

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 3: AAAI Technical Tracks

- Computer Vision II

Vancouver, Canada
20 – 27 February 2024

Part 1 of 2

ISBN: 979-8-3313-0085-2

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
Web: www.proceedings.com

TABLE OF CONTENTS

PART 1

AAAI TECHNICAL TRACK ON COMPUTER VISION II

Fine-Grained Multi-View Hand Reconstruction Using Inverse Rendering	1779
<i>Qijun Gan, Wentong Li, Jinwei Ren, Jianke Zhu</i>	
Attacking Transformers with Feature Diversity Adversarial Perturbation	1788
<i>Chenxing Gao, Hang Zhou, Junqing Yu, YuTeng Ye, Jiale Cai, Junle Wang, Wei Yang</i>	
Leveraging Imagery Data with Spatial Point Prior for Weakly Semi-Supervised 3D Object Detection	1797
<i>Hongzhi Gao, Zheng Chen, Zehui Chen, Lin Chen, Jiaming Liu, Shanghang Zhang, Feng Zhao</i>	
Dual-Prior Augmented Decoding Network for Long Tail Distribution in HOI Detection.....	1806
<i>Jiayi Gao, Kongming Liang, Tao Wei, Wei Chen, Zhanyu Ma, Jun Guo</i>	
LAMM: Label Alignment for Multi-Modal Prompt Learning	1815
<i>Jingsheng Gao, Jiacheng Ruan, Suncheng Xiang, Zefang Yu, Ke Ji, Mingye Xie, Ting Liu, Yuzhuo Fu</i>	
Frequency-Controlled Diffusion Model for Versatile Text-Guided Image-To-Image Translation.....	1824
<i>Xiang Gao, Zhengbo Xu, Junhan Zhao, Jiaying Liu</i>	
A General Implicit Framework for Fast NeRF Composition and Rendering	1833
<i>Xinyu Gao, Ziyi Yang, Yunlu Zhao, Yuxiang Sun, Xiaogang Jin, Changqing Zou</i>	
Multi-Scene Generalized Trajectory Global Graph Solver with Composite Nodes for Multiple Object Tracking	1842
<i>Yan Gao, Haojun Xu, Jie Li, Nannan Wang, Xinbo Gao</i>	
A Dual Stealthy Backdoor: From Both Spatial and Frequency Perspectives	1851
<i>Yudong Gao, Honglong Chen, Peng Sun, Junjian Li, Anqing Zhang, Zhibo Wang, Weifeng Liu</i>	
SoftCLIP: Softer Cross-Modal Alignment Makes CLIP Stronger.....	1860
<i>Yuting Gao, Jinfeng Liu, Zihan Xu, Tong Wu, Enwei Zhang, Ke Li, Jie Yang, Wei Liu, Xing Sun</i>	
Composite Sketch+Text Queries for Retrieving Objects with Elusive Names and Complex Interactions	1869
<i>Prajwal Gatti, Kshitij Parikh, Dhriti Prasanna Paul, Manish Gupta, Anand Mishra</i>	
Neuromorphic Event Signal-Driven Network for Video De-Raining.....	1878
<i>Chengjie Ge, Xueyang Fu, Peng He, Kunyu Wang, Chengzhi Cao, Zheng-Jun Zha</i>	
Beyond Prototypes: Semantic Anchor Regularization for Better Representation Learning	1887
<i>Yanqi Ge, Qiang Nie, Ye Huang, Yong Liu, Chengjie Wang, Feng Zheng, Wen Li, Lixin Duan</i>	
Learning Multi-Scale Video-Text Correspondence for Weakly Supervised Temporal Article Grounding.....	1896
<i>Wenjia Geng, Yong Liu, Lei Chen, Sujia Wang, Jie Zhou, Yansong Tang</i>	

