

# **38th AAAI Conference on Artificial Intelligence (AAAI-24), 36th Conference on Innovative Applications of Artificial Intelligence (IAAI-24), 14th Symposium on Educational Advances in Artificial Intelligence (EAAI-24)**

## **Volume 7: AAAI Technical Tracks**

- Computer Vision VI

Vancouver, Canada  
20 – 27 February 2024

Part 1 of 2

ISBN: 979-8-3313-0089-0

**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© (2024) by Association for the Advancement of Artificial Intelligence  
All rights reserved.

Printed with permission by Curran Associates, Inc. (2024)

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](mailto:curran@proceedings.com)  
Web: [www.proceedings.com](http://www.proceedings.com)

# TABLE OF CONTENTS

## PART 1

### AAAI TECHNICAL TRACK ON COMPUTER VISION VI

Hybrid-SORT: Weak Cues Matter for Online Multi-Object Tracking .....	6504
<i>Mingzhan Yang, Guangxin Han, Bin Yan, Wenhua Zhang, Jinqing Qi, Huchuan Lu, Dong Wang</i>	
Multi-Modal Prompting for Open-Vocabulary Video Visual Relationship Detection .....	6513
<i>Shuo Yang, Yongqi Wang, Xiaofeng Ji, Xinxiao Wu</i>	
Learning Dense Correspondence for NeRF-Based Face Reenactment .....	6522
<i>Songlin Yang, Wei Wang, Yushi Lan, Xiangyu Fan, Bo Peng, Lei Yang, Jing Dong</i>	
Motion Deblurring Via Spatial-Temporal Collaboration of Frames and Events.....	6531
<i>Wen Yang, Jinjian Wu, Jupao Ma, Leida Li, Guangming Shi</i>	
DGL: Dynamic Global-Local Prompt Tuning for Text-Video Retrieval .....	6540
<i>Xiangpeng Yang, Linchao Zhu, Xiaohan Wang, Yi Yang</i>	
Diverse and Stable 2D Diffusion Guided Text to 3D Generation with Noise Recalibration .....	6549
<i>Xiaofeng Yang, Fayao Liu, Yi Xu, Hanjing Su, Qingyao Wu, Guosheng Lin</i>	
Semantic Segmentation in Multiple Adverse Weather Conditions with Domain Knowledge Retention .....	6558
<i>Xin Yang, Wending Yan, Yuan Yuan, Michael Bi Mi, Robby T. Tan</i>	
Hyperspectral Image Reconstruction Via Combinatorial Embedding of Cross-Channel Spatio-Spectral Clues.....	6567
<i>Xingxing Yang, Jie Chen, Zaifeng Yang</i>	
Decomposing Semantic Shifts for Composed Image Retrieval.....	6576
<i>Xingyu Yang, Daqing Liu, Heng Zhang, Yong Luo, Chaoyue Wang, Jing Zhang</i>	
Gaze Target Detection by Merging Human Attention and Activity Cues .....	6585
<i>Yaokun Yang, Yihan Yin, Feng Lu</i>	
PM-INR: Prior-Rich Multi-Modal Implicit Large-Scale Scene Neural Representation.....	6594
<i>Yiyang Yang, Fukun Yin, Wen Liu, Jiayuan Fan, Xin Chen, Gang Yu, Tao Chen</i>	
FontDiffuser: One-Shot Font Generation Via Denoising Diffusion with Multi-Scale Content Aggregation and Style Contrastive Learning .....	6603
<i>Zhenhua Yang, Dezhi Peng, Yuxin Kong, Yuyi Zhang, Cong Yao, Lianwen Jin</i>	
Full-Body Motion Reconstruction with Sparse Sensing from Graph Perspective.....	6612
<i>Feiyu Yao, Zongkai Wu, Li Yi</i>	
FoSp: Focus and Separation Network for Early Smoke Segmentation .....	6621
<i>Lujian Yao, Haitao Zhao, Jingchao Peng, Zhongze Wang, Kaijie Zhao</i>	

