

39th AAAI Conference on Artificial Intelligence (AAAI-25), 37th Conference on Innovative Applications of Artificial Intelligence (IAAI-25), and 15th Symposium on Educational Advances in Artificial Intelligence (EAAI-25)

Volume 10: AAAI Technical Tracks

- Computer Vision IX

Philadelphia, Pennsylvania, USA
25 February - 4 March 2025

Part 1 of 2

ISBN: 979-8-3313-2520-6

Printed from e-media with permission by:

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



Some format issues inherent in the e-media version may also appear in this print version.

Copyright© (2025) by Association for the Advancement of Artificial Intelligence
All rights reserved.

Printed with permission by Curran Associates, Inc. (2025)

For permission requests, please contact Association for the Advancement of Artificial Intelligence
at the address below.

Association for the Advancement of Artificial Intelligence
2275 East Bayshore Road
Suite 160
Palo Alto, California 94303
USA

Phone: 1-650-328-3123
Fax: 1-650-321-4457

<https://aaai.org/Press/press.php>

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
Email: curran@proceedings.com
Web: www.proceedings.com

TABLE OF CONTENTS

PART 1

AAAI TECHNICAL TRACK ON COMPUTER VISION IX

Semi-supervised Infrared Small Target Detection with Thermodynamic-Inspired Uneven Perturbation and Confidence Adaptation.....	10013
<i>Mingjin Zhang, Wenteng Shang, Fei Gao, Qiming Zhang, FengQin Lu, Jing Zhang</i>	
MOCID: Motion Context and Displacement Information Learning for Moving Infrared Small Target Detection	10022
<i>Mingjin Zhang, Yuanjun Ouyang, Fei Gao, Jie Guo, Qiming Zhang, Jing Zhang</i>	
Decoupling Scattering: Pseudo-Label Guided NeRF for Scenes with Scattering Media	10031
<i>Mingyang Zhang, Junkang Zhang, Faming Fang, Guixu Zhang</i>	
PanoDiT: Panoramic Videos Generation with Diffusion Transformer	10040
<i>Muyang Zhang, Yuzhi Chen, Rongtao Xu, Changwei Wang, Jinming Yang, Weiliang Meng, Jianwei Guo, Huihuang Zhao, Xiaopeng Zhang</i>	
SIGraph: Saliency Image-Graph Network for Retinal Disease Classification in Fundus Image.....	10049
<i>Peng Zhang, Yuan Li, Haotian Song, Yankai Jiang, Yubo Tao, Hai Lin, Hongguang Cui</i>	
Visual Perturbation for Text-Based Person Search.....	10058
<i>Pengcheng Zhang, Xiaohan Yu, Xiao Bai, Jin Zheng</i>	
EDENet: Echo Direction Encoding Network for Place Recognition Based on Ground Penetrating Radar	10067
<i>Pengyu Zhang, Xieyuanli Chen, Yuwei Chen, Beizhen Bi, Zhuo Xu, Tian Jin, Xiaotao Huang, Liang Shen</i>	
View Transformation Robustness for Multi-View 3D Object Reconstruction with Reconstruction Error-Guided View Selection	10076
<i>Qi Zhang, Zhouhang Luo, Tao Yu, Hui Huang</i>	
Rethinking Pseudo-Label Guided Learning for Weakly Supervised Temporal Action Localization from the Perspective of Noise Correction	10085
<i>Quan Zhang, Yuxin Qi, Xi Tang, Rui Yuan, Xi Lin, Ke Zhang, Chun Yuan</i>	
Matching While Perceiving: Enhance Image Feature Matching with Applicable Semantic Amalgamation	10094
<i>Shihua Zhang, Zhenjie Zhu, Zizhuo Li, Tao Lu, Jiayi Ma</i>	
DiMSOD: A Diffusion-Based Framework for Multi-Modal Salient Object Detection	10103
<i>Shuo Zhang, Jiaming Huang, Wenbing Tang, Yan Wu, Terrence Hu, Xiaogang Xu, Jing Liu</i>	
Pragmatist: Multiview Conditional Diffusion Models for High-Fidelity 3D Reconstruction from Unposed Sparse Views	10112
<i>Songchun Zhang, Chunhui Zhao</i>	
Point Cloud Mamba: Point Cloud Learning via State Space Model.....	10121
<i>Tao Zhang, Haobo Yuan, Lu Qi, Jiangning Zhang, Qianyu Zhou, Shunping Ji, Shuicheng Yan, Xiangtai Li</i>	

