):	979-8-3315-5512-2
ISBN (Online):	979-8-3315-5511-5
ISSN:	2472-6737

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

DreamAnywhere: Object-Centric Panoramic 3D Scene Generation	1
<i>Edoardo A. Dominici, Jozef Hladky, Floor Verhoeven, Lukas Radl, Thomas Deixelberger, Stefan Ainetter Philipp Drescher, Stefan Hauswiesner, Arno Coomans, Giacomo Nazzaro, Konstantinos Vardis Markus Steinberger</i>	
ViSTA: Visual Storytelling using Multi-modal Adapters for Text-to-Image Diffusion Models	12
<i>Sibo Dong, Ismail Shaheen, Maggie Shen, Rupayan Mallick, Sarah Adel Bargal</i>	
Odo: Depth-Guided Diffusion for Identity-Preserving Body Reshaping	22
<i>Siddharth Khandelwal, Sridhar Kamath, Arjun Jain</i>	
BiPO: Bidirectional Partial Occlusion Network for Text-to-Motion Synthesis	32
<i>Seong-Eun Hong, SooBin Lim, JuYeong Hwang, Minwook Chang, HyeongYeop Kang</i>	
Reinforcement Learning-based Adaptive Control of Classifier-Free Guidance and Timestep Embeddings in Diffusion Models	43
<i>Haochen You, Baojing Liu, Hongyang He</i>	
TS-PCI: Point Cloud Frame Interpolation with Time-Aware Point Cloud Sampling and Self-Supervised Learning Strategy	54
<i>Kohei Matsuzaki, Keisuke Nonaka</i>	
Enhanced Back-Projection of Vision Features for 3D Symmetry Detection	66
<i>Isaac Aguirre, Ivan Sipiran</i>	
OracleGS: Grounding Generative Priors for Sparse-View Gaussian Splatting	77
<i>Atakan Topaloğlu, Kunyi Li, Michael Niemeyer, Nassir Navab, A. Murat Tekalp, Federico Tombari</i>	
UnderWater SLAM with Laser-light sectioning method using ST-GAT	88
<i>Heyang Gao, Kazuto Ichimaru, Takafumi Iwaguchi, Hiroshi Kawasaki</i>	
Leveraging Pretrained Representations for Cross-Modal Point Cloud Completion	97
<i>Kshitij Kale, Hrishikesh U, V Sreenidhe, Shylaja S S</i>	
Referring Change Detection in Remote Sensing Imagery	106
<i>Yilmaz Korkmaz, Jay N. Paranjape, Celso M. de Melo, Vishal M. Patel</i>	
Inpaint360GS: Efficient Object-Aware 3D Inpainting via Gaussian Splatting for 360° Scenes	117
<i>Shaoxiang Wang, Shihong Zhang, Christen Millerdurai, Rüdiger Westermann, Didier Stricker, Alain Pagani</i>	
A Multi-Agent Diffusion Approach for MRI Anomaly Segmentation via Modality-Specific LoRA Specialization	128
<i>Wafa Al Ghallabi, Muhammad Zaigham Zaheer, Ritesh Thawkar, Omkar Thawakar, Salman Khan Fahad Shahbaz Khan</i>	
GenHSI: Controllable Generation of Human-Scene Interaction Videos	138
<i>Zekun Li, Rui Zhou, Rahul Sajnani, Xiaoyan Cong, Daniel Ritchie, Srinath Sridhar</i>	
SymNet: A Multi-Task Network for Joint Radio Map Reconstruction and Transmitter Localization	150
<i>Lyuzhou Ye, Thanh Dat Le, Yan Huang</i>	

LooC: Effective Low-Dimensional Codebook for Compositional Vector Quantization	160
<i>Jie Li, Kwan-Yee K. Wong, Kai Han</i>	
End-to-End Fine-Tuning of 3D Texture Generation using Differentiable Rewards	171
<i>AmirHossein Zamani, Tianhao Xie, Amir G. Aghdam, Tiberiu Popa, Eugene Belilovsky</i>	
IDEAL-M3D: Instance Diversity-Enriched Active Learning for Monocular 3D Detection	181
<i>Johannes Meier, Florian Günther, Riccardo Marin, Oussema Dhaouadi, Jacques Kaiser, Daniel Cremers</i>	
FAE-Net: Fashion Attribute Editing via Disentangled Latent Conditioning in Diffusion Models	192
<i>P. Rajith Bhargav, Gaurab Bhattacharya, B S Vivek, Jayavardhana Gubbi</i>	
DiT-VTON: Diffusion Transformer Framework for Unified Multi-Category Virtual Try-On and Virtual Try-All with Integrated Image Editing	202
<i>Qi Li, Shuwen Qiu, Kee Kiat Koo, Julien Han, Karim Bouyarmane</i>	
MAESTRO: Masked AutoEncoders for Multimodal, Multitemporal, and Multispectral Earth Observation Data	212
<i>Antoine Labatie, Michael Vaccaro, Nina Lardiere, Anatol Garioud, Nicolas Gonthier</i>	
TA-Prompting: Enhancing Video Large Language Models for Dense Video Captioning via Temporal Anchors	225
<i>Wei-Yuan Cheng, Kai-Po Chang, Chi-Pin Huang, Fu-En Yang, Yu-Chiang Frank Wang</i>	
Towards Fast and Scalable Normal Integration using Continuous Components	236
<i>Francesco Milano, Jen Jen Chung, Lionel Ott, Roland Siegwart</i>	
A Framework for Real-Time Surgical Phase Recognition with Application to Robot-Assisted Partial Nephrectomy	245
<i>Marco Mezzina, Tom Vercauteren, Tinne Tuytelaars, Matthew B. Blaschko</i>	
A Woman with a Knife or A Knife with a Woman? Measuring Directional Bias Amplification in Image Captions	255
<i>Rahul Nair, Bhanu Tokas, Hannah Kerner</i>	
Forget Less by Learning Together through Concept Consolidation	265
<i>Arjun Ramesh Kaushik, Naresh Kumar Devulapally, Vishnu Suresh Lokhande, Nalini Ratha Venu Govindaraju</i>	
Action Anticipation at a Glimpse: To What Extent Can Multimodal Cues Replace Video?	276
<i>Manuel Benavent-Lledo, Konstantinos Bacharidis, Victoria Manousaki, Konstantinos Papoutsakis Antonis Argyros, Jose Garcia-Rodriguez</i>	
VISTA: A Vision and Intent-Aware Social Attention Framework for Multi-Agent Trajectory Prediction	287
<i>Stephane Da Silva Martins, Emanuel Aldea, Sylvie Le Hégarat-Masclé</i>	
ChameleonTuner: Automatic ISP Color Tuning in Subjective Scenarios	297
<i>Zijie Tan, Yuxin Yue, Bahador Rashidi</i>	
ART: Actor-Related Tubelet for Detecting Complex-shaped Action Tubes	308
<i>Jiaojiao Zhao</i>	
Training-free Multi-view 4D Human Motion Reconstruction Virtual Reality System	318
<i>Yijie Li, Ce Zheng, Yijie He, Joel Julin, Ryosuke Ichikari, Satoki Ogiso, Satoshi Nakae, Akihiro Sato Takeshi Kurata, László A. Jeni</i>	

EmojiDiff: Advanced Facial Expression Control with High Identity Preservation in Portrait Generation	328
<i>Liangwei Jiang, Ruida Li, Zhifeng Zhang, Shuo Fang, Chenguang Ma</i>	
OW-Rep: Open World Object Detection with Instance Representation Learning	339
<i>Sunoh Lee, Minsik Jeon, Jihong Min, Junwon Seo</i>	
Cluster-Guided Adversarial Perturbations for Robust Contrastive Learning	350
<i>Seongyun Seo, Sungmin Han, Jeonghyun Lee, Sangkyun Lee</i>	
A Universal Self-Attention Enhancement for Bridging Low-bit Quantization and Vision Transformers	360
<i>Jiahe Qian, Peisong Wang, Zhengyang Zhuge, Qinghao Hu, Jian Cheng</i>	
Overcoming Fine-Grained Visual Challenges in Animal Re-Identification via Semantic Feature Alignment	371
<i>Yihao Wu, Di Zhao, Yuzhuo Li, Matthew Alajas, Alistair S. Glen, Jingfeng Zhang, Gillian Dobbie, Daniel Wilson, Yun Sing Koh</i>	
M4U: Evaluating Multilingual Understanding and Reasoning for Large Multimodal Models	382
<i>Hongyu Wang, Jiayu Xu, Senwei Xie, Ruiping Wang, Jialin Li, Zhaojie Xie, Bin Zhang, Chuyan Xiong, Xilin Chen</i>	
RAT4D: Rig and Animate Objects without Surface Templates in 4D	393
<i>Mosam Dabhi, Simon Lucey, László A. Jeni</i>	
AFL-PRF: Adaptive Federated Learning for Low-Quality Data: Enhancing Performance, Robustness, and Fairness	402
<i>Pinrui Yu, Yiming Xie, Longtian Ye, Geng Yuan, Ningfang Mi, Xue Lin</i>	
Eff-GRot: Efficient and Generalizable Rotation Estimation with Transformers	412
<i>Fanis Mathioulakis, Gorjan Radevski, Tinne Tuytelaars</i>	
DreamMakeup: Face Makeup Customization using Latent Diffusion Models	422
<i>Geon Yeong Park, Inhwa Han, Serin Yang, Yeobin Hong, Seongmin Jeong, Heechan Jeon, Myeongjin Goh, Sung Won Yi, Jin Nam, Jong Chul Ye</i>	
Morphing Through Time: Diffusion-Based Bridging of Temporal Gaps for Robust Alignment in Change Detection	431
<i>Syedehanita Madani, Vishal M. Patel Johns</i>	
AdaptViG: Adaptive Vision GNN with Exponential Decay Gating	440
<i>Mustafa Munir, Md Mostafijur Rahman, Radu Marculescu</i>	
MooTrack360: A Novel Fisheye Camera Dataset for Robust Multi Dairy Cow Detection and Tracking	451
<i>Rasmus Gjerlund K. Christiansen, Toan Van Nguyen, Lasse Rose Malskær, Leon Bodenhagen, Dirk Kraft</i>	
MDUNet: Multimodal Decoding UNet for Passive Occluder-Aided Non-line-of-sight 3D Imaging	461
<i>Fadlullah Raji, John Murray-Bruce</i>	
Interleaved Vision-and-Language Generation via Generative Voken	472
<i>Kaizhi Zheng, Xuehai He, Xin Eric Wang</i>	
Root Completion from Intraoral Scans of Tooth Crowns using Diffusion with Patch Perturbation	483
<i>Yohan Jang, In-Seok Song, Seung Jun Baek</i>	

Systematic Analysis of the Unintentional CSAM-Generation-Potential of Text-to-Image Models	493
<i>Nicolas Göller, Martin Steinebach</i>	
SGPMIL: Sparse Gaussian Process Multiple Instance Learning	503
<i>Andreas Lolos, Stergios Christodoulidis, Aris L. Moustakas, Jose Dolz, Maria Vakalopoulou</i>	
Understanding Human-Like Biases in VLMs via Subjective Face Analytics	514
<i>Chaitanya Roygaga, Aparna Bharati</i>	
M-ErasureBench: A Comprehensive Multimodal Evaluation Benchmark for Concept Erasure in Diffusion Models	527
<i>Ju-Hsuan Weng, Jia-Wei Liao, Cheng-Fu Chou, Jun-Cheng Chen</i>	
Revisiting Layer Normalization for Point Cloud Test Time Adaptation	537
<i>Moslem Yazdanpanah, Ali Bahri, Mehrdad Noori, Sahar Dastani, Samuel Barbeau, David Osowiechi Gustavo Adolfo Vargas Hakim, Ismail Ben Ayed, Christian Desrosiers</i>	
Learning to Animate Images from A Few Videos to Portray Delicate Human Actions	547
<i>Haoxin Li, Yingchen Yu, Qilong Wu, Hanwang Zhang, Song Bai, Boyang Li</i>	
One-Shot Fine-Grained Re-Identification of Paint Marked Honey Bees using Vision Foundation Models	560
<i>Luke Meyers, Josué A. Rodríguez-Cordero, Rémi Mégret</i>	
From Street to Orbit: Training-Free Cross-View Retrieval via Location Semantics and LLM Guidance	570
<i>Jeongho Min, Dongyoung Kim, Jaehyup Lee</i>	
Mitigating Backdoor Attacks via Trigger Reconstruction and Model Hardening	580
<i>Guanhong Tao, Siyuan Cheng, Guangyu Shen, Yingqi Liu, Shengwei An, Zhuo Zhang, Zhenting Wang Hanxi Guo, Xiangyu Zhang</i>	
Network-agnostic distortion-robust projections for wide-angle image understanding	591
<i>Akshaya Athwale, Ola Ahmad, Jean-François Lalonde</i>	
4D-Animal: Freely Reconstructing Animatable 3D Animals from Videos	602
<i>Shanshan Zhong, Jiawei Peng, Zehan Zheng, Zhongzhan Huang, Wufei Ma, Guofeng Zhang, Qihao Liu Alan Yuille, Jieneng Chen</i>	
Text Slider: Efficient and Plug-and-Play Continuous Concept Control for Image/Video Synthesis via LoRA Adapters	613
<i>Pin-Yen Chiu, I-Sheng Fang, Jun-Cheng Chen</i>	
Cluster-based Pseudo-labeling for Semi-Supervised LiDAR Semantic Segmentation	623
<i>Qingju Guo, Shuang Li, Jing Geng, Binhui Xie, Jiawei Shan, Wei Li</i>	
Human Pose Aggregation for Multi-View Temporal Video Alignment	635
<i>Fabien Delattre, Tsung-Wei Huang, Guan-Ming Su, Erik Learned-Miller</i>	
Marshaled Learning: Bridging Large Neural Networks with Memory-Constrained Trusted Execution Environments in Federated Learning	647
<i>Shiwei Ding, Xiaoyong Yuan, Zhenlin Wang, Lan Emily Zhang, Giuseppe Ateniese</i>	
Feature-Disentangling RGB-NIR Fusion Network for Remote Driver Physiological Measurement	657
<i>Tayssir Bouraffa, Ziyuan Wang, Daniel Strüber</i>	

Prompt-OT: An Optimal Transport Regularization Paradigm for Knowledge Preservation in Vision-Language Model Adaptation	667
<i>Xiwen Chen, Wenhui Zhu, Peijie Qiu, Hao Wang, Huayu Li, Haiyu Wu, Aristeidis Sotiras, Yalin Wang Abolfazl Razi</i>	
Zero-Shot Video Deraining with Video Diffusion Models	677
<i>Tuomas Varanka, Juan Luis Gonzalez, Hyeonwoo Kim, Pablo Garrido, Xu Yao</i>	
Joint Modeling of Corruption-Driven and Information-Limited Uncertainty for Robust 3D Gaussian Splatting	688
<i>Zeji Hui, Amirali Khodadadian Gostar, WeiQin Chuah, Alireza Bab-Hadiashar, Ruwan Tennakoon</i>	
Data-Driven Lipschitz Continuity: A Cost-Effective Approach to Improve Adversarial Robustness	698
<i>Erh-Chung Chen, Pin-Yu Chen, I-Hsin Chung, Che-Rung Lee</i>	
CADe: Continual Weakly-supervised Video Anomaly Detection with Ensembles	708
<i>Satoshi Hashimoto, Tatsuya Konishi, Tomoya Kaichi, Kazunori Matsumoto, Mori Kurokawa</i>	
PaRaChute: Pathology-Radiology Cross-Modal Fusion for Missing-Modality-Robust Survival Prediction	718
<i>Pietro Caforio, Isabella Poles, Marco D. Santambrogio</i>	
How to Design and Train Your Implicit Neural Representation for Video Compression	729
<i>Matthew Gwilliam, Roy Zhang, Namitha Padmanabhan, Hongyang Du, Abhinav Shrivastava</i>	
Can We Challenge Open-Vocabulary Object Detectors with Generated Content in Street Scenes?	740
<i>Annika Mütze, Sadia Ilyas, Chrsitian Dörpelkus, Matthias Rottmann</i>	
Conjuring Positive Pairs for Efficient Unification of Representation Learning and Image Synthesis	751
<i>Imanol G. Estepa, Jesús M. Rodríguez-de-Vera, Ignacio Sarasúa, Bhalaji Nagarajan, Petia Radeva</i>	
DICE: Discrete Inversion Enabling Controllable Editing for Masked Generative Models	762
<i>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</i>	
Deepfake Detection that Generalizes Across Benchmarks	773
<i>Andrii Yermakov, Jan Cech, Jiri Matas, Mario Fritz</i>	
Unified Video Anomaly Detection Model for Detecting Different Anomaly Types	784
<i>Kijung Lee, Youngwan Jo, Sunghyun Ahn, Sanghyun Park</i>	
SCORP: Scene-Consistent Object Refinement via Proxy Generation and Tuning	795
<i>Ziwei Chen, Ziling Liu, Zitong Huang, Mingqi Gao, Feng Zheng</i>	
CountingDINO A Training-free Pipeline for Class-Agnostic Counting using Unsupervised Backbones	806
<i>Giacomo Pacini, Lorenzo Bianchi, Luca Ciampi, Nicola Messina, Giuseppe Amato, Fabrizio Falchi</i>	
ViGG: Robust RGB-D Point Cloud Registration using Visual-Geometric Mutual Guidance	816
<i>Congjia Chen, Shen Yan, Yufu Qu</i>	
Gaussian Representations for Video	827
<i>Sachin Shah, Anustup Choudhury, Guan-Ming Su, Jaclyn Pytlarz, Christopher A. Metzler, Trisha Mittal</i>	

MSRTrack: LLM-Powered Object Tracking with Motion and Semantic Reasoning	838
<i>Tong Shen, Di Wang, José M. F. Moura</i>	
Enabling High-Quality In-the-Wild Imaging from Severely Aberrated Metalens Bursts	849
<i>Debabrata Mandal, Zhihan Peng, Yujie Wang, Praneeth Chakravarthula</i>	
Boosting Medical Vision-Language Pretraining via Momentum Self-Distillation under Limited Computing Resources	860
<i>Phuc Pham, Nhu Pham, Ngoc Quoc Ly</i>	
Beyond Realism: Learning the Art of Expressive Composition with StickerNet	869
<i>Haoming Lu, David Kocharian, Humphrey Shi</i>	
Analysis of Text Accuracy and Visual Alignment in Vision-Language Models for Artistic Text Generation	879
<i>Fatima Alderazi, Motaz Alfarraj</i>	
MEDAL: multi-modal MEta-space Distillation and ALignment for Visual Compatibility Learning	888
<i>Dween Rabius Sanny, Vinay Kumar Verma, Prateek Sircar, Deepak Gupta</i>	
PS3: Part level instance segmentation in 3D	898
<i>Hong-Xuan Yen, Chiamin Chen, Yanqing Wang, Yu-Lun Liu, Min Sun</i>	
Saliency-Guided DETR for Moment Retrieval and Highlight Detection	907
<i>Aleksandr Gordeev, Vladimir Dokholyan, Irina Tolstykh, Maksim Kuprashevich</i>	
Accelerated Dose Generation in Gamma Knife Radiosurgery Using a Wavelet Diffusion Model for Sparse Representation	917
<i>Sangyoon Lee, Shubuendu Mishra, Yoichi Watanabe</i>	
DermEVAL: A Dermatologist-Reviewed Benchmark for Multimodal Large Language Models	927
<i>Hongjin Zhao, Weihao Li, Zhenyue Qin, Ge-Peng Ji, Yang Liu, Tom Gedeon, Nick Barnes</i>	
SynPlay: Large-Scale Synthetic Human Data with Real-World Diversity for Aerial-View Perception	938
<i>Jinsub Yim, Hyungtae Lee, Sungmin Eum, Yi-Ting Shen, Yan Zhang, Heesung Kwon Shuvra S. Bhattacharyya</i>	
Do generative video models understand physical principles?	948
<i>Saman Motamed, Laura Culp, Kevin Swersky, Priyank Jaini, Robert Geirhos</i>	
ZonUI-3B: Competitive GUI Grounding with a 3B VLM Trained on a Single Consumer GPU	959
<i>ZongHan Hsieh, ShengJing Yang, Tzer-Jen Wei</i>	
No MoCap Needed: Post-Training Motion Diffusion Models with Reinforcement Learning using Only Textual Prompts	967
<i>Macaluso Girolamo, Mandelli Lorenzo, Mirko Bicchierai, Stefano Berretti, Andrew D. Bagdanov</i>	
MIX-based Foreground and Background Patch Augmentation Guided by Physics and Material Properties for X-ray Detection	977
<i>Xintong Liu, Dongliang Chang, Yujun Tong, Zhanyu Ma</i>	
Reverse Personalization	988
<i>Han-Wei Kung, Tuomas Varanka, Nicu Sebe</i>	

Beyond the Encoder: Joint Encoder-Decoder Contrastive Pre-Training Improves Dense Prediction	1000
<i>Sébastien Quetin, Tapotosh Ghosh, Farhad Maleki</i>	
HyperPose: Hyper-pose Embeddings for 3D-Aware Generative Models with Self-Supervised Disentangling of Pose and Scene	1011
<i>Mijeong Kim, Namgi Kim, Bohyung Han</i>	
Learning Mask-Aware Offsets: Two-branch Deformable Attention Networks for Inpainting with Masked Region Avoidance	1022
<i>Hyeongseok Oh, Joonki Paik</i>	
Saliency-SGG: Enhancing Unbiased Scene Graph Generation with Iterative Saliency Estimation	1032
<i>Runfeng Qu, Ole Hall, Pia K Bideau, Julie Ouerfelli-Ethier, Martin Rolfs, Klaus Obermayer, Olaf Hellwich</i>	
SmokeBench: Evaluating Multimodal Large Language Models for Wildfire Smoke Detection	1043
<i>Tianye Qi, Weihao Li, Nick Barnes</i>	
Visibility guided Self-Supervised Occlusion-Resilient Human Pose Estimation	1054
<i>Arindam Dutta, Sarosij Bose, Rohit Kundu, Calvin-Khang Ta, Saketh Bachu, Konstantinos Karydis Amit K. Roy-Chowdhury</i>	
Diverse Sketch Colorization with Content-Enhanced Style Representation and Recolorization Distillation	1064
<i>Shuangming Mao, Haixiang Zhu</i>	
GateFusion: Hierarchical Gated Cross-Modal Fusion for Active Speaker Detection	1074
<i>Yu Wang, Juhyung Ha, Frangil M. Ramirez, Yuchen Wang, David J. Crandall</i>	
MR-Pruner: Training-free Multi-resolution Visual Token Pruning for Multi-modal Large Language Models	1084
<i>Seunghoon Han, Hyewon Lee, Soyoung Park, Jong-Ryul Lee, Sungsu Lim</i>	
SpikeRain: Towards Energy-Efficient Single Image Deraining with Spiking Neural Networks	1094
<i>Md Tanvir Islam, Inzamamul Alam, Sambit Bakshi, Khan Muhammad, Javier Del Ser, Sangtae Ahn</i>	
MBTI: Metric-Based Textual Inversion for Fine-Grained Image Generation	1106
<i>Byungkwan Chae, Youngjae Choi, Heewon Kim</i>	
ImageNet-sES: A First Systematic Study of Sensor–Environment Simulation Anchored by Real Recaptures	1117
<i>Ji-yoon Kim, Eunsu Baek, Hyung-Sin Kim</i>	
MomentMix Augmentation with Length-Aware DETR for Temporally Robust Moment Retrieval	1127
<i>Seojeong Park, Jiho Choi, Kyungjune Baek, Hyunjung Shim</i>	
From Lightweight CNNs to SpikeNets: Benchmarking Accuracy–Energy Tradeoffs with Pruned Spiking SqueezeNet	1137
<i>Radib Bin Kabir, Tawsif Tashwar Dipto, Mehedi Ahamed, Sabbir Ahmed, Md Hasanul Kabir</i>	
AFRAgent : An Adaptive Feature Renormalization Based High Resolution Aware GUI agent	1147
<i>Neeraj Anand, Rishabh Jain, Sohan Patnaik, Balaji Krishnamurthy, Mausoom Sarkar</i>	
BanglaProtha: Evaluating Vision Language Models in Underrepresented Long-tail Cultural Contexts	1159
<i>Md Fahim, Md Sakib Ul Rahman, Akm Moshir Rahman, Md Farhan Ishmam, Md Tasmim Rahman Fariha Tanjim Shifat, Fabiha Haider, Md Farhad Alam Bhuiyan</i>	