How to Evaluate the Generalization of Detection? a Benchmark for Comprehensive Open-Vocabulary Detection .....	6630
<i>Yiyang Yao, Peng Liu, Tiancheng Zhao, Qianqian Zhang, Jiajia Liao, Chunxin Fang, Kyusong Lee, Qing Wang</i>	
Diverse and Aligned Audio-To-Video Generation Via Text-To-Video Model Adaptation .....	6639
<i>Guy Yariv, Itai Gat, Sagie Benaim, Lior Wolf, Idan Schwartz, Yossi Adi</i>	
AltDiffusion: A Multilingual Text-To-Image Diffusion Model .....	6648
<i>Fulong Ye, Guang Liu, Xinya Wu, Ledell Wu</i>	
Mutual-Modality Adversarial Attack with Semantic Perturbation .....	6657
<i>Jingwen Ye, Ruonan Yu, Songhua Liu, Xinchao Wang</i>	
STDiff: Spatio-Temporal Diffusion for Continuous Stochastic Video Prediction .....	6666
<i>Xi Ye, Guillaume-Alexandre Bilodeau</i>	
DiffusionEdge: Diffusion Probabilistic Model for Crisp Edge Detection .....	6675
<i>Yunfan Ye, Kai Xu, Yuhang Huang, Renjiao Yi, Zhiping Cai</i>	
Dynamic Feature Pruning and Consolidation for Occluded Person Re-Identification .....	6684
<i>YuTeng Ye, Hang Zhou, Jiale Cai, Chenxing Gao, Youjia Zhang, Junle Wang, Qiang Hu, Junqing Yu, Wei Yang</i>	
Progressive Text-To-Image Diffusion with Soft Latent Direction .....	6693
<i>YuTeng Ye, Jiale Cai, Hang Zhou, Guanwen Li, Youjia Zhang, Zikai Song, Chenxing Gao, Junqing Yu, Wei Yang</i>	
UCMCTrack: Multi-Object Tracking with Uniform Camera Motion Compensation .....	6702
<i>Kefu Yi, Kai Luo, Xiaolei Luo, Jiangui Huang, Hao Wu, Rongdong Hu, Wei Hao</i>	
DiffRAW: Leveraging Diffusion Model to Generate DSLR-Comparable Perceptual Quality sRGB from Smartphone RAW Images .....	6711
<i>Mingxin Yi, Kai Zhang, Pei Liu, Tanli Zuo, Jingduo Tian</i>	
Efficient Look-Up Table from Expanded Convolutional Network for Accelerating Image Super-Resolution .....	6720
<i>Kai Yin, Jie Shen</i>	
CLIP-Gaze: Towards General Gaze Estimation Via Visual-Linguistic Model .....	6729
<i>Pengwei Yin, Guanzhong Zeng, Jingjing Wang, Di Xie</i>	
Point Deformable Network with Enhanced Normal Embedding for Point Cloud Analysis .....	6738
<i>Xingyilang Yin, Xi Yang, Liangchen Liu, Nannan Wang, Xinbo Gao</i>	
Revisiting Open-Set Panoptic Segmentation .....	6747
<i>Yufei Yin, Hao Chen, Wengang Zhou, Jiajun Deng, Haiming Xu, Houqiang Li</i>	
VQAttack: Transferable Adversarial Attacks on Visual Question Answering Via Pre-Trained Models .....	6755
<i>Ziyi Yin, Muchao Ye, Tianrong Zhang, Jiaqi Wang, Han Liu, Jinghui Chen, Ting Wang, Fenglong Ma</i>	
TF-CLIP: Learning Text-Free CLIP for Video-Based Person Re-Identification .....	6764
<i>Chenyang Yu, Xuehu Liu, Yingquan Wang, Pingping Zhang, Huchuan Lu</i>	

