

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 7: AAAI Technical Tracks

- Computer Vision VI

Philadelphia, Pennsylvania, USA
25 February - 4 March 2025

Part 1 of 2

ISBN: 979-8-3313-2517-6

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 VI

Eve: Efficient Multimodal Vision Language Models with Elastic Visual Experts	6694
<i>Miao Rang, Zhenni Bi, Chuanjian Liu, Yehui Tang, Kai Han, Yunhe Wang</i>	
Motif Guided Graph Transformers with Combinatorial Skeleton Prototype Learning for Skeleton- Based Person Re-Identification	6703
<i>Haocong Rao, Chunyan Miao</i>	
Click2Mask: Local Editing with Dynamic Mask Generation.....	6713
<i>Omer Regev, Omri Avrahami, Dani Lischinski</i>	
ISPDiffuser: Learning RAW-to-sRGB Mappings with Texture-Aware Diffusion Models and Histogram-Guided Color Consistency.....	6722
<i>Yang Ren, Hai Jiang, Menglong Yang, Wei Li, Shuaicheng Liu</i>	
Improving Integrated Gradient-based Transferable Adversarial Examples by Refining the Integration Path	6731
<i>Yuchen Ren, Zhengyu Zhao, Chenhao Lin, Bo Yang, Lu Zhou, Zhe Liu, Chao Shen</i>	
MM-CamObj: A Comprehensive Multimodal Dataset for Camouflaged Object Scenarios.....	6740
<i>Jiacheng Ruan, Wenzhen Yuan, Zehao Lin, Ning Liao, Zhiyu Li, Feiyu Xiong, Ting Liu, Yuzhuo Fu</i>	
GenHMR: Generative Human Mesh Recovery	6749
<i>Muhammad Usama Saleem, Ekkasit Pinyoanuntapong, Pu Wang, Hongfei Xue, Srijan Das, Chen Chen</i>	
FunEditor: Achieving Complex Image Edits via Function Aggregation with Diffusion Models	6758
<i>Mohammadreza Samadi, Fred X. Han, Mohammad Salameh, Hao Wu, Fengyu Sun, Chunhua Zhou, Di Niu</i>	
PVTree: Realistic and Controllable Palm Vein Generation for Recognition Tasks	6767
<i>Sheng Shang, Chenglong Zhao, Ruixin Zhang, Jianlong Jin, Jingyun Zhang, Rizen Guo, Shouhong Ding, Yunsheng Wu, Yang Zhao, Wei Jia</i>	
Video Summarization Using Denoising Diffusion Probabilistic Model.....	6776
<i>Zirui Shang, Yubo Zhu, Hongxi Li, Shuo Yang, Xinxiao Wu</i>	
Boosting Consistency in Story Visualization with Rich-Contextual Conditional Diffusion Models.....	6785
<i>Fei Shen, Hu Ye, Sibor Liu, Jun Zhang, Cong Wang, Xiao Han, Yang Wei</i>	
IMAGDressing-v1: Customizable Virtual Dressing	6795
<i>Fei Shen, Xin Jiang, Xin He, Hu Ye, Cong Wang, Xiaoyu Du, Zechao Li, Jinhui Tang</i>	
LDP: Generalizing to Multilingual Visual Information Extraction by Language Decoupled Pretraining	6805
<i>Huawen Shen, Gengluo Li, Jinwen Zhong, Yu Zhou</i>	