PoseGen: Learning to Generate 3D Human Pose Dataset with NeRF	1905
<i>Mohsen Gholami, Rabab Ward, Z. Jane Wang</i>	
SDAC: A Multimodal Synthetic Dataset for Anomaly and Corner Case Detection in Autonomous Driving	1914
<i>Lei Gong, Yu Zhang, Yingqing Xia, Yanyong Zhang, Jianmin Ji</i>	
ContactGen: Contact-Guided Interactive 3D Human Generation for Partners.....	1923
<i>Dongjun Gu, Jaehyeok Shim, Jaehoon Jang, Changwoo Kang, Kyungdon Joo</i>	
AnomalyGPT: Detecting Industrial Anomalies Using Large Vision-Language Models	1932
<i>Zhaopeng Gu, Bingke Zhu, Guibo Zhu, Yingying Chen, Ming Tang, Jinqiao Wang</i>	
SeqRank: Sequential Ranking of Salient Objects.....	1941
<i>Huankang Guan, Rynson W.H. Lau</i>	
Knowledge-Aware Neuron Interpretation for Scene Classification.....	1950
<i>Yong Guan, Freddy Lécué, Jiaoyan Chen, Ru Li, Jeff Z. Pan</i>	
Self-Supervised Representation Learning with Meta Comprehensive Regularization	1959
<i>Huijie Guo, Ying Ba, Jie Hu, Lingyu Si, Wenwen Qiang, Lei Shi</i>	
Graph Context Transformation Learning for Progressive Correspondence Pruning	1968
<i>Junwen Guo, Guobao Xiao, Shiping Wang, Jun Yu</i>	
Depth-Guided Robust and Fast Point Cloud Fusion NeRF for Sparse Input Views.....	1976
<i>Shuai Guo, Qiuwen Wang, Yijie Gao, Rong Xie, Li Song</i>	
Improving Panoptic Narrative Grounding by Harnessing Semantic Relationships and Visual Confirmation	1985
<i>Tianyu Guo, Haowei Wang, Yiwei Ma, Jiayi Ji, Xiaoshuai Sun</i>	
Learning to Manipulate Artistic Images	1994
<i>Wei Guo, Yuqi Zhang, De Ma, Qian Zheng</i>	
PICNN: A Pathway Towards Interpretable Convolutional Neural Networks.....	2003
<i>Wengang Guo, Jiayi Yang, Huilin Yin, Qijun Chen, Wei Ye</i>	
GSN: Generalisable Segmentation in Neural Radiance Field	2013
<i>Vinayak Gupta, Rahul Goel, Sirikonda Dhawal, P. J. Narayanan</i>	
AMD: Autoregressive Motion Diffusion.....	2022
<i>Bo Han, Hao Peng, Minjing Dong, Yi Ren, Yixuan Shen, Chang Xu</i>	
HuTuMotion: Human-Tuned Navigation of Latent Motion Diffusion Models with Minimal Feedback.....	2031
<i>Gaoge Han, Shaoli Huang, Mingming Gong, Jinglei Tang</i>	
MA-Net: Rethinking Neural Unit in the Light of Astrocytes	2040
<i>Mengqiao Han, Liyuan Pan, Xiabi Liu</i>	
Dual-Perspective Knowledge Enrichment for Semi-Supervised 3D Object Detection	2049
<i>Yucheng Han, Na Zhao, Weiling Chen, Keng Teck Ma, Hanwang Zhang</i>	
Exploiting the Social-Like Prior in Transformer for Visual Reasoning.....	2058
<i>Yudong Han, Yupeng Hu, Xueming Song, Haoyu Tang, Mingzhu Xu, Liqiang Nie</i>	