MM-Point: Multi-View Information-Enhanced Multi-Modal Self-Supervised 3D Point Cloud Understanding .....	6773
<i>Hai-Tao Yu, Mofei Song</i>	
Spatial Transform Decoupling for Oriented Object Detection .....	6782
<i>Hongtian Yu, Yunjie Tian, Qixiang Ye, Yunfan Liu</i>	
Step Vulnerability Guided Mean Fluctuation Adversarial Attack Against Conditional Diffusion Models .....	6791
<i>Hongwei Yu, Jiansheng Chen, Xinlong Ding, Yudong Zhang, Ting Tang, Huimin Ma</i>	
PaintHuman: Towards High-Fidelity Text-To-3D Human Texturing Via Denoised Score Distillation .....	6800
<i>Jianhui Yu, Hao Zhu, Liming Jiang, Chen Change Loy, Weidong Cai, Wayne Wu</i>	
CatFormer: Category-Level 6D Object Pose Estimation with Transformer .....	6808
<i>Sheng Yu, Di-Hua Zhai, Yuanqing Xia</i>	
DME: Unveiling the Bias for Better Generalized Monocular Depth Estimation .....	6817
<i>Songsong Yu, Yifan Wang, Yunzhi Zhuge, Lijun Wang, Huchuan Lu</i>	
DOCTR: Disentangled Object-Centric Transformer for Point Scene Understanding .....	6826
<i>Xiaoxuan Yu, Hao Wang, Weiming Li, Qiang Wang, Soonyong Cho, Younghun Sung</i>	
Discretization-Induced Dirichlet Posterior for Robust Uncertainty Quantification on Regression .....	6835
<i>Xuanlong Yu, Gianni Franchi, Jindong Gu, Emanuel Aldea</i>	
Attacks on Continual Semantic Segmentation by Perturbing Incremental Samples .....	6844
<i>Zhidong Yu, Wei Yang, Xike Xie, Zhenbo Shi</i>	
Data-Free Hard-Label Robustness Stealing Attack .....	6853
<i>Xiaojian Yuan, Kejian Chen, Wen Huang, Jie Zhang, Weiming Zhang, Nenghai Yu</i>	
Efficient Conditional Diffusion Model with Probability Flow Sampling for Image Super-Resolution.....	6862
<i>Yutao Yuan, Chun Yuan</i>	
SD-MVS: Segmentation-Driven Deformation Multi-View Stereo with Spherical Refinement and EM Optimization.....	6871
<i>Zhenlong Yuan, Jiakai Cao, Zhaoxin Li, Hao Jiang, Zhaoqi Wang</i>	
KeDuSR: Real-World Dual-Lens Super-Resolution Via Kernel-Free Matching.....	6881
<i>Huanjing Yue, Zifan Cui, Kun Li, Jingyu Yang</i>	
SurgicalSAM: Efficient Class Promptable Surgical Instrument Segmentation .....	6890
<i>Wenxi Yue, Jing Zhang, Kun Hu, Yong Xia, Jiebo Luo, Zhiyong Wang</i>	
Unveiling Details in the Dark: Simultaneous Brightening and Zooming for Low-Light Image Enhancement .....	6899
<i>Ziyu Yue, Jiaxin Gao, Zhixun Su</i>	
Weakly-Supervised Temporal Action Localization by Inferring Salient Snippet-Feature .....	6908
<i>Wulian Yun, Mengshi Qi, Chuanming Wang, Huadong Ma</i>	
Behavioral Recognition of Skeletal Data Based on Targeted Dual Fusion Strategy .....	6917
<i>Xiao Yun, Chenglong Xu, Kevin Riou, Kaiwen Dong, Yanjing Sun, Song Li, Kevin Subrin, Patrick Le Callet</i>	

Zero-Shot Aerial Object Detection with Visual Description Regularization .....	6926
<i>Zhengqing Zang, Chenyu Lin, Chenwei Tang, Tao Wang, Jiancheng Lv</i>	
Controllable Mind Visual Diffusion Model.....	6935
<i>Bohan Zeng, Shanglin Li, Xuhui Liu, Sicheng Gao, Xiaolong Jiang, Xu Tang, Yao Hu, Jianzhuang Liu, Baochang Zhang</i>	
MGQFormer: Mask-Guided Query-Based Transformer for Image Manipulation Localization.....	6944
<i>Kunlun Zeng, Ri Cheng, Weimin Tan, Bo Yan</i>	
Weakly-Supervised Mirror Detection Via Scribble Annotations .....	6953
<i>Mingfeng Zha, Yunqiang Pei, Guoqing Wang, Tianyu Li, Yang Yang, Wenbin Qian, Heng Tao Shen</i>	
Towards Compact 3D Representations Via Point Feature Enhancement Masked Autoencoders .....	6962
<i>Yaohua Zha, Huizhen Ji, Jinmin Li, Rongsheng Li, Tao Dai, Bin Chen, Zhi Wang, Shu-Tao Xia</i>	
Fine-Grained Knowledge Selection and Restoration for Non-Exemplar Class Incremental Learning.....	6971
<i>Jiang-Tian Zhai, Xialei Liu, Lu Yu, Ming-Ming Cheng</i>	
Multi-Prompts Learning with Cross-Modal Alignment for Attribute-Based Person Re-Identification .....	6979
<i>Yajing Zhai, Yawen Zeng, Zhiyong Huang, Zheng Qin, Xin Jin, Da Cao</i>	
Mono3DVG: 3D Visual Grounding in Monocular Images.....	6988
<i>Yang Zhan, Yuan Yuan, Zhitong Xiong</i>	
Amodal Scene Analysis Via Holistic Occlusion Relation Inference and Generative Mask Completion .....	6997
<i>Bowen Zhang, Qing Liu, Jianming Zhang, Yilin Wang, Liyang Liu, Zhe Lin, Yifan Liu</i>	
High-Quality Real-Time Rendering Using Subpixel Sampling Reconstruction.....	7006
<i>Boyu Zhang, Hongliang Yuan</i>	
Weakly Supervised Few-Shot Object Detection with DETR .....	7015
<i>Chenbo Zhang, Yinglu Zhang, Lu Zhang, Jiajia Zhao, Jihong Guan, Shuigeng Zhou</i>	
S2WAT: Image Style Transfer Via Hierarchical Vision Transformer Using Strips Window Attention .....	7024
<i>Chiyu Zhang, Xiaogang Xu, Lei Wang, Zaiyan Dai, Jun Yang</i>	
Synergistic Multiscale Detail Refinement Via Intrinsic Supervision for Underwater Image Enhancement .....	7033
<i>Dehuan Zhang, Jingchun Zhou, Chunle Guo, Weishi Zhang, Chongyi Li</i>	
W2P: Switching from Weak Supervision to Partial Supervision for Semantic Segmentation.....	7042
<i>Fangyuan Zhang, Tianxiang Pan, Jun-Hai Yong, Bin Wang</i>	
HyperEditor: Achieving Both Authenticity and Cross-Domain Capability in Image Editing Via Hypernetworks .....	7051
<i>Hai Zhang, Chunwei Wu, Guitao Cao, Hailing Wang, Wenming Cao</i>	
RadOcc: Learning Cross-Modality Occupancy Knowledge Through Rendering Assisted Distillation .....	7060
<i>Haiming Zhang, Xu Yan, Dongfeng Bai, Jiantao Gao, Pan Wang, Bingbing Liu, Shuguang Cui, Zhen Li</i>	
GSDD: Generative Space Dataset Distillation for Image Super-Resolution .....	7069
<i>Haiyu Zhang, Shaolin Su, Yu Zhu, Jinqiu Sun, Yanning Zhang</i>	