In2NeCT: Inter-class and Intra-class Neural Collapse Tuning for Semantic Segmentation of Imbalanced Remote Sensing Images	6814
<i>Junao Shen, Qiyun Hu, Tian Feng, Xinyu Wang, Hui Cui, Sensen Wu, Wei Zhang</i>	
Topology-Aware 3D Gaussian Splatting: Leveraging Persistent Homology for Optimized Structural Integrity	6823
<i>Tianqi Shen, Shaohua Liu, Jiaqi Feng, Ziyi Ma, Ning An</i>	
Fast Omni-Directional Image Super-Resolution: Adapting the Implicit Image Function with Pixel and Semantic-Wise Spherical Geometric Priors.....	6833
<i>Xuelin Shen, Yitong Wang, Silin Zheng, Kang Xiao, Wenhan Yang, Xu Wang</i>	
Medical Multimodal Model Stealing Attacks via Adversarial Domain Alignment	6842
<i>Yaling Shen, Zhixiong Zhuang, Kun Yuan, Maria-Irina Nicolae, Nassir Navab, Nicolas Padoy, Mario Fritz</i>	
ProcTag: Process Tagging for Assessing the Efficacy of Document Instruction Data.....	6851
<i>Yufan Shen, Chuwei Luo, Zhaoqing Zhu, Yang Chen, Qi Zheng, Zhi Yu, Jiajun Bu, Cong Yao</i>	
Free-Moving Object Reconstruction and Pose Estimation with Virtual Camera.....	6860
<i>Haixin Shi, Yinlin Hu, Daniel Koguciuk, Juan-Ting Lin, Mathieu Salzmann, David Ferstl</i>	
Normal-NeRF: Ambiguity-Robust Normal Estimation for Highly Reflective Scenes	6869
<i>Ji Shi, Xianghua Ying, Ruohao Guo, Bowei Xing, Wenzhen Yue</i>	
Neural Block Compression: Variable Bitrates Feature Blocks for Texture Representation.....	6878
<i>Rui Shi, Yishun Dou, Zhong Zheng, Xiangzhong Fang, Wenjun Zhang, Bingbing Ni</i>	
ResMaster: Mastering High-Resolution Image Generation via Structural and Fine-Grained Guidance.....	6887
<i>Shuwei Shi, Wenbo Li, Yuechen Zhang, Jingwen He, Biao Gong, Yinqiang Zheng</i>	
HS-FPN: High Frequency and Spatial Perception FPN for Tiny Object Detection	6896
<i>Zican Shi, Jing Hu, Jie Ren, Hengkang Ye, Xuyang Yuan, Yan Ouyang, Jia He, Bo Ji, Junyu Guo</i>	
Pixel Is Not a Barrier: An Effective Evasion Attack for Pixel-Domain Diffusion Models.....	6905
<i>Chun-Yen Shih, Li-Xuan Peng, Jia-Wei Liao, Ernie Chu, Cheng-Fu Chou, Jun-Cheng Chen</i>	
SdalsNet: Self-Distilled Attention Localization and Shift Network for Unsupervised Camouflaged Object Detection	6914
<i>Peiyao Shou, Yixiu Liu, Wei Wang, Yaoqi Sun, Zhigao Zheng, Shangdong Zhu, Chenggang Yan</i>	
OGP-Net: Optical Guidance Meets Pixel-Level Contrastive Distillation for Robust Multi-Modal and Missing Modality Segmentation	6922
<i>Aniruddh Sikdar, Jayant Teotia, Suresh Sundaram</i>	
SKI Models: Skeleton Induced Vision-Language Embeddings for Understanding Activities of Daily Living	6931
<i>Arkaprava Sinha, Dominick Reilly, Francois Bremond, Pu Wang, Srijan Das</i>	
JoVALE: Detecting Human Actions in Video Using Audiovisual and Language Contexts	6940
<i>Taein Son, Soo Won Seo, Jisong Kim, Seok Hwan Lee, Jun Won Choi</i>	
Fine-Grained Perception in Panoramic Scenes: A Novel Task, Dataset, and Method for Object Importance Ranking	6950
<i>Jia Song, Chenglizhao Chen, Xu Yu, Shanchen Pang</i>	