CommonForms: A Large, Diverse Dataset for Form Field Detection	1170
<i>Joe Barrow</i>	
Beyond Real Weights: Hypercomplex Representations for Stable Quantization	1180
<i>Jawad Ibn Ahad, Maisha Rahman, Amrijit Biswas, Muhammad Rafsan Kabir, Robin Krambroeckers</i>	
<i>Sifat Momen, Nabeel Mohammed, Shafin Rahman</i>	
mEOL: Training-Free Instruction-Guided Multimodal Embedder for Vector Graphics and Image Retrieval	1191
<i>Kyeong Seon Kim, Baek Seong-Eun, Lee Jung-Mok, Tae-Hyun Oh</i>	
Contrastive Integrated Gradients: A Feature Attribution-Based Method for Explaining Whole Slide Image Classification	1201
<i>Anh Mai Vu, Tuan L. Vo, Ngoc Lam Quang Bui, Nam N. B. Le, Akash Awasthi, Huy Q. Vo</i>	
<i>Thanh-Huy Nguyen, Zhu Han, Chandra Mohan, Hien Van Nguyen</i>	
PSA-MIL: A Probabilistic Spatial Attention-Based Multiple Instance Learning for Whole Slide Image Classification	1211
<i>Sharon Peled, Yosef E. Maruvka, Moti Freiman</i>	
Fully Unsupervised Self-debiasing of Text-to-Image Diffusion Models	1221
<i>Korada Sri Vardhana, Shrikrishna Lolla, Soma Biswas</i>	
Model-free Domain Adaptation for Concealed Multimodal Large-Language Models	1231
<i>Yu Mitsuzumi, Akisato Kimura, Hisashi Kashima</i>	
CAAC: Confidence-Aware Attention Calibration to Reduce Hallucinations in Large Vision-Language Models	1242
<i>Mehrdad Fazli, Bowen Wei, Ahmet Sari, Ziwei Zhu</i>	
LiDAR-DHMT: LiDAR-Adaptive Dual Hierarchical Mask Transformer for Robust Freespace Detection and Semantic Segmentation	1252
<i>Siyu Chen, Ting Han, Changshe Zhang, Xin Luo, Huan Chen, Meiliu Wu, Guorong Cai, Jinhe Su</i>	
Mixed Diffusion for 3D Indoor Scene Synthesis	1262
<i>Siyi Hu, Diego Martin Arroyo, Stephanie Debats, Fabian Manhardt, Luca Carlone, Federico Tombari</i>	
PiSA: A Self-Augmented Data Engine and Training Strategy for 3D Understanding with Large Models	1273
<i>Zilu Guo, Hongbin Lin, Zhihao Yuan, Chaoda Zheng, Pengshuo Qiu, Dongzhi Jiang, Renrui Zhang</i>	
<i>Chun-Mei Feng, Zhen Li</i>	
MARS: a Multimodal Alignment and Ranking System for Few-Shot Segmentation	1284
<i>Nico Catalano, Stefano Samele, Paoloertino, Matteo Matteucci</i>	
SilverLining: Data-First Mitigation of Spatial and Spectral Shortcuts Without Introducing New Confounders	1294
<i>Balagopal Unnikrishnan, Michael Brudno, Chris McIntosh</i>	
Distilling Diversity and Control in Diffusion Models	1304
<i>Rohit Gandikota, David Bau</i>	
Narrating For You: Prompt-guided Audio-visual Narrating Face Generation Employing Multi-entangled Latent Space	1314
<i>Aashish Chandra K, Aashutosh A V, Abhijit Das</i>	

Explaining the Unseen: Multimodal Vision-Language Reasoning for Situational Awareness in Underground Mining Disasters	1324
<i>Mizanur Rahman Jewel, Mohamed Elmahallawy, Sanjay Madria, Samuel Frimpong</i>	
FALCONEye: Finding Answers and Localizing Content in ONE-hour-long videos with multi-modal LLMs	1334
<i>Carlos Plou, Cesar Borja, Ruben Martinez-Cantin, Ana C. Murillo</i>	
Autocorrelation-based Fiducial Markers for Traceability	1345
<i>Ismail Bencheikh, Max Dunitz, Marie D’Autume, Enric Meinhardt-Llopis, Marc Pic, Gabriele Facciolo Pablo Musé</i>	
GrounDiff: Diffusion-Based Ground Surface Generation from Digital Surface Models	1355
<i>Oussema Dhaouadi, Johannes Meier, Jacques Kaiser, Daniel Cremers</i>	
QuadraNet V2: Efficient and Sustainable Training of High-Order Neural Networks with Quadratic Adaptation	1365
<i>Chenhui Xu, Fuxun Yu, Jinjun Xiong, Xiang Chen</i>	
AutoSew: A Geometric Approach to Stitching Prediction with Graph Neural Networks	1374
<i>Pablo Ríos-Navarro, Elena Garces, Jorge Lopez-Moreno</i>	
SaccadeX: Directed Acyclic Graph-based Semi-Supervised Learning of Continuous Ocular Dynamics from Sparse Neuromorphic Streams	1384
<i>Nuwan Bandara, Thivya Kandappu, Archan Misra</i>	
Federated Model Synchronization for Diagnostic Redefinition through a Novel Selective Parameter Unlearning	1395
<i>Mayank Kumar Kundalwal, Mamta, Deepak Mishra, Asif Ekbal</i>	
A Fast, Simple, and Flexible Scale Informative Feature Transform Module for Arbitrary Scale Image Super-Resolution	1405
<i>Aupendu Kar, Prabir Kumar Biswas</i>	
MageBench: Bridging Large Multimodal Models to Agents	1415
<i>Miaosen Zhang, Qi Dai, Yifan Yang, Jianmin Bao, Dongdong Chen, Kai Qiu, Chong Luo, Xin Geng Baining Guo</i>	
You May Speak Freely: Improving the Fine-Grained Visual Recognition Capabilities of Multimodal Large Language Models with Answer Extraction	1428
<i>Logan Lawrence, Oindrila Saha, Megan Wei, Chen Sun, Subhransu Maji, Grant Van Horn</i>	
InteracTalker: Prompt-Based Human-Object Interaction with Co-Speech Gesture Generation	1438
<i>Sreehari Rajan, Kunal Bhosikar, Charu Sharma</i>	
ITSELF: Attention Guided Fine-Grained Alignment for Vision–Language Retrieval	1448
<i>Tien-Huy Nguyen, Huu-Loc Tran, Thanh Duc Ngo</i>	
MarineEval: Assessing the Marine Intelligence of Vision-Language Models	1459
<i>Yuk-Kwan Wong, Tuan-An To, Jipeng Zhang, Ziqiang Zheng, Sai-Kit Yeung</i>	
Identity Verification from Human Scent using Channel Representation of 2D Gas Chromatography-Mass Spectrometry Data	1471
<i>Radim Spetlik, Jan Hlavsa, Jana Čechová, Petra Pojmanová, Jiří Matas, Štěpán Urban</i>	

milliMamba: Specular-Aware Human Pose Estimation via Dual mmWave Radar with Multi-Frame Mamba Fusion	1481
<i>Niraj Prakash Kini, Shiau-Rung Tsai, Guan-Hsun Lin, Wen-Hsiao Peng, Ching-Wen Ma, Jenq-Neng Hwang</i>	
OpenCowID: Zero-Shot Visual Identification of Dairy Cows	1491
<i>Omkar Prabhune, Younghyun Kim</i>	
QCFace: Image Quality Control for boosting Face Representation & Recognition	1501
<i>Duc-Phuong Doan-Ngo, Thanh-Dang Diep, Thanh Nguyen-Duc, Thanh-Sach LE, Nam Thoai</i>	
MMHOI: Modeling Complex 3D Multi-Human Multi-Object Interactions	1512
<i>Kaen Kogashi, Anoop Cherian, Meng-Yu Jennifer Kuo</i>	
BrightRate: Quality Assessment for User-Generated HDR Videos	1522
<i>Shreshth Saini, Bowen Chen, Yilin Wang, Neil Birkbeck, Balu Adsumilli, Alan C. Bovik</i>	
Reviving Unsupervised Optical Flow: Concept Reevaluation, Multi-Scale Advances and Full Open-Source Release	1533
<i>Azin Jahedi, Marc Rivinius, Noah Berenguel Senn, Andrés Bruhn</i>	
UniCoRN: Latent Diffusion-based Unified Controllable Image Restoration Network across Multiple Degradations	1543
<i>Debabrata Mandal, Soumitri Chattopadhyay, Guansen Tong, Praneeth Chakravarthula</i>	
DRWKV: Focusing on Object Edges for Low-Light Image Enhancement	1554
<i>Xuecheng Bai, Yuxiang Wang, Boyu Hu, Qinyuan Jie, Chuanzhi Xu, Kechen Li, Hongru Xiao, Vera Chung</i>	
Layout Anything: One Transformer for Universal Room Layout Estimation	1565
<i>Md Sohaq Mia, Muhammad Abdullah Adnan</i>	
BOP-Distrib: Revisiting 6D Pose Estimation Benchmarks for Better Evaluation under Visual Ambiguities	1575
<i>Boris Meden, Asma Brazi, Fabrice Mayran de Chamisso, Steve Bourgeois, Vincent Lepetit</i>	
Cosine Similarity is Almost All You Need (for Prototypical-Part Models)	1586
<i>Luke Moffett, Frank Willard, Maximillian Machado, Emmanuel Mokol, Jon Donnelly, Zhicheng Guo, Adam Costarino, Julia Yang, Giyoung Kim, Alina Jade Barnett, Cynthia Rudin</i>	
ORCA: Object Recognition and Comprehension for Archiving Marine Species	1597
<i>Yuk-Kwan Wong, Haixin Liang, Zeyu Ma, Yiwei Chen, Ziqiang Zheng, Rinaldi Gotama, Pascal Sebastian Lauren D. Sparks, Sai-Kit Yeung</i>	
Multimodal Medical Image Binding via Shared Text Embeddings	1610
<i>Yunhao Liu, Suyang Xi, Shiqi Liu, Hong Ding, Chicheng Jin, Chong Zhong, Junjun He, Catherine C. Liu, Yiqing Shen</i>	
PHYSPLAT: A Framework for Photorealistic Hybrid Simulation of Real and Synthetic Elements using 3D Gaussian Splatting	1621
<i>Mario Alfonso-Arsuaga, Henar Dominguez-Elvira, Jorge Casas-Guerrero, Andrea Castiella-Aguirrezabala, Lorenzo Costabile-Dominguez, Jorge Garcia-Gonzalez, Maria Naranjo-Almeida, Marc Comino-Trinidad, Jorge Lopez-Moreno</i>	
ArchitectHead: Continuous Level of Detail Control for 3D Gaussian Head Avatars	1632
<i>Peizhi Yan, Rabab Ward, Qiang Tang, Shan Du</i>	

Uncertainty-Aware Subset Selection for Robust Visual Explainability under Distribution Shifts	1643
<i>Madhav Gupta, Vishak Prasad, Ganesh Ramakrishnan</i>	
The Perceptual Observatory Characterizing Robustness and Grounding in MLLMs	1653
<i>Tejas Anvekar, Fenil Bardoliya, Pavan K. Turaga, Chitta Baral, Vivek Gupta</i>	
FAST-EQA: Efficient Embodied Question Answering with Global and Local Region Relevancy	1664
<i>Haochen Zhang, Nirav Savaliya, Faizan Siddiqui, Enna Sachdeva</i>	
RapidMV: Leveraging Spatio-Angular Latent Space for Efficient and Consistent Text-to-Multi-View Synthesis	1674
<i>Seungwook Kim, Yichun Shi, Kejie Li, Minsu Cho, Peng Wang</i>	
Diffusion-Based Authentication of Copy Detection Patterns: A Multimodal Framework with Printer Signature Conditioning	1685
<i>Bolutife Atoki, Iuliia Tkachenko, Bertrand Kerautret, Carlos Crispim-Junior</i>	
AD ² : Analysis and Detection of Adversarial Threats in Visual Perception for End-to-End Autonomous Driving Systems	1695
<i>Ishan Sahu, Somnath Hazra, Somak Aditya, Soumyajit Dey</i>	
SSplain: Sparse and Smooth Explainer for Retinopathy of Prematurity Classification	1705
<i>Elifnur Sunger, Tales Imbiriba, Peter Campbell, Deniz Erdogmus, Stratis Ioannidis, Jennifer Dy</i>	
Tables Guide Vision: Learning to See the Heart through Tabular Data	1716
<i>Marta Hasny, Maxime Di Folco, Keno Bressem, Julia Schnabel</i>	
SAFER-AiD: Saccade-Assisted Foveal-peripheral vision Enhanced Reconstruction for Adversarial Defense	1726
<i>Jiayang Liu, Daniel Ts'o, Yiming Bu, Qinru Qiu</i>	
Enhancing Object Detection Training via Joint Image-Annotation Generation	1736
<i>Roy Uziel, Oded Bialer</i>	
Descrip3D: Enhancing Large Language Model-based 3D Scene Understanding with Object-Level Text Descriptions	1746
<i>Jintang Xue, Ganning Zhao, Jie-En Yao, Hong-En Chen, Yue Hu, Meida Chen, Suyu You, C.-C. Jay Kuo</i>	
From SAM to DINOv2: Towards Distilling Foundation Models to Lightweight Baselines for Generalized Polyp Segmentation	1757
<i>Shivanshu Agnihotri, Snehashis Majhi, Deepak Ranjan Nayak, Debesh Jha</i>	
Not Like Transformers: Drop the Beat Representation for Dance Generation with Mamba-Based Diffusion Model	1767
<i>Sangjune Park, Inhyeok Choi, Donghyeon Soon, Youngwoo Jeon, Kyungdon Joo</i>	
BoxSplitGen: A Generative Model for 3D Part Bounding Boxes in Varying Granularity	1777
<i>Juil Koo, Wei-Tung Lin, Chanho Park, Chanhyeok Park, Minhyuk Sung</i>	
3D Gaussian Point Encoders	1788
<i>Jim James, Benjamin Wilson, Simon Lucey, James Hays</i>	
HumanGuideNet: Adapter-Based Alignment of Deep Neural Networks with Human Similarity Judgments	1798
<i>Xufu Liu, Yifan Yang, Zhengxin Zhang</i>	

Low-Rank Expert Merging for Multi-Source Domain Adaptation in Person Re-Identification	1809
<i>Taha Mustapha Nehdi, Nairouz Mrabah, Atif Belal, Marco Pedersoli, Eric Granger</i>	
MEGA-PCC: A Mamba-based Efficient Approach for Joint Geometry and Attribute Point Cloud Compression	1820
<i>Kai-Hsiang Hsieh, Monyneath Yim, Wen-Hsiao Peng, Jui-Chiu Chiang</i>	
HyPCA-Net: Advancing Multimodal Fusion in Medical Image Analysis	1831
<i>Joy Dhar, Manish Kumar Pandey, Debashis Das Chakladar, Maryam Haghighat, Azadeh Alavi Sajib Mistry, Nayyar Zaidi</i>	
CycleSL: Server-Client Cyclical Update Driven Scalable Split Learning	1841
<i>Mengdi Wang, Efe Bozkir, Enkelejda Kasneci</i>	
3DSceneEditor: Controllable 3D Scene Editing with Gaussian Splatting	1852
<i>Ziyang Yan, Yihua Shao, Minwen Liao, Siyu Chen, Nan Wang, Muyuan Lin, Jenq-Neng Hwang, Hao Zhao Fabio Remondino, Lei Li</i>	
Segmentation-Aware Latent Diffusion for Satellite Image Super-Resolution: Enabling Smallholder Farm Boundary Delineation	1864
<i>Aditi Agarwal, Anjali Jain, Nikita Saxena, Ishan Deshpande, Michal Kazmierski, Abigail Annkah Nadav Sherman, Karthikeyan Shanmugam, Alok Talekar, Vaibhav Rajan</i>	
mmWEAVER: Environment-Specific mmWave Signal Synthesis from a Photo and Activity Description	1875
<i>Mahathir Monjur, Shahriar Nirjon</i>	
Detection-Driven Object Count Optimization for Text-to-Image Diffusion Models	1885
<i>Oz Zafar, Yuval Cohen, Lior Wolf, Idan Schwartz</i>	
Roadside Monocular 3D Detection Prompted by 2D Detection	1895
<i>Yechi Ma, Wei Hua, Yanan Li, Shu Kong</i>	
UniCalib: Targetless LiDAR-camera Calibration via Probabilistic Flow on Unified Depth Representations	1906
<i>Shu Han, Xubo Zhu, Ji Wu, Ximeng Cai, Wen Yang, Huai Yu, Gui-Song Xia</i>	
Color Bind: Exploring Color Perception in Text-to-Image Models	1916
<i>Shay Shomer-Chai, Wenxuan Peng, Bharath Hariharan, Hadar Averbuch-Elor</i>	
ASC: Learning Augmentation Severity-Consistent Representations Improves Generalization via Augmentation Search	1926
<i>Amirhossein Alamdar, Hossein Jafarinia, Mahdi Noori, Mohammad Hossein Rohban</i>	
Detecting Out-of-Distribution Objects through Class-Conditioned Inpainting	1937
<i>Quang-Huy Nguyen, Jin Peng Zhou, Zhenzhen Liu, Khanh-Huyen Bui, Kilian Q. Weinberger, Wei-Lun Chao Dung D. Le</i>	
Snapmoji: Instant Generation of Animatable Dual-Stylized Avatars	1948
<i>Eric Ming Chen, Di Liu, Sizhuo Ma, Michael Vasilkovsky, Bing Zhou, Qiang Gao, Wenzhou Wang Jiahao Luo, Dimitris N. Metaxas, Vincent Sitzmann, Jian Wang</i>	
Seeing is Believing (and Predicting): Context-Aware Multi-Human Behavior Prediction with Vision Language Models	1959
<i>Utsav Panchal, Yuchen Liu, Luigi Palmieri, Ilche Georgievski, Marco Aiello</i>	

False Alarm Rectification for Early Smoke Segmentation	1969
<i>Hongjin Zhao, Weihao Li, Ge-Peng Ji, Nick Barnes</i>	
Interaction-via-Actions: Cattle Interaction Detection with Joint Learning of Action-Interaction Latent Space	1979
<i>Ren Nakagawa, Yang Yang, Risa Shinoda, Hiroaki Santo, Kenji Oyama, Fumio Okura, Takenao Ohkawa</i>	
Semi-Supervised Hierarchical Open-Set Classification	1989
<i>Erik Wallin, Fredrik Kahl, Lars Hammarstrand</i>	
ClusterMine: Robust Label-Free Visual Out-Of-Distribution Detection via Concept Mining from Text Corpora	1999
<i>Nikolas Adaloglou, Diana Petrusheva, Mohamed Asker, Felix Michels, Markus Kollmann</i>	
CURIO: Curvature-Aligned and Efficient OCR for Low-Resource Historical Manuscripts	2011
<i>Sai Madhusudan Gunda, Tathagata Ghosh, Simran Singh Sandral, Ravi Kiran Sarvadevabhatla</i>	
DoTA: Latent Distribution Conditioned Data Attribution for Diffusion Models	2022
<i>Ninad Joshi, Vivek Srivastava, Shirish Karande</i>	
Learnable Query-Enhanced Pose Transformation	2032
<i>Yi-Zhen Wang, Hong-Han Shuai</i>	
CAMP-VQA: Caption-Embedded Multimodal Perception for No-Reference Quality Assessment of Compressed Video	2042
<i>Xinyi Wang, Angeliki Katsenou, Junxiao Shen, David Bull</i>	
Evaluating Text-to-Image and Text-to-Video Synthesis with a Conditional Fréchet Distance	2052
<i>Jaywon Koo, Jefferson Hernandez, Moayed Haji-Ali, Ziyang Yang, Vicente Ordonez</i>	
From Bands to Depth: Understanding Bathymetry Decisions on Sentinel-2	2063
<i>Satyaki Roy Chowdhury, Aswathnarayan Radhakrishnan, Hari Subramoni</i>	
Pyramidal Spectrum: Frequency-based Hierarchically Vector Quantized VAE for Videos	2073
<i>Tushar Prakash, Onkar Susladkar, Inderjit S. Dhillon, Sparsh Mittal</i>	
PRISM-CAFO: Prior-conditioned Remote-sensing Infrastructure Segmentation and Mapping for CAFOs	2083
<i>Oishee Binte Hoque, Nibir Chandra Mandal, Kyle Luong, Amanda Wilson, Samarth Swarup Madhav Marathe, Abhijin Adiga</i>	
Dressing the Imagination: A Dataset for AI-Powered Translation of Text into Fashion Outfits and A Novel NeRA Adapter for Enhanced Feature Adaptation	2094
<i>Gayatri Deshmukh, Somsubhra De, Chirag Sehgal, Jishu Sen Gupta, Sparsh Mittal</i>	
EllipsianNet: Image-guided Sampling of 2D Gaussians for Gaussian Splatting	2104
<i>MyoungGon Kim, JeongHyeon Ahn, Seohyeon Park, Hyemi Kim, Seunghyun Park, Jung Ho Hwang JungHyun Han</i>	
Zero-Shot Coreset Selection via Iterative Subspace Sampling	2114
<i>Brent A. Griffin, Jacob Marks, Jason J. Corso</i>	
MorphXAI: An Explainable Framework for Morphological Analysis of Parasites in Blood Smear Images	2125
<i>Aqsa Yousaf, Sint Sint Win, Megan Coffee, Habeeb Olufowobi</i>	