DZAD: Diffusion-based Zero-shot Anomaly Detection.....	10131
<i>Tianrui Zhang, Liang Gao, Xinyu Li, Yiping Gao</i>	
Few-Shot Domain Adaptation for Learned Image Compression.....	10139
<i>Tianyu Zhang, Haotian Zhang, Yuqi Li, Li Li, Dong Liu</i>	
SVTformer: Spatial-View-Temporal Transformer for Multi-View 3D Human Pose Estimation	10148
<i>Wanruo Zhang, Mengyuan Liu, Hong Liu, Wenhao Li</i>	
Enhancing Implicit Neural Representations via Symmetric Power Transformation	10157
<i>WeiXiang Zhang, Shuzhao Xie, Chengwei Ren, Shijia Ge, Mingzi Wang, Zhi Wang</i>	
Bootstrapping Clustering of Gaussians for View-consistent 3D Scene Understanding.....	10166
<i>Wenbo Zhang, Lu Zhang, Ping Hu, Liqian Ma, Yunzhi Zhuge, Huchuan Lu</i>	
MTGA: Multi-View Temporal Granularity Aligned Aggregation for Event-Based Lip-Reading	10176
<i>Wenhao Zhang, Jun Wang, Yong Luo, Lei Yu, Wei Yu, Zheng He, Jialie Shen</i>	
Transtreaming: Adaptive Delay-aware Transformer for Real-time Streaming Perception	10185
<i>Xiang Zhang, Yufei Cui, Chenchen Fu, Zihao Wang, Yuyang Sun, Xue Liu, Weiwei Wu</i>	
Iterative Self-Training with Class-Aware Text-to-Image Synthesis for Visual Task Learning	10194
<i>Xiang Zhang, Wanqing Zhao, Pengyang Li, Ying Liu, Hangzai Luo, Sheng Zhong, Jinye Peng, Jianping Fan</i>	
Enhancing Multimodal Large Language Models Complex Reason via Similarity Computation.....	10203
<i>Xiaofeng Zhang, Fanshuo Zeng, Yihao Quan, Zheng Hui, Jiawei Yao</i>	
Toy-GS: Assembling Local Gaussians for Precisely Rendering Large-Scale Free Camera Trajectories	10212
<i>Xiaohan Zhang, Zhenyu Sun, Yukui Qiu, Junyan Su, Qi Liu</i>	
Beyond Spatial Domain: Cross-domain Promoted Fourier Convolution Helps Single Image Dehazing.....	10221
<i>Xiaozhe Zhang, Haidong Ding, Fengying Xie, Linpeng Pan, Yue Zi, Ke Wang, Haopeng Zhang</i>	
PhyCamo: A Robust Physical Camouflage via Contrastive Learning for Multi-View Physical Adversarial Attack	10230
<i>Ximin Zhang, Jinyin Chen, Haibin Zheng, Zhenguang Liu</i>	
CAMSIC: Content-aware Masked Image Modeling Transformer for Stereo Image Compression	10239
<i>Xinjie Zhang, Shenyuan Gao, Zhening Liu, Jiawei Shao, Xingtong Ge, Dailan He, Tongda Xu, Yan Wang, Jun Zhang</i>	
Pose Magic: Efficient and Temporally Consistent Human Pose Estimation with a Hybrid Mamba-GCN Network.....	10248
<i>Xinyi Zhang, Qiqi Bao, Qinpeng Cui, Wenming Yang, Qingmin Liao</i>	
SGDiff: Scene Graph Guided Diffusion Model for Image Collaborative SegCaptioning	10257
<i>Xu Zhang, Jin Yuan, Hanwang Zhang, Guojin Zhong, Yongsheng Zang, Jiacheng Lin, Zhiyong Li</i>	
VideoElevator: Elevating Video Generation Quality with Versatile Text-to-Image Diffusion Models	10266
<i>Yabo Zhang, Yuxiang Wei, Xianhui Lin, Zheng Hui, Peiran Ren, Xuansong Xie, Wangmeng Zuo</i>	

