


▼ Stop the war!

🔍 ✖
[+] [-] **ICCV 2023: Paris, France** ⬇️ 🔗 📄

> Home > Conferences and Workshops > ICCV

☰ Dagstuhl

- 📃 ⬇️ 🔗 🔗 🔗 IEEE/CVF International Conference on Computer Vision, ICCV 2023, Paris, France, October 1-6, 2023. IEEE 2023, ISBN 979-8-3503-0718-4
- 📃 ⬇️ 🔗 🔗 🔗 Xinyang Liu, Yijin Li, Yanbin Teng, Hujun Bao, Guofeng Zhang, Yinda Zhang, Zhaopeng Cui: **Multi-Modal Neural Radiance Field for Monocular Dense SLAM with a Light-Weight ToF Sensor.** 1-11
- 📃 ⬇️ 🔗 🔗 🔗 Chandan Yeshwanth, Yueh-Cheng Liu, Matthias Nießner, Angela Dai: **ScanNet++: A High-Fidelity Dataset of 3D Indoor Scenes.** 12-22
- 📃 ⬇️ 🔗 🔗 🔗 Jiachen Lu, Hongyang Li, Renyuan Peng, Feng Wen, Xinyue Cai, Wei Zhang, Hang Xu, Li Zhang: **Translating Images to Road Network: A Non-Autoregressive Sequence-to-Sequence Approach.** 23-33
- 📃 ⬇️ 🔗 🔗 🔗 Ruojin Cai, Joseph Tung, Qianqian Wang, Hadar Averbuch-Elor, Bharath Hariharan, Noah Snavely: **Doppelgangers: Learning to Disambiguate Images of Similar Structures.** 34-44
- 📃 ⬇️ 🔗 🔗 🔗 Jinjie Mai, Abdullah Hamdi , Silvio Giancola, Chen Zhao, Bernard Ghanem : **EgoLoc: Revisiting 3D Object Localization from Egocentric Videos with Visual Queries.** 45-57
- 📃 ⬇️ 🔗 🔗 🔗 Wenqiang Xu, Wenxin Du, Han Xue, Yutong Li, Ruolin Ye, Yan-Feng Wang, Cewu Lu: **ClothPose: A Real-world Benchmark for Visual Analysis of Garment Pose via An Indirect Recording Solution.** 58-68
-
- [-] **Refine list**
- showing all 2,157 records
- refine by search term**
- refine by author**
- Luc Van Gool (15)
 - Chen Change Loy (14)
 - Yi Yang 0001 (14)
 - Marc Pollefeys (13)
 - Jingdong Wang 0001 (13)
 - Ming-Hsuan Yang 0001 (12)
 - Zheng Shou 0001 (12)
 - Shuaicheng Liu (12)
 - Lei Zhang (11)
 - Deli Zhao (10)
 - Bernt Schiele (10)
 - 8,326 more options
- refine by access**
closed (2,156)



- Zijie Jiang, Masatoshi Okutomi:
EMR-MSF: Self-Supervised Recurrent Monocular Scene Flow Exploiting Ego-Motion Rigidity. 69-78
- Ruofan Liang, Huiting Chen, Chunlin Li, Fan Chen, Selvakumar Panneer, Nandita Vijaykumar:
ENVIDR: Implicit Differentiable Renderer with Neural Environment Lighting. 79-89
- Yihua Zhang, Ruisi Cai, Tianlong Chen, Guanhua Zhang, Huan Zhang, Pin-Yu Chen, Shiyu Chang, Zhangyang Wang, Sijia Liu:
Robust Mixture-of-Expert Training for Convolutional Neural Networks. 90-101
- Dong Lu, Zhiqiang Wang, Teng Wang, Weili Guan, Hongchang Gao, Feng Zheng:
Set-level Guidance Attack: Boosting Adversarial Transferability of Vision-Language Pre-training Models. 102-111
- Hritik Bansal, Fan Yin, Nishad Singh, Aditya Grover, Yu Yang, Kai-Wei Chang:
CleanCLIP: Mitigating Data Poisoning Attacks in Multimodal Contrastive Learning. 112-123
- Md Farhamdur Reza, Ali Rahmati, Tianfu Wu, Huaiyu Dai:
CGBA: Curvature-aware Geometric Black-box Attack. 124-133
- Minjong Lee, Dongwoo Kim:
Robust Evaluation of Diffusion-Based Adversarial Purification. 134-144
- Yao Ge, Yun Li, Keji Han, Junyi Zhu, Xianzhong Long:
Advancing Example Exploitation Can Alleviate Critical Challenges in Adversarial Training. 145-154
- Zixuan Zhu, Rui Wang, Cong Zou, Lihua Jing:
The Victim and The Beneficiary: Exploiting a Poisoned Model to Train a Clean Model on Poisoned Data. 155-164
- Indranil Sur, Karan Sikka, Matthew Walmer, Kaushik Koneripalli, Anirban Roy, Xiao Lin, Ajay Divakaran, Susmit Jha:
TIJO: Trigger Inversion with Joint Optimization for Defending Multimodal Backdoored Models. 165-175
- Yangru Huang, Peixi Peng, Yifan Zhao, Yunpeng Zhai, Haoran Xu, Yonghong Tian:
Simoun: Synergizing Interactive Motion-appearance Understanding for Vision-based Reinforcement Learning. 176-185
- Yiming Li, Qi Fang, Jiamu Bai, Siheng Chen, Felix Juefei-Xu, Chen Feng:
Among Us: Adversarially Robust Collaborative Perception by Consensus. 186-195
- Cristiano Saltori, Aljosa Osep, Elisa Ricci, Laura Leal-Taixé:
Walking Your LiDOG: A Journey Through Multiple Domains for LiDAR Semantic Segmentation. 196-206
- Yunpeng Zhai, Peixi Peng, Yifan Zhao, Yangru Huang, Yonghong Tian:
Stabilizing Visual Reinforcement Learning via Asymmetric Interactive Cooperation. 207-216

-      Yuanzhi Liang, Xiaohan Wang, Linchao Zhu, Yi Yang:
MAAL: Multimodality-Aware Autoencoder-based Affordance Learning for 3D Articulated Objects. 217-227
-     Lingdong Kong, Youquan Liu, Runnan Chen, Yuexin Ma, Xinge Zhu, Yikang Li, Yuenan Hou, Yu Qiao, Ziwei Liu:
Rethinking Range View Representation for LiDAR Segmentation. 228-240
-     Haitao Lin, Yanwei Fu, Xiangyang Xue:
PourIt!: Weakly-supervised Liquid Perception from a Single Image for Visual Closed-Loop Robotic Pouring. 241-251
-     Arthur Moreau, Nathan Piasco, Moussâb Bennehar, Dzmitry Tsishkou, Bogdan Stanciulescu, Arnaud de La Fortelle:
CROSSFIRE: Camera Relocalization On Self-Supervised Features from an Implicit Representation. 252-262
-     Hyesong Choi, Hunsang Lee, Seongwon Jeong, Dongbo Min:
Environment Agnostic Representation for Visual Reinforcement learning. 263-273
-     Qiongjie Cui, Huaijiang Sun, Jianfeng Lu, Weiqing Li, Bin Li, Hongwei Yi, Haofan Wang:
Test-time Personalizable Forecasting of 3D Human Poses. 274-283
-     Hao Xiang, Runsheng Xu, Jiaqi Ma:
HM-ViT: Hetero-modal Vehicle-to-Vehicle Cooperative Perception with Vision Transformer. 284-295
-     Antoine Mercier, Ruan Erasmus, Yashesh Savani, Manik Dhingra, Fatih Porikli, Guillaume Berger:
Efficient neural supersampling on a novel gaming dataset. 296-306
-     Hong-Wing Pang, Binh-Son Hua, Sai-Kit Yeung:
Locally Stylized Neural Radiance Fields. 307-316
-     Dongqing Wang, Tong Zhang, Sabine Süsstrunk:
NEMTO: Neural Environment Matting for Novel View and Relighting Synthesis of Transparent Objects. 317-327
-     Xiaoyang Kang, Tao Yang, Wenqi Ouyang, Peiran Ren, Lingzhi Li, Xuansong Xie:
DDColor: Towards Photo-Realistic Image Colorization via Dual Decoders. 328-338
-     Weicai Ye, Shuo Chen, Chong Bao, Hujun Bao, Marc Pollefeys, Zhaopeng Cui, Guofeng Zhang:
IntrinsicNeRF: Learning Intrinsic Neural Radiance Fields for Editable Novel View Synthesis. 339-351
-     Jiayi Liu, Ali Mahdavi-Amiri, Manolis Savva:
PARIS: Part-level Reconstruction and Motion Analysis for Articulated Objects. 352-363
-     Mingyuan Zhang, Xinying Guo, Liang Pan, Zhongang Cai, Fangzhou Hong, Huirong Li, Lei Yang, Ziwei Liu:
ReMoDiffuse: Retrieval-Augmented Motion Diffusion Model. 364-373