WWE-UIE: A Wavelet & White Balance Efficient Network for Underwater Image Enhancement	2135
<i>Ching-Heng Cheng, Jen-Wei Lee, Chia-Ming Lee, Chih-Chung Hsu</i>	
SPAR-Det: Segmentation-guided and Prior-Aided Routing for Small Object Detection	2146
<i>Seungchan Kwon, Gyuil Lim, Youngjoon Han</i>	
GeoHSAP: Geometric Hippocampus Shape Analysis Framework for Longitudinal Alzheimer's Disease Classification	2156
<i>Mubarak Olaoluwa, Heni Loukil, Arafet Sbei, Hassen Drira</i>	
BAFIS: Dataset + Framework to assess occupational Bias and Human Preference in modern Text-to-image Models	2168
<i>Thomas Klassert, Adrian Ulges, Biying Fu</i>	
Imitating the Functionality of Image-to-Image Models Using a Single Example	2178
<i>Nurit Spingarn, Tomer Michaeli</i>	
Beyond the Highlights: Video Retrieval with Salient and Surrounding Contexts	2188
<i>Jaehun Bang, Moon Ye-Bin, Tae-Hyun Oh, Kyungdon Joo</i>	
SCORE: Soft Label Compression-Centric Dataset Condensation via Coding Rate Optimization	2198
<i>Bowen Yuan, Yuxia Fu, Zijian Wang, Yadan Luo, Zi Huang</i>	
Sketch2Stitch: GANs for Abstract Sketch-Based Dress Synthesis	2209
<i>Faizan Farooq Khan, Eslam Abdelrahman Bakr, Davide Morelli, Marcella Cornia, Rita Cucchiara Mohamed Elhoseiny</i>	
AnyBald: Toward Realistic Diffusion-Based Hair Removal In-The-Wild	2220
<i>Yongjun Choi, Seungoh Han, Soomin Kim, Sumin Son, Mohsen Rohani, Edgar Maucourant, Dongbo Min Kyungdon Joo</i>	
Understanding Generative AI Capabilities in Everyday Image Editing Tasks	2231
<i>Brandon Collins, Mohammad Reza Taesiri, Logan Bolton, Viet Dac Lai, Franck DERNONCOURT, Trung Bui Anh Totti Nguyen</i>	
Reconstructing Realistic and Relightable Eyes	2242
<i>Wesley Khademi, Jogendra Kundu, Yatong An, Alexander Fix, David Colmenares</i>	
1LoRa: Summation Compression for Very Low-Rank Adaptation	2253
<i>Alessio Quercia, Zhuo Cao, Arya Bangun, Richard D. Paul, Abigail Morrison, Ira Assent, Hanno Schar</i>	
Illuminating Darkness: Learning to Enhance Low-light Images In-the-Wild	2263
<i>S. M. A. Sharif, Abdur Rehman, Zain Ul Abidin, Fayaz Ali Dharejo, Radu Timofte, Rizwan Ali Naqvi</i>	
DOODLE: Diffusion-based Out-of-Distribution Learning for Open-set LiDAR Semantic Segmentation	2273
<i>Changgyoon Oh, Hyeonseong Kim, Daehyun We, Jongoh Jeong, Yujeong Chae, Kuk-Jin Yoon</i>	
RobustFormer: Noise-Robust Pre-training for Images and Videos	2284
<i>Ashish Bastola, Nishant Luitel, Hao Wang, Danda Pani Paudel, Roshni Poudel, Abolfazl Razi</i>	
CVP: Central-Peripheral Vision-Inspired Multimodal Model for Spatial Reasoning	2295
<i>Zeyuan Chen, Xiang Zhang, Haiyang Xu, Jianwen Xie, Zhuowen Tu</i>	
Trajectory Tactics: When Transformers Learn Exploration to Generate Online Signature	2306
<i>Anurag Pandey, Aditya Nigam, Arnav Bhavsar, Ashutosh Sharma, Basu Verma, Divya Acharya, Mohd Amir</i>	

BrandFusion: Aligning Image Generation with Brand Styles	2316
<i>Parul Gupta, Varun Khurana, Yaman Kumar Singla, Balaji Krishnamurthy, Abhinav Dhall</i>	
Intra-Class Probabilistic Embeddings for Uncertainty Estimation in Vision-Language Models	2327
<i>Zhenxiang Lin, Maryam Haghighat, Will Browne, Dimity Miller</i>	
FARF-Net: Frequency-guided Adaptive Receptive Field Network for Edge-enhanced Polyp Segmentation	2338
<i>Xue Li, Aiwen Jiang, Hongqian Yu, Yang Xiao</i>	
Discrete Facial Encoding: A Framework for Data-driven Facial Display Discovery	2348
<i>Minh Tran, Maksim Siniukov, Zhangyu Jin, Mohammad Soleymani</i>	
Knowledge to Sight: Reasoning over Visual Attributes via Knowledge Decomposition for Abnormality Grounding	2359
<i>Jun Li, Che Liu, Wenjia Bai, Mingxuan Liu, Rossella Arcucci, Cosmin I. Bercea, Julia A. Schnabel</i>	
DenseBEV: Transforming BEV Grid Cells into 3D Objects	2370
<i>Marius Dähling, Sebastian Krebs, J. Marius Zöllner</i>	
Human knowledge integrated multi-modal learning for single source domain generalization	2380
<i>Ayan Banerjee, Kuntal Thakur, Sandeep Gupta</i>	
GraspDiffusion: Synthesizing Realistic Whole-body Hand-Object Interaction	2392
<i>Patrick Kwon, Chen Chen, Hanbyul Joo</i>	
ScoliGaitX: A Deep Multi-Modal Fusion Network for Scoliosis Assessment via Gait Video Analysis	2404
<i>Kaushik Vishwakarma, Aditya Nigam</i>	
Non-Contact Blood Pressure Estimation from Face Videos via Physiology-Aware Contrastive Learning	2414
<i>JaeHyuk Son, Young-Seok Choi</i>	
Ordinal-Aware Multimodal Engagement Recognition for Collaborative Learning	2424
<i>Nha Tran, Dat Ly, Phi Ta, Hung Nguyen, Hien D. Nguyen</i>	
DynaGSLAM: Real-Time Gaussian-Splatting SLAM for Online Rendering, Tracking, Motion Predictions of Moving Objects in Dynamic Scenes	2434
<i>Runfa Blark Li, Mahdi Shaghaghi, Keito Suzuki, Xinshuang Liu, Varun Moparthy, Bang Du, Walker Curtis Martin Renschler, Ki Myung Brian Lee, Nikolay Atanasov, Truong Nguyen</i>	
Test-Time Adaptation through Semantically-guided Feature Decomposition for Few-shot Chest X-ray Diagnosis	2445
<i>Jayant Mahawar, Angshuman Paul</i>	
RobustGait: Robustness Analysis for Appearance Based Gait Recognition	2455
<i>Reeshoon Sayera, Akash Kumar, Sirshapan Mitra, Prudvi Kamtam, Yogesh S Rawat</i>	
FlowMorph: Revealing an Optimizable Flow Latent Space for Controlled Image Morphing	2467
<i>Yan Zheng, Yi Yang, Lanqing Guo, Zhangyang Wang</i>	
PVeRA: Probabilistic Vector-Based Random Matrix Adaptation	2477
<i>Leo Fillioux, Enzo Ferrante, Paul-Henry Cournède, Maria Vakalopoulou, Stergios Christodoulidis</i>	
Harnessing Object Grounding for Time-Sensitive Video Understanding	2487
<i>Tz-Ying Wu, Sharath Nittur Sridhar, Subarna Tripathi</i>	

TaxonRL: Reinforcement Learning with Intermediate Rewards for Interpretable Fine-Grained Visual Reasoning	2497
<i>Maximilian von Klinski, Maximilian Schall</i>	
Revisiting Retentive Networks for Fast Range-View 3D LiDAR Semantic Segmentation	2511
<i>Simone Mosco, Daniel Fusaro, Wanmeng Li, Alberto Pretto</i>	
Zero-shot Hierarchical Plant Segmentation via Foundation Segmentation Models and Text-to-image Attention	2522
<i>Junhao Xing, Ryohei Miyakawa, Yang Yang, Xinpeng Liu, Risa Shinoda, Hiroaki Santo, Yosuke Toda, Fumio Okura</i>	
Moiré Zero: An Efficient and High-Performance Neural Architecture for Moiré Removal	2532
<i>Seungryong Lee, Woojeong Baek, Younghyun Kim, Eunwoo Kim, Haru Moon, Donggon Yoo, Eunbyung Park</i>	
LENVIZ: A High-Resolution Low-Exposure Night Vision Benchmark Dataset	2543
<i>Manjushree Aithal, Rosaura G. VidalMata, Manikandtan Kartha, Gong Chen, Eashan Adhikarla, Lucas N. Kristen, Zhicheng Fu, Nikhil A. Madusudhana, Joe Nasti</i>	
A-V Representation Learning via Audio Shift Prediction for Multimodal Deepfake Detection and Temporal Localization	2553
<i>Ashutosh Anshul, Eng Siong Chng, Deepu Rajan</i>	
CraftSVG: Multi-Object Text-to-SVG Synthesis via Layout Guided Diffusion	2564
<i>Ayan Banerjee, Nityanand Mathur, Josep Lladós, Umapada Pal, Anjan Dutta</i>	
Shift-Equivariant Complex-Valued Convolutional Neural Networks	2575
<i>Quentin Gabot, Teck-Yian Lim, Jérémy Fix, Joana Frontera-Pons, Chengfang Ren, Jean-Philippe Ovarlez</i>	
ObjectMeshDeform : Towards recovering precise 3D geometry of real objects via image-guided mesh deformation of 3D generative priors	2585
<i>Siddharth Katageri, Sanjana Sinha, Sourav Ghosh, Soumyadip Maity, Brojeshwar Bhowmick</i>	
Memory-Augmented Representation for Efficient Event-based Visuomotor Policy Learning with Adaptive Perception and Control	2596
<i>Uday Kamal, Saibal Mukhopadhyay</i>	
PADM: A Physics-aware Diffusion Model for Attenuation Correction	2606
<i>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</i>	
NerVast: Compression-Efficient Scaling of Implicit Neural Video Representations via Scene-based Parameter-sharing	2616
<i>Yunheon Lee, Juncheol Ye, Jaehong Kim, Dongsu Han</i>	
CineVerse: Consistent Keyframe Synthesis for Cinematic Scene Composition	2626
<i>Quynh Phung, Long Mai, Fabian David Caba Heilbron, Feng Liu, Jia-Bin Huang, Cusuh Ham</i>	
MMCM: Multimodality-aware Metric using Clustering-based Modes for Probabilistic Human Motion Prediction	2637
<i>Kyotaro Tokoro, Hiromu Taketsugu, Norimichi Ukita</i>	
Face-LLaVA: Facial Expression and Attribute Understanding through Instruction Tuning	2648
<i>Ashutosh Chaubey, Xulang Guan, Mohammad Soleymani</i>	

ConsensusXAI: A framework to examine class-wise agreement in medical imaging	2661
<i>Abbas Haider, David Wright, Ruth Hogg, Hui Wang, Tunde Peto, Richard Gault</i>	
Real-Time Tracking of Flexible Markers in Low-Contrast Fluoroscopy Using a Deep Neural Network Trained Solely on Synthetic Data	2670
<i>Tomoki Uchiyama, Yukinobu Sakata, Ryusuke Hirai, Hitoshi Ishikawa, Shinichiro Mori</i>	
OpenLVLM-MIA: A Controlled Benchmark Revealing the Limits of Membership Inference Attacks on Large Vision-Language Models	2680
<i>Ryoto Miyamoto, Xin Fan, Fuyuko Kido, Tsuneo Matsumoto, Hayato Yamana</i>	
DMAT: An End-to-End Framework for Joint Atmospheric Turbulence Mitigation and Object Detection	2690
<i>Paul Hill, Zhiming Liu, Alin Achim, David Bull, Nantheera Anantrasirichai</i>	
Divide and Refine: Enhancing Multimodal Representation and Explainability for Emotion Recognition in Conversation	2700
<i>Anh-Tuan Mai, Cam-Van Thi Nguyen, Duc-Trong Le</i>	
iMotion-LLM: Instruction-Conditioned Trajectory Generation	2710
<i>Abdulwahab Felemban, Nussair Hroub, Jian Ding, Eslam Abdelrahman, Xiaoqian Shen Abduallah Mohamed, Mohamed Elhoseiny</i>	
SasMamba: A Lightweight Structure-Aware Stride State Space Model for 3D Human Pose Estimation	2721
<i>Hu Cui, Wenqiang Hua, Renjing Huang, Shurui Jia, Tessai Hayama</i>	
SUGAR: A Sweeter Spot for Generative Unlearning of Many Identities	2731
<i>Dung Thuy Nguyen, Quang Nguyen, Preston K. Robinette, Eli Jiang, Taylor T. Johnson, Kevin Leach</i>	
Mitigating the Modality Gap: Few-Shot Out-of-Distribution Detection with Multi-modal Prototypes and Image Bias Estimation	2741
<i>Yimu Wang, Evelien Riddell, Adrian Chow, Sean Sedwards, Krzysztof Czarnecki</i>	
Matching Semantically Similar Non-Identical Objects	2752
<i>Yusuke Marumo, Kazuhiko Kawamoto, Satomi Tanaka, Shigenobu Hirano, Hiroshi Kera</i>	
UI-Styler: Ultrasound Image Style Transfer with Class-Aware Prompts for Cross-Device Diagnosis Using a Frozen Black-Box Inference Network	2765
<i>Nhat-Tuong Do-Tran, Ngoc-Hoang-Lam Le, Ching-Chun Huang</i>	
Tables Decoded: DELTA for Structure, TarQA for Understanding	2775
<i>Jahanvi Rajput, Dhruv Kudale, Saikiran Kasturi, Utkarsh Verma, Ganesh Ramakrishnan</i>	
What Happens When: Learning Temporal Orders of Events in Videos	2786
<i>Daechul Ahn, Yura Choi, Hyeonbeom Choi, Seongwon Cho, San Kim, Jonghyun Choi</i>	
UNO: Unifying One-stage Video Scene Graph Generation via Object-Centric Visual Representation Learning	2797
<i>Huy Le, Nhat Chung, Tung Kieu, Jingkang Yang, Ngan Le</i>	
Personalized Image Privacy Advisors via Federated Daisy-Chaining	2808
<i>Sourasekhar Banerjee, Vengateswaran Subramaniam, Debaditya Roy, Vigneshwaran Subbaraju Monowar Bhuyan</i>	
DiRe: Diversity-promoting Regularization for Dataset Condensation	2818
<i>Saumyanjan Mohanty, Aravind Reddy, Konda Reddy Mopuri</i>	

Vision-informed Semantic Text Alignment for Open-set Recognition in Remote Sensing	2828
<i>Siddhant Gole, Akash Pal, Ankit Jha, Subhasis Chaudhuri, Biplab Banerjee</i>	
Anatomy-VLM: A Fine-grained Vision-Language Model for Medical Interpretation	2838
<i>Difei Gu, Yunhe Gao, Mu Zhou, Dimitris Metaxas</i>	
Temporal Object Captioning for Street Scene Videos from LiDAR Tracks	2848
<i>Vignesh Gopinathan, Urs Zimmermann, Michael Arnold, Matthias Rottmann</i>	
STARS: Self-supervised Tuning for 3D Action Recognition in Skeleton Sequences	2858
<i>Soroush Mehraban, Mohammad Javad Rajabi, Andrea Iaboni, Babak Taati</i>	
RAVU: Retrieval Augmented Video Understanding with Compositional Reasoning over Graph	2869
<i>Sameer Malik, Ayush Singh, Moyuru Yamada, Dishank Aggarwal</i>	
Conditional Text-to-Image Generation with Reference Guidance	2879
<i>Taewook Kim, Ze Wang, Zhengyuan Yang, Jiang Wang, Lijuan Wang, Zicheng Liu, Qiang Qiu</i>	
Improved Wildfire Spread Prediction with Time-Series Data and the WSTS+ Benchmark	2890
<i>Saad Lahrichi, Jake Bova, Jesse Johnson, Jordan Malof</i>	
From Few-Shot to Zero-Shot Pallet Load Recognition: A Deployed Embedding-Based Vision System for Industrial Logistics	2901
<i>Juan Jesús Losada Del Olmo, Emilio Pardo Ballesteros, Pedro E. López-de-Teruel, Alberto Ruiz</i>	
Graph-Based Spectral Attention with Multi-Spectral Images for Illuminant Estimation	2912
<i>Dong-Hoon Kang, Seung-Yeop Baek, Jong-Ok Kim</i>	
Fast Vision Mamba: Pooling Spatial Dimensions for Accelerated Processing	2923
<i>Saarthak Kapse, Robin Betz, Srinivasan Sivanandan</i>	
Extreme Amodal Face Detection	2934
<i>Changlin Song, Yunzhong Hou, Michael Randall Barnes, Rahul Shome, Dylan Campbell</i>	
ENCORE: A Neural Collapse Perspective on Out-of-Distribution Detection in Deep Neural Networks	2944
<i>A.Q.M. Sazzad Sayyed, Nathaniel D. Bastian, Francesco Restuccia</i>	
Performance of Conformal Prediction in Capturing Aleatoric Uncertainty	2954
<i>Misgina Tsighe Hagos, Claes Lundström</i>	
Scalpel: Fine-Grained Alignment of Attention Activation Manifolds via Mixture Gaussian Bridges to Mitigate Multimodal Hallucination	2964
<i>Ziqiang Shi, Rujie Liu, Shanshan Yu, Satoshi Munakata, Koichi Shirahata</i>	
Unified Alignment Protocol: Making Sense of the Unlabeled Data in New Domains	2974
<i>Sabbir Ahmed, Mamshad Nayeem Rizve, Abdullah Al Arafat, Jacqueline Tiffany Liu, Rahim Hossain Mohaiminul Al Nahian, Adnan Siraj Rakin</i>	
Feedback Alignment Meets Low-Rank Manifolds: A Structured Recipe for Local Learning	2984
<i>Arani Roy, Marco P. Apolinario, Shruti Das Biswas, Kaushik Roy</i>	
Learning from Unknown for Open-Set Test-Time Adaptation	2993
<i>Taki Hasan Rafi, Amit Agarwal, Hitesh L. Patel, Dong-Kyu Chae</i>	
Streaming Real-Time Trajectory Prediction Using Endpoint-Aware Modeling	3005
<i>Alexander Prutsch, David Schinagl, Horst Possegger</i>	

StreetView-Waste: A Multi-Task Dataset for Urban Waste Management	3015
<i>Diogo J. Paulo, João Martins, Hugo Proença, João C. Neves</i>	
AnyAnomaly: Zero-Shot Customizable Video Anomaly Detection with LVLM	3026
<i>Sunghyun Ahn, Youngwan Jo, Kijung Lee, Sein Kwon, Inpyo Hong, Sanghyun Park</i>	
SkelSplat: Robust Multi-view 3D Human Pose Estimation with Differentiable Gaussian Rendering	3036
<i>Laura Bragagnolo, Leonardo Barcellona, Stefano Ghidoni</i>	
Evaluating the Capability of Video Question Generation for Expert Knowledge Elicitation	3047
<i>Huaying Zhang, Atsushi Hashimoto, Tosho Hirasawa</i>	
Dragonite: Single-Step Drag-based Image Editing with Geometric-Semantic Guidance	3057
<i>Meng-Ting Zhong, Tai-Ming Huang, Shung-Fu Chen, Wen-Huang Cheng, Kai-Lung Hua</i>	
Augmenting with NeRFs: Fast Relocalization on Densified Datasets	3067
<i>Michael Tomadakis, Rebecca Borissova, Yuxuan Zhang, Sanjeev Koppal</i>	
GASP: Unifying Geometric and Semantic Self-Supervised Pre-training for Autonomous Driving	3077
<i>William Ljungbergh, Adam Lilja, Adam Tonderski, Arvid Laveno Ling, Carl Lindström, Willem Verbeke Junsheng Fu, Christoffer Petersson, Lars Hammarstrand, Michael Felsberg</i>	
Towards Reliable Test-Time Adaptation: Style Invariance as a Correctness Likelihood	3088
<i>Gilhyun Nam, Taewon Kim, Joonhyun Jeong, Eunho Yang</i>	
TalkingPose: Efficient Face and Gesture Animation with Feedback-guided Diffusion Model	3098
<i>Alireza Javanmardi, Pragati Jaiswal, Tewodros Amberbir Habtegebrial, Christen Millerdurai Shaoxiang Wang, Alain Pagani, Didier Stricker</i>	
Large Sign Language Models: Toward 3D American Sign Language Translation	3109
<i>Sen Zhang, Xiaoxiao He, Di Liu, Zhaoyang Xia, Mingyu Zhao, Chaowei Tan, Vivian Li, Bo Liu Dimitris N. Metaxas, Mubbasir Kapadia</i>	
Crafting Descriptive Information for a Zero-shot Method to Improve Knowledge-Based Visual Question Answering Performance	3120
<i>Mohammad Mahdi Moradi, Sudhir Mudur</i>	
Learning Action Hierarchies via Hybrid Geometric Diffusion	3129
<i>Arjun Ramesh Kaushik, Nalini K. Ratha, Venu Govindaraju</i>	
Learning Group Actions In Disentangled Latent Image Representations	3140
<i>Farhana Hossain Swarnali, Miaomiao Zhang, Tonmoy Hossain</i>	
GAITGen: Disentangled Motion-Pathology Impaired Gait Generative Model – Bringing Motion Generation to the Clinical Domain	3150
<i>Vida Adeli, Soroush Mehraban, Majid Mirmehdi, Alan Whone, Benjamin Filtjens Amirhossein Dadashzadeh, Alfonso Fasano, Andrea Iaboni, Babak Taati</i>	
Gradient-Free Classifier Guidance for Diffusion Model Sampling	3162
<i>Rahul Shenoy, Zhihong Pan, Kaushik Balakrishnan, Qisen Cheng, Yongmoon Jeon, Heejune Yang Jaewon Kim</i>	
Mitigating Object and Action Hallucinations in Multimodal LLMs via Self-Augmented Contrastive Alignment	3172
<i>Kai-Po Chang, Wei-Yuan Cheng, Chi-Pin Huang, Fu-En Yang, Yu-Chiang Frank Wang</i>	