Track the Answer: Extending TextVQA from Image to Video with Spatio-Temporal Clues	10275
<i>Yan Zhang, Gangyan Zeng, Huawen Shen, Daiqing Wu, Yu Zhou, Can Ma</i>	
Category Prompt Mamba Network for Nuclei Segmentation and Classification	10284
<i>Ye Zhang, Zijie Fang, Yifeng Wang, Lingbo Zhang, Xianchao Guan, Yongbing Zhang</i>	
Improving Generalization of Universal Adversarial Perturbation via Dynamic Maximin Optimization	10293
<i>Yechao Zhang, Yingzhe Xu, Junyu Shi, Leo Yu Zhang, Shengshan Hu, Minghui Li, Yanjun Zhang</i>	
Cross-Modal Few-Shot Learning with Second-Order Neural Ordinary Differential Equations.....	10302
<i>Yi Zhang, Chun-Wun Cheng, Junyi He, Zhihai He, Carola-Bibiane Schönlieb, Yuyan Chen, Angelica I Aviles-Rivero</i>	
InstantSticker: Realistic Decal Blending via Disentangled Object Reconstruction.....	10311
<i>Yi Zhang, Xiaoyang Huang, Yishun Dou, Yue Shi, Rui Shi, Ye Chen, Bingbing Ni, Wenjun Zhang</i>	
DiffGrasp: Whole-Body Grasping Synthesis Guided by Object Motion Using a Diffusion Model.....	10320
<i>Yonghao Zhang, Qiang He, Yanguang Wan, Yinda Zhang, Xiaoming Deng, Cuixia Ma, Hongan Wang</i>	
Partial Point Cloud Registration with Multi-view 2D Image Learning	10329
<i>Yue Zhang, Yue Wu, Wenping Ma, Maoguo Gong, Hao Li, Biao Hou</i>	
RP-PGD: Boosting Segmentation Robustness with a Region-and-Prototype Based Adversarial Attack	10338
<i>Yuxuan Zhang, Zhenbo Shi, Shuchang Wang, Wei Yang, Shaowei Wang, Yinxing Xue</i>	
Stable-Hair: Real-World Hair Transfer via Diffusion Model	10348
<i>Yuxuan Zhang, Qing Zhang, Yiren Song, Jichao Zhang, Hao Tang, Jiaming Liu</i>	
Zero-Shot Learning in Industrial Scenarios: New Large-Scale Benchmark, Challenges and Baseline.....	10357
<i>Zekai Zhang, Qinghui Chen, Maomao Xiong, Shijiao Ding, Zhanzhi Su, Xinjie Yao, Yiming Sun, Cong Bai, Jinglin Zhang</i>	
Leveraging Consistent Spatio-Temporal Correspondence for Robust Visual Odometry.....	10367
<i>Zhaoxing Zhang, Junda Cheng, Gangwei Xu, Xiaoxiang Wang, Can Zhang, Xin Yang</i>	
Training-Free Image Manipulation Localization Using Diffusion Models	10376
<i>Zhenfei Zhang, Ming-Ching Chang, Xin Li</i>	
Multi-scale Activation, Refinement, and Aggregation: Exploring Diverse Cues for Fine-Grained Bird Recognition	10385
<i>Zhicheng Zhang, Hao Tang, Jinhui Tang</i>	
Single-view Image to Novel-view Generation for Hand-Object Interactions	10394
<i>Zhongqun Zhang, Yihua Cheng, Eduardo Pérez-Pellitero, Yiren Zhou, Jiankang Deng, Hyung Jin Chang, Jifei Song</i>	
Cross-PCR: A Robust Cross-Source Point Cloud Registration Framework	10403
<i>Guiyu Zhao, Zhentao Guo, Zewen Du, Hongbin Ma</i>	
DriveDreamer-2: LLM-Enhanced World Models for Diverse Driving Video Generation	10412
<i>Guosheng Zhao, Xiaofeng Wang, Zheng Zhu, Xinze Chen, Guan Huang, Xiaoyi Bao, Xingang Wang</i>	