CSL: Class-Agnostic Structure-Constrained Learning for Segmentation Including the Unseen .....	7078
<i>Hao Zhang, Fang Li, Lu Qi, Ming-Hsuan Yang, Narendra Ahuja</i>	
A Robust Mutual-Reinforcing Framework for 3D Multi-Modal Medical Image Fusion Based on Visual-Semantic Consistency .....	7087
<i>Hao Zhang, Xuhui Zuo, Huabing Zhou, Tao Lu, Jiayi Ma</i>	
Learning Task-Aware Language-Image Representation for Class-Incremental Object Detection .....	7096
<i>Hongquan Zhang, Bin-Bin Gao, Yi Zeng, Xudong Tian, Xin Tan, Zhizhong Zhang, Yanyun Qu, Jun Liu, Yuan Xie</i>	
Identification of Necessary Semantic Undertakers in the Causal View for Image-Text Matching.....	7105
<i>Huatian Zhang, Lei Zhang, Kun Zhang, Zhendong Mao</i>	
HR-Pro: Point-Supervised Temporal Action Localization Via Hierarchical Reliability Propagation.....	7115
<i>Huaxin Zhang, Xiang Wang, Xiaohao Xu, Zhiwu Qing, Changxin Gao, Nong Sang</i>	
AvatarVerse: High-Quality & Stable 3D Avatar Creation from Text and Pose.....	7124
<i>Huichao Zhang, Bowen Chen, Hao Yang, Liao Qu, Xu Wang, Li Chen, Chao Long, Feida Zhu, Daniel Du, Min Zheng</i>	
Improving the Adversarial Transferability of Vision Transformers with Virtual Dense Connection.....	7133
<i>Jianping Zhang, Yizhan Huang, Zhuoer Xu, Weibin Wu, Michael R. Lyu</i>	
Curvature-Invariant Adversarial Attacks for 3D Point Clouds .....	7142
<i>Jianping Zhang, Wenwei Gu, Yizhan Huang, Zhihan Jiang, Weibin Wu, Michael R. Lyu</i>	
Cross-Modal Feature Distribution Calibration for Few-Shot Visual Question Answering.....	7151
<i>Jing Zhang, Xiaoqiang Liu, Mingzhe Chen, Zhe Wang</i>	
Robust 3D Tracking with Quality-Aware Shape Completion.....	7160
<i>Jingwen Zhang, Zikun Zhou, Guangming Lu, Jiandong Tian, Wenjie Pei</i>	
Neighborhood-Enhanced 3D Human Pose Estimation with Monocular LiDAR in Long-Range Outdoor Scenes.....	7169
<i>Jingyi Zhang, Qihong Mao, Guosheng Hu, Siqu Shen, Cheng Wang</i>	
NeRF-LiDAR: Generating Realistic LiDAR Point Clouds with Neural Radiance Fields.....	7178
<i>Junge Zhang, Feihu Zhang, Shaochen Kuang, Li Zhang</i>	
Point Cloud Part Editing: Segmentation, Generation, Assembly, and Selection .....	7187
<i>Kaiyi Zhang, Yang Chen, Ximing Yang, Weizhong Zhang, Cheng Jin</i>	
CatmullRom Splines-Based Regression for Image Forgery Localization .....	7196
<i>Li Zhang, Mingliang Xu, Dong Li, Jianming Du, Rujing Wang</i>	
Deep Semantic Graph Transformer for Multi-View 3D Human Pose Estimation.....	7205
<i>Lijun Zhang, Kangkang Zhou, Feng Lu, Xiang-Dong Zhou, Yu Shi</i>	