CtrlAvatar: Controllable Avatars Generation via Disentangled Invertible Networks	6959
<i>Wenfeng Song, Yang Ding, Fei Hou, Shuai Li, Aimin Hao, Xia Hou</i>	
ERL-MPP: Evolutionary Reinforcement Learning with Multi-head Puzzle Perception for Solving Large-scale Jigsaw Puzzles of Eroded Gaps.....	6968
<i>Xingke Song, Xiaoying Yang, Chenglin Yao, Jianfeng Ren, Ruibin Bai, Xin Chen, Xudong Jiang</i>	
Temporal Coherent Object Flow for Multi-Object Tracking.....	6978
<i>Zikai Song, Run Luo, Lintao Ma, Ying Tang, Yi-Ping Phoebe Chen, Junqing Yu, Wei Yang</i>	
Toward Improving Robustness and Accuracy in Unsupervised Domain Adaptation	6987
<i>Aishwarya Soni, Tanim Datta</i>	
Hierarchical Vector Quantization for Unsupervised Action Segmentation.....	6996
<i>Federico Spurio, Emad Bahrami, Gianpiero Francesca, Juergen Gall</i>	
Enhancing Noise-Robust Losses for Large-Scale Noisy Data Learning	7006
<i>Max Staats, Matthias Thamm, Bernd Rosenow</i>	
RI-MAE: Rotation-Invariant Masked AutoEncoders for Self-Supervised Point Cloud Representation Learning.....	7015
<i>Kunming Su, Qiuxia Wu, Panpan Cai, Xiaogang Zhu, Xuequan Lu, Zhiyong Wang, Kun Hu</i>	
Can We Get Rid of Handcrafted Feature Extractors? SparseViT: Nonsemantics-Centered, Parameter-Efficient Image Manipulation Localization Through Spare-Coding Transformer	7024
<i>Lei Su, Xiaochen Ma, Xuekang Zhu, Chaoqun Niu, Zeyu Lei, Ji-Zhe Zhou</i>	
EigenSR: Eigenimage-Bridged Pre-Trained RGB Learners for Single Hyperspectral Image Super- Resolution.....	7033
<i>Xi Su, Xiangfei Shen, Mingyang Wan, Jing Nie, Lihui Chen, Haijun Liu, Xichuan Zhou</i>	
Prior-guided Hierarchical Harmonization Network for Efficient Image Dehazing	7042
<i>Xiongfei Su, Siyuan Li, Yuning Cui, Miao Cao, Yulun Zhang, Zheng Chen, Zongliang Wu, Zedong Wang, Yuanlong Zhang, Xin Yuan</i>	
Dual-branch Graph Feature Learning for NLOS Imaging.....	7051
<i>Xiongfei Su, Tianyi Zhu, Lina Liu, Zheng Chen, Yulun Zhang, Siyuan Li, Juntian Ye, Feihu Xu, Xin Yuan</i>	
Learning Fine-Grained Alignment for Aerial Vision-Dialog Navigation	7060
<i>Yifei Su, Dong An, Kehan Chen, Weichen Yu, Baiyang Ning, Yonggen Ling, Yan Huang, Liang Wang</i>	
Explicit Relational Reasoning Network for Scene Text Detection.....	7069
<i>Yuchen Su, Zhineng Chen, Yongkun Du, Zhilong Ji, Kai Hu, Jinfeng Bai, Xieping Gao</i>	
3D Annotation-Free Learning by Distilling 2D Open-Vocabulary Segmentation Models for Autonomous Driving	7078
<i>Boyi Sun, Yuhang Liu, Xingxia Wang, Bin Tian, Long Chen, Fei-Yue Wang</i>	
Ultra-High Resolution Segmentation via Boundary-Enhanced Patch-Merging Transformer.....	7087
<i>Haopeng Sun, Yingwei Zhang, Lumin Xu, Sheng Jin, Yiqiang Chen</i>	
NeuralFlix: A Simple While Effective Framework for Semantic Decoding of Videos from Non- invasive Brain Recordings.....	7096
<i>Jingyuan Sun, Mingxiao Li, Marie-Francine Moens</i>	