Cobra: Extending Mamba to Multi-Modal Large Language Model for Efficient Inference	10421
<i>Han Zhao, Min Zhang, Wei Zhao, Pengxiang Ding, Siteng Huang, Donglin Wang</i>	
Adaptive Wavelet-Positional Encoding for High-Frequency Information Learning in Implicit Neural Representation	10430
<i>Hongxu Zhao, Zelin Gao, Yue Wang, Rong Xiong, Yu Zhang</i>	
MagicNaming: Consistent Identity Generation by Finding a “Name Space” in T2I Diffusion Models	10439
<i>Jing Zhao, Heliang Zheng, Chaoyue Wang, Long Lan, Wanrong Huang, Yuhua Tang</i>	
Multimodal Class-aware Semantic Enhancement Network for Audio-Visual Video Parsing	10448
<i>Pengcheng Zhao, Jinxing Zhou, Yang Zhao, Dan Guo, Yanxiang Chen</i>	
Excluding the Impossible for Open Vocabulary Semantic Segmentation.....	10457
<i>Shiyuan Zhao, Baodi Liu, Yu Bai, Weifeng Liu, Shuai Shao</i>	
KALAHASH: Knowledge-Anchored Low-Resource Adaptation for Deep Hashing	10465
<i>Shu Zhao, Tan Yu, Xiaoshuai Hao, Wenchao Ma, Vijaykrishnan Narayanan</i>	
Training-free Open-Vocabulary Semantic Segmentation via Diverse Prototype Construction and Sub-region Matching	10474
<i>Xuanpu Zhao, Dianmo Sheng, Zhentao Tan, Zhiwei Zhao, Tao Gong, Qi Chu, Bin Liu, Nenghai Yu</i>	
Audio-Visual Adaptive Fusion Network for Question Answering Based on Contrastive Learning	10483
<i>Xujian Zhao, Yixin Wang, Peiquan Jin</i>	
Local Conditional Controlling for Text-to-Image Diffusion Models.....	10492
<i>Yibo Zhao, Liang Peng, Yang Yang, Zekai Luo, Hengjia Li, Yao Chen, Zheng Yang, Xiaofei He, Wei Zhao, Qinglin Lu, Wei Liu, Boxi Wu</i>	
Hierarchical Cross-Modal Alignment for Open-Vocabulary 3D Object Detection	10501
<i>Youjun Zhao, Jiaying Lin, Rynson W. H. Lau</i>	
ESEG: Event-Based Segmentation Boosted by Explicit Edge-Semantic Guidance.....	10510
<i>Yucheng Zhao, Gengyu Lyu, Ke Li, Zihao Wang, Hao Chen, Zhen Yang, Yongjian Deng</i>	
NightReID: A Large-Scale Nighttime Person Re-Identification Benchmark	10519
<i>Yuxuan Zhao, Weijian Ruan, He Li, Mang Ye</i>	
HFF-Tracker: A Hierarchical Fine-grained Fusion Tracker for Referring Multi-Object Tracking	10528
<i>Zeyong Zhao, Yanchao Hao, Minghao Zhang, Qingbin Liu, Bo Li, Dianbo Sui, Shizhu He, Xi Chen</i>	
Heterogeneous Prompt-Guided Entity Inferring and Distilling for Scene-Text Aware Cross-Modal Retrieval	10537
<i>Zhiqian Zhao, Liang Li, Jiehua Zhang, Yaoqi Sun, Xichun Sheng, Haibing Yin, Shaowei Jiang</i>	
Multi-concept Model Immunization through Differentiable Model Merging	10546
<i>Amber Yijia Zheng, Raymond A. Yeh</i>	

PART 2

PHR-DIFF: Portrait Highlights Removal via Patch-aware Diffusion Model	10555
<i>Hongsheng Zheng, Zhongyun Bao, Gang Fu, Xuze Jiao, Chunxia Xiao</i>	