## PART 2

Brush Your Text: Synthesize Any Scene Text on Images Via Diffusion Model .....	7215
<i>Lingjun Zhang, Xinyuan Chen, Yaohui Wang, Yue Lu, Yu Qiao</i>	

IRPruneDet: Efficient Infrared Small Target Detection Via Wavelet Structure-Regularized Soft Channel Pruning .....	7224
<i>Mingjin Zhang, Handi Yang, Jie Guo, Yunsong Li, Xinbo Gao, Jing Zhang</i>	
M2Doc: A Multi-Modal Fusion Approach for Document Layout Analysis .....	7233
<i>Ning Zhang, Hiuyi Cheng, Jiayu Chen, Zongyuan Jiang, Jun Huang, Yang Xue, Lianwen Jin</i>	
Multi-View People Detection in Large Scenes Via Supervised View-Wise Contribution Weighting.....	7242
<i>Qi Zhang, Yunfei Gong, Daijie Chen, Antoni B. Chan, Hui Huang</i>	
Aligning Geometric Spatial Layout in Cross-View Geo-Localization Via Feature Recombination.....	7251
<i>Qingwang Zhang, Yingying Zhu</i>	
MobileInst: Video Instance Segmentation on the Mobile.....	7260
<i>Renhong Zhang, Tianheng Cheng, Shusheng Yang, Haoyi Jiang, Shuai Zhang, Jiancheng Lyu, Xin Li, Xiaowen Ying, Dashan Gao, Wenyu Liu, Xinggong Wang</i>	
Scalable Geometric Fracture Assembly Via Co-Creation Space Among Assemblers .....	7269
<i>Ruiyuan Zhang, Jiayang Liu, Zexi Li, Hao Dong, Jie Fu, Chao Wu</i>	
S3A: Towards Realistic Zero-Shot Classification Via Self Structural Semantic Alignment .....	7278
<i>Sheng Zhang, Muzammal Naseer, Guangyi Chen, Zhiqiang Shen, Salman Khan, Kun Zhang, Fahad Shahbaz Khan</i>	
A Computation-Aware Shape Loss Function for Point Cloud Completion .....	7287
<i>Shunran Zhang, Xiubo Zhang, Tsz Nam Chan, Shenghui Zhang, Leong Hou U</i>	
Vision-Language Pre-Training with Object Contrastive Learning for 3D Scene Understanding.....	7296
<i>Taolin Zhang, Sunan He, Tao Dai, Zhi Wang, Bin Chen, Shu-Tao Xia</i>	
Transformer-Based Selective Super-Resolution for Efficient Image Refinement .....	7305
<i>Tianyi Zhang, Kishore Kasichainula, Yaoxin Zhuo, Baoxin Li, Jae-Sun Seo, Yu Cao</i>	
Exploring Base-Class Suppression with Prior Guidance for Bias-Free One-Shot Object Detection.....	7314
<i>Wenwen Zhang, Yun Hu, Hangguan Shan, Eryun Liu</i>	
HEAP: Unsupervised Object Discovery and Localization with Contrastive Grouping.....	7323
<i>Xin Zhang, Jinheng Xie, Yuan Yuan, Michael Bi Mi, Robby T. Tan</i>	
Scribble Hides Class: Promoting Scribble-Based Weakly-Supervised Semantic Segmentation with Its Class Label .....	7332
<i>Xinliang Zhang, Lei Zhu, Hangzhou He, Lujia Jin, Yanye Lu</i>	
Negative Pre-Aware for Noisy Cross-Modal Matching.....	7341
<i>Xu Zhang, Hao Li, Mang Ye</i>	
Compositional Inversion for Stable Diffusion Models.....	7350
<i>Xulu Zhang, Xiao-Yong Wei, Jinlin Wu, Tianyi Zhang, Zhaoxiang Zhang, Zhen Lei, Qing Li</i>	
Cross-Modal Match for Language Conditioned 3D Object Grounding .....	7359
<i>Yachao Zhang, Runze Hu, Ronghui Li, Yanyun Qu, Yuan Xie, Xiu Li</i>	
MotionGPT: Finetuned LLMs Are General-Purpose Motion Generators .....	7368
<i>Yaqi Zhang, Di Huang, Bin Liu, Shixiang Tang, Yan Lu, Lu Chen, Lei Bai, Qi Chu, Nenghai Yu, Wanli Ouyang</i>	