Guided Model Merging for Hybrid Data Learning: Leveraging Centralized Data to Refine Decentralized Models	3182
<i>Junyi Zhu, Ruicong Yao, Taha Ceritli, Savas Ozkan, Matthew B. Blaschko, Eunchung Noh, Jeongwon Min Cho Jung Min, Mete Ozay</i>	
Fused Similarity Measure Based Alignment with Dual-Scale Adaptive Selection for Weakly Supervised Video Anomaly Detection	3193
<i>Yue-Gao Lu, Hong-Jie Xing, Chun-Guo Li</i>	
PointNet4D: A Lightweight 4D Point Cloud Video Backbone for Online and Offline Perception in Robotic Applications	3203
<i>Yunze Liu, Zifan Wang, Peiran Wu, Jiayang Ao</i>	
Cross-Modal Event Encoder: Bridging Image–Text Knowledge to Event Streams	3213
<i>Sunghoon Jeong, Hanning Chen, Sanggeon Yun, Suhyeon Cho, Wenjun Huang, Xiangjian Liu Mohsen Imani</i>	
SAIL: Self-supervised Learning of Lighting-Invariant Representations from Real Images with Latent Diffusion	3223
<i>Hala Djeghim, Céline Loscos, Désiré Sidibé</i>	
PSDiffusion: Harmonized Multi-Layer Image Generation via Layout and Appearance Alignment	3233
<i>Dingbang Huang, Wenbo Li, Yifei Zhao, Xinyu Pan, Yanhong Zeng, Bo Dai</i>	
Nerve: Neighbourhood & Entropy-guided Random-walk for training free open-Vocabulary sEgmentation	3243
<i>Kunal Mahatha, Jose Dolz, Christian Desrosiers</i>	
CLIP’s Visual Embedding Projector is a Few-shot Cornucopia	3254
<i>Mohammad Fahes, Tuan-Hung Vu, Andrei Bursuc, Patrick Pérez, Raoul de Charette</i>	
RegionAligner: Bridging Ego-Exo Views for Object Correspondence via Unified Text-Visual Learning	3265
<i>Yuhao Su, Ehsan Elhamifar</i>	
Towards Streaming LiDAR Object Detection with Point Clouds as Egocentric Sequences	3275
<i>Mellon M. Zhang, Glen Chou, Saibal Mukhopadhyay</i>	
Show Me: Unifying Instructional Image and Video Generation with Diffusion Models	3285
<i>Yujiang Pu, Zhanbo Huang, Vishnu Boddeti, Yu Kong</i>	
VOCAL: Visual Odometry via ContrAstive Learning	3297
<i>Chi-Yao Huang, Zeel Bhatt, Yezhou Yang</i>	
Pose-Diverse Multi-View Virtual Try-on from a Single Frontal Image via Diffusion Transformer	3310
<i>Seonghee Han, Minchang Chung, Gyeongsu Cho, Kyungdon Joo, Taehwan Kim</i>	
SimForce: Force and Surface Electromyography from Full Body Video with Graph Neural Nets	3320
<i>Esha Dasgupta, Boeun Kim, Sang Hoon Yeo, Hyung Jin Chang</i>	
SCATR: Mitigating New Instance Suppression in LiDAR-based Tracking-by-Attention via Second Chance Assignment and Track Query Dropout	3330
<i>Brian Cheong, Letian Wang, Sandro Papais, Steven L. Waslander</i>	

AirLock ⁺ : Scaling UAV-to-Satellite Image Registration for Target Geolocalization and Geospatial Augmented Reality	3340
<i>Zhiyun Deng, Austin Case, Luis Sentis</i>	
Blur2Sharp: Human Novel Pose and View Synthesis with Generative Prior Refinement	3350
<i>Chia-Hern Lai, I-Hsuan Lo, Yen-Ku Yeh, Thanh-Nguyen Truong, Ching-Chun Huang</i>	
Power of Boundary and Reflection: Semantic Transparent Object Segmentation using Pyramid Vision Transformer with Transparent Cues	3360
<i>Tuan-Anh Vu, Hai Nguyen-Truong, Ziqiang Zheng, Binh-Son Hua, Qing Guo, Ivor W. Tsang, Sai-Kit Yeung</i>	
HEART-PFL: Stable Personalized Federated Learning under Heterogeneity with Hierarchical Directional Alignment and Adversarial Knowledge Transfer	3370
<i>Minjun Kim, Minje Kim</i>	
Detecting Social Engagement of Elderly From Lifelog Image-streams to Identify Effective Cues for Autobiographic Recall	3380
<i>Vengateswaran Subramaniam, Vigneshwaran Subbaraju, Debaditya Roy, Pramath Krishna Thivya Kandappu, Qianli Xu</i>	
HistoMILKD: A Multiple Instance Learning based Multi-Teacher Knowledge Distillation Framework for Whole Slide Image Classification	3390
<i>Mayur Mallya, Ali Khajegili Mirabadi, Hossein Farahani, Ali Bashashati</i>	
Stroke Modeling Enables Vectorized Character Generation with Large Vectorized Glyph Model	3401
<i>Xinyue Zhang, Haolong Li, Jiawei Ma, Chen Ye</i>	
CasTex: Cascaded Text-to-Texture Synthesis via Explicit Texture Maps and Physically-Based Shading	3411
<i>Mishan Aliev, Dmitry Baranchuk, Kirill Struminsky</i>	
DATTA: Domain-Adversarial Test-Time Adaptation for Cross-Domain WiFi-Based Human Activity Recognition	3421
<i>Julian Strohmayer, Rafael Sterzinger, Matthias Wödlinger, Martin Kampel</i>	
MixER: From Cross-Modal to Mixed-Modal Visible-Infrared Re-Identification	3431
<i>Mahdi Alehdaghi, Rajarshi Bhattacharya, Pourya Shamsolmoali, Rafael M. O. Cruz, Eric Granger</i>	
LighthouseGS: Indoor Structure-aware 3D Gaussian Splatting for Panorama-Style Mobile Captures	3441
<i>Seungoh Han, Jaehoon Jang, Hyunsu Kim, Jaeheung Surh, Junhyung Kwak, Hyowon Ha, Kyungdon Joo</i>	
FLARES: Fast and Accurate LiDAR Multi-Range Semantic Segmentation	3451
<i>Bin Yang, Alexandru Paul Condurache</i>	
Language Integration in Fine-Tuning Multimodal Large Language Models for Image-Based Regression	3462
<i>Roy H. Jennings, Genady Paikin, Roy Shaul, Evgeny Soloveichik</i>	
MANTA: Physics-Informed Generalized Underwater Object Tracking	3472
<i>Suhas Srinath, Hemang Jamadagni, Aditya Chandrasekar, Prathosh A P</i>	
KD360-VoxelBEV: LiDAR and 360-degree Camera Cross Modality Knowledge Distillation for Bird's-Eye-View Segmentation	3483
<i>Wenke E, Yixin Sun, Jiaxu Liu, Hubert P. H. Shum, Amir Atapour-Abarghouei, Toby P. Breckon</i>	
MAFM ³ : Modular Adaptation of Foundation Models for Multi-Modal Medical AI	3494
<i>Mohammad Areeb Qazi, Munachiso S Nwadike, Ibrahim Almakky, Mohammad Yaqub, Numan Saeed</i>	

SPOC: Spatially-Progressing Object State Change Segmentation in Video	3504
<i>Priyanka Mandikal, Tushar Nagarajan, Alex Stoken, Zihui Xue, Kristen Grauman</i>	
Self-Supervised Visual Prompting for Cross-Domain Road Damage Detection	3514
<i>Xi Xiao, Zhuxuanzi Wang, Mingqiao Mo, Chen Liu, Chenrui Ma, Yanshu Li, Smita Krishnaswamy</i>	
<i>Xiao Wang, Tianyang Wang</i>	
Data-Driven Loss Functions for Inference-Time Optimization in Text-to-Image	3525
<i>Sapir Esther Yiflach, Yuval Atzmon, Gal Chechik</i>	
Multi-Modal Soccer Scene Analysis with Masked Pre-Training	3536
<i>Marc Peral, Guillem Capellera, Luis Ferraz, Antonio Rubio, Antonio Agudo</i>	
Visual Detector Compression via Location-Aware Discriminant Analysis	3546
<i>Qizhen Lan, Jung Im Choi, Qing Tian</i>	
Latent Uncertainty-Aware Multi-View SDF Scan Completion	3556
<i>Faezeh Zakeri, Lukas Ruppert, Raphael Braun, Hendrik P.A. Lensch</i>	
Comp4D: Compositional 4D Scene Generation	3567
<i>Hanwen Liang, Dejia Xu, Neel P. Bhatt, Hezhen Hu, Hanxue Liang, Konstantinos N. Plataniotis</i>	
SENCA-st: Integrating Spatial Transcriptomics and Histopathology with Cross Attention Shared Encoder for Region Identification in Cancer Pathology	3578
<i>Shanaka Liyanaarachchi, Chathurya Wijethunga, Shihab Aaqil Ahamed, Akthas Absar, Ranga Rodrigo</i>	
Towards High-Fidelity, Identity-Preserving Real-Time Makeup Transfer: Decoupling Style Generation	3588
<i>Lydia Chau, Zhi Yu, Ruowei Jiang</i>	
Feature Inversion as a Lens on Vision Encoders	3598
<i>Eduard Allakhverdov, Dmitrii Tarasov, Elizaveta Goncharova, Andrey Kuznetsov</i>	
MuSACo: Multimodal Subject-Specific Selection and Adaptation for Expression Recognition with Co-Training	3606
<i>Muhammad Osama Zeeshan, Natacha Gillet, Alessandro Lameiras Koerich, Marco Pedersoli</i>	
<i>Francois Bremond, Eric Granger</i>	
SCALEX: Scalable Concept and Latent Exploration for Diffusion Models	3617
<i>E. Zhixuan Zeng, Yuhao Chen, Alexander Wong</i>	
DCSHARP: 3D Gaussian Splatting with Direction Cosine Spherical Harmonics and Shape-Aware Pruning	3628
<i>Ahmed Hasssan, Jian Meng, Yuanbo Xiangli, Jae-sun Seo</i>	
DOTGraph: CLIP-Driven Feature Disentanglement and Optimal Transport based Graph Learning for Few-Shot Segmentation	3638
<i>Shreya Biswas, Zhaozheng Yin</i>	
Delta-LLaVA: Base-then-Specialize Alignment for Token-Efficient Vision-Language Models	3648
<i>Mohamad Zamini, Diksha Shukla</i>	
Sketch-guided Cage-based 3D Gaussian Splatting Deformation	3658
<i>Tianhao Xie, Noam Aigerman, Eugene Belilovsky, Tiberiu Popa</i>	

Autoregressive Styled Text Image Generation, but Make it Reliable	3668
<i>Carmine Zaccagnino, Fabio Quattrini, Vittorio Pippi, Silvia Cascianelli, Alessio Tonioni, Rita Cucchiara</i>	
Universal Neural Architecture Space: Covering ConvNets, Transformers and Everything in Between	3679
<i>Ondřej Týbl, Lukáš Neumann</i>	
Sea-CLIP: Mining Semantic-Aware Representations for Few-Shot Anomaly Detection with CLIP	3689
<i>Xiao Guo, Zhimin Chen, Carlos D. Castillo, Hongcheng Wang, Xiaoming Liu</i>	
CLIP-IT: CLIP-based Pairing of Histology Images with Privileged Textual Information	3700
<i>Banafsheh Karimian, Giulia Avanzato, Soufiane Belharbi, Alexis Guichemerre, Luke McCaffrey Mohammadhadi Shateri, Eric Granger</i>	
LightGazeNet: A Lightweight GNN-based Architecture for Gaze Estimation	3710
<i>Heena Patel, Anirban Chowdhury, Pooja Jigar Choksy, Samiksha Pradeep Pachade, Ajinkya Puar</i>	
Codebook Knowledge with Mamba-Transformer For Low-Light Image Enhancement	3720
<i>Runhua Deng, Aiwon Jiang, Long Peng, Qiuhai Yan</i>	
TICLS: Tightly Coupled Language Text Spotter	3730
<i>Leeje Jang, Yijun Lin, Yao-Yi Chiang, Jerod Weinman</i>	
Perception-Inspired Color Space Design for Photo White Balance Editing	3741
<i>Yang Cheng, Ziteng Cui, Lin Gu, Shenghan Su, Zenghui Zhang</i>	
CLUE: Bringing Machine Unlearning to Mobile Devices	3750
<i>A.Q.M. Sazzad Sayyed, Nathaniel D. Bastian, Michael De Lucia, Ananthram Swami, Francesco Restuccia</i>	
FairScene: Learning Class-Disentangled 2D/3D Representations for Semantic Scene Completion	3760
<i>Dian Jia, Pei Yu, Wei Tang</i>	
Frequency Is What You Need: Considering Word Frequency When Text Masking Benefits Vision-Language Model Pre-training	3771
<i>Mingliang Liang, Martha Larson</i>	
Event-based Graph Representation with Spatial and Motion Vectors for Asynchronous Object Detection	3781
<i>Aayush Atul Verma, Arpitsinh Vaghela, Bharatesh Chakravarthi, Kaustav Chanda, Yezhou Yang</i>	
ART-ASyn: Anatomy-aware Realistic Texture-based Anomaly Synthesis Framework for Chest X-Rays	3792
<i>Qinyi Cao, Jianan Fan, Weidong Cai</i>	
High-Rate Mixout: Revisiting Mixout for Robust Domain Generalization	3803
<i>Masih Aminbeidokhti, Heitor Rapela Medeiros, Srikanth Muralidharan, Eric Granger, Marco Pedersoli</i>	
MuseDance: A Diffusion-based Music-Driven Image Animation System	3813
<i>Zhikang Dong, Weituo Hao, Ju-Chiang Wang, Peng Zhang, Pawel Polak</i>	
A Little More Like This: Text-to-Image Retrieval with Vision-Language Models Using Relevance Feedback	3825
<i>Bulat Khaertdinov, Mirela Popa, Nava Tintarev</i>	
Bi-ICE: An Inner Interpretable Framework for Image Classification via Bi-directional Interactions between Concept and Input Embeddings	3835
<i>Jinyung Hong, Yearim Kim, Keun Hee Park, Sangyu Han, Nojun Kwak, Theodore P. Pavlic</i>	

DTMIR-Pro: Domain Translation with Prompt-based Latent-Space Generalization for Multi-Weather Image Restoration	3846
<i>Ashutosh Kulkarni, Prashant W. Patil, Santosh Kumar Vipparthi, Subrahmanyam Murala Balasubramanian Raman</i>	
ObjectCore – Efficient Few-shot Logical Anomaly Detection using Object Representations	3857
<i>Matic Fučka, Vitjan Zavrtanik, Danijel Skočaj</i>	
A Novel Metric for Detecting Memorization in Generative Models for Brain MRI Synthesis	3868
<i>Antonio Scardace, Lemuel Puglisi, Francesco Guarnera, Sebastiano Battiato, Daniele Ravi</i>	
Unsupervised Modular Adaptive Region Growing and RegionMix Classification for Wind Turbine Segmentation	3878
<i>Raül Pérez-Gonzalo, Riccardo Magro, Andreas Espersen, Antonio Agudo</i>	
Gen-AFFECT: Generation of Avatar Fine-grained Facial Expressions with Consistent identity	3889
<i>Hao Yu, Rupayan Mallick, Margrit Betke, Sarah Adel Bargal</i>	
FlowEO: Generative Unsupervised Domain Adaptation for Earth Observation	3900
<i>Georges Le Bellier, Nicolas Audebert</i>	
Efficient Vision Transformers via Token Merging with Head-Wise Attention Correction	3908
<i>Yuki Ichikawa, Masato Motomura, Thiem Van Chu, Daichi Fujiki</i>	
2S-CEDiff: A Two-Stage Diffusion Framework for Generating High-Fidelity Contrast-Enhanced CT Images from Non-Contrast Scans	3918
<i>Yibang Wu, Tzung-Dau Wang, Shang-Hong Lai</i>	
JOCA: Task-Driven Joint Optimisation of Camera Hardware and Adaptive Camera Control Algorithms	3928
<i>Chengyang Yan, Mitch Bryson, Donald G. Dansereau</i>	
CAPE: A CLIP-Aware Pointing Ensemble of Complementary Heatmap Cues for Embodied Reference Understanding	3939
<i>Fevziye Irem Eyiokur, Dogucan Yaman, Hazım Kemal Ekenel, Alexander Waibel</i>	
Online Episodic Memory Visual Query Localization with Egocentric Streaming Object Memory	3951
<i>Zaira Manigrasso, Matteo Dunnhofer, Antonino Furnari, Moritz Nottebaum, Antonio Finocchiario Davide Marana, Rosario Forte, Giovanni Maria Farinella, Christian Micheloni</i>	
Mobile-Oriented Video Diffusion: Enabling Text-to-Video Generation on Mobile Devices Without Retraining, Compression, or Pruning	3961
<i>Bosung Kim, Kyuhwan Lee, Isu Jeong, Jungmin Cheon, Yeojin Lee, Seulki Lee</i>	
WorkZone3D: A Multimodal Dataset for 3D Work Zone Perception in Autonomous Driving	3972
<i>Shounak Sural, Nishad Sahu, Ragunathan Raj Rajkumar</i>	
Unconditional Priors Matter! Improving Conditional Generation of Fine-Tuned Diffusion Models	3982
<i>Prin Phunyaphibarn, Phillip Y. Lee, Jaihoon Kim, Minhyuk Sung</i>	
Training-free Detection of Text-to-video Generations via Over-coherence	3993
<i>Jonathan Brokman, Oren Rachmil, Omer Hofman, Roy Betser, Amit Giloni, Roman Vainshtein Hisashi Kojima</i>	
ReBrain: Brain MRI Reconstruction from Sparse CT Slice via Retrieval-Augmented Diffusion	4004
<i>Junming Liu, Yifei Sun, Weihua Chen, Yujin Kang, Yirong Chen, Ding Wang, Guosun Zeng</i>	

Efficient Text-Guided Convolutional Adapter for the Diffusion Model	4015
<i>Aryan Das, Koushik Biswas, Swalpa Kumar Roy, Badri Narayana Patro, Vinay Kumar Verma</i>	
FLoMo-Net: A Novel Task-Adaptive Mixture of Experts Routing Framework with Frequency and Uncertainty Correction for Medical Image Segmentation	4025
<i>Md Rayhan Ahmed, Patricia Lasserre</i>	
From Detection to Anticipation: Online Understanding of Struggles across Various Tasks and Activities	4036
<i>Shijia Feng, Michael Wray, Walterio Mayol-Cuevas</i>	
CaRS: A Causal Intervention Segmentation Framework and Benchmark Dataset for Autonomous Driving under Transitional Weather Conditions	4046
<i>Kondapally Madhavi, K Naveen Kumar, C Krishna Mohan, Sobhan Babu</i>	
Domain Generalizing DINO for Visual Regression via Latent Distractor Subspace Consistency	4057
<i>Nikhil Reddy, Chetan Arora, Mahsa Baktashmotlagh</i>	
WALDO: Where Unseen Model-based 6D Pose Estimation Meets Occlusion	4067
<i>Sajjad Pakdamansavoji, Yintao Ma, Amir Rasouli, Tongtong Cao</i>	
Decoupling Shape and Texture in SAM-2 via Controlled Texture Replacement	4077
<i>Inbal Cohen, Boaz Meivar, Peihan Tu, Shai Avidan, Gal Oren</i>	
HumanBench: Two Heads, No Legs, But Mostly Human, the State of Generative Capabilities in T2I Models	4087
<i>Anubhooti Jain, Mayank Vatsa, Richa Singh</i>	
SOAF: Scene Occlusion-aware Neural Acoustic Field	4097
<i>Huiyu Gao, Jiahao Ma, David Ahmedt-Aristizabal, Chuong Nguyen, Miaomiao Liu</i>	
FedSCAL: Leveraging Server and Client Alignment for Unsupervised Federated Source-Free Domain Adaptation	4108
<i>M. Yashwanth, Sampath Koti, Arunabh Singh, Shyam Marjit, Anirban Chakraborty</i>	
Structure-Aware Feature Rectification with Region Adjacency Graphs for Training-Free Open-Vocabulary Semantic Segmentation	4118
<i>Qiming Huang, Hao Ai, Jianbo Jiao</i>	
One Model, Many Behaviors: Training-Induced Effects on Out-of-Distribution Detection	4128
<i>Gerhard Krumpl, Henning Avenhaus, Horst Possegger</i>	
TalkingHeadBench: A Multi-Modal Benchmark & Analysis of Talking-Head DeepFake Detection	4139
<i>Xinqi Xiong, Prakrut Patel, Qingyuan Fan, Amisha Wadhwa, Sarathy Selvam, Xiao Guo, Luchao Qi Xiaoming Liu, Roni Sengupta</i>	
Rethinking Real Image Editing: Unleashing Diverse Editing Operators via Multi-Objective Optimization	4150
<i>Siyuan Wang, Xi Yang, Zihao Zhou, Huiru Shao, Rui Zhang, Qiufeng Wang, Guangliang Cheng Kaizhu Huang</i>	
Decomposition Sampling for Efficient Region Annotations in Active Learning	4160
<i>Jingna Qiu, Frauke Wilm, Mathias Öttl, Jonas Utz, Maja Schlereth, Moritz Schillinger, Marc Aubreville Katharina Breininger</i>	

DNA: Dual-branch Network with Adaptation for Open-Set Online Handwriting Generation	4170
<i>Tsai-Ling Huang, Nhat-Tuong Do-Tran, Ngoc-Hoang-Lam Le, Hong-Han Shuai, Ching-Chun Huang</i>	
STEG-AIW: Spatio-Temporal Gating and Adaptive-Timestep Inference for Efficient Spiking Neural Networks	4180
<i>Gulfam Ahmed Saju, Anton Spirkin, Felipe Marcelino, Yuchou Chang</i>	
Enhancing Visual Planning with Auxiliary Tasks and Multi-token Prediction	4190
<i>Ce Zhang, Yale Song, Ruta Desai, Michael Louis Iuzzolino, Joseph Tighe, Gedas Bertasius, Satwik Kottur</i>	
Self-Supervised Compression and Artifact Correction for Streaming Underwater Imaging Sonar	4201
<i>Rongsheng Qian, Chi Xu, Xiaoqiang Ma, Hao Fang, Yili Jin, William I. Atlas, Jiangchuan Liu</i>	
Perceptually Guided 3DGS Streaming and Rendering for Mixed Reality	4212
<i>Yunxiang Zhang, Sai Harsha Mupparaju, Kenneth Chen, Jenna Kang, Xinyu Zhang, Maito Omori, Kazuyuki Arimatsu, Qi Sun</i>	
Unlocking Vision-Language Models for Video Anomaly Detection via Fine-Grained Prompting	4223
<i>Shu Zou, Xinyu Tian, Lukas Wesemann, Fabian Waschkowski, Zhaoyuan Yang, Jing Zhang</i>	
Unsupervised Memorability Modeling from Tip-of-the-Tongue Retrieval Queries	4234
<i>Sree Bhattacharyya, Yaman K. Singla, Sudhir Yarram, Somesh Singh, Harini Si, James Z. Wang</i>	
Learning Compact Video Representations for Efficient Long-form Video Understanding in Large Multimodal Models	4242
<i>Yuxiao Chen, Jue Wang, Zhikang Zhang, Jingru Yi, Xu Zhang, Yang Zou, Zhaowei Cai, Jianbo Yuan, Xinyu Li, Hao Yang, Davide Modolo</i>	
Guided Texture Segmentation via Coordinate-Aware Class-Ratio Mapping	4253
<i>Bishal Ranjan Swain, Kyung Joo Cheoi, Jaepil Ko</i>	
Empowering Source-Free Domain Adaptation via MLLM-Guided Reliability-Based Curriculum Learning	4262
<i>Dongjie Chen, Kartik Patwari, Zhengfeng Lai, Xiaoguang Zhu, Sen-Ching Cheung, Chen-Nee Chuah</i>	
ExDDV: A New Dataset for Explainable Deepfake Detection in Video	4273
<i>Vlad Hondru, Eduard Hoge, Darian Onchiş, Radu Tudor Ionescu</i>	
AGENet: Adaptive Edge-aware Geodesic Distance Learning for Few-Shot Medical Image Segmentation	4285
<i>Ziyuan Gao</i>	
CONSTANT: Towards High-Quality One-Shot Handwriting Generation with Patch Contrastive Enhancement and Style-Aware Quantization	4295
<i>Anh-Duy Le, Van-Linh Pham, Thanh-Nam Vo, Xuan Toan Mai, Tuan-Anh Tran</i>	
DCText: Scheduled Attention Masking for Visual Text Generation via Divide-and-Conquer Strategy	4305
<i>Jaewoo Song, Jooyoung Choi, Kanghyun Baek, Sangyub Lee, Daemin Park, Sungroh Yoon</i>	
VFace: A Training-Free Approach for Diffusion-Based Video Face Swapping	4315
<i>Sanoojan Baliah, Yohan Abeysinghe, Rusiru Thushara, Khan Muhammad, Abhinav Dhall, Karthik Nandakumar, Muhammad Haris Khan</i>	
VividAnimator: An End-to-End Audio and Pose-driven Half-Body Human Animation Framework	4325
<i>Donglin Huang, Yongyuan Li, Tianhang Liu, Junming Huang, Xiaoda Yang, Chi Wang, Weiwei Xu</i>	