Breaking Information Isolation: Accelerating MRI via Inter-sequence Mapping and Progressive Masking	10564
<i>Jianwei Zheng, Xiaomin Yao, Guojiang Shen, Wei Li, Jiawei Jiang</i>	
Supportive Negatives Spectral Augmentation for Source-Free Cross-Domain Segmentation	10573
<i>Kexin Zheng, Haifeng Xia, Siyu Xia, Ming Shao, Zhengming Ding</i>	
Anti-Diffusion: Preventing Abuse of Modifications of Diffusion-Based Models	10582
<i>Li Zheng, Liangbin Xie, Jiantao Zhou, Xintao Wang, Haiwei Wu, Jinyu Tian</i>	
Seg2Box: 3D Object Detection by Point-Wise Semantics Supervision	10591
<i>Maoji Zheng, Ziyu Xu, Qiming Xia, Hai Wu, Chenglu Wen, Cheng Wang</i>	
When Shadow Removal Meets Intrinsic Image Decomposition: A Joint Learning Framework Using Unpaired Data.....	10599
<i>Rongjia Zheng, Qing Zhang, Yongwei Nie, Wei-Shi Zheng</i>	
A New Adversarial Perspective for LiDAR-based 3D Object Detection	10608
<i>Shijun Zheng, Weiquan Liu, Yu Guo, Yu Zang, Siqi Shen, Cheng Wang</i>	
ContextualStory: Consistent Visual Storytelling with Spatially-Enhanced and Storyline Context	10617
<i>Sixiao Zheng, Yanwei Fu</i>	
OODML: Whole Slide Image Classification Meets Online Pseudo-Supervision and Dynamic Mutual Learning	10626
<i>Tingting Zheng, Kui Jiang, Hongxun Yao, Yi Xiao, Zhongyuan Wang</i>	
Decoupled Spatio-Temporal Consistency Learning for Self-Supervised Tracking	10635
<i>Yaozong Zheng, Bineng Zhong, Qihua Liang, Ning Li, Shuxiang Song</i>	
Universal Domain Adaptive Object Detection via Dual Probabilistic Alignment	10644
<i>Yuanfan Zheng, Jinlin Wu, Wuyang Li, Zhen Chen</i>	
MMPF: Multi-Modal Perception Framework for Abnormal Medical Condition Detection.....	10653
<i>Chuyi Zhong, Dingkan Yang, Peng Zhai, Lihua Zhang</i>	
DECIDER: Difference-aware Contrastive Diffusion Model with Adversarial Perturbations for Image Change Captioning	10662
<i>Guojin Zhong, Jinhong Hu, Jiajun Chen, Jin Yuan, Wenbo Pan</i>	
IPVTOn: Image-based 3D Virtual Try-on with Image Prompt Adapter	10671
<i>Xiaojing Zhong, Zhonghua Wu, Xiaofeng Yang, Guosheng Lin, Qingyao Wu</i>	
VioTGPT: Learning to Schedule Vision Tools Towards Intelligent Video Internet of Things	10680
<i>Yaoyao Zhong, Mengshi Qi, Rui Wang, Yuhao Qiu, Yang Zhang, Huadong Ma</i>	
PointCFormer: A Relation-Based Progressive Feature Extraction Network for Point Cloud Completion	10689
<i>Yi Zhong, Weize Quan, Dong-Ming Yan, Jie Jiang, Yingmei Wei</i>	
SlerpFace: Face Template Protection via Spherical Linear Interpolation.....	10698
<i>Zhizhou Zhong, Yuxi Mi, Yuge Huang, Jianqing Xu, Guodong Mu, Shouhong Ding, Jingyun Zhang, Rizen Guo, Yunsheng Wu, Shuigeng Zhou</i>	

UrBench: A Comprehensive Benchmark for Evaluating Large Multimodal Models in Multi-View Urban Scenarios	10707
<i>Baichuan Zhou, Haote Yang, Dairong Chen, Junyan Ye, Tianyi Bai, Jinhua Yu, Songyang Zhang, Dahua Lin, Conghui He, Weijia Li</i>	
Polarization Guided Mask-Free Shadow Removal.....	10716
<i>Chu Zhou, Chao Xu, Boxin Shi</i>	
Controllable Distortion-Perception Tradeoff Through Latent Diffusion for Neural Image Compression.....	10725
<i>Chuan Zhou, Guo Lu, Jiangchuan Li, Xiangyu Chen, Zhengxue Cheng, Li Song, Wenjun Zhang</i>	
Position-Aware Guided Point Cloud Completion with CLIP Model.....	10734
<i>Feng Zhou, Qi Zhang, Ju Dai, Lei Li, Qing Fan, Junliang Xing</i>	
TrackGo: A Flexible and Efficient Method for Controllable Video Generation.....	10743
<i>Haitao Zhou, Chuang Wang, Rui Nie, Jinlin Liu, Dongdong Yu, Qian Yu, Changhu Wang</i>	
Low-Light Image Enhancement via Generative Perceptual Priors.....	10752
<i>Han Zhou, Wei Dong, Xiaohong Liu, Yulun Zhang, Guangtao Zhai, Jun Chen</i>	
Video Anomaly Detection with Motion and Appearance Guided Patch Diffusion Model	10761
<i>Hang Zhou, Jiale Cai, Yuteng Ye, Yonghui Feng, Chenxing Gao, Junqing Yu, Zikai Song, Wei Yang</i>	
Core-to-Global Reasoning for Compositional Visual Question Answering	10770
<i>Hao Zhou, Tingjin Luo, Zhangqi Jiang</i>	
Joint Class-level and Instance-level Relationship Modeling for Novel Class Discovery.....	10779
<i>Jiaying Zhou, Qingchao Chen</i>	
GaussianPainter: Painting Point Cloud into 3D Gaussians with Normal Guidance	10788
<i>Jingqiu Zhou, Lue Fan, Xuesong Chen, Linjiang Huang, Si Liu, Hongsheng Li</i>	
All-in-One: Transferring Vision Foundation Models into Stereo Matching.....	10797
<i>Jingyi Zhou, Haoyu Zhang, Jiakang Yuan, Peng Ye, Tao Chen, Hao Jiang, Meiya Chen, Yangyang Zhang</i>	
SceneX: Procedural Controllable Large-Scale Scene Generation	10806
<i>Mengqi Zhou, Yuxi Wang, Jun Hou, Shougao Zhang, Yiwei Li, Chuanchen Luo, Junran Peng, Zhaoxiang Zhang</i>	
GLIC: General Format Learned Image Compression.....	10815
<i>MingSheng Zhou, MingMing Kong</i>	
Mitigating Feature Gap for Adversarial Robustness by Feature Disentanglement.....	10825
<i>Nuoyan Zhou, Dawei Zhou, Decheng Liu, Nannan Wang, Xinbo Gao</i>	
Spatiotemporal-Aware Neural Fields for Dynamic CT Reconstruction	10834
<i>Qinyang Zhou, Yunfan Ye, Zhiping Cai</i>	
3DMambaIPF: A State Space Model for Iterative Point Cloud Filtering via Differentiable Rendering	10843
<i>Qingyuan Zhou, Weidong Yang, Ben Fei, Jingyi Xu, Rui Zhang, Keyi Liu, Yeqi Luo, Ying He</i>	