VE-Bench: Subjective-Aligned Benchmark Suite for Text-Driven Video Editing Quality Assessment	7105
<i>Shangkun Sun, Xiaoyu Liang, Songlin Fan, Wenxu Gao, Wei Gao</i>	
Guided and Variance-Corrected Fusion with One-shot Style Alignment for Large-Content Image Generation	7114
<i>Shoukun Sun, Min Xian, Tiankai Yao, Fei Xu, Luca Capriotti</i>	
Hierarchically-Structured Open-Vocabulary Indoor Scene Synthesis with Pre-trained Large Language Model.....	7122
<i>Weilin Sun, Xinran Li, Manyi Li, Kai Xu, Xiangxu Meng, Lei Meng</i>	
Diversifying Query: Region-Guided Transformer for Temporal Sentence Grounding	7131
<i>Xiaolong Sun, Liushuai Shi, Le Wang, Sanping Zhou, Kun Xia, Yabing Wang, Gang Hua</i>	
M2Flow: A Motion Information Fusion Framework for Enhanced Unsupervised Optical Flow Estimation in Autonomous Driving.....	7140
<i>Xunpei Sun, Gang Chen, Zuoxun Hou</i>	
Leveraging Large Vision-Language Model as User Intent-Aware Encoder for Composed Image Retrieval	7149
<i>Zelong Sun, Dong Jing, Guoxing Yang, Nanyi Fei, Zhiwu Lu</i>	
RealisID: Scale-Robust and Fine-Controllable Identity Customization via Local and Global Complementation	7158
<i>Zhaoyang Sun, Fei Du, Weihua Chen, Fan Wang, Yaxiong Chen, Yi Rong, Shengwu Xiong</i>	
SoundBrush: Sound as a Brush for Visual Scene Editing.....	7167
<i>Kim Sung-Bin, Kim Jun-Seong, Junseok Ko, Yewon Kim, Tae-Hyun Oh</i>	
First Line of Defense: A Robust First Layer Mitigates Adversarial Attacks.....	7176
<i>Janani Suresh, Nancy Nayak, Sheetal Kalyani</i>	
C2P-CLIP: Injecting Category Common Prompt in CLIP to Enhance Generalization in Deepfake Detection	7184
<i>Chuangchuang Tan, Renshuai Tao, Huan Liu, Guanghua Gu, Baoyuan Wu, Yao Zhao, Yunchao Wei</i>	
Modality-Aware Shot Relating and Comparing for Video Scene Detection.....	7193
<i>Jiawei Tan, Hongxing Wang, Kang Dang, Jiaxin Li, Zhilong Ou</i>	
Beyond Human Data: Aligning Multimodal Large Language Models by Iterative Self-Evolution	7202
<i>Wentao Tan, Qiong Cao, Yibing Zhan, Chao Xue, Changxing Ding</i>	
ALLVB: All-in-One Long Video Understanding Benchmark	7211
<i>Xichen Tan, Yuanjing Luo, Yunfan Ye, Fang Liu, Zhiping Cai</i>	
Neighbor Does Matter: Density-Aware Contrastive Learning for Medical Semi-supervised Segmentation	7220
<i>Feilong Tang, Zhongxing Xu, Ming Hu, Wenxue Li, Peng Xia, Yiheng Zhong, Hanjun Wu, Jionglong Su, Zongyuan Ge</i>	
A Training-free Synthetic Data Selection Method for Semantic Segmentation	7229
<i>Hao Tang, Siyue Yu, Jian Pang, Bingfeng Zhang</i>	
MUSE: Mamba Is Efficient Multi-scale Learner for Text-video Retrieval	7238
<i>Haoran Tang, Meng Cao, Jinfa Huang, Ruyang Liu, Peng Jin, Ge Li, Xiaodan Liang</i>	