Fine-grained Defocus Blur Control for Generative Image Models	4335
<i>Ayush Shrivastava, Connelly Barnes, Xuaner Zhang, Lingzhi Zhang, Andrew Owens, Sohrab Amirghodsi Eli Shechtman</i>	
CalibBEV: LiDAR-Camera Calibration via BEV Alignment	4345
<i>Filippo D’Addeo, Lorenzo Cipelli, Adriano Cardace, Emanuele Ghelfi, Andrea Zinelli, Massimo Bertozzi</i>	
X-JEPA: A Novel Joint Learning Cross-Modal Predictive Alignment Framework for Remote Sensing Image Retrieval	4355
<i>Shabnam Choudhury, Yash Salunkhe, Vaibhav Rajan, Subhasis Chaudhuri, Biplab Banerjee</i>	
SSMRadNet : A Sample-wise State-Space Framework for Efficient and Ultra-Light Radar Segmentation and Object Detection	4365
<i>Anuvab Sen, Mir Sayeed Mohammad, Saibal Mukhopadhyay</i>	
Rank-based Geographical Regularization: Revisiting Contrastive Self-Supervised Learning for Multispectral Remote Sensing Imagery	4375
<i>Tom Burgert, Leonard Hackel, Paolo Rota, Begüm Demir</i>	
OMeGa: Joint Optimization of Explicit Meshes and Gaussian Splats for Robust Scene-Level Surface Reconstruction	4386
<i>Yuhang Cao, Haojun Yan, Danya Yao</i>	
Confidence Through Parallel Attention for Depth and Uncertainty Estimation in Dynamic Environments	4396
<i>Onkar Susladkar, Rohit Pawar, Chirag Sehgal, Samaksh Ujjawal, Sparsh Mittal</i>	
BiNAR: A Bi-Modal Framework for Non-Aligned RGB-IR 3D Reconstruction via Gaussian Splatting	4407
<i>Zhongwen Wang, Han Ling, Weihao Zhang, Yinghui Sun, Quansen Sun</i>	
Spec-Gloss Surfels and Normal-Diffuse Priors for Relightable Glossy Objects	4417
<i>Georgios Kouros, Minye Wu, Tinne Tuytelaars</i>	
Occlusion Boundary and Depth: Mutual Enhancement via Multi-Task Learning	4427
<i>Lintao Xu, Yinghao Wang, Chaohui Wang</i>	
Ego-EXTRA: video-language Egocentric Dataset for EXpert-TRAinee assistance	4438
<i>Francesco Ragusa, Michele Mazzamuto, Rosario Forte, Irene D’Ambra, James Fort, Jakob Engel Antonino Furnari, Giovanni Maria Farinella</i>	
Similarity-aware Probabilistic Embeddings Modeling for Video-Text Retrieval	4451
<i>Yulian Huang, Pengxu Wei, Zhicheng Dong, Liang Lin</i>	
PromptGAR: Flexible Promptive Group Activity Recognition	4461
<i>Zhangyu Jin, Andrew Feng, Ankur Chemburkar, Celso M. De Melo</i>	
Spacewalk-18: A Benchmark for Multimodal and Long-form Procedural Video Understanding in Novel Domains	4472
<i>Zitian Tang, Rohan Myer Krishnan, Zhiqiu Yu, Chen Sun</i>	
Broadcast2Pitch: Game State Reconstruction from Unconstrained Soccer Videos	4483
<i>Yin May Oo, Yewon Hwang, Muhammad Amrulloh Robbani, Vanyi Chao, Ankhzaya Jamsrandorj Hoang Quoc Nguyen, Kyung-Ryoul Mun, Jinwook Kim</i>	

VLMs Guided Interpretable Decision Making for Autonomous Driving*	4494
<i>Xin Hu, Taotao Jing, Renran Tian, Zhengming Ding</i>	
DARB-Splatting: Generalizing Splatting with Decaying Anisotropic Radial Basis Functions	4504
<i>Hashiru Pramuditha, Vinasirajan Viruthshaan, Vishagar Arunan, Saeedha Nazar, Sameera Ramasinghe</i> <i>Simon Lucey, Ranga Rodrigo</i>	
Surgical Gaussian Surfels: Highly Accurate Real-time Surgical Scene Rendering using Gaussian Surfels	4515
<i>Idris O. Sunmola, Zhenjun Zhao, Samuel Schmidgall, Yumeng Wang, Paul Maria Scheikl, Viet Pham</i> <i>Axel Krieger</i>	
Gated Temporal Fusion Transformers for Robust Multi-Object Tracking	4525
<i>Jinho Kim, Kuk-Jin Yoon</i>	
SIAM: Synchronous Interaction Attention for Human Mesh Recovery	4535
<i>Niaz Ahmad, Saif Ullah, Youngmoon Lee, Guanghui Wang</i>	
SynchroRaMa : Lip-Synchronized and Emotion-Aware Talking Face Generation via Multi-Modal Emotion Embedding	4546
<i>Phyo Thet Yee, Dimitrios Kollias, Sudeepta Mishra, Abhinav Dhall</i>	
Are All Marine Species Created Equal? Performance Disparities in Underwater Object Detection	4556
<i>Melanie Wille, Tobias Fischer, Scarlett Raine</i>	
LVM-Lite: Training Large Vision Models with Efficient Sequential Modeling	4566
<i>Xianhang Li, Hongru Zhu, Sucheng Ren, Linjie Yang, Peng Wang, Heng Wang, Xiaohui Shen, Qing Liu</i> <i>Cihang Xie</i>	
HiGlassRM: Learning to Remove High-prescription Glasses via Synthetic Dataset Generation	4577
<i>Sebin Lee, Heewon Kim</i>	
Transformer-Based Inpainting for Real-Time 3D Streaming in Sparse Multi-Camera Setups	4587
<i>Leif Van Holland, Domenic Zingsheim, Mana Takhsa, Hannah Dröge, Patrick Stotko, Markus Plack</i> <i>Reinhard Klein</i>	
SeaClips: A Video Dataset for Maritime Object Detection	4599
<i>Franziska Denk, Christian Rankl, Shaban Almouahed, David Moser, Robert Sablatnig</i>	
UniDiff: Parameter-Efficient Adaptation of Diffusion Models for Land Cover Classification with Multi-Modal Remotely Sensed Imagery and Sparse Annotations	4611
<i>Yuzhen Hu, Saurabh Prasad</i>	
CropAT: Leveraging Diffusion-Generated Target-Like Cropped Objects for Pseudo-Label Refinement in Domain-Adaptive Object Detection	4621
<i>Chen-Che Huang, Tzuhsuan Huang, Jun-Cheng Chen</i>	
Beyond Faces: A Multimodal Person Clustering for Unconstrained Environments	4631
<i>Sahngmin Yoo, Sangwon Lee, Seongin Jo</i>	
Eye-for-an-eye: Appearance Transfer with Dense Semantic Correspondence in Diffusion Models	4641
<i>Sooyeon Go, Kyungmook Choi, Minjung Shin, Youngjung Uh</i>	
Towards Photorealistic Style Transfer with Multimodal Guidance and Robustness to Content Images in Arbitrary Styles	4651
<i>Ruikai Zhou, Yating Liu, Yi Xu</i>	

Hierarchical Adaptive networks with Task vectors for Test-Time Adaptation	4661
<i>Sameer Ambekar, Marta Hasny, Laura Daza, Daniel M. Lang, Julia A. Schnabel</i>	
Automated Pore Detection from In-Situ FDM 3D Printing Video: A Comparative Evaluation of Modern Segmentation Models	4673
<i>Abdullah Al Ahad Khan, Md Shariful Islam, Lin Li, Lai Jiang, Noushin Ghaffari</i>	
Safe Vision-Language Models via Unsafe Weights Manipulation	4682
<i>Moreno D’Incà, Elia Peruzzo, Xingqian Xu, Humphrey Shi, Nicu Sebe, Massimiliano Mancini</i>	
SOPHY: Generating Simulation-Ready Objects with PHYsical Materials	4693
<i>Junyi Cao, Evangelos Kalogerakis</i>	
Generalization of Real World Video Deblurring by Image-to-image Translation	4705
<i>Kassymzhomart Aitbek, Seungjoon Yang</i>	
A Dataset and Framework for Learning State-invariant Object Representations	4715
<i>Rohan Sarkar, Avinash Kak</i>	
Better Safe Than Sorry? Overreaction Problem of Vision Language Models in Visual Emergency Recognition	4724
<i>Dasol Choi, Seunghyun Lee, Youngsook Song</i>	
Optimizing LVLMs with On-Policy Data for Effective Hallucination Mitigation	4733
<i>Chengzhi Yu, Yifan Xu, Yifan Chen, Wenyi Zhang</i>	
Countering Multi-modal Representation Collapse through Rank-targeted Fusion	4744
<i>Seulgi Kim, Kiran Kokilepersaud, Mohit Prabhushankar, Ghassan AlRegib</i>	
Learning Unified Spatio-temporal Representations for Efficient Compressed Video Understanding	4755
<i>Shristi Das Biswas, Efstathia Soufleri, Arani Roy, Kaushik Roy</i>	
More Than Memory Savings: Zeroth-Order Optimization Mitigates Forgetting in Continual Learning	4766
<i>Wanhao Yu, Zheng Wang, Shuteng Niu, Sen Lin, Li Yang</i>	
Global Focal and Radial Distortion Averaging from Radial Fundamental Matrices for Robust Self-Calibration	4777
<i>Sergei Solonets, Daniil Sinitsyn, Daniel Cremers</i>	
UCDSC: Open Set UnCertainty aware Deep Simplex Classifier for Medical Image Datasets	4787
<i>Arnav Aditya, Nitin Kumar, Saurabh Shigwan</i>	
DODA: Adapting Object Detectors to Dynamic Agricultural Environments in Real-Time with Diffusion	4797
<i>Shuai Xiang, Pieter M. Blok, James Burridge, Haozhou Wang, Wei Guo</i>	
Structured Context Learning for Generic Event Boundary Detection	4808
<i>Xin Gu, Congcong Li, Xinyao Wang, Dexiang Hong, Libo Zhang, Tiejian Luo, Longyin Wen, Heng Fan</i>	
Sun-E: Dataset and Benchmark for Event-Based Sun Sensing	4818
<i>Sydney Dolan, Alessandro Golkar</i>	
FCC: Fully Connected Correlation for One-Shot Segmentation	4827
<i>Seonghyeon Moon, Haein Kong, Muhammad Haris Khan, Mubbasir Kapadia, Yuewei Lin</i>	
MoRe: Monocular Geometry Refinement via Graph Optimization for Cross-View Consistency	4838
<i>Dongki Jung, Jaehoon Choi, Yonghan Lee, Sungmin Eum, Heesung Kwon, Dinesh Manocha</i>	

ProSkill: Segment-Level Skill Assessment in Procedural Videos	4849
<i>Michele Mazzamuto, Daniele Di Mauro, Gianpiero Francesca, Giovanni Maria Farinella, Antonino Furnari</i>	
MergeSlide: Continual Model Merging and Task-to-Class Prompt-Aligned Inference for Lifelong Learning on Whole Slide Images	4859
<i>Doanh C. Bui, Ba Hung Ngo, Hoai Luan Pham, Khang Nguyen, Mai K. Nguyen, Yasuhiko Nakashima</i>	
CoL ² A: Convolution-free Local Linear Attention for SpatioTemporal Event Processing	4869
<i>Yusuke Sekikawa, Jun Nagata, Itsumi Araki, Andreu Girbau</i>	
Dronaquatics: Real-time Swimming Analytics Using Drone Captured Imagery	4881
<i>Thu Tran, Harold Abraham Joseph, Kichang Lee, Kenny Tsu Wei Choo, Dong Ma, Shaohui Foong Thivya Kandappu, Jeonggil Ko, Rajesh Balan</i>	
Histopath-C: Towards Realistic Domain Shifts for Histopathology Vision-Language Adaptation	4890
<i>Mehrdad Noori, Gustavo A. Vargas Hakim, David Osowiechi, Fereshteh Shakeri, Ali Bahri Moslem Yazdanpanah, Sahar Dastani, Ismail Ben Ayed, Christian Desrosiers</i>	
Robust Multimodal Emotion Recognition from Incomplete Modalities via Query-Based Unimodal and Cross-Modal Learning	4901
<i>Ryo Miyoshi, Mayu Otani, Yuki Okafuji</i>	
WiSE-OD: Benchmarking Robustness in Infrared Object Detection	4912
<i>Heitor R. Medeiros, Atif Belal, Masih Aminbeidokhti, Eric Granger, Marco Pedersoli</i>	
Patch-wise Retrieval: A Bag of Practical Techniques for Instance-level Matching	4922
<i>Wonseok Choi, Sohwi Lim, Nam Hyeon-Woo, Moon Ye-Bin, Dong-Ju Jeong, Jinyoung Hwang Tae-Hyun Oh</i>	
Gaussian Swaying $\square$: Surface-Based Framework for Aerodynamic Simulation with 3D Gaussians	4932
<i>Hongru Yan, Xiang Zhang, Zeyuan Chen, Fangyin Wei, Zhuowen Tu</i>	
Restora-Flow: Mask-Guided Image Restoration with Flow Matching	4943
<i>Arnela Hadzic, Franz Thaler, Lea Bogensperger, Simon Johannes Joham, Martin Urschler</i>	
PrevMatch: Revisiting and Maximizing Temporal Knowledge in Semi-Supervised Semantic Segmentation	4953
<i>Wooseok Shin, Hyun Joon Park, Jin Sob Kim, Juan Yun, Se Hong Park, Sung Won Han</i>	
One-shot Portrait Stylization via Geometric Alignment	4964
<i>Xinrui Wang, Zilin Guo, Zhuoru Li, Jinze Yu, Heng Zhang, Yusuke Iwasawa, Yutaka Matsuo, Jiaxian Guo</i>	
Zero-Shot Table Extraction in Business Documents: A Unified Benchmark with Error Taxonomy and Ecological Analysis	4974
<i>Elliott Thomas, Mickael Coustaty, Aurélie Joseph, Tri-Cong Pham, Gaspar Deloin, Elodie Carel Vincent Poulain D'Andecy, Jean-Marc Ogier</i>	
SegMango: Early Deep Mango Yield Prediction based on Flower Segmentation and Weather Data	4984
<i>Janaksinh Ven, Charu Sharma, Azeemuddin Syed</i>	
Geo3DVQA: Evaluating Vision-Language Models for 3D Geospatial Reasoning from Aerial Imagery	4994
<i>Mai Tsujimoto, Junjue Wang, Weihao Xuan, Naoto Yokoya</i>	
PoseGaussian: Pose-Driven Novel View Synthesis for Robust 3D Human Reconstruction	5004
<i>Ju Shen, Chen Chen, Tam V. Nguyen, Vijayan K. Asari</i>	

Timestamp Query Transformer for Temporal Action Segmentation	5016
<i>Tieqiao Wang, Sinisa Todorovic</i>	
QC-SF: Improving Computer Vision for Airborne LiDAR Point Clouds of Boreal Forests with Quebec Simulated Forest Dataset	5026
<i>Olivier Stocker, Reza Mahmoudi Kouhi, Omid Reisi Gahruei, Thierry Badard, Eric Guilbert</i>	
RemEdit: Efficient Diffusion Editing with Riemannian Geometry	5037
<i>Eashan Adhikarla, Brian D. Davison</i>	
GrowTAS: Progressive Expansion from Small to Large Subnets for Efficient ViT Architecture Search	5047
<i>Hyunju Lee, Youngmin Oh, Jeimin Jeon, Donghyeon Baek, Bumsub Ham</i>	
Subspace-Guided Knowledge Distillation for Efficient Model Transfer	5057
<i>Zeeshan Hayder, Ali Cheraghian, Lars Petersson, Mehrtash Harandi</i>	
TimeRefine: Temporal Grounding with Time Refining Video LLM	5067
<i>Xizi Wang, Feng Cheng, Ziyang Wang, Huiyu Wang, Md Mohaiminul Islam, Lorenzo Torresani Mohit Bansal, Gedas Bertasius, David Crandall</i>	
Curve Skeletonization in Continuous domain for Meshes and Point Clouds	5079
<i>Jai Bardhan, Ramya Hebbalaguppe, Aravind Udupa</i>	
Gene-DML: Dual-Pathway Multi-Level Discrimination for Gene Expression Prediction from Histopathology Images	5090
<i>Yaxuan Song, Jianan Fan, Hang Chang, Weidong Cai</i>	
Color Preserving CMOS-SPAD Fusion for Multi-Frame HDR	5100
<i>Aleksi Suonsivu, Lauri Salmela, Lassi Helin, Leevi Uosukainen, Giacomo Boracchi</i>	
Unsupervised Segmentation by Diffusing, Walking and Cutting	5110
<i>Daniela Ivanova, Marco Aversa, Paul Henderson, John Williamson</i>	
Learning spatio-temporal feature representations for video-based gaze estimation	5121
<i>Alexandre Personnic, Mihai Băce</i>	
ImageChain: Advancing Sequential Image-to-Text Reasoning in Multimodal Large Language Models	5131
<i>Danae Sánchez Villegas, Ingo Ziegler, Desmond Elliott</i>	
Edge-Aware Image Manipulation via Diffusion Models with a Novel Structure-Preservation Loss	5142
<i>Minsu Gong, Nuri Ryu, Jungseul Ok, Sunghyun Cho</i>	
3D Superquadric Splatting	5154
<i>Daniel MacSwayne, Aleš Leonardis, Jianbo Jiao</i>	
Controllable Long-term Motion Generation with Extended Joint Targets	5164
<i>Eunjong Lee, Eunhee Kim, Sanghoon Hong, Eunho Jung, Jihoon Kim</i>	
ST-Think: How Multimodal Large Language Models Reason About 4D Worlds from Ego-Centric Videos	5174
<i>Peiran Wu, Yunze Liu, Miao Liu, Junxiao Shen</i>	
RPT-SR: Regional Prior attention Transformer for infrared image Super-Resolution	5184
<i>Youngwan Jin, Incheol Park, Yagiz Nalcakan, Hyeongjin Ju, Sanghyeop Yeo, Shiho Kim</i>	

LASOR: Towards Clinically Transparent and Explainable Ophthalmic Report Generation via Lesion-Aware Segmentation	5194
<i>Jian Park, Hyunseon Won, JeeEun Kim, Joon Seo Hwang, Jeong Mo Han, Ji In Park Daniel Duck-Jin Hwang, Jinyoung Han</i>	
Quantifying the Limits of Segmentation Foundation Models: Modeling Challenges in Segmenting Tree-Like and Low-Contrast Objects	5205
<i>Yixin Zhang, Nicholas Konz, Kevin Kramer, Maciej A. Mazurowski</i>	
Lorentz Entailment Cone for Semantic Segmentation	5216
<i>Zahid Hasan, Masud Ahmed, Nirmalya Roy</i>	
WarpRF: Multi-View Consistency for Training-Free Uncertainty Quantification and Applications in Radiance Fields	5226
<i>Sadra Safadoust, Fabio Tosi, Fatma Güney, Matteo Poggi</i>	
GAEA: A Geolocation Aware Conversational Assistant	5236
<i>Ron Campos, Ashmal Vayani, Parth Parag Kulkarni, Rohit Gupta, Aizan Zafar, Aritra Dutta, Mubarak Shah</i>	
WSSSP-Net: Weakly Supervised Semantic Segmentation Plugin Network for Face Anti-Spoofing	5247
<i>Krzysztof Galus, Piotr Syga, Piotr Kawa</i>	
Concord: Concept-Informed Diffusion for Dataset Distillation	5258
<i>Jianyang Gu, Haonan Wang, Ruoxi Jia, Saeed Vahidian, Vyacheslav Kungurtsev, Wei Jiang, Yiran Chen</i>	
Improving Out-of-Distribution Detection using Segmented Images and Cross-View Attention Fusion	5269
<i>Alexander Politowicz, Sahisnu Mazumder, Bing Liu</i>	
An improved architecture for part-based animal re-identification through semantic segmentation distillation	5280
<i>Eugênio Dias Ribeiro Neto, Marc Chaumont, Gérard Subsol, Michel de Garine-Wichatitsky, Hélène Guis</i>	
FB-4D: Spatial-Temporal Coherent Dynamic 3D Content Generation with Feature Banks	5290
<i>Jinwei Li, Huan-Ang Gao, Wenyi Li, Haohan Chi, Chenyu Liu, Chenxi Du, Yiqian Liu, Mingju Gao Guiyu Zhang, Zongzheng Zhang, Li Yi, Yao Yao, Jingwei Zhao, Hongyang Li, Yikai Wang, Hao Zhao</i>	
Hestia: Voxel-Face-Aware Hierarchical Next-Best-View Acquisition for Efficient 3D Reconstruction	5302
<i>Cheng-You Lu, Zhuoli Zhuang, Nguyen Thanh Trung Le, Da Xiao, Yu-Cheng Chang, Thomas Do Srinath Sridhar, Chin-Teng Lin</i>	
R ³ : Reconstruction, Raw, and Rain: Deraining Directly in the Bayer Domain	5313
<i>Nate Rothschild, Moshe Kimhi, Avi Mendelson, Chaim Baskin</i>	
An Efficient Multi-Rater Setup Towards Personalized and Diversified Medical Image Segmentation	5322
<i>Sajed Almorsy, Ayman Khalafallah, Marwan Torki</i>	
HiMix : Hierarchical Visual-Textual Mixing Network for Lesion Segmentation	5332
<i>Soojin Hwang, Jaeyoon Sim, Won Hwa Kim</i>	
FSP-DETR: Few-Shot Prototypical Parasitic Ova Detection	5342
<i>Shubham Trehan, Udhav Ramachandran, Akash Rao, Ruth Scimeca, Sathyanarayanan N. Aakur</i>	
DF-Mamba: Deformable State Space Modeling for 3D Hand Pose Estimation in Interactions	5352
<i>Yifan Zhou, Takehiko Ohkawa, Guwenxiao Zhou, Kanoko Goto, Takumi Hirose, Yusuke Sekikawa Nakamasa Inoue</i>	