Test-Time Adaptation on Noisy Data via Model-Pruning-Based Filtering and Flatness-Aware Entropy Minimization	10852
<i>Xingzhi Zhou, Zhiliang Tian, Boyang Zhang, Yibo Zhang, Ka Chun Cheung, Simon See, Hao Yang, Yun Zhou, Nevin L. Zhang</i>	
Dynamic Contrastive Knowledge Distillation for Efficient Image Restoration	10861
<i>Yunshuai Zhou, Junbo Qiao, Jincheng Liao, Wei Li, Simiao Li, Jiao Xie, Yunhang Shen, Jie Hu, Shaohui Lin</i>	
Improving Generalization of Deep Neural Networks by Optimum Shifting.....	10870
<i>Yuyan Zhou, Ye Li, Lei Feng, Sheng-Jun Huang</i>	
Achieving Ensemble-Like Performance in a Single Model: A Feature Diversification Framework for Image-Text Matching.....	10879
<i>Zhao Zhou, Yiqun Wang, Weizhong Zhang, Yingbin Zheng, Xiangcheng Du, Cheng Jin</i>	
Expanding the Scope of Negatives: Boosting Image-Text Matching with Negatives Distribution Guided Learning.....	10887
<i>Zhao Zhou, Weizhong Zhang, Xiangcheng Du, Yingbin Zheng, Cheng Jin</i>	
An Exemplar-based Framework for Chinese Text Recognition	10896
<i>Zhao Zhou, Xiangcheng Du, Yingbin Zheng, Xingjiao Wu, Cheng Jin</i>	
Dense Audio-Visual Event Localization Under Cross-Modal Consistency and Multi-Temporal Granularity Collaboration.....	10905
<i>Ziheng Zhou, Jinxing Zhou, Wei Qian, Shengeng Tang, Xiaojun Chang, Dan Guo</i>	
GoHD: Gaze-oriented and Highly Disentangled Portrait Animation with Rhythmic Poses and Realistic Expressions.....	10914
<i>Ziqi Zhou, Weize Quan, Hailin Shi, Wei Li, Lili Wang, Dong-Ming Yan</i>	
MultiBooth: Towards Generating All Your Concepts in an Image from Text	10923
<i>Chenyang Zhu, Kai Li, Yue Ma, Chunming He, Xiu Li</i>	
Adaptive Prototype Replay for Class Incremental Semantic Segmentation.....	10932
<i>Guilin Zhu, Dongyue Wu, Changxin Gao, Runmin Wang, Weidong Yang, Nong Sang</i>	
Optimizing Label Assignment for Weakly Supervised Person Search	10941
<i>Haiyang Zhu, Xi Yang, Nannan Wang</i>	
TAMER: Tree-Aware Transformer for Handwritten Mathematical Expression Recognition.....	10950
<i>Jianhua Zhu, Wenqi Zhao, Yu Li, Xingjian Hu, Liangcai Gao</i>	
Two-stream Beats One-stream: Asymmetric Siamese Network for Efficient Visual Tracking	10959
<i>Jiawen Zhu, Huayi Tang, Xin Chen, Xinying Wang, Dong Wang, Huchuan Lu</i>	
A Lottery Ticket Hypothesis Approach with Sparse Fine-tuning and MAE for Image Forgery Detection and Localization.....	10968
<i>Jiaying Zhu, Dong Li, Xueyang Fu, Gege Shi, Jie Xiao, Aiping Liu, Zheng-Jun Zha</i>	
Large Images Are Gaussians: High-Quality Large Image Representation with Levels of 2D Gaussian Splatting.....	10977
<i>Lingting Zhu, Guying Lin, Jinnan Chen, Xinjie Zhang, Zhenchao Jin, Zhao Wang, Lequan Yu</i>	
A Comprehensive Overhaul of Multimodal Assistant with Small Language Models	10986
<i>Minjie Zhu, Yichen Zhu, Ning Liu, Xin Liu, Zhiyuan Xu, Chaomin Shen, Yaxin Peng</i>	