-      Maham Tanveer, Yizhi Wang, Ali Mahdavi-Amiri, Hao Zhang:
DS-Fusion: Artistic Typography via Discriminated and Stylized Diffusion.
374-384
-     Yi-Ling Qiao, Alexander Gao, Yiran Xu, Yue Feng, Jia-Bin Huang, Ming C. Lin:
Dynamic Mesh-Aware Radiance Fields. 385-396
-     Wenzhang Sun, Yunlong Che, Yandong Guo, Han Huang:
Neural Reconstruction of Relightable Human Model from Monocular Video.
397-407
-     Alexander Mai, Dor Verbin, Falko Kuester, Sara Fridovich-Keil:
Neural Microfacet Fields for Inverse Rendering. 408-418
-     Ishit Mehta, Manmohan Chandraker, Ravi Ramamoorthi:
A Theory of Topological Derivatives for Inverse Rendering of Geometry.
419-429
-     Eta Sella, Gal Fiebelman, Peter Hedman, Hadar Averbuch-Elor:
Vox-E: Text-guided Voxel Editing of 3D Objects. 430-440
-     Chenxin Li, Brandon Y. Feng , Zhiwen Fan, Panwang Pan, Zhangyang Wang:
StegaNeRF: Embedding Invisible Information within Neural Radiance Fields. 441-453
-     Liu He, Daniel G. Aliaga:
GlobalMapper: Arbitrary-Shaped Urban Layout Generation. 454-464
-     Fan Lu, Yan Xu, Guang Chen, Hongsheng Li, Kwan-Yee Lin, Changjun Jiang:
Urban Radiance Field Representation with Deformable Neural Mesh Primitives. 465-476
-     Barbara Roessle, Matthias Nießner:
End2End Multi-View Feature Matching with Differentiable Pose Optimization. 477-487
-     Chen Geng , Hong-Xing Yu, Sharon Zhang, Maneesh Agrawala, Jiajun Wu:
Tree-Structured Shading Decomposition. 488-498
-     Dominique Piché-Meunier, Yannick Hold-Geoffroy, Jianming Zhang, Jean-François Lalonde:
Lens Parameter Estimation for Realistic Depth of Field Modeling. 499-508
-     Chongyang Zhong, Lei Hu, Zihao Zhang, Shihong Xia:
AttT2M: Text-Driven Human Motion Generation with Multi-Perspective Attention Mechanism. 509-519
-     Manuel Ladron De Guevara, Yannick Hold-Geoffroy, Jose Echevarria, Cameron Smith, Yijun Li, Daichi Ito:
Cross-modal Latent Space Alignment for Image to Avatar Translation. 520-529
-     Yibo Yang, Stephan Mandt:
Computationally-Efficient Neural Image Compression with Shallow Decoders. 530-540

-      Salwa K. Al Khatib , Mohamed El Amine Boudjoghra, Jean Lahoud, Fahad Shahbaz Khan:
3D Instance Segmentation via Enhanced Spatial and Semantic Supervision.
541-550
-     Zhijie Deng, Yucen Luo:
Learning Neural Eigenfunctions for Unsupervised Semantic Segmentation.
551-561
-     Weiguang Zhao, Yuyao Yan, Chaolong Yang, Jianan Ye, Xi Yang, Kaizhu Huang:
Divide and Conquer: 3D Point Cloud Instance Segmentation With Point-Wise Binarization.
562-571
-     Wentong Li, Yuqian Yuan, Song Wang, Jianke Zhu, Jianshu Li, Jian Liu, Lei Zhang:
Point2Mask: Point-supervised Panoptic Segmentation via Optimal Transport.
572-581
-     Sina Gholamian, Ali Vahdat:
Handwritten and Printed Text Segmentation: A Signature Case Study.
582-592
-     Siheyon Kim, Juyeon Ko, Minseok Joo, Juhan Cha, Jaewon Lee, Hyunwoo J. Kim:
Semantic-Aware Implicit Template Learning via Part Deformation Consistency.
593-603
-     Yunze Liu, Junyu Chen, Zekai Zhang, Jingwei Huang, Li Yi:
LeaF: Learning Frames for 4D Point Cloud Sequence Understanding.
604-613
-     Sanghyun Jo, In-Jae Yu, Kyungsu Kim:
MARS: Model-agnostic Biased Object Removal without Additional Supervision for Weakly-Supervised Semantic Segmentation.
614-623
-     Zelin Peng, Guanchun Wang, Lingxi Xie, Dongsheng Jiang, Wei Shen, Qi Tian:
USAGE: A Unified Seed Area Generation Paradigm for Weakly Supervised Semantic Segmentation.
624-634
-     Maksym Bekuzarov, Ariana Bermudez, Joon-Young Lee, Hao Li:
XMem++: Production-level Video Segmentation From Few Annotated Frames.
635-644
-     Maolin Gao, Paul Roetzer, Marvin Eisenberger, Zorah Lähner, Michael Möller, Daniel Cremers, Florian Bernard:
SIGMA: Scale-Invariant Global Sparse Shape Matching.
645-654
-     Qianxiong Xu, Wenting Zhao, Guosheng Lin, Cheng Long:
Self-Calibrated Cross Attention Network for Few-Shot Segmentation.
655-665
-     Kehan Li, Yian Zhao, Zhennan Wang, Zesen Cheng, Peng Jin, Xiangyang Ji, Li Yuan, Chang Liu, Jie Chen:
Multi-granularity Interaction Simulation for Unsupervised Interactive Segmentation.
666-676
-     Sunghwan Kim, Dae-Hwan Kim, Hoseong Kim:
Texture Learning Domain Randomization for Domain Generalized

-      **Segmentation.** 677-687
Tiankang Su, Huihui Song, Dong Liu, Bo Liu, Qingshan Liu:
Unsupervised Video Object Segmentation with Online Adversarial Self-Tuning. 688-698
-     Jun Chen, Deyao Zhu , Guocheng Qian , Bernard Ghanem , Zhicheng Yan, Chenchen Zhu, Fanyi Xiao, Sean Chang Culatana, Mohamed Elhoseiny :
Exploring Open-Vocabulary Semantic Segmentation from CLIP Vision Encoder Distillation Only. 699-710
-     Nazir Nayal, Misra Yavuz, João F. Henriques, Fatma Güney:
RbA: Segmenting Unknown Regions Rejected by All. 711-722
-     Sriram Ravindran, Debraj Basu:
Sempart: Self-supervised Multi-resolution Partitioning of Image Semantics. 723-733
-     Sadra Safadoust, Fatma Güney:
Multi-Object Discovery by Low-Dimensional Object Motion. 734-744
-     Enxu Li, Sergio Casas, Raquel Urtasun:
MemorySeg: Online LiDAR Semantic Segmentation with a Latent Memory. 745-754
-     Changwei Wang, Rongtao Xu, Shibiao Xu, Weiliang Meng, Xiaopeng Zhang:
Treating Pseudo-labels Generation as Image Matting for Weakly Supervised Semantic Segmentation. 755-765
-     Rui Yang, Lin Song, Yixiao Ge, Xiu Li:
BoxSnake: Polygonal Instance Segmentation with Box Supervision. 766-776
-     Quan Tang, Bowen Zhang, Jiajun Liu, Fagui Liu, Yifan Liu:
Dynamic Token Pruning in Plain Vision Transformers for Semantic Segmentation. 777-786
-     Yichen Liu, Benran Hu, Junkai Huang, Yu-Wing Tai, Chi-Keung Tang:
Instance Neural Radiance Field. 787-796
-     Kunyang Han, Yong Liu, Jun Hao Liew, Henghui Ding, Jiajun Liu, Yitong Wang, Yansong Tang, Yujiu Yang, Jiashi Feng, Yao Zhao, Yunchao Wei:
Global Knowledge Calibration for Fast Open-Vocabulary Segmentation. 797-807
-     Duo Peng, Ping Hu, Qiuhong Ke, Jun Liu:
Diffusion-based Image Translation with Label Guidance for Domain Adaptive Semantic Segmentation. 808-820
-     Yuhe Liu, Chuanjian Liu, Kai Han, Quan Tang, Zengchang Qin:
Boosting Semantic Segmentation from the Perspective of Explicit Class Embeddings. 821-831
-     Hala Lamdouar, Weidi Xie, Andrew Zisserman:
The Making and Breaking of Camouflage. 832-842
-     Zekang Zhang, Guangyu Gao, Jianbo Jiao , Chi Harold Liu, Yunchao Wei:
CoinSeg: Contrast Inter- and Intra- Class Representations for Incremental