Context-Preserving Dermoscopic Editing: Mask-Guided Lesion-Aware Diffusion for Attribute Modification	5364
<i>Tao Sun, Yun Jiang, Yarong Jin, Huanting Guo, Zequn Zhang</i>	
How I Met Your Bias: Investigating Bias Amplification in Diffusion Models	5374
<i>Nathan Roos, Ekaterina Iakovleva, Ani Gjergji, Vito Paolo Pastore, Enzo Tartaglione</i>	
BREEN: Bridge Data-Efficient Encoder-Free Multimodal Learning with Learnable Queries	5384
<i>Tianle Li, Yongming Rao, Winston Hu, Yu Cheng</i>	
ProtoGMVAE: A Variational Auto-Encoder with True Gaussian Mixture Prior for Prototypical-based Self-Explainability	5396
<i>Martin Blanchard, Christophe Ducottet, Damien Muselet, Olivier Delézy</i>	
AEON: Adaptive Embedding Optimized Noise for Robust Watermarking in Diffusion Models	5406
<i>Muhammad Shahid Muneer, Simon S. Woo</i>	
Revisiting Vision–Language Foundations for No-Reference Image Quality Assessment	5416
<i>Ankit Yadav, Ta Duc Huy, Lingqiao Liu</i>	
Semi-supervised Domain Adaptation via Mutual Alignment through Joint Error	5426
<i>Dexuan Zhang, Thomas Westfechtel, Tatsuya Harada</i>	
Unified Control for Inference-Time Guidance of Denoising Diffusion Models	5437
<i>Maurya Goyal, Anuj Singh, Hadi Jamali-Rad</i>	
Learning Subglacial Bed Topography from Sparse Radar with Physics-Guided Residuals	5447
<i>Bayu Adhi Tama, Jianwu Wang, Vandana Janeja, Mostafa Cham</i>	
4D Multimodal Co-attention Fusion Network with Latent Contrastive Alignment for Alzheimer’s Diagnosis	5457
<i>Yuxiang Wei, Yanteng Zhang, Xi Xiao, Tianyang Wang, Xiao Wang, Vince D. Calhoun</i>	
One-Cycle Structured Pruning via Stability-Driven Subnetwork Search	5467
<i>Deepak Ghimire, Dayoung Kil, Seonghwan Jeong, Jaesik Park, Seong-Heum Kim</i>	
PredMapNet: Future and Historical Reasoning for Consistent Online HD Vectorized Map Construction	5477
<i>Bo Lang, Nirav Savaliya, Zhihao Zheng, Jinglun Feng, Zheng-Hang Yeh, Mooi Choo Chuah</i>	
Multi-view stereo with multiple projectors for oneshot entire shape scan based on Neural SDF and DSSS demultiplexing	5488
<i>Kota Nishihara, Ryo Furukawa, Ryusuke Sagawa, Hiroshi Kawasaki</i>	
FreeCond: Free Lunch in the Input Conditions of Text-Guided Inpainting	5498
<i>Teng-Fang Hsiao, Bo-Kai Ruan, Sung-Lin Tsai, Yi-Lun Wu, Hong-Han Shuai</i>	
ControlEvents: Controllable Synthesis of Event Camera Data with Foundational Prior from Image Diffusion Models	5509
<i>Yixuan Hu, Yuxuan Xue, Simon Klenk, Daniel Cremers, Gerard Pons-Moll</i>	
DPBridge: Latent Diffusion Bridge for Dense Prediction	5520
<i>Haorui Ji, Taojun Lin, Hongdong Li</i>	
PatchEAD: Unifying Industrial Visual Prompting Frameworks for Patch-Exclusive Anomaly Detection	5531
<i>Po-Han Huang, Jeng-Lin Li, Po-Hsuan Huang, Ming-Ching Chang, Wei-Chao Chen</i>	

SurfDist: Interpretable Three-Dimensional Instance Segmentation Using Curved Surface Patches	5541
<i>Jackson Borchardt, Saul Kato</i>	
CRISP: Cylindrical Rendering for In-Stream Point Clouds	5550
<i>Hyungwoo Kan, Seonyoung Jang, Yeojun Yoon, Byung Tae Oh</i>	
INRetouch: Context Aware Implicit Neural Representation for Photography Retouching	5560
<i>Omar Elezabi, Marcos V. Conde, Zongwei Wu, Radu Timofte</i>	
Line Art Colorization with Offset Prior-based Diffusion Model	5570
<i>Xuan Zhu, Miao Cao, Fang-Lue Zhang, Yu-Kun Lai, Paul L Rosin</i>	
Food Image Generation on Multi-Noun Categories	5581
<i>Xinyue Pan, Yuhao Chen, Jiangpeng He, Fengqing Zhu</i>	
RobuMTL: Enhancing Multi-Task Learning Robustness Against Weather Conditions	5591
<i>Tasneem Shaffee, Sherief Reda</i>	
EndoPBR: Photorealistic Synthetic Data for Surgical 3D Vision via Physically-based Rendering	5601
<i>John J. Han, Jie Ying Wu</i>	
Inpainting of Sparse Depth Maps from Monocular Depth-from-Focus on Pixel Processor Arrays	5612
<i>Maciej Lewandowski, Piotr Dudek</i>	
DMS2F-HAD: A Dual-branch Mamba-based Spatial-Spectral Fusion Network for Hyperspectral Anomaly Detection	5623
<i>Aayushma Pant, Lakpa Tamang, Tsz-Kwan Lee, Sunil Aryal</i>	
F-ViT: Foundation Model Guided Visible-to-Infrared Translation	5633
<i>Jay Nitin Paranjape, Celso M de Melo, Vishal M. Patel</i>	
KFS-Bench: Comprehensive Evaluation of Key Frame Sampling in Long Video Understanding	5643
<i>Zongyao Li, Kengo Ishida, Satoshi Yamazaki, Xiaotong Ji, Jianquan Liu</i>	
FujiView: Multimodal Late-Fusion for Predicting Scenic Visibility	5653
<i>Bryceton Bible, Nehal Hasnaeen, Hairong Qi</i>	
Improving Animal Pose Estimation through Species Similarity Measures and Rigorous Label Definition	5662
<i>Medhashree Parhy, Shaan Chanchani, Claire Kim, Josh Mansky, Zian Pan, Parth Thakre, Haoyu Chen Amy R. Reibman</i>	
Grounding Descriptions in Images informs Zero-Shot Visual Recognition	5672
<i>Shaunak Halbe, Junjiao Tian, K J Joseph, James Seale Smith, Katherine Stevo, Vineeth N Balasubramanian Zsolt Kira</i>	
Semi-supervised Key-Point Estimation for Echocardiography Video	5682
<i>Seok-Hwan Oh, Hyeon-Jik Lee, Guil Jung, Myeong-Gee Kim, Young-Min Kim, Hyuksool Kwon Hyeon-Min Bae</i>	
Anatomically-guided masked autoencoder pre-training for aneurysm detection	5693
<i>Alberto M. Ceballos Arroyo, Jisoo Kim, Chu-Hsuan Lin, Lei Qin, Geoffrey S. Young, Huaizu Jiang</i>	
Style-Friendly SNR Sampler for Style-Driven Generation	5703
<i>Jooyoung Choi, Chaehun Shin, Yeongtak Oh, Heeseung Kim, Jungbeom Lee, Sungroh Yoon</i>	

Lose Your Self (LoYS): an adversarial entropy-based unsupervised approach for model debiasing	5714
<i>Vito Paolo Pastore, Massimiliano Ciranni, Vittorio Murino</i>	
MagicDrive3D: Controllable 3D Generation for Any-View Rendering in Street Scenes	5724
<i>Ruiyuan Gao, Kai Chen, Zhihao Li, Lanqing Hong, Zhenguo Li, Qiang Xu</i>	
Grounding Degradations in Natural Language for All-In-One Video Restoration	5734
<i>Muhammad Kamran Janjua, Amirhosein Ghasemabadi, Kunlin Zhang, Mohammad Salameh, Chao Gao Di Niu</i>	
ControlVP: Interactive Geometric Refinement of AI-Generated Images with Consistent Vanishing Points	5744
<i>Ryota Okumura, Kaede Shiohara, Toshihiko Yamasaki</i>	
Enhancing Vision Language Corruption Robustness using Cross-Distribution & Prompted Denoisers	5754
<i>Sameer Shafayet Latif, Sadab Shiper, K. M. Rahiduzzaman Kiran, Md Farhan Ishmam, Md Azam Hossain Abu Raihan Mostofa Kamal, Md Hamjajul Ashmafee</i>	
BlendCLIP: Bridging Synthetic and Real Domains for Zero-Shot 3D Object Classification with Multimodal Pretraining	5766
<i>Ajinkya Khoche, Gergő László Nagy, Maciej Wozniak, Thomas Gustafsson, Patric Jensfelt</i>	
Towards Egocentric 3D Hand Pose Estimation in Unseen Domains	5776
<i>Wiktor Mucha, Michael Wray, Martin Kampel</i>	
Direct Visual Grounding by Directing Attention of Visual Tokens	5787
<i>Parsa Esmaeilkhani, Longin Jan Latecki</i>	
Motion-Aware Graph Fusion Network for 3D Human Pose Estimation	5798
<i>Yen Pham, Xiaohui Yuan, Chengyuan Zhuang</i>	
UniGaze: Towards Universal Gaze Estimation via Large-scale Pre-Training	5809
<i>Jiawei Qin, Xucong Zhang, Yusuke Sugano</i>	
Unsupervised Discovery of Long-Term Spatiotemporal Periodic Workflows in Human Activities	5821
<i>Fan Yang, Quanting Xie, Atsunori Moteki, Shoichi Masui, Shan Jiang, Kanji Uchino, Yonatan Bisk Graham Neubig</i>	
VAST-ReID: A Low-Light Benchmark Dataset for Person Re-Identification with Visual and Attribute-Rich Semantic Tracking	5833
<i>Hammad Khan, Rakesh Kumar Giri, Kamalakar Vijay Thakare, Heeseung Choi, Hyunjoo Jung Debi Prosad Dogra, Ig-Jae Kim</i>	
DexAvatar: 3D Sign Language Reconstruction with Hand and Body Pose Priors	5842
<i>Kaustubh Kundu, Hrishav Bakul Barua, Lucy Robertson-Bell, Zhixi Cai, Kalin Stefanov</i>	
DREAM: Dynamic Prompts and GuidedMix for Efficient Continual Adaptation of Visual-Language Models	5853
<i>Evelyn Chee, Mong Li Lee, Wynne Hsu</i>	
brat : Aligned Multi-View Embeddings for Brain MRI Analysis	5864
<i>Maxime Kayser, Maksim Gridnev, Wanting Wang, Max Bain, Aneesh Rangnekar, Avijit Chatterjee Aleksandr Petrov, Harini Veeraraghavan, Nathaniel C. Swinburne</i>	

Towards Fine-Grained Adaptation of CLIP via a Self-Trained Alignment Score	5875
<i>Eman Ali, Sathira Silva, Chetan Arora, Muhammad Haris Khan</i>	
Advancing Multimodal LLMs by Large-Scale 3D Visual Instruction Dataset Generation	5886
<i>Liu He, Xiao Zeng, Yizhi Song, Albert Y. C. Chen, Lu Xia, Shashwat Verma, Sankalp Dayal, Min Sun</i>	
<i>Cheng-Hao Kuo, Daniel Aliaga</i>	
CLIP-UP: CLIP-Based Unanswerable Problem Detection for Visual Question Answering	5898
<i>Ben Vardi, Oron Nir, Ariel Shamir</i>	
Isolating the Role of Temporal Information in Video Saliency: A Controlled Experimental Analysis	5909
<i>Peter El-Jiz, Matthias Kuemmerer, Matthias Tangemann, Matthias Bethge, Andreas Bartels</i>	
<i>Michael Mario Bannert</i>	
Diffusion-Based Action Recognition Generalizes to Untrained Domains	5919
<i>Rogério Guimarães, Frank Xiao, Pietro Perona, Markus Marks</i>	
CORA: Consistency-Guided Semi-Supervised Framework for Reasoning Segmentation	5934
<i>Prantik Howlader, Hoang Nguyen-Canh, Srijan Das, Jingyi Xu, Hieu Le, Dimitris Samaras</i>	
Hierarchical Instance Tracking to Balance Privacy Preservation with Accessible Information	5945
<i>Neelima Prasad, Jarek Reynolds, Neel Karsanbhai, Tanusree Sharma, Lotus Zhang, Abigale Stangl</i>	
<i>Yang Wang, Leah Findlater, Danna Gurari</i>	
Causality-Driven Audits of Model Robustness	5956
<i>Nathan Drenkow, William Paul, Chris Ribaudo, Mathias Unberath</i>	
BAFLE-DCT: Bypassing Adversarial Filters via Frequency-Selective Embedding in the DCT Domain	5967
<i>Thilina Mendis, Farah Kandah, Sathyanarayanan N. Aakur</i>	
Logit-Adjusted Test-Time Adaptation under Partial Class Imbalance	5977
<i>Thilina Weerasinghe, Ruwan Tennakoon, WeiQin Chuah, Alireza Bab-Hadiashar</i>	
Test Time Adaptation Using Adaptive Quantile Recalibration	5986
<i>Paria Mehrbod, Pedro Vianna, Geraldin Nanfack, Guy Wolf, Eugene Belilovsky</i>	
OSEG: Improving Diffusion sampling through Orthogonal Smoothed Energy Guidance	5996
<i>Masud An Nur Islam Fahim, Nazmus Saqib, Joon-Min Gil</i>	
Hymavi : A Hybrid Mamba-Attention Network in Multi-View Framework for Volumetric Medical Image Segmentation	6006
<i>Sy Dat Tran, Jin Kyu Gahm</i>	
RoadBench: A Vision-Language Foundation Model and Benchmark for Road Damage Understanding	6016
<i>Xi Xiao, Yunbei Zhang, Janet Wang, Lin Zhao, Yuxiang Wei, Hengjia Li, Yanshu Li, Xiao Wang</i>	
<i>Swalpa Kumar Roy, Hao Xu, Tianyang Wang</i>	
IMKD: Intensity-Aware Multi-Level Knowledge Distillation for Camera-Radar Fusion	6027
<i>Shashank Mishra, Karan Patil, Didier Stricker, Jason Rambach</i>	
LogicCBMs: Logic-Enhanced Concept-Based Learning	6039
<i>Deepika SN Vemuri, Gautham Bellamkonda, Aditya Pola, Vineeth N Balasubramanian</i>	
Understanding the Visual Projection Space of Multimodal LLMs	6049
<i>Suncheon Jeong, Yoojeong Song, Hyungjoon Kim</i>	

NoHumansRequired: Autonomous High-Quality Image Editing Triplet Mining	6059
<i>Maksim Kuprashevich, Grigorii Alekseenko, Irina Tolstykh, Georgii Fedorov, Bulat Suleimanov Vladimir Dokholyan, Aleksandr Gordeev</i>	
SSMT-Net: A Semi-Supervised Multitask Transformer-Based Network for Thyroid Nodule Segmentation in Ultrasound Images	6069
<i>Muhammad Umar Farooq, Abd Ur Rehman, Azka Rehman, Muhammad Usman, Dong-Kyu Chae</i>	
Gaussian Splatting Map Registration with Orthographic Bird’s-Eye-View Renderings	6080
<i>H. Leblond, G. Simon, R. Martins, C. Demonceaux, M.-O. Berger</i>	
MUSE: Model-based Uncertainty-aware Similarity Estimation for zero-shot 2D Object Detection and Segmentation	6090
<i>Sungmin Cho, Sungbum Park, Insoo Oh</i>	
MVAT: Multi-View Aware Teacher for Weakly Supervised 3D Object Detection	6101
<i>Saad Lahlali, Alexandre Fournier-Mongieux, Nicolas Granger, Hervé Le Borgne, Quoc-Cuong Pham</i>	
ODE _t (ODE _l): Shortcutting the Time and the Length in Diffusion and Flow Models for Faster Sampling	6111
<i>Denis Gudovskiy, Wenzhao Zheng, Tomoyuki Okuno, Yohei Nakata, Kurt Keutzer</i>	
TM-Adapter: Temporal Merge Adapter for Efficient Global Temporal Modeling	6121
<i>Woo Joo Hahm, Seungwoo Jang, Hyeon Tak Kim, Daeun Lee, Kwangsu Kim</i>	
Diagnose Like A REAL Pathologist: An Uncertainty-Focused Approach for Trustworthy Multi-Resolution Multiple Instance Learning	6132
<i>Sungrae Hong, Sol Lee, Jisu Shin, Jiwon Jeong, Mun Yong Yi</i>	
Align Video Diffusion Model with Online Video-Centric Preference Optimization	6142
<i>Jiacheng Zhang, Jie Wu, Weifeng Chen, Yatai Ji, Xuefeng Xiao, Weilin Huang, Kai Han</i>	
SceneProp: Combining Neural Network and Markov Random Field for Scene-Graph Grounding	6153
<i>Keita Otani, Tatsuya Harada</i>	
GHOST: Getting to the Bottom of Hallucinations with A Multi-round Consistency Benchmark	6163
<i>Vibashan VS, Nadine Chang, Jenny Schmalfluss, Vishal M. Patel, Zhiding Yu, Jose M. Alvarez</i>	
GeneVA: A Dataset of Human Annotations for Generative Text to Video Artifacts	6174
<i>Jenna Kang, Maria Beatriz Silva, Patsorn Sangkloy, Kenneth Chen, Niall L. Williams, Qi Sun</i>	
Zero-Shot Domain Generalisation via Prompt-Driven Feature Refinement	6184
<i>Tingrui Qiao, Di Zhao, Caroline Walker, Chris Cunningham, Yun Sing Koh</i>	
Can Image Splicing and Copy-Move Forgery Be Detected by the Same Model? Forensim: An Attention-Based State-Space Approach	6194
<i>Soumyaroop Nandi, Prem Natarajan</i>	
Distilling Offline Action Detection Models into Real-Time Streaming Models	6205
<i>Deep Patel, Yasunori Babazaki, Yasuto Nagase, Iain Melvin, Martin Renqiang Min</i>	
AuthGuard: Generalizable Deepfake Detection via Language Guidance	6215
<i>Guangyu Shen, Zhihua Li, Xiang Xu, Tianchen Zhao, Zheng Zhang, Dongsheng An, Zhuowen Tu, Yifan Xing Qin Zhang</i>	

GroupPortrait: Multi-ID Portrait Generation with High Identity Preservation and Fine-Grained Control	6226
<i>Meijia Huang, Ruida Li, Bing Ma, Liangwei Jiang, Shuo Fang, Chenguang Ma</i>	
GFT-GCN: Privacy-Preserving 3D Face Mesh Recognition with Spectral Diffusion	6236
<i>Hichem Felouat, Hanrui Wang, Isao Echizen</i>	
Denoise, Divide, Distill, and Predict D ³ P: Towards Forecasting Long-horizon Real-world Anomaly from Normalcy	6246
<i>Quentin M��rilleau, Snehashis Majhi, Antitza Dantcheva, Quan Kong, Lorenzo Garattoni Gianpiero Francesca, Fran��ois Br��mond</i>	
See, Record, Do: Automated Generation of UI Workflows from Tutorial Videos	6256
<i>Adam Beauchaine, Craig Shue</i>	
Video and Language Alignment in 2D Systems for 3D Multi-object Scenes with Multi-Information Derivative-Free Control	6266
<i>Jason Armitage, Rico Sennrich</i>	
FocalComm: Hard Instance-Aware Multi-Agent Perception	6277
<i>Dereje Shenkut, Vijayakumar Bhagavatula</i>	
SFMNet: Sparse Focal Modulation for 3D Object Detection	6287
<i>Oren ShROUT, Ayellet Tal</i>	
MoSCo: Real-time and Efficient Text-to-Motion Synthesis via Delta Training	6298
<i>Zhiyuan Zhang, Lingqiao Liu</i>	
VLMDiff: Leveraging Vision-Language Models for Multi-Class Anomaly Detection with Diffusion	6309
<i>Samet Hicsonmez, Abd El Rahman Shabayek, Djamila Aouada</i>	
Predicting Task fMRI Contrasts from Resting-State fMRI Using Sparse 3D Convolutions	6320
<i>Ivan Sviridov, Maria Boyko, Maksim Sharaev</i>	
SceneEdited: A City-Scale Benchmark for 3D HD Map Updating via Image-Guided Change Detection	6330
<i>Chun-Jung Lin, Tat-Jun Chin, Sourav Garg, Feras Dayoub</i>	
Overcoming Small Data Limitations in Video-Based Infant Respiration Estimation	6340
<i>Liyang Song, Hardik Bishnoi, Sai Kumar Reddy Manne, Sarah Ostadabbas, Briana J. Taylor, Michael Wan</i>	
Non-Aligned Reference Image Quality Assessment for Novel View Synthesis	6350
<i>Abhijay Ghildyal, Rajesh Suredi, Nabajeet Barman, Saman Zadtootaghaj, Alan C Bovik</i>	
Pointmap-Conditioned Diffusion for Consistent Novel View Synthesis	6360
<i>Thang-Anh-Quan Nguyen, Laurent Caraffa, Jean-Philippe Tarel, Roland Br��mond</i>	
TED-4DGS: Temporally Activated and Embedding-based Deformation for 4DGS Compression	6371
<i>Cheng-Yuan Ho, He-Bi Yang, Jui-Chiu Chiang, Yu-Lun Liu, Wen-Hsiao Peng</i>	
QUOTA: Quantifying Objects with Text-to-Image Models for Any Domain	6381
<i>Wenfang Sun, Yingjun Du, Gaowen Liu, Yefeng Zheng, Cees G. M. Snoek</i>	
Single-step Diffusion for Image Compression at Ultra-Low Bitrates	6391
<i>Chanung Park, Joo Chan Lee, Jong Hwan Ko</i>	
Virtually Unrolling the Herculaneum Papyri by Diffeomorphic Spiral Fitting	6401
<i>Paul Henderson</i>	