Concept-Guided Prompt Learning for Generalization in Vision-Language Models .....	7377
<i>Yi Zhang, Ce Zhang, Ke Yu, Yushun Tang, Zhihai He</i>	
ISP-Teacher:Image Signal Process with Disentanglement Regularization for Unsupervised Domain Adaptive Dark Object Detection .....	7387
<i>Yin Zhang, Yongqiang Zhang, Zian Zhang, Man Zhang, Rui Tian, Mingli Ding</i>	
ArtBank: Artistic Style Transfer with Pre-Trained Diffusion Model and Implicit Style Prompt Bank .....	7396
<i>Zhanjie Zhang, Quanwei Zhang, Wei Xing, Guangyuan Li, Lei Zhao, Jiakai Sun, Zehua Lan, Junsheng Luan, Yiling Huang, Huaizhong Lin</i>	
A New Benchmark and Model for Challenging Image Manipulation Detection.....	7405
<i>Zhenfei Zhang, Mingyang Li, Ming-Ching Chang</i>	
TMFormer: Token Merging Transformer for Brain Tumor Segmentation with Missing Modalities .....	7414
<i>Zheyu Zhang, Gang Yang, Yueyi Zhang, Huanjing Yue, Aiping Liu, Yunwei Ou, Jian Gong, Xiaoyan Sun</i>	
FaceRSA: RSA-Aware Facial Identity Cryptography Framework.....	7423
<i>Zhongyi Zhang, Tianyi Wei, Wenbo Zhou, Hanqing Zhao, Weiming Zhang, Nenghai Yu</i>	
Spatial-Contextual Discrepancy Information Compensation for GAN Inversion .....	7432
<i>Ziqiang Zhang, Yan Yan, Jing-Hao Xue, Hanzi Wang</i>	
Self-Distillation Regularized Connectionist Temporal Classification Loss for Text Recognition: A Simple Yet Effective Approach.....	7441
<i>Ziyin Zhang, Ning Lu, Minghui Liao, Yongshuai Huang, Cheng Li, Min Wang, Wei Peng</i>	
PNerFLoc: Visual Localization with Point-Based Neural Radiance Fields .....	7450
<i>Boming Zhao, Luwei Yang, Mao Mao, Hujun Bao, Zhaopeng Cui</i>	
SimDistill: Simulated Multi-Modal Distillation for BEV 3D Object Detection.....	7460
<i>Haimei Zhao, Qiming Zhang, Shanshan Zhao, Zhe Chen, Jing Zhang, Dacheng Tao</i>	
Large Occluded Human Image Completion Via Image-Prior Cooperating .....	7469
<i>Hengrun Zhao, Yu Zeng, Huchuan Lu, Lijun Wang</i>	
Recognizing Ultra-High-Speed Moving Objects with Bio-Inspired Spike Camera .....	7478
<i>Junwei Zhao, Shiliang Zhang, Zhaofei Yu, Tiejun Huang</i>	
Rethinking Two-Stage Referring Expression Comprehension: A Novel Grounding and Segmentation Method Modulated by Point .....	7487
<i>Peizhi Zhao, Shiyi Zheng, Wenye Zhao, Dongsheng Xu, Pijian Li, Yi Cai, Qingbao Huang</i>	
Optical Flow for Spike Camera with Hierarchical Spatial-Temporal Spike Fusion .....	7496
<i>Rui Zhao, Ruiqin Xiong, Jian Zhang, Xinfeng Zhang, Zhaofei Yu, Tiejun Huang</i>	
Towards Fine-Grained HBOE with Rendered Orientation Set and Laplace Smoothing .....	7505
<i>Ruisi Zhao, Mingming Li, Zheng Yang, Binbin Lin, Xiaohui Zhong, Xiaobo Ren, Deng Cai, Boxi Wu</i>	
No Head Left Behind – Multi-Head Alignment Distillation for Transformers.....	7514
<i>Tianyang Zhao, Kunwar Yashraj Singh, Srikar Appalaraju, Peng Tang, Vijay Mahadevan, R. Manmatha, Ying Nian Wu</i>	
SFC: Shared Feature Calibration in Weakly Supervised Semantic Segmentation.....	7525
<i>Xinqiao Zhao, Feilong Tang, Xiaoyang Wang, Jimin Xiao</i>	