**Segmentation.** 843-853

Xueyi Liu, Bin Wang, He Wang, Li Yi:

Few-Shot Physically-Aware Articulated Mesh Generation via Hierarchical Deformation. 854-864

Fenggen Yu, Yiming Qian, Francisca Gil-Ureta, Brian Jackson, Eric Bennett, Hao Zhang:

HAL3D: Hierarchical Active Learning for Fine-Grained 3D Part Labeling. 865-875

Tianyi Shi, Xiaohuan Ding, Liang Zhang, Xin Yang:

FreeCOS: Self-Supervised Learning from Fractals and Unlabeled Images for Curvilinear Object Segmentation. 876-886

Xin Xu, Tianyi Xiong, Zheng Ding, Zhuowen Tu:

MasQCLIP for Open-Vocabulary Universal Image Segmentation. 887-898

Kaining Ying, Qing Zhong, Weian Mao, Zhenhua Wang, Hao Chen, Lin Yuanbo Wu, Yifan Liu, Chengxiang Fan, Yunzhi Zhuge, Chunhua Shen:

CTVIS: Consistent Training for Online Video Instance Segmentation. 899-908

Ting Chen, Lala Li, Saurabh Saxena, Geoffrey E. Hinton, David J. Fleet:

A Generalist Framework for Panoptic Segmentation of Images and Videos.

909-919



Bo Miao, Mohammed Bennamoun, Yongsheng Gao, Ajmal Mian:

Spectrum-guided Multi-granularity Referring Video Object Segmentation.

920-930



Changqi Wang, Haoyu Xie, Yuhui Yuan, Chong Fu, Xiangyu Yue:

Space Engage: Collaborative Space Supervision for Contrastive-based Semi-Supervised Semantic Segmentation. 931-942

Hoyoung Kim, Minhyeon Oh, Sehyun Hwang, Suha Kwak, Jungseul Ok:

Adaptive Superpixel for Active Learning in Semantic Segmentation. 943-953

Yuxin Mao, Jing Zhang, Mochu Xiang, Yiran Zhong, Yuchao Dai:

Multimodal Variational Auto-encoder based Audio-Visual Segmentation.

954-965



Yichen Yuan, Yifan Wang, Lijun Wang, Xiaoqi Zhao, Huchuan Lu, Yu Wang, Weibo Su, Lei Zhang:

Isomer: Isomeric Transformer for Zero-shot Video Object Segmentation.

966-976



Cheng-Kun Yang, Min-Hung Chen, Yung-Yu Chuang, Yen-Yu Lin:

2D-3D Interlaced Transformer for Point Cloud Segmentation with Scene-Level Supervision. 977-987

Mischa Dombrowski, Hadrien Reynaud, Matthew Baugh, Bernhard Kainz

Foreground-Background Separation through Concept Distillation from Generative Image Foundation Models. 988-998

Muzhi Zhu, Hengtao Li, Hao Chen, Chengxiang Fan, Weian Mao, Chenchen Jing, Yifan Liu, Chunhua Shen:

**SegPrompt: Boosting Open-world Segmentation via Category-level Prompt Learning.** 999-1008

Boyang Li, Yingqian Wang, Longguang Wang, Fei Zhang, Ting Liu, Zaiping Lin, Wei An, Yulan Guo:

Monte Carlo Linear Clustering with Single-Point Supervision is Enough for Infrared Small Target Detection. 1009-1019

Hao Zhang, Feng Li, Xueyan Zou, Shilong Liu, Chunyuan Li, Jianwei Yang, Lei Zhang:

A Simple Framework for Open-Vocabulary Segmentation and Detection. 1020-1031

Zongwei Wu, Danda Pani Paudel , Deng-Ping Fan, Jingjing Wang, Shuo Wang, Cédric Demonceaux, Radu Timofte, Luc Van Gool:

Source-free Depth for Object Pop-out. 1032-1042

Amit Kumar Rana, Sabarinath Mahadevan, Alexander Hermans, Bastian Leibe:
DynaMITe: Dynamic Query Bootstrapping for Multi-object Interactive Segmentation Transformer. 1043-1052

Junzhang Chen, Xiangzhi Bai:
Atmospheric Transmission and Thermal Inertia Induced Blind Road Segmentation with a Large-Scale Dataset TBRSD. 1053-1063

Yuxi Wang, Jian Liang, Jun Xiao, Shuqi Mei, Yuran Yang, Zhaoxiang Zhang:
Informative Data Mining for One-shot Cross-Domain Semantic Segmentation. 1064-1074

Shan Wang, Chuong Nguyen, Jiawei Liu, Kaihao Zhang, Wenhan Luo, Yanhao Zhang, Sundaram Muthu, Fahira Afzal Maken, Hongdong Li:
Homography Guided Temporal Fusion for Road Line and Marking Segmentation. 1075-1085

Cong Han, Yujie Zhong, Dengjie Li, Kai Han, Lin Ma:
Open-Vocabulary Semantic Segmentation with Decoupled One-Pass Network. 1086-1096

Junlong Li, Bingyao Yu, Yongming Rao, Jie Zhou, Jiwen Lu:
TCOVIS: Temporally Consistent Online Video Instance Segmentation. 1097-1107

Liyi Chen, Chenyang Lei, Ruihuang Li, Shuai Li, Zhaoxiang Zhang, Lei Zhang:
FPR: False Positive Rectification for Weakly Supervised Semantic Segmentation. 1108-1118

Lukas Zbinden, Lars Doorenbos, Theodoros Pissas, Adrian Thomas Huber, Raphael Sznitman, Pablo Márquez-Neila:
Stochastic Segmentation with Conditional Categorical Diffusion Models. 1119-1129

Xinlong Wang, Xiaosong Zhang, Yue Cao, Wen Wang, Chunhua Shen, Tiejun Huang:
SegGPT: Towards Segmenting Everything In Context. 1130-1140