Improving Audio-Visual Segmentation with Bidirectional Generation.....	2067
<i>Dawei Hao, Yuxin Mao, Bowen He, Xiaodong Han, Yuchao Dai, Yiran Zhong</i>	
Hand-Centric Motion Refinement for 3D Hand-Object Interaction Via Hierarchical Spatial-Temporal Modeling	2076
<i>Yuze Hao, Jianrong Zhang, Tao Zhuo, Fuan Wen, Hehe Fan</i>	
Progressive Feature Self-Reinforcement for Weakly Supervised Semantic Segmentation	2085
<i>Jingxuan He, Lechao Cheng, Chaowei Fang, Zunlei Feng, Tingting Mu, Mingli Song</i>	
Prompting Multi-Modal Image Segmentation with Semantic Grouping.....	2094
<i>Qibin He</i>	
Low-Latency Space-Time Supersampling for Real-Time Rendering.....	2103
<i>Ruian He, Shili Zhou, Yuqi Sun, Ri Cheng, Weimin Tan, Bo Yan</i>	
Collaborative Weakly Supervised Video Correlation Learning for Procedure-Aware Instructional Video Analysis.....	2112
<i>Tianyao He, Huabin Liu, Yuxi Li, Xiao Ma, Cheng Zhong, Yang Zhang, Weiyao Lin</i>	
Frequency-Adaptive Pan-Sharpening with Mixture of Experts.....	2121
<i>Xuanhua He, Keyu Yan, Rui Li, Chengjun Xie, Jie Zhang, Man Zhou</i>	
Enhancing RAW-To-SRGB with Decoupled Style Structure in Fourier Domain	2130
<i>Xuanhua He, Tao Hu, Guoli Wang, Zejin Wang, Run Wang, Qian Zhang, Keyu Yan, Ziyi Chen, Rui Li, Chengjun Xie, Jie Zhang, Man Zhou</i>	
A User-Friendly Framework for Generating Model-Preferred Prompts in Text-To-Image Synthesis.....	2139
<i>Nailei Hei, Qianyu Guo, Zihao Wang, Yan Wang, Haofen Wang, Wenqiang Zhang</i>	
Optimize & Reduce: A Top-Down Approach for Image Vectorization	2148
<i>Or Hirschorn, Amir Jevnisek, Shai Avidan</i>	
MotionMix: Weakly-Supervised Diffusion for Controllable Motion Generation.....	2157
<i>Nhat M. Hoang, Kehong Gong, Chuan Guo, Michael Bi Mi</i>	
Commonsense for Zero-Shot Natural Language Video Localization	2166
<i>Meghana Holla, Ismini Lourentzou</i>	
Learning Subject-Aware Cropping by Outpainting Professional Photos.....	2175
<i>James Hong, Lu Yuan, Michaël Gharbi, Matthew Fisher, Kayvon Fatahalian</i>	
High-Fidelity Diffusion-Based Image Editing	2184
<i>Chen Hou, Guoqiang Wei, Zhibo Chen</i>	
Domain-Hallucinated Updating for Multi-Domain Face Anti-Spoofing.....	2193
<i>Chengyang Hu, Ke-Yue Zhang, Taiping Yao, Shice Liu, Shouhong Ding, Xin Tan, Lizhuang Ma</i>	
QI-IRA: Quantum-Inspired Interactive Ranking Aggregation for Person Re-Identification.....	2202
<i>Chunyu Hu, Hong Zhang, Chao Liang, Hao Huang</i>	
SpaceGTN: A Time-Agnostic Graph Transformer Network for Handwritten Diagram Recognition and Segmentation	2211
<i>Haoxiang Hu, Cangjun Gao, Yaokun Li, Xiaoming Deng, YuKun Lai, Cuixia Ma, Yong-Jin Liu, Hongan Wang</i>	