Diffusion Noise Optimization for Synthetic VLM Training	6412
<i>Ren Ohkubo, Rintaro Yanagi, Hirokatsu Kataoka, Yutaka Satoh</i>	
Histogram Assisted Quality Aware Generative Model for Resolution Invariant NIR Image Colorization	6422
<i>Abhinav Attri, Rajeev Ranjan Dwivedi, Samiran Das, Vinod Kumar Kurmi</i>	
Polymorph: Energy-Efficient Multi-Label Classification for Video Streams on Embedded Devices	6432
<i>Saeid Ghafouri, Mohsen Fayyaz, Xiangchen Li, Deepu John, Bo Ji, Dimitrios S. Nikolopoulos</i> <i>Hans Vandierendonck</i>	
A Unified Diffusion-Based Framework for Multi-Agent Trajectory Prediction Integrating Structured Multi-Modal Representations	6442
<i>Chenxi Yang, Suyang Xi, Hong Ding, Yiqing Shen, Yunhao Liu</i>	
ChartQA-X: Generating Explanations for Visual Chart Reasoning	6453
<i>Shamanthak Hegde, Pooyan Fazli, Hasti Seifi</i>	
Distribution Highlighted Reference-based Label Distribution Learning for Facial Age Estimation	6464
<i>Satoshi Suzuki, Shin'Ya Yamaguchi, Shoichiro Takeda, Takuhiro Kaneko, Shota Orihashi, Ryo Masumura</i>	
T2VWorldBench: A Benchmark for Evaluating World Knowledge in Text-to-Video Generation	6474
<i>Yubin Chen, Xuyang Guo, Zhenmei Shi, Zhao Song, Jiahao Zhang</i>	
UniTabBank: A Large Scale Multi-Lingual, Multi-Layout, Multi-Type, Multi-Format Dataset for Table Detection	6486
<i>Ajoy Mondal, Saumya Mundra, Avijit Dasgupta, C. V. Jawahar</i>	
R-MMA: Enhancing Vision-Language Models with Recurrent Adapters for Few-Shot and Cross-Domain Generalization	6496
<i>Md Fahim, Md Farhan Ishmam, Mir Sazzat Hossain, M Ashraful Amin, Amin Ahsan Ali</i> <i>AKM Mahbubur Rahman</i>	
Hybrid State Representation for Video Procedure Planning	6507
<i>Woo Suk Choi, Youwon Jang, Minsu Lee, Byoung-Tak Zhang</i>	
Training-Free Few-Shot Segmentation via Vision-Language Guided Prompting	6517
<i>Euihyun Yoon, Taejin Park, Jaekoo Lee</i>	
High-Level Semantics and Low-Level Features Fusion for Multi-Scale Object Detection in Dynamic Construction Environments	6527
<i>Mahdi Bonyani, Maryam Soleymani, Chao Wang</i>	
Mean-Shift Distillation for Diffusion Mode Seeking	6537
<i>Vikas Thamizharasan, Nikitas Chatzis, Iliyan Georgiev, Matthew Fisher, Evangelos Kalogerakis, Difan Liu</i> <i>Nanxuan Zhao, Michal Lukáč</i>	
TacticalCalib: End-to-End 6-DoF Camera Pose Regression for Tactical Camera Calibration	6547
<i>Liang Fan, Xiaoqian Liu, Zhi Chen, Ling kai Yang</i>	
F-INR: Functional Tensor Decomposition for Implicit Neural Representations	6557
<i>Sai Karthikeya Vemuri, Tim Büchner, Joachim Denzler</i>	
V2XScene: Multi-View Consistent 3D Scene Simulation for Collaborative Perception	6569
<i>Yanfei Li, Yi Gong, Yuan Zeng</i>	

WiSAR3D - Aerial LiDAR dataset for 3D object detection	6580
<i>Oren ShROUT, Ori Nizan, Yizhak Ben-Shabat, Ayellet Tal</i>	
A Deep Network for Object Detection on Inland Waters	6590
<i>Dennis Griesser, Bastian Goldluecke, Matthias O. Franz, Georg Umlauf</i>	
CoreCaption: Core Caption based Text-to-Video Retrieval	6600
<i>Junkyu Jang</i>	
ZebraPose: Zebra Detection and Pose Estimation using only Synthetic Data	6611
<i>Elia Bonetto, Amir Ahmad</i>	
OPFormer: Object Pose Estimation leveraging foundation model with geometric encoding	6621
<i>Artem Moroz, Vít Zeman, Martin Mikšík, Elizaveta Isianova, Miroslav David, Pavel Burget, Varun Burde</i>	
FAIR-SIGHT: Fairness Assurance in Image Recognition via Simultaneous Conformal Thresholding and Dynamic Output Repair	6633
<i>Arya Fayyazi, Mehdi Kamal, Massoud Pedram</i>	
GDoFS: Gaussian DoF Separation for Plausible 3D Geometry in Sparse-View 3DGS	6643
<i>Yongsung Kim, Jooyoung Choi, Sungroh Yoon</i>	
Learning Beyond Labels: Self-Supervised Handwritten Text Recognition	6653
<i>Shree Mitra, Ajoy Mondal, C.V. Jawahar</i>	
Guiding What Not to Generate: Automated Negative Prompting for Text-Image Alignment	6664
<i>Sangha Park, Eunji Kim, Yeongtak Oh, Jooyoung Choi, Sungroh Yoon</i>	
Neural Geometry Image-Based Representations with Optimal Transport (OT)	6676
<i>Xiang Gao, Yuanpeng Liu, Jiazhi Li, Xinmu Wang, Minghao Guo, Yu Guo, Xiyun Song, Heather Yu Zhiqiang Lao, Xianfeng David Gu</i>	
RealDroneVision: Dataset and Architecture Advancements for Small-Object Drone Detection	6687
<i>Arun Kumar Sivapuram, Pranav R T Peddinti, Harish Puppala, Komuravelli Prashanth, Jaladi Sri Harsha Rama Krishna Sai Gorthi</i>	
PerVL-Bench: Benchmarking Multimodal Personalization for Large Vision–Language Models	6696
<i>Minsung Kim</i>	
TRACE: Confounder-free Adversarial Fine-tuning for Robust Object Detection	6705
<i>Wonho Lee, Jisu Lee, Hyunsik Na, Sohee Park, Daeseon Choi</i>	
Zero-LEAD: Source-Free Universal Domain Adaptation for Abdominal Multi-Organ Segmentation	6715
<i>Ahmed El-Sayed, Marwan Torki</i>	
Correcting and Quantifying Systematic Errors in 3D Box Annotations for Autonomous Driving	6724
<i>Alexandre Justo Miro, Ludvig Af Klinteberg, Bogdan Timus, Aron Asefaw, Ajinkya Khoche Thomas Gustafsson, Sina Sharif Mansouri, Masoud Daneshtalab</i>	
FastHMR: Accelerating Human Mesh Recovery via Token and Layer Merging with Diffusion Decoding	6733
<i>Soroush Mehraban, Andrea Iaboni, Babak Taati</i>	

Point2Pose: A Generative Framework for 3D Human Pose Estimation with Multi-View Point Cloud Dataset	6744
<i>Hyunsoo Lee, Daeum Jeon, Hyeokjae Oh</i>	
UniVid: Unifying Vision Tasks with Pre-trained Video Generation Models	6754
<i>Lan Chen, Yuchao Gu, Qi Mao</i>	
Optimal Transport for Rectified Flow Image Editing: Unifying Inversion-Based and Direct Methods	6764
<i>Marian Lupaşcu, Mihai Sorin Stupariu</i>	
Synthesizing Compositional Videos from Text Description	6775
<i>Prajwal Singh, Kuldeep Kulkarni, Shanmuganathan Raman, Harsh Rangwani</i>	
S2O: Static to Openable Enhancement for Articulated 3D Objects	6785
<i>Denys Iliash, Hanxiao Jiang, Yiming Zhang, Manolis Savva, Angel X. Chang</i>	
HodgeFormer: Transformers for Learnable Operators on Triangular Meshes through Data-Driven Hodge Matrices	6796
<i>Akis Nousias, Stavros Nousias</i>	
Exploring Automated Recognition of Instructional Activity and Discourse from Multimodal Classroom Data	6806
<i>Ivo Bueno, Ruikun Hou, Babette Bühler, Tim Fütterer, James Drimalla, Jonathan K. Foster, Peter Youngs, Peter Gerjets, Ulrich Trautwein, Enkelejda Kasneci</i>	
From Prompt to Production: Automating Brand-Safe Marketing Imagery with Text-to-Image Models	6818
<i>Parmida Atighehchian, Henry Wang, Andrei Kapustin, Boris Lerner, Tiancheng Jiang, Taylor Jensen, Negin Sokhandan</i>	
Equivariant Sampling for Improving Diffusion Model-based Image Restoration	6827
<i>Chenxu Wu, Qingpeng Kong, Peiang Zhao, Wendi Yang, Wenxin Ma, Fenghe Tang, Zihang Jiang, S.Kevin Zhou</i>	
PoseAdapt: Sustainable Human Pose Estimation via Continual Learning Benchmarks and Toolkit	6840
<i>Muhammad Saif Ullah Khan, Didier Stricker</i>	
SAVeD: Learning to Denoise Low-SNR Video for Improved Downstream Performance	6851
<i>Suzanne Stathatos, Michael Hobbey, Pietro Perona, Markus Marks</i>	
ATM: Enhanced Alignment for Text-to-Motion Generation	6862
<i>Ke Han, Yueming Lyu, Weichen Yu, Nicu Sebe</i>	
Memoire: Learning User Personas from Gallery Tags for Personalized Photo Curation	6873
<i>Praful Mathur, Mohsin Iftekhhar, Aman Sharma, Sarvesh Tiwari, Meghali Deka, Sathish Cherukuri, K Roopa Sheshadri, Rakesh Valusa</i>	
CSGaussian: Progressive Rate-Distortion Compression and Segmentation for 3D Gaussian Splatting	6883
<i>Yu-Jen Tseng, Chia-Hao Kao, Jing-Zhong Chen, Alessandro Gnutti, Shao-Yuan Lo, Yen-Yu Lin, Wen-Hsiao Peng</i>	
Test-Time Adaptation for Video Highlight Detection Using Meta-Auxiliary Learning and Cross-Modality Hallucinations	6893
<i>Zahidul Islam, Sujoy Paul, Mrigank Rochan</i>	

SD-CSFL: A Synthetic Data-Driven Conformity Scoring Framework for Robust Federated Learning	6903
<i>Ebtisaam Alharbi, Abdulrahman Kerim, Leandro Soriano Marcolino, Qiang Ni</i>	
Alignment and Distillation: A Robust Framework for Multimodal Domain Generalizable Human Action Recognition	6913
<i>Hyeonbin Ji, Juyeob Lee, Eunil Park</i>	
Procedure Learning via Regularized Gromov-Wasserstein Optimal Transport	6925
<i>Syed Ahmed Mahmood, Ali Shah Ali, Umer Ahmed, Fawad Javed Fateh, M. Zeeshan Zia, Quoc-Huy Tran</i>	
Clear Sights on Site: A Spatial-Adaptive Channel Network for Deblurring Construction Site Images	6936
<i>Mahdi Bonyani, Maryam Soleymani, Chao Wang</i>	
SegMo: Segment-aligned Text to 3D Human Motion Generation	6946
<i>Bowen Dang, Lin Wu, Xiaohang Yang, Zheng Yuan, Zhixiang Chen</i>	
From Darkness to Detail: Frequency-Aware SSMs for Low-Light Vision	6956
<i>Eashan Adhikarla, Kai Zhang, Gong Chen, John Nicholson, Brian D. Davison</i>	
Multimodal Adversarial Defense for Vision-Language Models by Leveraging One-To-Many Relationships	6968
<i>Futa Waseda, Antonio Tejero-de-Pablos, Isao Echizen</i>	
ForestSplats: Deformable transient field for Gaussian Splatting in the Wild	6978
<i>Wongi Park, Myeongseok Nam, Siwon Kim, Sangwoo Jo, Soomok Lee</i>	
Graph Query Networks for Object Detection with Automotive Radar	6988
<i>Loveneet Saini, Hasan Tercan, Tobias Meisen</i>	
FlowCLAS: Enhancing Normalizing Flow-Based Anomaly Segmentation Via Contrastive Learning	6998
<i>Chang Won Lee, Selina Leveugle, Paul Grouchy, Chris Langley, Svetlana Stolpner, Jonathan Kelly Steven L. Waslander</i>	
ScoreNet: Netting Lightweight Quality Scores for Better Visual Assessment with Large Multi-Modality Models	7008
<i>Bahador Rashidi, Kiarash Aghakasiri, Shupeí Zhang, Amirmohsen Sattarifard, Yue Zhang, Chao Gao</i>	
VectorSynth: Fine-Grained Satellite Image Synthesis with Structured Semantics	7019
<i>Daniel Cher, Brian Wei, Srikumar Sastry, Nathan Jacobs</i>	
Digital Forensic AI You Can Explain: A Case Study on Video Source Camera Identification	7030
<i>Maryna Veksler, Kemal Akkaya, Selcuk Uluagac</i>	
Modeling and Learning Multiple Hypotheses for Monocular 3D Object Detection	7040
<i>Hyeonjeong Park, Peixi Xiong, Pei Yu, Wei Tang</i>	
SGD-Mix: Enhancing Domain-Specific Image Classification with Label-Preserving Data Augmentation	7051
<i>Yixuan Dong, Fang-Yi Su, Jung-Hsien Chiang</i>	
DreamCatcher: Efficient Multi-Concept Customization via Representation Finetuning	7062
<i>Jungwon Lee, Changhun Lee, Eunhyeok Park</i>	
Enhancing Monocular 3D Hand Reconstruction with Learned Texture Priors	7073
<i>Giorgos Karvounas, Nikolaos Kyriazis, Iason Oikonomidis, Georgios Pavlakos, Antonis A. Argyros</i>	

PALMS+: Modular Image-Based Floor Plan Localization Leveraging Depth Foundation Model	7084
<i>Yunqian Cheng, Benjamin Princen, Roberto Manduchi</i>	
IPCD: Intrinsic Point-Cloud Decomposition	7094
<i>Shogo Sato, Takuhiro Kaneko, Shoichiro Takeda, Tomoyasu Shimada, Kazuhiko Murasaki, Taiga Yoshida Ryuichi Tanida, Akisato Kimura</i>	
Multimodal Graph Representation Learning over Arbitrary Sets of Modalities	7104
<i>Santosh Patapati, Trisanth Srinivasan</i>	
FNOpt: Resolution-Agnostic, Self-Supervised Cloth Simulation using Meta-Optimization with Fourier Neural Operators	7116
<i>Ruochen Chen, Thuy Tran, Shaifali Parashar</i>	
D2Mamba: Dual Domain Guided Informed Search in State Space Model for Underwater Image Enhancement	7126
<i>Alik Pramanick, Soumajit Roy, Arijit Sur</i>	
Image-Guided Semantic Pseudo-LiDAR Point Generation for 3D Object Detection	7137
<i>Minseung Lee, Seokha Moon, Seung Joon Lee, Reza Mahjourian, Jinkyu Kim</i>	
HABIT: Human Action Benchmark for Interactive Traffic in CARLA	7148
<i>Mohan Ramesh, Mark Azer, Fabian Flohr</i>	
EVTP-IVS: Effective Visual Token Pruning For Unifying Instruction Visual Segmentation In Multi-Modal Large Language Models	7158
<i>Wenhui Zhu, Xiwen Chen, Zhipeng Wang, Shao Tang, Sayan Ghosh, Xuanzhao Dong, Rajat Koner Yalin Wang</i>	
AuViRe: Audio-visual Speech Representation Reconstruction for Deepfake Temporal Localization	7168
<i>Christos Koutlis, Symeon Papadopoulos</i>	
Exploring the Boundaries of Diffusion Models for Offline Writer Identification with Sparse and Intra-Variable Data	7178
<i>Aritra Dey, Chandranath Adak, Kumari Priya, Soumi Chattopadhyay, Sukalpa Chanda</i>	
CLoCKDistill: Consistent Location and Context aware Knowledge Distillation for DETRs	7188
<i>Qizhen Lan, Qing Tian</i>	
MaxInfo: A Training-Free Key-Frame Selection Method Using Maximum Volume for Enhanced Video Understanding	7198
<i>Pengyi Li, Irina Abdullaeva, Alexander Gambashidze, Andrey Kuznetsov, Ivan Oseledets</i>	
Cycle-Consistent Multi-Graph Matching for Self-Supervised Annotation of C. Elegans	7208
<i>Sebastian Stricker, Christoph Karg, Lisa Hutschenreiter, Bogdan Savchynskyy, Dagmar Kainmueller</i>	
Automated Suturing Skill Assessment in Robot-assisted Surgery from Endoscopic Videos using Clinically-guided Evaluation Criteria	7218
<i>Atharva Sunil Deo, Ujjwal Pasupulety, Nicholas Matsumoto, Jay Moran, Cherine Yang, Jeanine Kim Rafal Dariusz Kocielnik, Aurash Naser-Tavakolian, Andrew Hung</i>	
Deep Image Decomposition for Medical Imaging Anonymization and Curation	7229
<i>Yael Elkin, Gal Ben-Arie, Tammy Riklin-Raviv</i>	

Intraoperative 2D/3D Registration via Spherical Similarity Learning and Differentiable Levenberg-Marquardt Optimization	7239
<i>Minheng Chen, Youyong Kong</i>	
ACuRE: Accurate Continuity-Regularized SpO ₂ Estimation Using Liquid Time-Constant Networks	7250
<i>Shahzad Ahmad, Divya Mishra, Sania Bano, Sukalpa Chanda, Yogesh Singh Rawat</i>	
CAST: Evaluating Multi-Object Trackers with Context-Aware Switch and Transfer Scores	7260
<i>Jin Bai, Gregory D. Hager</i>	
Advancing Player Identification and Tracking with Global ID Fusion (GIF)	7269
<i>Karol Wojtulewicz, Minxing Liu, Niklas Carlsson</i>	
Distilling What and Why: Enhancing Driver Intention Prediction with MLLMs	7281
<i>Sainithin Artham, Avijit Dasgupta, Shankar Gangisetty, C. V. Jawahar</i>	
LASER: Lip Landmark Assisted Speaker Detection for Robustness	7291
<i>Le Thien Phuc Nguyen, Zhuoran Yu, Yong Jae Lee</i>	
VADER: Towards Causal Video Anomaly Understanding with Relation-Aware Large Language Models	7301
<i>Ying Cheng, Yu-Ho Lin, Min-Hung Chen, Fu-En Yang, Shang-Hong Lai</i>	
SCAdapter: Content-Style Disentanglement for Diffusion Style Transfer	7312
<i>Luan Thanh Trinh</i>	
T2LF: LLM-Guided Multimodal Diffusion for Text-to-Light Field Synthesis	7322
<i>Soyoung Yoon, Namhyuk Ahn, In Kyu Park</i>	
VideoSketcher: A Training-Free Approach for Coherent Video Sketch Transfer	7333
<i>Huining Li, Bangzhen Liu, Rui Yang, Yang Zhou, Chenshu Xu, Xufang Pang, Shengfeng He</i>	
Zero-Shot Audio-Visual Editing via Cross-Modal Delta Denoising	7344
<i>Yan-Bo Lin, Kevin Lin, Zhengyuan Yang, Linjie Li, Jianfeng Wang, Chung-Ching Lin, Xiaofei Wang, Gedas Bertasius, Lijuan Wang</i>	
SceneEval: Evaluating Semantic Coherence in Text-Conditioned 3D Indoor Scene Synthesis	7355
<i>Hou In Ivan Tam, Hou In Derek Pun, Austin T. Wang, Angel X. Chang, Manolis Savva</i>	
IPTQ-ViT: Post-Training Quantization of Non-linear Functions for Integer-only Vision Transformers	7366
<i>Gihwan Kim, Jemin Lee, Hyungshin Kim</i>	
MM-TS: Multi-Modal Temperature and Margin Schedules for Contrastive Learning with Long-Tail Data	7376
<i>Siarhei Sheludsko, Dhimitrios Duka, Bernt Schiele, Hilde Kuehne, Anna Kukleva</i>	
Boosting Unsupervised Video Instance Segmentation with Automatic Quality-Guided Self-Training	7387
<i>Kaixuan Lu, Mehmet Onurcan Kaya, Dim P. Papadopoulos</i>	
Locally Explaining Prediction Behavior via Gradual Interventions and Measuring Property Gradients	7398
<i>Niklas Penzel, Joachim Denzler</i>	
Beyond Paired Data: Self-Supervised UAV Geo-Localization from Reference Imagery Alone	7409
<i>Tristan Amadei, Enric Meinhardt-Llopis, Benedicte Bascle, Corentin Abgrall, Gabriele Facciolo</i>	

Where is the Watermark? Interpretable Watermark Detection at the Block Level	7420
<i>Maria Bulychev, Neil G. Marchant, Benjamin I. P. Rubinstein</i>	
PointSt3R: Point Tracking through 3D Grounded Correspondence	7430
<i>Rhodri Guerrier, Adam W. Harley, Dima Damen</i>	
Uplifting Table Tennis: A Robust, Real-World Application for 3D Trajectory and Spin Estimation	7440
<i>Daniel Kienzle, Katja Ludwig, Julian Lorenz, Shin'Ichi Satoh, Rainer Lienhart</i>	
FairVLM: Enhancing Fairness and Prompt Sensitivity in Vision Language Models for Medical Image Segmentation	7450
<i>Md Motiur Rahman, Saeka Rahman, Smriti Bhatt, Miad Faezipour</i>	
SHaSaM: Submodular Hard Sample Mining for Fair Facial Attribute Recognition	7461
<i>Anay Majee, Rishabh Iyer</i>	
DM ³ Net: Dual-Camera Super-Resolution via Domain Modulation and Multi-scale Matching	7472
<i>Cong Guan, Jiacheng Ying, Yuya Ieiri, Osamu Yoshie</i>	
SuperRivolution: Fine-Scale Rivers from Coarse Temporal Satellite Imagery	7482
<i>Rangel Daroya, Subhransu Maji</i>	
Adversarial Pseudo-replay for Exemplar-free Class-incremental Learning	7493
<i>Hiroto Honda</i>	
DiffRegCD: Integrated Registration and Change Detection with Diffusion Features	7503
<i>Syedehanita Madani, Rama Chellappa, Vishal M. Patel</i>	
HDR Reconstruction Boosting with Training-Free and Exposure-Consistent Diffusion	7513
<i>Yo-Tin Lin, Su-Kai Chen, Hou-Ning Hu, Yen-Yu Lin, Yu-Lun Liu</i>	
Optimizing against Infeasible Inclusions from Data for Semantic Segmentation through Morphology	7524
<i>Shamik Basu, Luc Van Gool, Christos Sakaridis</i>	
3D Cell Oversegmentation Correction via Geo-Wasserstein Divergence	7534
<i>Peter Chen, Bryan Chang, Olivia A Creasey, Julie Beth Sneddon, Zev J Gartner, Yining Liu</i>	
TopoRec: Point Cloud Recognition Using Topological Data Analysis	7544
<i>Anirban Ghosh, Iliya Kulbaka, Ian Dahlin, Ayan Dutta</i>	
MapleGrasp: Mask-guided Feature Pooling for Language-driven Efficient Robotic Grasping	7554
<i>Vineet Bhat, Naman Patel, Prashanth Krishnamurthy, Ramesh Karri, Farshad Khorrami</i>	
Remote Sensing Forestry Similarity Convolution	7565
<i>Shikuan Wang, Yuangong Chen, Jianzhou Gong, Lingyi Meng, Mengquan Wu, Longxing Liu, Haiwei Yuan Mingbin Guo</i>	
RampWatch: An In-the-Wild Dataset and Text-Guided Detection Framework for Recreational Vessels	7576
<i>Malik Muhammad Asim, Claire B. Smallwood, Abdullah Tariq, Johnny Lo, Syed Zulqarnain Gilani</i>	
Enhancing Reverse Distillation with Core Exemplar Learning for Unified Multi-Class Anomaly Detection	7586
<i>Heechul Lim, Min-Soo Kim, Hyun-Boo Lee, Suk-Ju Kang, Kang-Wook Chon, Haeyun Lee</i>	
Leveraging Sparsity for Privacy in Collaborative Inference	7596
<i>Maximilian Andreas Hoefler, Karsten Mueller, Wojciech Samek</i>	