PART 2

Promptable Representation Distribution Learning and Data Augmentation for Gigapixel Histopathology WSI Analysis	7247
<i>Kunming Tang, Zhiguo Jiang, Jun Shi, Wei Wang, Haibo Wu, Yushan Zheng</i>	
Compressing Streamable Free-Viewpoint Videos to 0.1 MB per Frame	7257
<i>Luyang Tang, Jiayu Yang, Rui Peng, Yongqi Zhai, Shihe Shen, Ronggang Wang</i>	
Sign-IDD: Iconicity Disentangled Diffusion for Sign Language Production	7266
<i>Shengeng Tang, Jiayi He, Dan Guo, Yanyan Wei, Feng Li, Richang Hong</i>	
BEV-TSR: Text-Scene Retrieval in BEV Space for Autonomous Driving.....	7275
<i>Tao Tang, Dafeng Wei, Zhengyu Jia, Tian Gao, Changwei Cai, Chengkai Hou, Peng Jia, Kun Zhan, Haiyang Sun, Fan JingChen, Yixing Zhao, Xiaodan Liang, Xianpeng Lang, Yang Wang</i>	
More Text, Less Point: Towards 3D Data-Efficient Point-Language Understanding	7284
<i>Yuan Tang, Xu Han, Xianzhi Li, Qiao Yu, Jinfeng Xu, Yixue Hao, Long Hu, Min Chen</i>	
Empowering LLMs with Pseudo-Untrimmed Videos for Audio-Visual Temporal Understanding	7293
<i>Yunlong Tang, Daiki Shimada, Jing Bi, Mingqian Feng, Hang Hua, Chenliang Xu</i>	
CaRDiff: Video Salient Object Ranking Chain of Thought Reasoning for Saliency Prediction with Diffusion.....	7302
<i>Yunlong Tang, Gen Zhan, Li Yang, Yiting Liao, Chenliang Xu</i>	
RAGG: Retrieval-Augmented Grasp Generation Model.....	7311
<i>Zhenhua Tang, Bin Zhu, Yanbin Hao, Chong-Wah Ngo, Richang Hong</i>	
Cycle3D: High-quality and Consistent Image-to-3D Generation via Generation-Reconstruction Cycle.....	7320
<i>Zhenyu Tang, Junwu Zhang, Xinhua Cheng, Wangbo Yu, Chaoran Feng, Yatian Pang, Bin Lin, Li Yuan</i>	
From Representation Space to Prognostic Insights: Whole Slide Image Generation with Hierarchical Diffusion Model for Survival Prediction	7329
<i>Zhihao Tang, Xi Zhang, Chaozhuo Li</i>	
3D ² -Actor: Learning Pose-Conditioned 3D-Aware Denoiser for Realistic Gaussian Avatar Modeling	7338
<i>Zichen Tang, Hongyu Yang, Hanchen Zhang, Jiaxin Chen, Di Huang</i>	
Kernel-Aware Graph Prompt Learning for Few-Shot Anomaly Detection.....	7347
<i>Fenfang Tao, Guo-Sen Xie, Fang Zhao, Xiangbo Shu</i>	
Relieving Universal Label Noise for Unsupervised Visible-Infrared Person Re-Identification by Inferring from Neighbors	7356
<i>Xiao Teng, Long Lan, Dingyao Chen, Kele Xu, Nan Yin</i>	
Stitch, Contrast, and Segment: Learning a Human Action Segmentation Model Using Trimmed Skeleton Videos	7365
<i>Haitao Tian, Pierre Payeur</i>	
DrivingForward: Feed-forward 3D Gaussian Splatting for Driving Scene Reconstruction from Flexible Surround-view Input.....	7374
<i>Qijian Tian, Xin Tan, Yuan Xie, Lizhuang Ma</i>	