Learning Explicit Contact for Implicit Reconstruction of Hand-Held Objects from Monocular Images	2220
<i>Junxing Hu, Hongwen Zhang, Zerui Chen, Mengcheng Li, Yunlong Wang, Yebin Liu, Zhenan Sun</i>	
DALDet: Depth-Aware Learning Based Object Detection for Autonomous Driving	2229
<i>Ke Hu, Tongbo Cao, Yuan Li, Song Chen, Yi Kang</i>	
COMMA: Co-Articulated Multi-Modal Learning.....	2238
<i>Lianyu Hu, Liqing Gao, Zekang Liu, Chi-Man Pun, Wei Feng</i>	
Latent Space Editing in Transformer-Based Flow Matching.....	2247
<i>Vincent Tao Hu, Wei Zhang, Meng Tang, Pascal Mettes, Deli Zhao, Cees Snoek</i>	
BLIVA: A Simple Multimodal LLM for Better Handling of Text-Rich Visual Questions	2256
<i>Wenbo Hu, Yifan Xu, Yi Li, Weiyue Li, Zeyuan Chen, Zhuowen Tu</i>	
A Dynamic Learning Method Towards Realistic Compositional Zero-Shot Learning.....	2265
<i>Xiaoming Hu, Zilei Wang</i>	
LF-ViT: Reducing Spatial Redundancy in Vision Transformer for Efficient Image Recognition	2274
<i>Youbing Hu, Yun Cheng, Anqi Lu, Zhiqiang Cao, Dawei Wei, Jie Liu, Zhijun Li</i>	
O ² -Recon: Completing 3D Reconstruction of Occluded Objects in the Scene with a Pre-Trained 2D Diffusion Model	2285
<i>Yubin Hu, Sheng Ye, Wang Zhao, Matthieu Lin, Yuze He, Yu-Hui Wen, Ying He, Yong-Jin Liu</i>	
Arbitrary-Scale Video Super-Resolution Guided by Dynamic Context	2294
<i>Cong Huang, Jiahao Li, Lei Chu, Dong Liu, Yan Lu</i>	
Dynamic Weighted Combiner for Mixed-Modal Image Retrieval	2303
<i>Fuxiang Huang, Lei Zhang, Xiaowei Fu, Suqi Song</i>	
NeuSurf: On-Surface Priors for Neural Surface Reconstruction from Sparse Input Views	2312
<i>Han Huang, Yulun Wu, Junsheng Zhou, Ge Gao, Ming Gu, Yu-Shen Liu</i>	
Seeing Dark Videos Via Self-Learned Bottleneck Neural Representation	2321
<i>Haofeng Huang, Wenhan Yang, Lingyu Duan, Jiaying Liu</i>	

PART 2

Combinatorial CNN-Transformer Learning with Manifold Constraints for Semi-Supervised Medical Image Segmentation.....	2330
<i>Huimin Huang, Yawen Huang, Shiao Xie, Lanfen Lin, Ruofeng Tong, Yen-Wei Chen, Yuexiang Li, Yefeng Zheng</i>	
Sparse Bayesian Deep Learning for Cross Domain Medical Image Reconstruction.....	2339
<i>Jiaxin Huang, Qi Wu, Yazhou Ren, Fan Yang, Aodi Yang, Qianqian Yang, Xiaorong Pu</i>	
UniCell: Universal Cell Nucleus Classification Via Prompt Learning.....	2348
<i>Junjia Huang, Haofeng Li, Xiang Wan, Guanbin Li</i>	
SC-NeuS: Consistent Neural Surface Reconstruction from Sparse and Noisy Views.....	2357
<i>Shi-Sheng Huang, Zixin Zou, Yichi Zhang, Yan-Pei Cao, Ying Shan</i>	