-      Xi Chen, Shuang Li, Ser-Nam Lim, Antonio Torralba, Hengshuang Zhao:
Open-vocabulary Panoptic Segmentation with Embedding Modulation.
1141-1150
-     Yuyuan Liu, Choubo Ding, Yu Tian, Guansong Pang, Vasileios Belagiannis, Ian D. Reid, Gustavo Carneiro:
Residual Pattern Learning for Pixel-wise Out-of-Distribution Detection in Semantic Segmentation. 1151-1161
-     Pitchaporn Rewatbowornwong, Nattanat Chatthee, Ekapol Chuangsawanich, Supasorn Suwajanakorn:
Zero-guidance Segmentation Using Zero Segment Labels. 1162-1172
-     Jiawei Liu, Changkun Ye, Shan Wang, Ruihai Cui, Jing Zhang, Kaihao Zhang, Nick Barnes :
Model Calibration in Dense Classification with Adaptive Label Perturbation. 1173-1184
-     Jie Ma, Chuan Wang, Yang Liu, Liang Lin, Guanbin Li:
Enhanced Soft Label for Semi-Supervised Semantic Segmentation. 1185-1195
-     Kaixin Cai, Pengzhen Ren , Yi Zhu, Hang Xu, Jianzhuang Liu, Changlin Li, Guangrun Wang, Xiaodan Liang:
MixReorg: Cross-Modal Mixed Patch Reorganization is a Good Mask Learner for Open-World Semantic Segmentation. 1196-1205
-     Weijia Wu, Yuzhong Zhao, Mike Zheng Shou, Hong Zhou, Chunhua Shen:
DiffuMask: Synthesizing Images with Pixel-level Annotations for Semantic Segmentation Using Diffusion Models. 1206-1217
-     Rui Sun, Yuan Wang, Huayu Mai, Tianzhu Zhang, Feng Wu:
Alignment Before Aggregation: Trajectory Memory Retrieval Network for Video Object Segmentation. 1218-1228
-     Peixia Li, Pulak Purkait, Thalaiyasingam Ajanthan, Majid Abdolshah, Ravi Garg, Hisham Husain, Chenchen Xu, Stephen Gould, Wanli Ouyang, Anton van den Hengel:
Semi-Supervised Semantic Segmentation under Label Noise via Diverse Learning Groups. 1229-1238
-     Cody Simons, Dripta S. Raychaudhuri, Sk Miraj Ahmed, Suya You, Konstantinos Karydis, Amit K. Roy-Chowdhury:
SUMMIT: Source-Free Adaptation of Uni-Modal Models to Multi-Modal Targets. 1239-1249
-     Yu-Hsing Hsieh, Guan-Sheng Chen, Shun-Xian Cai, Ting-Yun Wei, Huei-Fang Yang, Chu-Song Chen:
Class-incremental Continual Learning for Instance Segmentation with Image-level Weak Supervision. 1250-1261
-     Jianxiong Gao, Xuelin Qian, Yikai Wang, Tianjun Xiao, Tong He, Zheng Zhang, Yanwei Fu:
Coarse-to-Fine Amodal Segmentation with Shape Prior. 1262-1271

-      Ke Fan, Jingshi Lei, Xuelin Qian, Miaopeng Yu, Tianjun Xiao, Tong He, Zheng Zhang, Yanwei Fu:
Rethinking Amodal Video Segmentation from Learning Supervised Signals with Object-centric Representation. 1272-1281
-     Tao Zhang, Xingye Tian, Yu Wu, Shunping Ji, Xuebo Wang, Yuan Zhang, Pengfei Wan:
DVIS: Decoupled Video Instance Segmentation Framework. 1282-1291
-     Ayça Takmaz, Jonas Schult, Irem Kaftan, Mertcan Akçay, Bastian Leibe, Robert W. Sumner, Francis Engelmann, Siyu Tang :
3D Segmentation of Humans in Point Clouds with Synthetic Data. 1292-1304
-     Shijie Lian, Hua Li, Runmin Cong, Suqi Li, Wei Zhang, Sam Kwong:
WaterMask: Instance Segmentation for Underwater Imagery. 1305-1315
-     Ho Kei Cheng, Seoung Wug Oh, Brian Price, Alexander G. Schwing, Joon-Young Lee:
Tracking Anything with Decoupled Video Segmentation. 1316-1326
-     Chenming Li, Daoan Zhang, Wenjian Huang, Jianguo Zhang:
Cross Contrasting Feature Perturbation for Domain Generalization. 1327-1337
-     Lei Fan, Bo Liu, Haoxiang Li, Ying Wu, Gang Hua:
Flexible Visual Recognition by Evidential Modeling of Confusion and Ignorance. 1338-1347
-     Rabab Abdelfattah, Qing Guo, Xiaoguang Li, Xiaofeng Wang, Song Wang:
CDUL: CLIP-Driven Unsupervised Learning for Multi-Label Image Classification. 1348-1357
-     Jongyoun Noh, Hyekang Park, Junghyup Lee, Bumsuk Ham:
RankMixup: Ranking-Based Mixup Training for Network Calibration. 1358-1368
-     Yang Lu, Yiliang Zhang, Bo Han, Yiu-Ming Cheung, Hanzi Wang:
Label-Noise Learning with Intrinsically Long-Tailed Data. 1369-1378
-     Xingyu Liu, Sanping Zhou, Le Wang, Gang Hua:
Parallel Attention Interaction Network for Few-Shot Skeleton-based Action Recognition. 1379-1388
-     Jiangning Zhang, Xiangtai Li, Jian Li, Liang Liu, Zhucun Xue, Boshen Zhang, Zhengkai Jiang, Tianxin Huang, Yabiao Wang, Chengjie Wang:
Rethinking Mobile Block for Efficient Attention-based Models. 1389-1400
-     Dongjun Lee, Seokwon Song, Jihee Suh, Joonmyeong Choi, Sanghyeok Lee, Hyunwoo J. Kim:
Read-only Prompt Optimization for Vision-Language Few-shot Learning. 1401-1411
-     Zhongzhan Huang, Mingfu Liang, Jinghui Qin, Shanshan Zhong, Liang Lin:
Understanding Self-attention Mechanism via Dynamical System Perspective. 1412-1422



Wenqiao Zhang, Changshuo Liu, Lingze Zeng, Beng Chin Ooi, Siliang Tang, Yueling Zhuang:
Learning in Imperfect Environment: Multi-Label Classification with Long-Tailed Distribution and Partial Labels. 1423-1432

Shunxin Wang, Raymond N. J. Veldhuis, Christoph Brune , Nicola Strisciuglio:
What do neural networks learn in image classification? A frequency shortcut perspective. 1433-1442

Tong Liang, Jim Davis:
Inducing Neural Collapse to a Fixed Hierarchy-Aware Frame for Reducing Mistake Severity. 1443-1452

Reza Averly, Wei-Lun Chao:
Unified Out-Of-Distribution Detection: A Model-Specific Perspective. 1453-1463

Myeongho Jeon, Myungjoo Kang, Joonseok Lee:
A Unified Framework for Robustness on Diverse Sampling Errors. 1464-1472

Xuelin Zhu, Jian Liu, Weijia Liu, Jiawei Ge, Bo Liu, Jiuxin Cao:
Scene-Aware Label Graph Learning for Multi-Label Image Classification. 1473-1482

Xiaobo Xia, Jiankang Deng, Wei Bao, Yuxuan Du, Bo Han, Shiguang Shan, Tongliang Liu:
Holistic Label Correction for Noisy Multi-Label Classification. 1483-1493

Guiping Cao, Shengda Luo, Wenjian Huang, Xiangyuan Lan, Dongmei Jiang, Yaowei Wang, Jianguo Zhang:
Strip-MLP: Efficient Token Interaction for Vision MLP. 1494-1504

Ke Xu, Lei Han, Ye Tian, Shangshang Yang, Xingyi Zhang:
EQ-Net: Elastic Quantization Neural Networks. 1505-1514

Renrong Shao, Wei Zhang, Jianhua Yin, Jun Wang:
Data-free Knowledge Distillation for Fine-grained Visual Categorization. 1515-1525

Xilin He, Qinliang Lin, Cheng Luo, Weicheng Xie, Siyang Song, Feng Liu, Linlin Shen:
Shift from Texture-bias to Shape-bias: Edge Deformation-based Augmentation for Robust Object Recognition. 1526-1535

Isack Lee, Eungi Lee, Seok Bong Yoo:
Latent-OFER: Detect, Mask, and Reconstruct with Latent Vectors for Occluded Facial Expression Recognition. 1536-1546

Nan Zhou, Jiaxin Chen, Di Huang:
DR-Tune: Improving Fine-tuning of Pretrained Visual Models by Distribution Regularization with Semantic Calibration. 1547-1556

Jaewoo Park, Jacky Chen Long Chai, Jaeho Yoon, Andrew Beng Jin Teoh:
Understanding the Feature Norm for Out-of-Distribution Detection. 1557-1567