Unsupervised Self-Prior Embedding Neural Representation for Iterative Sparse-View CT Reconstruction.....	7383
<i>Xuanyu Tian, Lixuan Chen, Qing Wu, Chenhe Du, Jingjing Shi, Hongjiang Wei, Yuyao Zhang</i>	
AI-generated Image Quality Assessment in Visual Communication.....	7392
<i>Yu Tian, Yixuan Li, Baoliang Chen, Hanwei Zhu, Shiqi Wang, Sam Kwong</i>	
ChatterBox: Multimodal Referring and Grounding with Chain-of-Questions	7401
<i>Yunjie Tian, Tianren Ma, Lingxi Xie, Qixiang Ye</i>	
Anywhere: A Multi-Agent Framework for User-Guided, Reliable, and Diverse Foreground-Conditioned Image Generation.....	7410
<i>Xie Tianyidan, Rui Ma, Qian Wang, Xiaoqian Ye, Feixuan Liu, Ying Tai, Zhenyu Zhang, Lanjun Wang, Zili Yi</i>	
G-VEval: A Versatile Metric for Evaluating Image and Video Captions Using GPT-4o.....	7419
<i>Tony Cheng Tong, Sirui He, Zhiwen Shao, Dit-Yan Yeung</i>	
Improving Transformer Based Line Segment Detection with Matched Predicting and Re-ranking	7428
<i>Xin Tong, Shi Peng, Baojie Tian, Yufei Guo, Xuhui Huang, Zhe Ma</i>	
Collaborative Learning for 3D Hand-Object Reconstruction and Compositional Action Recognition from Egocentric RGB Videos Using Superquadrics.....	7437
<i>Tze Ho Elden Tse, Runyang Feng, Linfang Zheng, Jiho Park, Yixing Gao, Jihie Kim, Ales Leonardis, Hyung Jin Chang</i>	
Memory-Augmented Re-Completion for 3D Semantic Scene Completion	7446
<i>Yu-Wen Tseng, Sheng-Ping Yang, Jhih-Ciang Wu, I-Bin Liao, Yung-Hui Li, Hong-Han Shuai, Wen-Huang Cheng</i>	
TextToucher: Fine-Grained Text-to-Touch Generation	7455
<i>Jiahang Tu, Hao Fu, Fengyu Yang, Hanbin Zhao, Chao Zhang, Hui Qian</i>	
Query-centric Audio-Visual Cognition Network for Moment Retrieval, Segmentation and Step-Captioning.....	7464
<i>Yunbin Tu, Liang Li, Li Su, Qingming Huang</i>	
Watch Video, Catch Keyword: Context-aware Keyword Attention for Moment Retrieval and Highlight Detection	7473
<i>Sung Jin Um, Dongjin Kim, Sangmin Lee, Jung Uk Kim</i>	
VOILA: Complexity-Aware Universal Segmentation of CT Images by Voxel Interacting with Language.....	7482
<i>Zishuo Wan, Yu Gao, Wanyuan Pang, Dawei Ding</i>	
ParGo: Bridging Vision-Language with Partial and Global Views.....	7491
<i>An-Lan Wang, Bin Shan, Wei Shi, Kun-Yu Lin, Xiang Fei, Guozhi Tang, Lei Liao, Jingqun Tang, Can Huang, Wei-Shi Zheng</i>	
Overcoming Heterogeneous Data in Federated Medical Vision-Language Pre-training: A Triple-Embedding Model Selector Approach.....	7500
<i>Aowen Wang, Zhiwang Zhang, Dongang Wang, Fanyi Wang, Haotian Hu, Jinyang Guo, Yipeng Zhou, Chaoyi Pang, Shiting Wen</i>	
RealisHuman: A Two-Stage Approach for Refining Malformed Human Parts in Generated Images	7509
<i>Benzhi Wang, Jingkai Zhou, Jingqi Bai, Yang Yang, Weihua Chen, Fan Wang, Zhen Lei</i>	