MFTN: Multi-Level Feature Transfer Network Based on MRI-Transformer for MR Image Super-Resolution.....	2366
<i>Shuying Huang, Ge Chen, Yong Yang, Xiaozheng Wang, Chenbin Liang</i>	
SDGAN: Disentangling Semantic Manipulation for Facial Attribute Editing.....	2374
<i>Wenmin Huang, Weiqi Luo, Jiwu Huang, Xiaochun Cao</i>	
Frozen CLIP Transformer is an Efficient Point Cloud Encoder	2382
<i>Xiaoshui Huang, Zhou Huang, Sheng Li, Wentao Qu, Tong He, Yuenan Hou, Yifan Zuo, Wanli Ouyang</i>	
G2L-CariGAN: Caricature Generation from Global Structure to Local Features	2391
<i>Xin Huang, Yunfeng Bai, Dong Liang, Feng Tian, Jinyuan Jia</i>	
3D Visibility-Aware Generalizable Neural Radiance Fields for Interacting Hands.....	2400
<i>Xuan Huang, Hanhui Li, Zejun Yang, Zhisheng Wang, Xiaodan Liang</i>	
Sunshine to Rainstorm: Cross-Weather Knowledge Distillation for Robust 3D Object Detection	2409
<i>Xun Huang, Hai Wu, Xin Li, Xiaoliang Fan, Chenglu Wen, Cheng Wang</i>	
Structure-CLIP: Towards Scene Graph Knowledge to Enhance Multi-Modal Structured Representations	2417
<i>Yufeng Huang, Jiji Tang, Zhuo Chen, Rongsheng Zhang, Xinfeng Zhang, Weijie Chen, Zeng Zhao, Zhou Zhao, Tangjie Lv, Zhipeng Hu, Wen Zhang</i>	
Voxel Or Pillar: Exploring Efficient Point Cloud Representation for 3D Object Detection.....	2426
<i>Yuhao Huang, Sanping Zhou, Junjie Zhang, Jinpeng Dong, Nanning Zheng</i>	
COMBAT: Alternated Training for Effective Clean-Label Backdoor Attacks	2436
<i>Tran Huynh, Dang Nguyen, Tung Pham, Anh Tran</i>	
MagiCapture: High-Resolution Multi-Concept Portrait Customization.....	2445
<i>Junha Hyung, Jaeyo Shin, Jaegul Choo</i>	
Rethinking Peculiar Images by Diffusion Models: Revealing Local Minima’s Role.....	2454
<i>Jinhyeok Jang, Chan-Hyun Youn, Minsu Jeon, Changha Lee</i>	
ProxyDet: Synthesizing Proxy Novel Classes Via Classwise Mixup for Open-Vocabulary Object Detection	2462
<i>Joonhyun Jeong, Geondo Park, Jayeon Yoo, Hyungsik Jung, Heesu Kim</i>	
A Diffusion Model with State Estimation for Degradation-Blind Inverse Imaging	2471
<i>Liya Ji, Zhefan Rao, Sinno Jialin Pan, Chenyang Lei, Qifeng Chen</i>	
SSMG: Spatial-Semantic Map Guided Diffusion Model for Free-Form Layout-To-Image Generation	2480
<i>Chengyou Jia, Minnan Luo, Zhuohang Dang, Guang Dai, Xiaojun Chang, Mengmeng Wang, Jingdong Wang</i>	
TiMix: Text-Aware Image Mixing for Effective Vision-Language Pre-Training	2489
<i>Chaoya Jiang, Wei Ye, Haiyang Xu, Qinghao Ye, Ming Yan, Ji Zhang, Shikun Zhang</i>	
Revealing the Proximate Long-Tail Distribution in Compositional Zero-Shot Learning.....	2498
<i>Chenyi Jiang, Haofeng Zhang</i>	