Thin-Plate Spline-based Interpolation for Animation Line Inbetweening	10995
<i>Tianyi Zhu, Wei Shang, Dongwei Ren</i>	
Mesh Watermark Removal Attack and Mitigation: A Novel Perspective of Function Space	11004
<i>Xingyu Zhu, Guanhui Ye, Chengdong Dong, Xiapu Luo, Shiyao Zhang, Xuetao Wei</i>	
DreaMark: Rooting Watermark in Score Distillation Sampling Generated Neural Radiance Fields	11013
<i>Xingyu Zhu, Xiapu Luo, Xuetao Wei</i>	
Mesoscopic Insights: Orchestrating Multi-Scale & Hybrid Architecture for Image Manipulation Localization	11022
<i>Xuekang Zhu, Xiaochen Ma, Lei Su, Zhuohang Jiang, Bo Du, Xiwen Wang, Zeyu Lei, Wentao Feng, Chi-Man Pun, Ji-Zhe Zhou</i>	
Exploring Task-Level Optimal Prompts for Visual In-Context Learning	11031
<i>Yan Zhu, Huan Ma, Changqing Zhang</i>	
MUC: Mixture of Uncalibrated Cameras for Robust 3D Human Body Reconstruction	11040
<i>Yitao Zhu, Sheng Wang, Mengjie Xu, Zixu Zhuang, Zhixin Wang, Kaidong Wang, Han Zhang, Qian Wang</i>	
ST3: Accelerating Multimodal Large Language Model by Spatial-Temporal Visual Token Trimming	11049
<i>Jiedong Zhuang, Lu Lu, Ming Dai, Rui Hu, Jian Chen, Qiang Liu, Haoji Hu</i>	
Dynamic Entity-Masked Graph Diffusion Model for Histopathology Image Representation Learning	11058
<i>Zhenfeng Zhuang, Min Cen, Yanfeng Li, Fangyu Zhou, Lequan Yu, Baptiste Magnier, Liansheng Wang</i>	
CoCoCo: Improving Text-Guided Video Inpainting for Better Consistency, Controllability and Compatibility	11067
<i>Bojia Zi, Shihao Zhao, Xianbiao Qi, Jianan Wang, Yukai Shi, Qianyu Chen, Bin Liang, Rong Xiao, Kam-Fai Wong, Lei Zhang</i>	
RhythmMamba: Fast, Lightweight, and Accurate Remote Physiological Measurement	11077
<i>Bochao Zou, Zizheng Guo, Xiaocheng Hu, Huimin Ma</i>	
AdaCo: Overcoming Visual Foundation Model Noise in 3D Semantic Segmentation via Adaptive Label Correction	11086
<i>Pufan Zou, Shijia Zhao, Weijie Huang, Qiming Xia, Chenglu Wen, Wei Li, Cheng Wang</i>	
L-Man: A Large Multi-modal Model Unifying Human-centric Tasks	11095
<i>Jialong Zuo, Ying Nie, Tianyu Guo, Huaxin Zhang, Jiahao Hong, Nong Sang, Changxin Gao, Kai Han</i>	

Author Index