FakeDiffer: Distributional Disparity Learning on Differentiated Reconstruction for Face Forgery Detection	7518
<i>Bo Wang, Zhao Zhang, Suiyi Zhao, Xianming Ye, Haijun Zhang, Meng Wang</i>	
Taylor Series-Inspired Local Structure Fitting Network for Few-shot Point Cloud Semantic Segmentation	7527
<i>Changshuo Wang, Shuting He, Xiang Fang, Meiqing Wu, Siew-Kei Lam, Prayag Tiwari</i>	
Focus on Local: Finding Reliable Discriminative Regions for Visual Place Recognition	7536
<i>Changwei Wang, Shunpeng Chen, Yukun Song, Rongtao Xu, Zherui Zhang, Jiguang Zhang, Haoran Yang, Yu Zhang, Kexue Fu, Shide Du, Zhiwei Xu, Longxiang Gao, Li Guo, Shibiao Xu</i>	
Explore In-Context Segmentation via Latent Diffusion Models	7545
<i>Chaoyang Wang, Xiangtai Li, Henghui Ding, Lu Qi, Jiangning Zhang, Yunhai Tong, Chen Change Loy, Shuicheng Yan</i>	
Geometry-Aware 3D Salient Object Detection Network.....	7554
<i>Chen Wang, Liyuan Zhang, Le Hui, Qi Liu, Yuchao Dai</i>	
QCS:Feature Refining from Quadruplet Cross Similarity for Facial Expression Recognition	7563
<i>Chengpeng Wang, Li Chen, Lili Wang, Zhaofan Li, Xuebin Lv</i>	
RHandS: Refining Malformed Hands for Generated Images with Decoupled Structure and Style Guidance.....	7573
<i>Chengrui Wang, Pengfei Liu, Min Zhou, Ming Zeng, Xubin Li, Tiezheng Ge, Bo Zheng</i>	
Multi-clue Consistency Learning to Bridge Gaps Between General and Oriented Object in Semi-supervised Detection	7582
<i>Chenxu Wang, Chunyan Xu, Xiang Li, YuXuan Li, Xu Guo, Ziqi Gu, Zhen Cui</i>	
RA-GAR: A Richly Annotated Benchmark for Gait Attribute Recognition	7591
<i>Chenye Wang, Saihui Hou, Aoqi Li, Qingyuan Cai, Yongzhen Huang</i>	
Towards Efficient Object Re-Identification with a Novel Cloud-Edge Collaborative Framework	7600
<i>Chuanming Wang, Yuxin Yang, Mengshi Qi, Huanhuan Zhang, Huadong Ma</i>	
Intra and Inter Parser-Prompted Transformers for Effective Image Restoration	7609
<i>Cong Wang, Jinshan Pan, Liyan Wang, Wei Wang</i>	
Target-Driven Distillation: Consistency Distillation with Target Timestep Selection and Decoupled Guidance.....	7619
<i>Cunzheng Wang, Ziyuan Guo, Yuxuan Duan, Huaxia Li, Nemo Chen, Xu Tang, Yao Hu</i>	
ADBA: Approximation Decision Boundary Approach for Black-Box Adversarial Attacks	7628
<i>Feiyang Wang, Xingquan Zuo, Hai Huang, Gang Chen</i>	
A Black-Box Evaluation Framework for Semantic Robustness in Bird's Eye View Detection	7637
<i>Fu Wang, Yanghao Zhang, Xiangyu Yin, Guangliang Cheng, Zeyu Fu, Xiaowei Huang, Wenjie Ruan</i>	
Scene Graph-Grounded Image Generation.....	7646
<i>Fuyun Wang, Tong Zhang, Yuanzhi Wang, Xiaoya Zhang, Xin Liu, Zhen Cui</i>	
S ³ -Mamba: Small-Size-Sensitive Mamba for Lesion Segmentation	7655
<i>Gui Wang, Yuexiang Li, Wenting Chen, Meidan Ding, Wooi Ping Cheah, Rong Qu, Jianfeng Ren, Linlin Shen</i>	