MWSIS: Multimodal Weakly Supervised Instance Segmentation with 2D Box Annotations for Autonomous Driving	2507
<i>Guangfeng Jiang, Jun Liu, Yuzhi Wu, Wenlong Liao, Tao He, Pai Peng</i>	
Transferable Video Moment Localization by Moment-Guided Query Prompting	2516
<i>Hao Jiang, Yang Yizhang, Yadong Mu</i>	
In-Hand 3D Object Reconstruction from a Monocular RGB Video	2525
<i>Shijian Jiang, Qi Ye, Rengan Xie, Yuchi Huo, Xiang Li, Yang Zhou, Jiming Chen</i>	
AACP: Aesthetics Assessment of Children’s Paintings Based on Self-Supervised Learning	2534
<i>Shiqi Jiang, Ning Li, Chen Shi, Liping Guo, Changbo Wang, Chenhui Li</i>	
Exploring Self- And Cross-Triplet Correlations for Human-Object Interaction Detection	2543
<i>Weibo Jiang, Weihong Ren, Jiandong Tian, Liangqiong Qu, Zhiyong Wang, Honghai Liu</i>	
Comprehensive Visual Grounding for Video Description	2552
<i>Wenhui Jiang, Yibo Cheng, Linxin Liu, Yuming Fang, Yuxin Peng, Yang Liu</i>	
Far3D: Expanding the Horizon for Surround-View 3D Object Detection.....	2561
<i>Xiaohui Jiang, Shuailin Li, Yingfei Liu, Shihao Wang, Fan Jia, Tiancai Wang, Lijin Han, Xiangyu Zhang</i>	
Delving into Multimodal Prompting for Fine-Grained Visual Classification.....	2570
<i>Xin Jiang, Hao Tang, Junyao Gao, Xiaoyu Du, Shengfeng He, Zechao Li</i>	
MCA: Moment Channel Attention Networks.....	2579
<i>Yangbo Jiang, Zhiwei Jiang, Le Han, Zenan Huang, Nenggan Zheng</i>	
Towards Robust Image Stitching: An Adaptive Resistance Learning Against Compatible Attacks.....	2589
<i>Zhiying Jiang, Xingyuan Li, Jinyuan Liu, Xin Fan, Risheng Liu</i>	
Instance-Aware Multi-Camera 3D Object Detection with Structural Priors Mining and Self-Boosting Learning	2598
<i>Yang Jiao, Zequn Jie, Shaoxiang Chen, Lechao Cheng, Jingjing Chen, Lin Ma, Yu-Gang Jiang</i>	
PromptMRG: Diagnosis-Driven Prompts for Medical Report Generation.....	2607
<i>Haibo Jin, Haoxuan Che, Yi Lin, Hao Chen</i>	
PCE-Palm: Palm Crease Energy Based Two-Stage Realistic Pseudo-Palmprint Generation.....	2616
<i>Jianlong Jin, Lei Shen, Ruixin Zhang, Chenglong Zhao, Ge Jin, Jingyun Zhang, Shouhong Ding, Yang Zhao, Wei Jia</i>	
SwiftPillars: High-Efficiency Pillar Encoder for Lidar-Based 3D Detection	2625
<i>Xin Jin, Kai Liu, Cong Ma, Ruining Yang, Fei Hui, Wei Wu</i>	
DeS3: Adaptive Attention-Driven Self and Soft Shadow Removal Using ViT Similarity	2634
<i>Yeying Jin, Wei Ye, Wenhan Yang, Yuan Yuan, Robby T. Tan</i>	
AMD: Anatomical Motion Diffusion with Interpretable Motion Decomposition and Fusion.....	2643
<i>Beibei Jing, Youjia Zhang, Zikai Song, Junqing Yu, Wei Yang</i>	
Retrieval-Augmented Primitive Representations for Compositional Zero-Shot Learning.....	2652
<i>Chenchen Jing, Yukun Li, Hao Chen, Chunhua Shen</i>	