Unifying Multi-Modal Uncertainty Modeling and Semantic Alignment for Text-To-Image Person Re-Identification.....	7534
<i>Zhiwei Zhao, Bin Liu, Yan Lu, Qi Chu, Nenghai Yu</i>	
Mining Gaze for Contrastive Learning Toward Computer-Assisted Diagnosis .....	7543
<i>Zihao Zhao, Sheng Wang, Qian Wang, Dinggang Shen</i>	
Quad Bayer Joint Demosaicing and Denoising Based on Dual Encoder Network with Joint Residual Learning.....	7552
<i>Bolun Zheng, Haoran Li, Quan Chen, Tingyu Wang, Xiaofei Zhou, Zhenghui Hu, Chenggang Yan</i>	
End-To-End RGB-D Image Compression Via Exploiting Channel-Modality Redundancy .....	7562
<i>Huiming Zheng, Wei Gao</i>	
Any-Size-Diffusion: Toward Efficient Text-Driven Synthesis for Any-Size HD Images .....	7571
<i>Qingping Zheng, Yuanfan Guo, Jiankang Deng, Jianhua Han, Ying Li, Songcen Xu, Hang Xu</i>	
Spatio-Temporal Fusion for Human Action Recognition Via Joint Trajectory Graph.....	7579
<i>Yaolin Zheng, Hongbo Huang, Xiuying Wang, Xiaoxu Yan, Longfei Xu</i>	
ODTrack: Online Dense Temporal Token Learning for Visual Tracking .....	7588
<i>Yaozong Zheng, Bineng Zhong, Qihua Liang, Zhiyi Mo, Shengping Zhang, Xianxian Li</i>	
PVALane: Prior-Guided 3D Lane Detection with View-Agnostic Feature Alignment.....	7597
<i>Zewen Zheng, Xuemin Zhang, Yongqiang Mou, Xiang Gao, Chengxin Li, Guoheng Huang, Chi-Man Pun, Xiaochen Yuan</i>	
SpFormer: Spatio-Temporal Modeling for Scanpaths with Transformer .....	7605
<i>Wenqi Zhong, Linzhi Yu, Chen Xia, Junwei Han, Dingwen Zhang</i>	
ExpCLIP: Bridging Text and Facial Expressions Via Semantic Alignment .....	7614
<i>Yicheng Zhong, Huawei Wei, Peiji Yang, Zhisheng Wang</i>	
Learning Image Demoiréing from Unpaired Real Data .....	7623
<i>Yunshan Zhong, Yuyao Zhou, Yuxin Zhang, Fei Chao, Rongrong Ji</i>	
Lifting by Image – Leveraging Image Cues for Accurate 3D Human Pose Estimation .....	7632
<i>Feng Zhou, Jianqin Yin, Peiyang Li</i>	
NavGPT: Explicit Reasoning in Vision-And-Language Navigation with Large Language Models.....	7641
<i>Gengze Zhou, Yicong Hong, Qi Wu</i>	
Novel Class Discovery in Chest X-Rays Via Paired Images and Text .....	7650
<i>Jiaying Zhou, Yang Liu, Qingchao Chen</i>	
AMSP-UOD: When Vortex Convolution and Stochastic Perturbation Meet Underwater Object Detection .....	7659
<i>Jingchun Zhou, Zongxin He, Kin-Man Lam, Yudong Wang, Weishi Zhang, Chunle Guo, Chongyi Li</i>	
SOGDet: Semantic-Occupancy Guided Multi-View 3D Object Detection .....	7668
<i>Qiu Zhou, Jinming Cao, Hanchao Leng, Yifang Yin, Yu Kun, Roger Zimmermann</i>	
Test-Time Adaptation Via Style and Structure Guidance for Histological Image Registration .....	7677
<i>Shenglong Zhou, Zhiwei Xiong, Feng Wu</i>	