-      Ruoyi Du, Wenqing Yu, Heqing Wang, Ting-En Lin, Dongliang Chang, Zhanyu Ma:
Multi-View Active Fine-Grained Visual Recognition. 1568-1578
-     Ruiyuan Gao, Chenchen Zhao, Lanqing Hong, Qiang Xu:
DiffGuard: Semantic Mismatch-Guided Out-of-Distribution Detection using Pre-trained Diffusion Models. 1579-1589
-     Yurong Guo, Ruoyi Du, Yuan Dong, Timothy M. Hospedales, Yi-Zhe Song, Zhanyu Ma:
Task-aware Adaptive Learning for Cross-domain Few-shot Learning. 1590-1599
-     Qidong Huang, Xiaoyi Dong, Dongdong Chen, Yinpeng Chen, Lu Yuan, Gang Hua, Weiming Zhang, Nenghai Yu:
Improving Adversarial Robustness of Masked Autoencoders via Test-time Frequency-domain Prompting. 1600-1610
-     Shouwen Wang, Qian Wan, Xiang Xiang, Zhigang Zeng:
Saliency Regularization for Self-Training with Partial Annotations. 1611-1620
-     Lanyun Zhu, Tianrun Chen, Jianxiong Yin, Simon See, Jun Liu:
Learning Gabor Texture Features for Fine-Grained Recognition. 1621-1631
-     Kunchang Li, Yali Wang, Yinan He, Yizhuo Li, Yi Wang, Limin Wang, Yu Qiao:
UniFormerV2: Unlocking the Potential of Image ViTs for Video Understanding. 1632-1643
-     Ziyi Zhang, Weikai Chen, Chaowei Fang, Zhen Li, Lechao Chen, Liang Lin, Guanbin Li:
RankMatch: Fostering Confidence and Consistency in Learning with Noisy Labels. 1644-1654
-     Yanan Wu, Zhixiang Chi, Yang Wang, Songhe Feng:
MetaGCD: Learning to Continually Learn in Generalized Category Discovery. 1655-1665
-     Zhiqiang Shen:
FerKD: Surgical Label Adaptation for Efficient Distillation. 1666-1675
-     Chengxin Liu, Hao Lu, Zhiguo Cao, Tongliang Liu:
Point-Query Quadtree for Crowd Counting, Localization, and More. 1676-1685
-     Jaewoo Park, Yoon Gyo Jung, Andrew Beng Jin Teoh:
Nearest Neighbor Guidance for Out-of-Distribution Detection. 1686-1695
-     Hyunjae Lee, Heon Song, Hyeonsoo Lee, Gihyeon Lee, Suyeong Park, Donggeun Yoo:
Bayesian Optimization Meets Self-Distillation. 1696-1705
-     Yu-Ming Tang, Yi-Xing Peng, Wei-Shi Zheng:
When Prompt-based Incremental Learning Does Not Meet Strong Pretraining. 1706-1716
-     Chengkai Hou, Jieyu Zhang, Tianyi Zhou:
When to Learn What: Model-Adaptive Data Augmentation Curriculum.



1717-1728

- Florent Chiaroni, Jose Dolz, Imtiaz Masud Ziko, Amar Mitiche, Ismail Ben Ayed:
Parametric Information Maximization for Generalized Category Discovery.
1729-1739
- Jiazheng Xing, Mengmeng Wang, Yudi Ruan, Bofan Chen, Yaowei Guo, Boyu Mu, Guang Dai, Jingdong Wang, Yong Liu:
Boosting Few-shot Action Recognition with Graph-guided Hybrid Matching.
1740-1750
- Liang Chen, Yong Zhang, Yibing Song, Anton van den Hengel, Lingqiao Liu:
Domain Generalization via Rationale Invariance. 1751-1760
- Ziqing Wang, Yuetong Fang, Jiahang Cao, Qiang Zhang, Zhongrui Wang, Renjing Xu:
Masked Spiking Transformer. 1761-1771
- Wuxuan Shi, Mang Ye:
Prototype Reminiscence and Augmented Asymmetric Knowledge Aggregation for Non-Exemplar Class-Incremental Learning. 1772-1781
- Yun Li, Zhe Liu, Saurav Jha, Lina Yao:
Distilled Reverse Attention Network for Open-world Compositional Zero-Shot Learning. 1782-1791
- Shuo He, Guowu Yang, Lei Feng:
Candidate-aware Selective Disambiguation Based On Normalized Entropy for Instance-dependent Partial-label Learning. 1792-1801
- Hualiang Wang, Yi Li, Huifeng Yao, Xiaomeng Li:
CLIPN for Zero-Shot OOD Detection: Teaching CLIP to Say No. 1802-1812
- Benzhi Wang, Yang Yang, Jinlin Wu, Guo-Jun Qi, Zhen Lei:
Self-similarity Driven Scale-invariant Learning for Weakly Supervised Person Search. 1813-1822
- Chanho Ahn, Kikyung Kim, Ji-Won Baek, Jongin Lim, Seungju Han:
Sample-wise Label Confidence Incorporation for Learning with Noisy Labels. 1823-1832
- Xiaobo Xia, Bo Han, Yibing Zhan, Jun Yu, Mingming Gong, Chen Gong, Tongliang Liu:
Combating Noisy Labels with Sample Selection by Mining High-Discrepancy Examples. 1833-1843
- Pingyu Wu, Wei Zhai, Yang Cao, Jiebo Luo, Zheng-Jun Zha:
Spatial-Aware Token for Weakly Supervised Object Localization. 1844-1854
- Sriram Balasubramanian, Soheil Feizi:
Towards Improved Input Masking for Convolutional Neural Networks.
1855-1865
- Robert van der Klis, Stephan Alaniz, Massimiliano Mancini , Cássio Fraga Dantas, Dino Ienco, Zeynep Akata, Diego Marcos:
PDiscoNet: Semantically consistent part discovery for fine-grained recognition. 1866-1876

-     Divyansh Srivastava, Tuomas P. Oikarinen, Tsui-Wei Weng:
Corrupting Neuron Explanations of Deep Visual Features. 1877-1886
-     Dawid Rymarczyk, Joost van de Weijer, Bartosz Zielinski, Bartłomiej Twardowski:
ICICLE: Interpretable Class Incremental Continual Learning. 1887-1898
-     Uddeshya Upadhyay, Shyamgopal Karthik, Massimiliano Mancini , Zeynep Akata:
ProbVLM: Probabilistic Adapter for Frozen Vision-Language Models. 1899-1910
-     Julia Hornauer, Adrian Holzbock, Vasileios Belagiannis:
Out-of-Distribution Detection for Monocular Depth Estimation. 1911-1921
-     Sukrut Rao , Moritz Böhle, Amin Parchami-Araghi, Bernt Schiele:
Studying How to Efficiently and Effectively Guide Models with Explanations. 1922-1933
-     Amil David, Yossi Gandelsman, Alexei A. Efros, Assaf Shocher:
Rosetta Neurons: Mining the Common Units in a Model Zoo. 1934-1943
-     Nanne van Noord :
Prototype-based Dataset Comparison. 1944-1954
-     Haozhe Liu, Mingchen Zhuge, Bing Li, Yuhui Wang, Francesco Faccio, Bernard Ghanem , Jürgen Schmidhuber:
Learning to Identify Critical States for Reinforcement Learning from Videos. 1955-1965
-     Alexandros Stergiou , Nikos Deligiannis:
Leaping Into Memories: Space-Time Deep Feature Synthesis. 1966-1976
-     Yifei Zhang, Siyi Gu, Yuyang Gao, Bo Pan, Xiaofeng Yang, Liang Zhao:
MAGI: Multi-Annotated Explanation-Guided Learning. 1977-1987
-     Wei Huang, Xingyu Zhao , Gaojie Jin, Xiaowei Huang:
SAFARI: Versatile and Efficient Evaluations for Robustness of Interpretability. 1988-1998
-     Hang Li, Jindong Gu, Rajat Koner, Sahand Sharifzadeh, Volker Tresp:
Do DALL-E and Flamingo Understand Each Other? 1999-2010
-     Qihan Huang, Mengqi Xue, Wenqi Huang, Haofei Zhang , Jie Song, Yongcheng Jing, Mingli Song:
Evaluation and Improvement of Interpretability for Self-Explainable Part-Prototype Networks. 2011-2020
-     Jingwei Zhang, Farzan Farnia:
MoreauGrad: Sparse and Robust Interpretation of Neural Networks via Moreau Envelope. 2021-2030
-     Kelu Yao, Jin Wang, Boyu Diao, Chao Li:
Towards Understanding the Generalization of Deepfake Detectors from a Game-Theoretical View. 2031-2041
-     Xue Wang, Zhibo Wang, Haiqin Weng, Hengchang Guo, Zhifei Zhang, Lu Jin, Tao Wei, Kui Ren:



Counterfactual-based Saliency Map: Towards Visual Contrastive Explanations for Neural Networks. 2042-2051



Giyoung Jeon, Haedong Jeong, Jaesik Choi :

Beyond Single Path Integrated Gradients for Reliable Input Attribution via Randomized Path Sampling. 2052-2061



Chong Wang, Yuyuan Liu, Yuanhong Chen, Fengbei Liu, Yu Tian, Davis J. McCarthy, Helen Frazer, Gustavo Carneiro:

Learning Support and Trivial Prototypes for Interpretable Image Classification. 2062-2072



Oren Barkan, Yehonatan Elisha, Yuval Asher, Amit Eshel, Noam Koenigstein: **Visual Explanations via Iterated Integrated Attributions.** 2073-2084



Nan Liu, Yilun Du, Shuang Li, Joshua B. Tenenbaum, Antonio Torralba: **Unsupervised Compositional Concepts Discovery with Text-to-Image Generative Models.** 2085-2095



Xiaoshi Wu, Keqiang Sun, Feng Zhu, Rui Zhao, Hongsheng Li:

Human Preference Score: Better Aligning Text-to-image Models with Human Preference. 2096-2105



Elad Levi, Eli Brosh, Mykola Mykhailych, Meir Perez:

DLT: Conditioned layout generation with Joint Discrete-Continuous Diffusion Layout Transformer. 2106-2115



Thanh Van Le, Hao Phung, Thuan Hoang Nguyen, Quan Dao, Ngoc N. Tran, Anh Tuan Tran:

Anti-DreamBooth: Protecting users from personalized text-to-image synthesis. 2116-2127



Michal J. Tyškiewicz, Pascal Fua, Eduard Trulls:

GECCO: Geometrically-Conditioned Point Diffusion Models. 2128-2138



Shengqu Cai, Eric Ryan Chan, Songyou Peng, Mohamad Shahbazi, Anton Obukhov, Luc Van Gool, Gordon Wetzstein:

DiffDreamer: Towards Consistent Unsupervised Single-view Scene Extrapolation with Conditional Diffusion Models. 2139-2150



Korrawe Karunratanakul, Konpat Preechakul, Supasorn Suwajanakorn, Siyu Tang :

Guided Motion Diffusion for Controllable Human Motion Synthesis. 2151-2162



Yanzhao Zheng, Yunzhou Shi, Yuhao Cui, Zhongzhou Zhao, Zhiling Luo, Wei Zhou:

COOP: Decoupling and Coupling of Whole-Body Grasping Pose Generation. 2163-2173



Guillaume Couairon, Marlène Careil, Matthieu Cord, Stéphane Lathuilière, Jakob Verbeek:

Zero-shot spatial layout conditioning for text-to-image diffusion models. 2174-2183

-      Aibek Alanov, Vadim Titov, Maksim Nakhodnov, Dmitry P. Vetrov:
StyleDomain: Efficient and Lightweight Parameterizations of StyleGAN for One-shot and Few-shot Domain Adaptation. 2184-2194
-     Jianfeng Xiang, Jiaolong Yang, Yu Deng, Xin Tong:
GRAM-HD: 3D-Consistent Image Generation at High Resolution with Generative Radiance Manifolds. 2195-2205
-     Alexander C. Li, Mihir Prabhudesai, Shivam Duggal, Ellis Brown , Deepak Pathak:
Your Diffusion Model is Secretly a Zero-Shot Classifier. 2206-2217
-     Jiali Cui, Ying Nian Wu, Tian Han:
Learning Hierarchical Features with Joint Latent Space Energy-Based Prior. 2218-2227
-     Liang Xu, Ziyang Song, Dongliang Wang, Jing Su, Zhicheng Fang, Chenjing Ding, Weihao Gan, Yichao Yan, Xin Jin, Xiaokang Yang, Wenjun Zeng, Wei Wu:
ActFormer: A GAN-based Transformer towards General Action-Conditioned 3D Human Motion Generation. 2228-2238
-     Ruoshi Liu, Chengzhi Mao, Purva Tendulkar, Hao Wang, Carl Vondrick:
Landscape Learning for Neural Network Inversion. 2239-2250
-     Martin Nicolas Everaert , Marco Bocchio, Sami Arpa, Sabine Süsstrunk, Radhakrishna Achanta:
Diffusion in Style. 2251-2261
-     Gene Chou, Yuval Bahat, Felix Heide:
Diffusion-SDF: Conditional Generative Modeling of Signed Distance Functions. 2262-2272
-     Xuanmeng Zhang, Jianfeng Zhang, Rohan Chacko, Hongyi Xu, Guoxian Song, Yi Yang, Jiashi Feng:
GETAvatar: Generative Textured Meshes for Animatable Human Avatars. 2273-2282
-     Aishwarya Agarwal, Srikrishna Karanam, K. J. Joseph, Apoorv Saxena, Koustava Goswami, Balaji Vasan Srinivasan:
A-STAR: Test-time Attention Segregation and Retention for Text-to-image Synthesis. 2283-2293
-     Shilin Lu, Yanzhu Liu, Adams Wai-Kin Kong:
TF-ICON: Diffusion-Based Training-Free Cross-Domain Image Composition. 2294-2305
-     Yijun Qian, Jack Urbanek, Alexander G. Hauptmann, Jungdam Won:
Breaking The Limits of Text-conditioned 3D Motion Synthesis with Elaborative Descriptions. 2306-2316
-     Germán Barquero, Sergio Escalera , Cristina Palmero:
BeLFusion: Latent Diffusion for Behavior-Driven Human Motion Prediction. 2317-2327
-     Amir Hertz, Kfir Aberman, Daniel Cohen-Or:
Delta Denoising Score. 2328-2337

-  [Xingyu Chen, Yu Deng, Baoyuan Wang:](#) **Mimic3D: Thriving 3D-Aware GANs via 3D-to-2D Imitation.** 2338-2348
-  [Amit Raj, Srinivas Kaza, Ben Poole, Michael Niemeyer, Nataniel Ruiz, Ben Mildenhall, Shiran Zada, Kfir Aberman, Michael Rubinstein, Jonathan T. Barron, Yuanzhen Li, Varun Jampani:](#) **DreamBooth3D: Subject-Driven Text-to-3D Generation.** 2349-2359
-  [Shuang Song, Yuanbang Liang, Jing Wu, Yu-Kun Lai, Yipeng Qin:](#) **Feature Proliferation - the "Cancer" in StyleGAN and its Treatments.** 2360-2370
-  [Berkay Kicanaoglu, Pablo Garrido, Gaurav Bharaj:](#) **Unsupervised Facial Performance Editing via Vector-Quantized StyleGAN Representations.** 2371-2382
-  [Jianfeng Xiang, Jiaolong Yang, Binbin Huang, Xin Tong:](#) **3D-aware Image Generation using 2D Diffusion Models.** 2383-2393
-  [Ganghun Lee, Minji Kim, Yunsu Lee, Minsu Lee, Byoung-Tak Zhang:](#) **Neural Collage Transfer: Artistic Reconstruction via Material Manipulation.** 2394-2405
-  [Teng Hu, Jiangning Zhang, Liang Liu, Ran Yi, Siqi Kou, Haokun Zhu, Xu Chen, Yabiao Wang, Chengjie Wang, Lizhuang Ma:](#) **Phasic Content Fusing Diffusion Model with Directional Distribution Consistency for Few-Shot Model Adaption.** 2406-2415
-  [Hansheng Chen, Jiatao Gu, Anpei Chen, Wei Tian, Zhuowen Tu, Lingjie Liu, Hao Su:](#) **Single-Stage Diffusion NeRF: A Unified Approach to 3D Generation and Reconstruction.** 2416-2425
-  [Rohit Gandikota, Joanna Materzynska, Jaden Fiotto-Kaufman, David Bau ID:](#) **Erasing Concepts from Diffusion Models.** 2426-2436
-  [Ziyang Yuan, Yiming Zhu, Yu Li, Hongyu Liu, Chun Yuan:](#) **Make Encoder Great Again in 3D GAN Inversion through Geometry and Occlusion-Aware Encoding.** 2437-2447
-  [Seunggyu Chang, Gihoon Kim, Hayeon Kim:](#) **HairNeRF: Geometry-Aware Image Synthesis for Hairstyle Transfer.** 2448-2458
-  [Yuanze Lin, Chen Wei, Huiyu Wang, Alan L. Yuille, Cihang Xie:](#) **SMAUG: Sparse Masked Autoencoder for Efficient Video-Language Pre-training.** 2459-2469
-  [Peng Jin, Hao Li, Zesen Cheng, Kehan Li, Xiangyang Ji, Chang Liu, Li Yuan, Jie Chen:](#) **DiffusionRet: Generative Text-Video Retrieval with Diffusion Model.** 2470-2481
-  [Anwen Hu, Shizhe Chen, Liang Zhang, Qin Jin:](#) **Explore and Tell: Embodied Visual Captioning in 3D Environments.** 2482-2491