CrossBind: Collaborative Cross-Modal Identification of Protein Nucleic-Acid-Binding Residues.....	2661
<i>Linglin Jing, Sheng Xu, Yifan Wang, Yuzhe Zhou, Tao Shen, Zhigang Ji, Hui Fang, Zhen Li, Siqu Sun</i>	
X4D-SceneFormer: Enhanced Scene Understanding on 4D Point Cloud Videos Through Cross-Modal Knowledge Transfer	2670
<i>Linglin Jing, Ying Xue, Xu Yan, Chaoda Zheng, Dong Wang, Ruimao Zhang, Zhigang Wang, Hui Fang, Bin Zhao, Zhen Li</i>	
VVS: Video-To-Video Retrieval with Irrelevant Frame Suppression.....	2679
<i>Won Jo, Geuntaek Lim, Gwangjin Lee, Hyunwoo Kim, Byungsoo Ko, Yukyung Choi</i>	
Rethinking Robustness of Model Attributions	2688
<i>Sandesh Kamath, Sankalp Mittal, Amit Deshpande, Vineeth N Balasubramanian</i>	
Cross-Constrained Progressive Inference for 3D Hand Pose Estimation with Dynamic Observer-Decision-Adjuster Networks	2697
<i>Zehan Kan, Xueting Hu, Zihan Liao, Ke Yu, Zhihai He</i>	
Catch-Up Mix: Catch-Up Class for Struggling Filters in CNN.....	2705
<i>Minsoo Kang, Minkoo Kang, Suhyun Kim</i>	
VLCounter: Text-Aware Visual Representation for Zero-Shot Object Counting	2714
<i>Seunggu Kang, WonJun Moon, Euiyeon Kim, Jae-Pil Heo</i>	
StegFormer: Rebuilding the Glory of Autoencoder-Based Steganography.....	2723
<i>Xiao Ke, Huanqi Wu, Wenzhong Guo</i>	
Expediting Contrastive Language-Image Pretraining Via Self-Distilled Encoders.....	2732
<i>Bumsoo Kim, Jinhyung Kim, Yeonsik Jo, Seung Hwan Kim</i>	
Weakly Supervised Semantic Segmentation for Driving Scenes.....	2741
<i>Dongseob Kim, Seungho Lee, Junsuk Choe, Hyunjung Shim</i>	
FPRF: Feed-Forward Photorealistic Style Transfer of Large-Scale 3D Neural Radiance Fields	2750
<i>GeonU Kim, Kim Youwang, Tae-Hyun Oh</i>	
Let There Be Sound: Reconstructing High Quality Speech from Silent Videos.....	2759
<i>Ji-Hoon Kim, Jaehun Kim, Joon Son Chung</i>	
Expand-And-Quantize: Unsupervised Semantic Segmentation Using High-Dimensional Space and Product Quantization	2768
<i>Jiyoung Kim, Kyuhong Shim, Insu Lee, Byonghyo Shim</i>	
Sync-NeRF: Generalizing Dynamic NeRFs to Unsynchronized Videos.....	2777
<i>Seoha Kim, Jeongmin Bae, Youngsik Yun, Hahyun Lee, Gun Bang, Youngjung Uh</i>	
Improving Open Set Recognition Via Visual Prompts Distilled from Common-Sense Knowledge	2786
<i>Seongyeop Kim, Hyung-Il Kim, Yong Man Ro</i>	
Gaussian Mixture Proposals with Pull-Push Learning Scheme to Capture Diverse Events for Weakly Supervised Temporal Video Grounding.....	2795
<i>Sunoh Kim, Jungchan Cho, Joonsang Yu, YoungJoon Yoo, Jin Young Choi</i>	
PARSAC: Accelerating Robust Multi-Model Fitting with Parallel Sample Consensus	2804
<i>Florian Kluger, Bodo Rosenhahn</i>	

Distribution Matching for Multi-Task Learning of Classification Tasks: A Large-Scale Study on Faces & Beyond	2813
<i>Dimitrios Kollias, Viktoriia Sharmanska, Stefanos Zafeiriou</i>	
Block Image Compressive Sensing with Local and Global Information Interaction	2822
<i>Xiaoyu Kong, Yongyong Chen, Feng Zheng, Zhenyu He</i>	
QDETRv: Query-Guided DETR for One-Shot Object Localization in Videos	2831
<i>Yogesh Kumar, Saswat Mallick, Anand Mishra, Sowmya Rasipuram, Anutosh Maitra, Roshni Ramnani</i>	
LaViP: Language-Grounded Visual Prompting	2840
<i>Nilakshan Kunanathaseelan, Jing Zhang, Mehrtash Harandi</i>	
Towards More Faithful Natural Language Explanation Using Multi-Level Contrastive Learning in VQA	2849
<i>Chengen Lai, Shengli Song, Shiqi Meng, Jingyang Li, Sitong Yan, Guangneng Hu</i>	
MatchDet: A Collaborative Framework for Image Matching and Object Detection	2858
<i>Jinxiang Lai, Wenlong Wu, Bin-Bin Gao, Jun Liu, Jiawei Zhan, Congchong Nie, Yi Zeng, Chengjie Wang</i>	
ViTree: Single-Path Neural Tree for Step-Wise Interpretable Fine-Grained Visual Categorization.....	2866
<i>Danning Lao, Qi Liu, Jiazi Bu, Junchi Yan, Wei Shen</i>	
MaskDiff: Modeling Mask Distribution with Diffusion Probabilistic Model for Few-Shot Instance Segmentation	2874
<i>Minh-Quan Le, Tam V. Nguyen, Trung-Nghia Le, Thanh-Toan Do, Minh N. Do, Minh-Triet Tran</i>	

Author Index