Reducing Spatial Fitting Error in Distillation of Denoising Diffusion Models .....	7686
<i>Shengzhe Zhou, Zejian Li, Shengyuan Zhang, Lefan Hou, Changyuan Yang, Guang Yang, Zhiyuan Yang, Lingyun Sun</i>	
SAMFlow: Eliminating Any Fragmentation in Optical Flow with Segment Anything Model .....	7695
<i>Shili Zhou, Ruiian He, Weimin Tan, Bo Yan</i>	
Efficient Lightweight Image Denoising with Triple Attention Transformer .....	7704
<i>Yubo Zhou, Jin Lin, Fangchen Ye, Yanyun Qu, Yuan Xie</i>	
Intentional Evolutionary Learning for Untrimmed Videos with Long Tail Distribution .....	7713
<i>Yuxi Zhou, Xiujie Wang, Jianhua Zhang, Jiajia Wang, Jie Yu, Hao Zhou, Yi Gao, Shengyong Chen</i>	
SasWOT: Real-Time Semantic Segmentation Architecture Search WithOut Training .....	7722
<i>Chendi Zhu, Lujun Li, Yuli Wu, Zhengxing Sun</i>	
Enhance Sketch Recognition's Explainability Via Semantic Component-Level Parsing .....	7731
<i>Guangming Zhu, Siyuan Wang, Tianci Wu, Liang Zhang</i>	
Learning Discriminative Noise Guidance for Image Forgery Detection and Localization .....	7739
<i>Jiaying Zhu, Dong Li, Xueyang Fu, Gang Yang, Jie Huang, Aiping Liu, Zheng-Jun Zha</i>	
Video Frame Prediction from a Single Image and Events .....	7748
<i>Juanjuan Zhu, Zhexiong Wan, Yuchao Dai</i>	
Finding Visual Saliency in Continuous Spike Stream .....	7757
<i>Lin Zhu, Xianzhang Chen, Xiao Wang, Hua Huang</i>	
SEER: Backdoor Detection for Vision-Language Models Through Searching Target Text and Image Trigger Jointly .....	7766
<i>Liuwan Zhu, Rui Ning, Jiang Li, Chunsheng Xin, Hongyi Wu</i>	
Text Image Inpainting Via Global Structure-Guided Diffusion Models .....	7775
<i>Shipeng Zhu, Pengfei Fang, Chenjie Zhu, Zuoyan Zhao, Qiang Xu, Hui Xue</i>	
Rethinking Mesh Watermark: Towards Highly Robust and Adaptable Deep 3D Mesh Watermarking .....	7784
<i>Xingyu Zhu, Guanhui Ye, Xiapu Luo, Xuetao Wei</i>	
Boosting Few-Shot Learning Via Attentive Feature Regularization .....	7793
<i>Xingyu Zhu, Shuo Wang, Jinda Lu, Yanbin Hao, Haifeng Liu, Xiangnan He</i>	
Memory-Efficient Prompt Tuning for Incremental Histopathology Classification .....	7802
<i>Yu Zhu, Kang Li, Lequan Yu, Pheng Ann Heng</i>	
SPGroup3D: Superpoint Grouping Network for Indoor 3D Object Detection .....	7811
<i>Yun Zhu, Le Hui, Yaqi Shen, Jin Xie</i>	
SEIT: Structural Enhancement for Unsupervised Image Translation in Frequency Domain .....	7820
<i>Zhifeng Zhu, Yaochen Li, Yifan Li, Jinhua Yang, Peijun Chen, Yuehu Liu</i>	
A Pre-Convolved Representation for Plug-And-Play Neural Illumination Fields .....	7828
<i>Yiyu Zhuang, Qi Zhang, Xuan Wang, Hao Zhu, Ying Feng, Xiaoyu Li, Ying Shan, Xun Cao</i>	
IPRemover: A Generative Model Inversion Attack Against Deep Neural Network Fingerprinting and Watermarking .....	7837
<i>Wei Zong, Yang-Wai Chow, Willy Susilo, Joonsang Baek, Jongkil Kim, Seyit Camtepe</i>	

DiffBEV: Conditional Diffusion Model for Bird’s Eye View Perception.....	7846
<i>Jiayu Zou, Kun Tian, Zheng Zhu, Yun Ye, Xingang Wang</i>	
Cross-Covariate Gait Recognition: A Benchmark .....	7855
<i>Shinan Zou, Chao Fan, Jianbo Xiong, Chuanfu Shen, Shiqi Yu, Jin Tang</i>	
Towards Efficient Diffusion-Based Image Editing with Instant Attention Masks.....	7864
<i>Siyu Zou, Jiji Tang, Yiyi Zhou, Jing He, Chaoyi Zhao, Rongsheng Zhang, Zhipeng Hu, Xiaoshuai Sun</i>	
VQCNR: Clearer Night Image Restoration with Vector-Quantized Codebook .....	7873
<i>Wenbin Zou, Hongxia Gao, Tian Ye, Liang Chen, Weipeng Yang, Shasha Huang, Hongsheng Chen, Sixiang Chen</i>	
Enhancing Neural Radiance Fields with Adaptive Multi-Exposure Fusion: A Bilevel Optimization Approach for Novel View Synthesis .....	7882
<i>Yang Zou, Xingyuan Li, Zhiying Jiang, Jinyuan Liu</i>	
Improved MLP Point Cloud Processing with High-Dimensional Positional Encoding .....	7891
<i>Yanmei Zou, Hongshan Yu, Zhengeng Yang, Zechuan Li, Naveed Akhtar</i>	
Sparse3D: Distilling Multiview-Consistent Diffusion for Object Reconstruction from Sparse Views .....	7900
<i>Zixin Zou, Weihao Cheng, Yan-Pei Cao, Shi-Sheng Huang, Ying Shan, Song-Hai Zhang</i>	
CEDFlow: Latent Contour Enhancement for Dark Optical Flow Estimation .....	7909
<i>Fengyuan Zuo, Zhaolin Xiao, Haiyan Jin, Haonan Su</i>	

**Author Index**