-      Xuanlin Li, Yunhao Fang, Minghua Liu, Zhan Ling, Zhuowen Tu, Hao Su:
Distilling Large Vision-Language Model with Out-of-Distribution Generalizability. 2492-2503
-     Xu Yang, Zhangzikang Li, Haiyang Xu, Hanwang Zhang, Qinghao Ye, Chenliang Li, Ming Yan, Yu Zhang, Fei Huang, Songfang Huang:
Learning Trajectory-Word Alignments for Video-Language Tasks. 2504-2514
-     Dizhan Xue, Shengsheng Qian, Changsheng Xu:
Variational Causal Inference Network for Explanatory Visual Question Answering. 2515-2525
-     Moon Ye-Bin, Jisoo Kim, Hongyeob Kim, Kilho Son, Tae-Hyun Oh:
TextManiA: Enriching Visual Feature by Text-driven Manifold Augmentation. 2526-2537
-     Jiannan Wu, Yi Jiang, Bin Yan, Huchuan Lu, Zehuan Yuan, Ping Luo:
Segment Every Reference Object in Spatial and Temporal Spaces. 2538-2550
-     Juncheng Li, Minghe Gao, Longhui Wei, Siliang Tang, Wenqiao Zhang, Mengze Li, Wei Ji, Qi Tian, Tat-Seng Chua, Yueling Zhuang:
Gradient-Regulated Meta-Prompt Learning for Generalizable Vision-Language Models. 2551-2562
-     Bumsoo Kim, Yeonsik Jo, Jinhyung Kim, Seung-Hwan Kim:
Misalign, Contrast then Distill: Rethinking Misalignments in Language-Image Pretraining. 2563-2572
-     Yifeng Zhang , Shi Chen, Qi Zhao:
Toward Multi-Granularity Decision-Making: Explicit Visual Reasoning with Hierarchical Knowledge. 2573-2583
-     Junyu Bi, Daixuan Cheng, Ping Yao, Bochen Pang, Yuefeng Zhan, Chuanguang Yang, Yujing Wang, Hao Sun, Weiwei Deng, Qi Zhang:
VL-Match: Enhancing Vision-Language Pretraining with Token-Level and Instance-Level Matching. 2584-2593
-     Ioana Croitoru, Simion-Vlad Bogolin, Samuel Albanie, Yang Liu, Zhaowen Wang, Seunghyun Yoon, Franck Dernoncourt, Hailin Jin, Trung Bui:
Moment Detection in Long Tutorial Videos. 2594-2604
-     Xiangyang Zhu, Renrui Zhang, Bowei He , Aojun Zhou, Dong Wang, Bin Zhao, Peng Gao:
Not All Features Matter: Enhancing Few-shot CLIP with Adaptive Prior Refinement. 2605-2615
-     Nitzan Bitton Guetta, Yonatan Bitton, Jack Hessel, Ludwig Schmidt, Yuval Elovici, Gabriel Stanovsky, Roy Schwartz:
Breaking Common Sense: WHOOPS! A Vision-and-Language Benchmark of Synthetic and Compositional Images. 2616-2627
-     Yixuan Wu, Zhao Zhang , Chi Xie, Feng Zhu, Rui Zhao:
Advancing Referring Expression Segmentation Beyond Single Image. 2628-2638

-      Xiangyang Zhu, Renrui Zhang, Bowei He , Ziyu Guo, Ziyao Zeng, Zipeng Qin, Shanghang Zhang, Peng Gao:
PointCLIP V2: Prompting CLIP and GPT for Powerful 3D Open-world Learning. 2639-2650
-     Weizhen He, Weijie Chen, Binbin Chen, Shicai Yang, Di Xie, Luojun Lin, Donglian Qi, Yueting Zhuang:
Unsupervised Prompt Tuning for Text-Driven Object Detection. 2651-2661
-     Zehan Wang, Haifeng Huang, Yang Zhao, Linjun Li, Xize Cheng, Yichen Zhu, Aoxiong Yin, Zhou Zhao:
Distilling Coarse-to-Fine Semantic Matching Knowledge for Weakly Supervised 3D Visual Grounding. 2662-2671
-     Sophia Gu, Christopher Clark, Aniruddha Kembhavi:
I can't believe there's no images! : Learning Visual Tasks Using Only Language Supervision. 2672-2683
-     Guanghui Li, Mingqi Gao, Heng Liu, Xiantong Zhen, Feng Zheng:
Learning Cross-Modal Affinity for Referring Video Object Segmentation Targeting Limited Samples. 2684-2693
-     Henghui Ding, Chang Liu, Shuting He, Xudong Jiang , Chen Change Loy:
MeViS: A Large-scale Benchmark for Video Segmentation with Motion Expressions. 2694-2703
-     Chun-Mei Feng, Kai Yu, Yong Liu, Salman Khan, Wangmeng Zuo:
Diverse Data Augmentation with Diffusions for Effective Test-time Prompt Tuning. 2704-2714
-     Xi Tian, Yong-Liang Yang, Qi Wu:
ShapeScaffolder: Structure-Aware 3D Shape Generation from Text. 2715-2724
-     Vishaal Udandarao, Ankush Gupta, Samuel Albanie:
SuS-X: Training-Free Name-Only Transfer of Vision-Language Models. 2725-2736
-     Yiwei Ma, Huawei Wang, Xiaoqing Zhang, Guannan Jiang, Xiaoshuai Sun, Weilin Zhuang, Jiayi Ji, Rongrong Ji:
X-Mesh: Towards Fast and Accurate Text-driven 3D Stylization via Dynamic Textual Guidance. 2737-2748
-     Dongming Wu, Tiancai Wang, Yuang Zhang, Xiangyu Zhang, Jianbing Shen:
OnlineRefer: A Simple Online Baseline for Referring Video Object Segmentation. 2749-2758
-     Yifan Yang, Weiquan Huang, Yixuan Wei, Houwen Peng, Xinyang Jiang, Huiqiang Jiang, Fangyun Wei, Yin Wang, Han Hu, Lili Qiu, Yuqing Yang:
Attentive Mask CLIP. 2759-2769
-     Jiangtong Li, Li Niu, Liqing Zhang:
Knowledge Proxy Intervention for Deconfounded Video Question Answering. 2770-2781