SSC-VAE: Structured Sparse Coding Based Variational Autoencoder for Detail Preserved Image Reconstruction.....	7665
<i>Hao Wang, Lu Wang, Zhongyu Wang, Lixin Ma, Ye Luo</i>	
BLS-GAN: A Deep Layer Separation Framework for Eliminating Bone Overlap in Conventional Radiographs.....	7674
<i>Haolin Wang, Yafei Ou, Prasoon Ambalathankandy, Gen Ota, Pengyu Dai, Masayuki Ikebe, Kenji Suzuki, Tamotsu Kamishima</i>	
MV-VTON: Multi-View Virtual Try-On with Diffusion Models	7682
<i>Haoyu Wang, Zhilu Zhang, Donglin Di, Shiliang Zhang, Wangmeng Zuo</i>	
EMControl: Adding Conditional Control to Text-to-Image Diffusion Models via Expectation-Maximization.....	7691
<i>He Wang, Longquan Dai, Jinhui Tang</i>	
Dual Conditioned Motion Diffusion for Pose-Based Video Anomaly Detection	7700
<i>Hongsong Wang, Andi Xu, Pinle Ding, Jie Gui</i>	
M2OST: Many-to-one Regression for Predicting Spatial Transcriptomics from Digital Pathology Images	7709
<i>Hongyi Wang, Xiuju Du, Jing Liu, Shuyi Ouyang, Yen-Wei Chen, Lanfen Lin</i>	
SpotActor: Training-Free Layout-Controlled Consistent Image Generation.....	7718
<i>Jiahao Wang, Caixia Yan, Weizhan Zhang, Haonan Lin, Mengmeng Wang, Guang Dai, Tieliang Gong, Hao Sun, Jingdong Wang</i>	
RAP-SR: RestorAtion Prior Enhancement in Diffusion Models for Realistic Image Super-Resolution.....	7727
<i>Jiangang Wang, Qingnan Fan, Jinwei Chen, Hong Gu, Feng Huang, Wenqi Ren</i>	
FaceA-Net: Facial Attribute-Driven ID Preserving Image Generation Network.....	7736
<i>Jiayu Wang, Yue Yu, Jingjing Chen, Qi Dai, Yu-Gang Jiang</i>	
MM-Mixing: Multi-Modal Mixing Alignment for 3D Understanding.....	7744
<i>Jiaze Wang, Yi Wang, Ziyu Guo, Renrui Zhang, Donghao Zhou, Guangyong Chen, Anfeng Liu, Pheng-Ann Heng</i>	
Efficient Self-Supervised Video Hashing with Selective State Spaces.....	7753
<i>Jinpeng Wang, Niu Lian, Jun Li, Yuting Wang, Yan Feng, Bin Chen, Yongbing Zhang, Shu-Tao Xia</i>	
OV-DQUO: Open-Vocabulary DETR with Denoising Text Query Training and Open-World Unknown Objects Supervision	7762
<i>Junjie Wang, Bin Chen, Bin Kang, Yulin Li, Weizhi Xian, Yichi Chen, Yong Xu</i>	
InpDiffusion: Image Inpainting Localization via Conditional Diffusion Models.....	7771
<i>Kai Wang, Shaozhang Niu, Qixian Hao, Jiwei Zhang</i>	
Alignment-Free RGB-T Salient Object Detection: A Large-Scale Dataset and Progressive Correlation Network.....	7780
<i>Kunpeng Wang, Keke Chen, Chenglong Li, Zhengzheng Tu, Bin Luo</i>	
Tracking Everything Everywhere across Multiple Cameras	7789
<i>Li-Heng Wang, YuJu Cheng, Tyng-Luh Liu</i>	

Author Index