Improvise, Adapt, Overcome — Telescopic Adapters for Efficient Fine-tuning of Vision Language Models in Medical Imaging	7605
<i>Ujjwal Mishra, Vinita Shukla, Praful Hambarde, Amit Shukla</i>	
SVD-Det: A Lightweight Framework for Video Forgery Detection Using Semantic and Visual Defect Cues	7616
<i>Tsung-Shan Yang, Tianyu Zhang, Feng Qian, Bing Yan, C.-C. Jay Kuo</i>	
Joint Optimization of Camera Model and Deep Neural Network for Image Recognition	7626
<i>Youta Noboru, Yuko Ozasa, Masayuki Tanaka</i>	
Training-free Conditional Image Embedding Framework Leveraging Large Vision Language Models	7636
<i>Masayuki Kawarada, Kosuke Yamada, Antonio Tejero-de-Pablos, Naoto Inoue</i>	
ReFineVQA: Iterative Refinement of Video Description via Feedback Generation for Video Question Answering	7647
<i>Jeongwan Shin, Chan Hur, Seongmin Cho, Jaeho Choi, Hyeyoung Park</i>	
MIST: Multilingual Incidental Dataset for Scene Text Detection	7658
<i>Saumya Mundra, Ajoy Mondal, C.V Jawahar</i>	
Multi-Grained Text-Guided Image Fusion for Multi-Exposure and Multi-Focus Scenarios	7668
<i>Mingwei Tang, Jiahao Nie, Guang Yang, Ziqing Cui, Jie Li</i>	
MemeTAG: Keyword-Driven Meme Classification through Tag Embedding Reconstruction	7679
<i>Akshit Sharma, Prashant W Patil</i>	
Leveraging Semantic Attribute Binding for Free-Lunch Color Control in Diffusion Models	7689
<i>Héctor Laria, Alexandra Gomez-Villa, Jiang Qin, Muhammad Atif Butt, Bogdan Raducanu Javier Vazquez-Corral, Joost van de Weijer, Kai Wang</i>	
MedPEFT-CL: Dual-Phase Parameter-Efficient Continual Learning with Medical Semantic Adapter and Bidirectional Memory Consolidation	7699
<i>Ziyuan Gao, Philippe Morel</i>	
Splatter Layout: Geometry-embedded 3D Reconstruction via Surface Unfolding	7709
<i>Bryan Heryanto, Tackgeun You, Chanwoo Kim, Hwasup Lim</i>	
HOLO: Holistic Lightweight Optimization for Scene Understanding with Auto-Annotation and Multimodal Learning	7719
<i>Xiaoyun Hu, Xiaohan Yan, Nan Wang, Gang Wei, Zhicheng Wang</i>	
KMOPS: Keypoint-Driven Method for Multi-Object Pose and Metric Size Estimation from Stereo Images	7730
<i>Ying-Kun Wu, Yi Shen, Tzuhsuan Huang, I-Sheng Fang, Jun-Cheng Chen</i>	
PDV: Prompt Directional Vectors for Zero-shot Composed Image Retrieval	7740
<i>Osman Tursun, Sinan Kalkan, Simon Denman, Clinton Fookes</i>	
GRAPE (Gaussian Rendering for Accelerated Pixel Enhancement) Brings Fast and Lightweight Arbitrary Super-Resolution	7750
<i>Jung In Jang, Kyong Hwan Jin</i>	

Fetal and Neonatal Cortical Surface Reconstruction with Anatomical Normal-guidance and Perceptual Enhancements	7759
<i>Jiyang Lee, Woori Bae, U-Geun Ji, Hanyeol Yang, Jong-Min Lee</i>	
View-aware Cross-modal Distillation for Multi-view Action Recognition	7769
<i>Trung Thanh Nguyen, Yasutomo Kawanishi, Vijay John, Takahiro Komamizu, Ichiro Ide</i>	
NRGMark: Localized Watermarking for Energy Transparency in Images	7779
<i>Shruti Agarwal, Élie Michel, Vishal Asnani, Tania Mathern, John Collomosse</i>	
Test-Time Consistency in Vision Language Models	7789
<i>Shih-Han Chou, Shivam Chandhok, James J. Little, Leonid Sigal</i>	
Revisiting an Old Perspective Projection for Monocular 3D Morphable Models Regression	7799
<i>Toby Chong, Ryota Nakajima</i>	
General and Domain-Specific Zero-shot Detection of Generated Images via Conditional Likelihood	7809
<i>Roy Betser, Omer Hofman, Roman Vainshtein, Guy Gilboa</i>	
Co-STAR: Collaborative Curriculum Self-Training with Adaptive Regularization for Source-Free Video Domain Adaptation	7821
<i>Amirhossein Dadashzadeh, Parsa Esmati, Majid Mirmehdi</i>	
Being Positive about Negative Queries: Exclusion Aware Multimodal Retrieval using Disentangled Representations	7832
<i>Prachi Jha, Sumit Bhatia, Srikanta Bedathur</i>	
DualRes: Production-ready Dynamic Object Detection	7842
<i>Jibril El Hassani, Thomas Verelst</i>	
Semantic Map Guided Bird's-Eye View Learning for Online HD Map Construction	7852
<i>Huantao Ren, Hesham M. Eraqi, ABM Musa, Mohamed Moustafa</i>	
ISALux: Illumination and Semantics-Aware Transformer Employing Mixture of Experts for Low Light Image Enhancement	7862
<i>Raul Balmez, Alexandru Brateanu, Ciprian Orhei, Codruta O. Ancuti, Cosmin Ancuti</i>	
FastPose-ViT: A Vision Transformer for Real-Time Spacecraft Pose Estimation	7873
<i>Pierre Ancey, Andrew Price, Saqib Javed, Mathieu Salzmann</i>	
QuEENet: Quantum-Enhanced Expressive Network for Image Classification	7883
<i>Shashank Bayal, Rushikesh Govind Dawane, Komal, Santosh Kumar Vipparthi, Subrahmanyam Murala</i>	
SOLAR: Switchable Output Layer for Accuracy and Robustness in Once-for-All Training	7893
<i>Shaharyar Ahmed Khan Tareen, Lei Fan, Xiaojing Yuan, Qin Lin, Bin Hu</i>	
FG-Tracer: Tracing Information Flow in Multimodal Large Language Models in Free-Form Generation	7903
<i>Alessia Saporita, Vittorio Pipoli, Federico Bolelli, Lorenzo Baraldi, Andrea Acquaviva, Elisa Ficarra</i>	
Patch Your Matcher: Correspondence-Aware Image-to-Image Translation Unlocks Cross-Modal Matching via Single-Modality Priors	7913
<i>Anton Frolov, Volker Rodehorst</i>	
Diversity Preserving Coresets for Image Quality Assessment	7925
<i>Arpita Nema, Hanwei Zhu, Xi Zhang, Weisi Lin</i>	

SAVE: Sparse Autoencoder-Driven Visual Information Enhancement for Mitigating Object Hallucination	7935
<i>Sangha Park, Seungryong Yoo, Jisoo Mok, Sungroh Yoon</i>	
NavMapFusion: Diffusion-based Fusion of Navigation Maps for Online Vectorized HD Map Construction	7945
<i>Thomas Monninger, Zihan Zhang, Steffen Staab, Sihao Ding</i>	
GFT: Graph Feature Tuning for Efficient Point Cloud Analysis	7955
<i>Manish Dhakal, Venkat R. Dasari, Rajshekhar Sunderraman, Yi Ding</i>	
QAL: A Loss for Recall–Precision Balance in 3D Reconstruction	7965
<i>Pranay Meshram, Yash Turkar, Kartikeya Singh, Praveen Raj Masilamani, Charuvahan Adhivarahan Karthik Dantu</i>	
Meta-YOLO: Metadata-Guided Real-Time Object Detector in Aerial Imagery	7975
<i>Deukryeol Yoon, Seonghak Kim, Young Hwa Sung, Jinho Jung</i>	
Q-Former Autoencoder: A Modern Framework for Medical Anomaly Detection	7985
<i>Francesco Dalmonte, Emirhan Bayar, Emre Akbas, Mariana-Iuliana Georgescu</i>	
AusSmoke meets MultiNatSmoke: a fully-labelled diverse smoke segmentation dataset	7996
<i>Weihaoli, Hongjin Zha, Gao Zhu, Ge-Peng Ji, Nicholas Wilson, Marta Yebra, Nick Barnes</i>	
NAPP: Noise-Adaptive Prototype Perturbation for Few-Shot Learning	8007
<i>Ilhwan Kim, Sangwoo Yun, Dongheon Lee, Seongsu Kim, Joonki Paik</i>	
GaussianHeadTalk: Wobble-Free 3D Talking Heads with Audio Driven Gaussian Splatting	8017
<i>Madhav Agarwal, Mingtian Zhang, Laura Sevilla-Lara, Steven McDonagh</i>	
Splannequin: Freezing Monocular Mannequin-Challenge Footage with Dual-Detection Splatting	8028
<i>Hao-Jen Chien, Yi-Chuan Huang, Chung-Ho Wu, Wei-Lun Chao, Yu-Lun Liu</i>	
Optimization-Free Style Transfer for 3D Gaussian Splats	8041
<i>Raphael Du Sablon, David Hart</i>	
PEaRL: Pathway-Enhanced Representation Learning for Gene and Pathway Expression Prediction from Histology	8052
<i>Sejuti Majumder, Saarthak Kapse, Moinak Bhattacharya, Xuan Xu, Alisa Yurovsky, Prateek Prasanna</i>	
Flood-LDM: Generalizable Latent Diffusion Models for rapid and accurate zero-shot High-Resolution Flood Mapping	8063
<i>Sun Han Neo, Sachith Seneviratne, Herath Mudiyanse Viraj Vidura Herath, Abhishek Saha Sanka Rasnayaka, Lucy Amanda Marshall</i>	
LangPose: Language-Aligned Motion for Robust 3D Human Pose Estimation	8073
<i>Longyun Liao, Rong Zheng</i>	
SphereEdit: Spherical Semantic Editing in Diffusion Models	8084
<i>Salamata Konate, Hassan Hamidi, Elham Dolatabadi, Frank Rudzicz, Laleh Seyyed-Kalantri</i>	
FedEFC: Federated Learning Using Enhanced Forward Correction Against Noisy Labels	8094
<i>Seunghun Yu, Jin-Hyun Ahn, Joonhyuk Kang</i>	

Photo Dating by Facial Age Aggregation	8103
<i>Jakub Paplham, Vojtěch Franc</i>	
Scalable Video Action Anticipation with Cross Linear Attentive Memory	8113
<i>Zeyun Zhong, Manuel Martin, David Schneider, David J. Lerch, Chengzhi Wu, Frederik Diederichs</i>	
<i>Juergen Gall, Jürgen Beyerer</i>	
DUDA: Distilled Unsupervised Domain Adaptation for Lightweight Semantic Segmentation	8124
<i>Beomseok Kang, Niluthpol Chowdhury Mithun, Abhinav Rajvanshi, Han-Pang Chiu, Supun Samarasekera</i>	
Generalizing Sports Feedback Generation by Watching Competitions and Reading Books: A Rock Climbing Case Study	8136
<i>Arushi Rai, Adriana Kovashka</i>	
START: Spatial and Textual Learning for Chart Understanding	8146
<i>Zhuoming Liu, Xiaofeng Gao, Feiyang Niu, Qiaozi Gao, Liu Liu, Robinson Piramuthu</i>	
VRAgent: Self-Refining Agent for Zero-Shot Multimodal Video Retrieval	8157
<i>Ketul Shah, Pankaj Nathani, Rama Chellappa, Fabian Caba Heilbron</i>	
MapVerse: A Benchmark for Geospatial Question Answering on Diverse Real-World Maps	8168
<i>Sharat Bhat, Harshita Khandelwal, Tushar Kataria, Vivek Gupta</i>	
IMPACT: Interpretable Most Important Person Analysis and Classification using Transformer-based Models	8179
<i>Akshat Rampuria, Kamakshya Prasad Nayak, Kamalakar Vijay Thakare, Tushar Joshi</i>	
<i>Aditya Dhananjay Singh, Haesol Park, Heeseung Choi, Hyunjoo Jung, Debi Prosad Dogra, Ig-Jae Kim</i>	
SurgXBench: Explainable Vision-Language Model Benchmark for Surgery	8188
<i>Jiajun Cheng, Xianwu Zhao, Sainan Liu, Xiaofan Yu, Ravi Prakash, Patrick J. Codd, Jonathan Elliott Katz</i>	
<i>Shan Lin</i>	
SeqFeedNet: Sequential Feature Feedback Network for Background Subtraction	8199
<i>Yu-Shun Huang, Jing-Ming Guo, Yi-Xiang Yang</i>	
Robust Scene Coordinate Regression via Geometrically-Consistent Global Descriptors	8209
<i>Son Tung Nguyen, Alejandro Fontan, Michael Milford, Tobias Fischer</i>	
VitaGlyph: Vitalizing Artistic Typography with Flexible Dual-branch Diffusion Models	8220
<i>Kailai Feng, Yabo Zhang, Haodong Yu, Zhilong Ji, Jinfeng Bai, Hongzhi Zhang, Wangmeng Zuo</i>	
CaFlow: Enhancing Long-Term Action Quality Assessment with Causal Counterfactual Flow	8231
<i>Ruisheng Han, Kanglei Zhou, Shuang Chen, Amir Atapour-Abarghouei, Hubert P. H. Shum</i>	
AortaDiff: A Unified Multitask Diffusion Framework for Contrast-Free AAA Imaging	8242
<i>Yuxuan Ou, Ning Bi, Jiazheng Pan, Jiancheng Yang, Boliang Yu, Usama Zidan, Regent Lee, Vicente Grau</i>	
DirectDrag: High-Fidelity, Mask-Free, Prompt-Free Drag-based Image Editing via Readout-Guided Feature Alignment	8252
<i>Sheng-Hao Liao, Shang-Fu Chen, Tai-Ming Huang, Wen-Huang Cheng, Kai-Lung Hua</i>	
Bridging the Domain Gap in Small Multimodal Models: A Dual-level Alignment Perspective	8262
<i>Aveen Dayal, Peketì Divya, Nidhi Tiwari, Linga Reddy Cenkeramaddi, C Krishna Mohan, Abhinav Kumar</i>	

Crash2DocAI: Automated Integration of Post-Crash Car Part Images into Technical Reports	8272
<i>Václav Diviš, Jessica Giovagnola, Khalil Ben Chikha, Marek Hruží</i>	
Conversational Image Generation: Towards Multi-Round Personalized Generation with Multi-Modal Language Models	8282
<i>Haochen Zhang, Animesh Sinha, Felix Juefei-Xu, Haoyu Ma, Kunpeng Li, Zhipeng Fan, Xiaoliang Dai Tingbo Hou, Peizhao Zhang, Zecheng He</i>	
CSF-Net: Context-Semantic Fusion Network for Large Mask Inpainting	8292
<i>Chae-Yeon Heo, Yeong-Jun Cho</i>	
SceneShine: Illumination-aware Human Scene Gaussian Re-Splatting from Mobile Device Video	8302
<i>Xuqian Ren, Wenjia Wang, Mai Ngoc Nguyen, Juho Kannala, Esa Rahtu</i>	
See, Think, Learn: A Self-Taught Multimodal Reasoner	8313
<i>Sourabh Sharma, Sonam Gupta, Sadbhawna</i>	
SpecGen: Neural Spectral BRDF Generation via Spectral-Spatial Tri-plane Aggregation	8323
<i>Zhenyu Jin, Wenjie Li, Zhanyu Ma, Heng Guo</i>	
Pretraining Helps When Capacity Allows: Evidence from Ultra-Small ConvNets	8333
<i>Srikanth Muralidharan, Heitor R. Medeiros, Masih Aminbeidokhti, Eric Granger, Marco Pedersoli</i>	
Mem-MLP: Real-Time 3D Human Motion Generation from Sparse Inputs	8343
<i>Sinan Mutlu, Georgios F. Angelis, Savas Ozkan, Paul Wisbey, Anastasios Drosou, Mete Ozay</i>	
UltraClean: A Simple Framework to Train Robust Neural Networks against Backdoor Attacks	8353
<i>Bingyin Zhao, Yingjie Lao</i>	
GorillaWatch: An Automated System for In-the-Wild Gorilla Re-Identification and Population Monitoring	8364
<i>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</i>	
Reciprocal Teaching: Dynamic Multi-Model Teacher-Student Learning for Multiple Noisy Annotations	8376
<i>Wenjie Ai, Cuong C. Nguyen, Adrian Hilton, Gustavo Carneiro</i>	
DuPLUS: Dual-Prompt Vision-Language Framework for Universal Medical Image Segmentation and Prognosis	8386
<i>Numan Saeed, Tausifa Jan Saleem, Fadillah Maani, Muhammad Ridzuan, Hu Wang, Mohammad Yaqub</i>	
TriaGS: Differentiable Triangulation-Guided Geometric Consistency for 3D Gaussian Splatting	8396
<i>Quan Tran, Tuan Dang</i>	
SDT-6D: Fully Sparse Depth-Transformer for Staged End-to-End 6D Pose Estimation in Industrial Multi-View Bin Picking	8406
<i>Nico Leuze, Maximilian Hoh, Samed Doğan, Nicolas R. Peña, Alfred Schoettl</i>	
Generalized Category Discovery for LiDAR Semantic Segmentation	8416
<i>Minseok Kim, Jiyong Boo, Kuk-Jin Yoon</i>	
ICONIC-444: A 3.1-Million-Image Dataset for OOD Detection Research	8427
<i>Gerhard Kruppl, Henning Avenhaus, Horst Possegger</i>	

Any Detector Can Detect Anything	8437
<i>Thomas E. Huang, Siyuan Li, Martin Danelljan, Henghui Ding, Luc Van Gool, Fisher Yu</i>	
Towards Unconstrained Cross-View Pose Estimation	8448
<i>Alexander Wollam, Kyle Ashley, Maxim Shugaev, Oliver Arend, Ilya Semenov, Hadis Dashtestani</i> <i>Sumved Ravi, Nathan Jacobs</i>	
Disentangle and Regularize: Sign Language Production with Articulator-Based Disentanglement and Channel-Aware Regularization	8458
<i>Sümeyye Meryem Taşyürek, Tuğçe Kızıltepe, Hacer Yalim Keles</i>	
SmoothDiffusion-VE: Real-time Generative Video Editing Using Adaptive Feature Cache	8468
<i>Mustafa Munir, Sophia Zalewski, Shiqiu Liu, David Tarjan, Sushmitha Beled, Anjul Patney</i> <i>Radu Marculescu</i>	
SafeguardGS: 3D Gaussian Primitive Pruning While Avoiding Catastrophic Scene Destruction	8479
<i>Yongjae Lee, Zhaoliang Zhang, Deliang Fan</i>	
Uncertainty-Aware Vision-Language Segmentation for Medical Imaging	8490
<i>Aryan Das, Tanishq Rachamalla, Koushik Biswas, Swalpa Kumar Roy, Vinay Kumar Verma</i>	
Stabilizing Direct Training of Spiking Neural Networks: Membrane Potential Initialization and Threshold-robust Surrogate Gradient	8500
<i>Hyunho Kook, Byeongho Yu, Jeong Min Oh, Eunhyeok Park</i>	
DocWaveDiff: A Predict-and-Refine approach for Document Image Enhancement with Wavelet U-Nets and Diffusion models	8511
<i>Matteo Marulli, Marco Bertini</i>	
Chain-of-Look Spatial Reasoning for Dense Surgical Instrument Counting	8521
<i>Rishikesh Bhyri, Brian R Quaranto, Junsong Yuan, Peter C W Kim, Nan Xi</i>	
Rethinking Latent Variable in Learned Image Compression	8531
<i>Fangzhou Yi, Zhicheng Gong, Hui Zeng</i>	
AugMapNet: Improving Spatial Latent Structure via BEV Grid Augmentation for Enhanced Vectorized Online HD Map Construction	8541
<i>Thomas Monninger, Md Zafar Anwar, Stanislaw Antol, Steffen Staab, Sihao Ding</i>	
From Cognitive Priors to Instance Semantics: A Unified Framework for Multi-task Affective Computing	8551
<i>Guanyu Hu, Dimitrios Kollias, Xinyu Yang</i>	
FuLLaMa: Training-free Diffusion-based Object Removal with Context Preservation	8563
<i>İlke Demir, Umur Aybars Çiftçi</i>	
STRinGS: Selective Text Refinement in Gaussian Splatting	8574
<i>Abhinav Raundhal, Gaurav Behera, P. J. Narayanan, Ravi Kiran Sarvadevabhatla, Makarand Tapaswi</i>	
VIZOR: Viewpoint-Invariant Zero-Shot Scene Graph Generation for 3D Scene Reasoning	8584
<i>Vivek Madhavaram, Vartika Sengar, Arkadipta De, Charu Sharma</i>	
Relevance-aware Multi-context Contrastive Decoding for Retrieval-augmented Visual Question Answering	8596
<i>Jongha Kim, Byungoh Ko, Jeehye Na, Jinsung Yoon, Hyunwoo J. Kim</i>	

CanKD: Cross-Attention-based Non-local operation for Feature-based Knowledge Distillation	8606
<i>Shizhe Sun, Wataru Ohyama</i>	
FlyPose: Towards Robust Human Pose Estimation From Aerial Views	8617
<i>Hassaan Farooq, Marvin Brenner, Peter Stütz</i>	
MedROV: Towards Real-Time Open-Vocabulary Detection Across Diverse Medical Imaging Modalities	8628
<i>Tooba Tehreem Sheikh, Jean Lahoud, Rao Muhammad Anwer, Fahad Shahbaz Khan, Salman Khan Hisham Cholakkal</i>	
Exploiting Label-Independent Regularization from Spatial Patterns for Whole Slide Image Analysis	8639
<i>Weiye Wu, Xinwen Xu, Chongyang Gao, Xingjian Diao, Siting Li, Jiang Gui</i>	
SVS-GAN for Semantic Synthesis of Traffic Videos for Autonomous Driving	8650
<i>Khaled M. Seyam, Julian Wiederer, Markus Braun, Bin Yang</i>	
Integrating Multi-scale and Multi-filtration Topological Features for Medical Image Classification*	8660
<i>Pengfei Gu, Huimin Li, Haoteng Tang, Dongkuan Xu, Erik Enriquez, DongChul Kim, Bin Fu Danny Z. Chen</i>	
NeuroBridge: Few-Shot Cross-Modal Neuron Re-identification via Dual-Channel Deep Metric Learning	8670
<i>Wenwei Li, Mingwei Liao, Lingyi Cai, Anan Li</i>	
Sketch3R: Rapid and Realistic 3D VR Sketch Creation to Shape Retrieval	8680
<i>Mritunjoy Halder, Shivam Ashok Shukla, Lokender Tiwari, Raghav Mittal, Brojeshwar Bhowmick</i>	
PhyEduVideo: A Benchmark for Evaluating Text-to-Video Models for Physics Education	8690
<i>Megha Mariam K M, Aditya Arun, Zakaria Laskar, C.V. Jawahar</i>	
Dual-Domain Multimodal Hyperbolic Fusion for Cardiopulmonary Disease Diagnosis in Emergency Care	8700
<i>Ke Nan, Maggie Samaan, Benjamin Burns, Xia Ning, Yuchi Han, Yuan Xue</i>	