-      Kevin Qinghong Lin, Pengchuan Zhang, Joya Chen, Shraman Pramanick, Difei Gao, Alex Jinpeng Wang, Rui Yan, Mike Zheng Shou:
UniVTG: Towards Unified Video-Language Temporal Grounding. 2782-2792
-     Yunbin Tu, Liang Li, Li Su, Zheng-Jun Zha, Chenggang Yan, Qingming Huang:
Self-supervised Cross-view Representation Reconstruction for Change Captioning. 2793-2803
-     Ziyang Wang, Yi-Lin Sung, Feng Cheng, Gedas Bertasius, Mohit Bansal:
Unified Coarse-to-Fine Alignment for Video-Text Retrieval. 2804-2815
-     Yang Liu, Jiahua Zhang, Qingchao Chen, Yuxin Peng:
Confidence-aware Pseudo-label Learning for Weakly Supervised Visual Grounding. 2816-2826
-     Chengyang Zhao, Yikang Shen, Zhenfang Chen, Mingyu Ding, Chuang Gan:
TextPSG: Panoptic Scene Graph Generation from Textual Descriptions. 2827-2838
-     Wei Lin, Leonid Karlinsky, Nina Shvetsova, Horst Possegger, Mateusz Kozinski, Rameswar Panda, Rogério Feris, Hilde Kuehne, Horst Bischof:
MATch, eXpand and Improve: Unsupervised Finetuning for Zero-Shot Action Recognition with Language Knowledge. 2839-2850
-     Yaowei Li, Bang Yang, Xuxin Cheng, Zhihong Zhu, Hongxiang Li, Yuexian Zou:
Unify, Align and Refine: Multi-Level Semantic Alignment for Radiology Report Generation. 2851-2862
-     Devaansh Gupta, Siddhant Kharbanda, Jiawei Zhou, Wanhu Li, Hanspeter Pfister, Donglai Wei:
CLIPTrans: Transferring Visual Knowledge with Pre-trained Models for Multimodal Machine Translation. 2863-2874
-     Morris Alper, Hadar Averbuch-Elor:
Learning Human-Human Interactions in Images from Weak Textual Supervision. 2875-2887
-     Chaoya Jiang, Haiyang Xu, Wei Ye, Qinghao Ye, Chenliang Li, Ming Yan, Bin Bi, Shikun Zhang, Fei Huang, Songfang Huang:
BUS : Efficient and Effective Vision-language Pre-training with Bottom-Up Patch Summarization. 2888-2898
-     Ziyu Zhu, Xiaojian Ma, Yixin Chen, Zhidong Deng, Siyuan Huang, Qing Li:
3D-VisTA: Pre-trained Transformer for 3D Vision and Text Alignment. 2899-2909
-     Kaicheng Yang, Jiankang Deng, Xiang An, Jiawei Li, Ziyong Feng, Jia Guo, Jing Yang, Tongliang Liu:
ALIP: Adaptive Language-Image Pre-training with Synthetic Caption. 2910-2919
-     Cheng Shi, Sibei Yang:
LoGoPrompt: Synthetic Text Images Can Be Good Visual Prompts for Vision-Language Models. 2920-2929

-      Wooyoung Kang, Jonghwan Mun, Sungjun Lee, Byungseok Roh:
Noise-aware Learning from Web-crawled Image-Text Data for Image Captioning. 2930-2940
-     Zi Qian, Xin Wang, Xuguang Duan, Pengda Qin, Yuhong Li, Wenwu Zhu:
Decouple Before Interact: Multi-Modal Prompt Learning for Continual Visual Question Answering. 2941-2950
-     Yushi Hu, Hang Hua, Zhengyuan Yang, Weijia Shi, Noah A. Smith, Jiebo Luo:
PromptCap: Prompt-Guided Image Captioning for VQA with GPT-3. 2951-2963
-     Yu Wu, Yana Wei, Haozhe Wang, Yongfei Liu, Sibei Yang, Xuming He:
Grounded Image Text Matching with Mismatched Relation Reasoning. 2964-2975
-     Mohamed Ashraf Abdelsalam, Samrudhdhi B. Rangrej, Isma Hadji, Nikita Dvornik, Konstantinos G. Derpanis, Afsaneh Fazly:
GePSAn: Generative Procedure Step Anticipation in Cooking Videos. 2976-2985
-     Chan Hee Song, Brian M. Sadler, Jiaman Wu, Wei-Lun Chao, Clayton Washington, Yu Su:
LLM-Planner: Few-Shot Grounded Planning for Embodied Agents with Large Language Models. 2986-2997
-     Zi-Yuan Hu, Yanyang Li, Michael R. Lyu, Liwei Wang:
VL-PET: Vision-and-Language Parameter-Efficient Tuning via Granularity Control. 2998-3008
-     Manuele Barraco, Sara Sarto, Marcella Cornia, Lorenzo Baraldi, Rita Cucchiara:
With a Little Help from your own Past: Prototypical Memory Networks for Image Captioning. 3009-3019
-     Jaemin Cho, Abhay Zala, Mohit Bansal:
DALL-EVAL: Probing the Reasoning Skills and Social Biases of Text-to-Image Generation Models. 3020-3031
-     Yicong Hong, Yang Zhou, Ruiyi Zhang, Franck Dernoncourt, Trung Bui, Stephen Gould, Hao Tan:
Learning Navigational Visual Representations with Semantic Map Supervision. 3032-3044
-     Jiajin Tang, Ge Zheng, Jingyi Yu, Sibei Yang:
CoTDet: Affordance Knowledge Prompting for Task Driven Object Detection. 3045-3055
-     Nan Xi, Jingjing Meng, Junsong Yuan:
Open Set Video HOI detection from Action-centric Chain-of-Look Prompting. 3056-3066
-     An Yan, Yu Wang, Yiwu Zhong, Chengyu Dong, Zexue He, Yujie Lu, William Yang Wang, Jingbo Shang, Julian J. McAuley:
Learning Concise and Descriptive Attributes for Visual Recognition. 3067-3077



Dohwan Ko, Ji Soo Lee, Miso Choi, Jaewon Chu, Jihwan Park, Hyunwoo J. Kim:
Open-Vocabulary Video Question Answering: A New Benchmark for Evaluating the Generalizability of Video Question Answering Models. 3078-3089



Thomas Mensink, Jasper R. R. Uijlings, Lluís Castrejón, Arushi Goel, Felipe Cadar, Howard Zhou , Fei Sha, André Araújo, Vittorio Ferrari:
Encyclopedic VQA: Visual questions about detailed properties of fine-grained categories. 3090-3101



Daechul Ahn, Daneul Kim, Gwangmo Song, Seung Hwan Kim, Honglak Lee, Dongyeop Kang, Jonghyun Choi:
Story Visualization by Online Text Augmentation with Context Memory. 3102-3112



Junjie Fei, Teng Wang, Jinrui Zhang, Zhenyu He, Chengjie Wang, Feng Zheng:
Transferable Decoding with Visual Entities for Zero-Shot Image Captioning. 3113-3123



Alex Jinpeng Wang, Kevin Qinghong Lin, David Junhao Zhang, Stan Weixian Lei, Mike Zheng Shou:
Too Large; Data Reduction for Vision-Language Pre-Training. 3124-3134



Weihan Wang, Zhen Yang, Bin Xu, Juanzi Li, Yankui Sun:
ViLTA: Enhancing Vision-Language Pre-training through Textual Augmentation. 3135-3146



Roni Paiss, Ariel Ephrat, Omer Tov, Shiran Zada, Inbar Mosseri, Michal Irani, Tali Dekel:
Teaching CLIP to Count to Ten. 3147-3157



Junsheng Zhou, Baorui Ma, Shujuan Li, Yu-Shen Liu, Zhizhong Han:
Learning a More Continuous Zero Level Set in Unsigned Distance Fields through Level Set Projection. 3158-3169



Wenyan Cong, Hanxue Liang, Peihao Wang, Zhiwen Fan, Tianlong Chen, Mukund Varma T, Yi Wang, Zhangyang Wang:
Enhancing NeRF akin to Enhancing LLMs: Generalizable NeRF Transformer with Mixture-of-View-Experts. 3170-3181



Yixuan Li, Lihan Jiang, Lining Xu, Yuanbo Xiangli, Zhenzhi Wang, Dahua Lin, Bo Dai:
MatrixCity: A Large-scale City Dataset for City-scale Neural Rendering and Beyond. 3182-3192



Aron Schmied, Tobias Fischer, Martin Danelljan, Marc Pollefeys, Fisher Yu:
R3D3: Dense 3D Reconstruction of Dynamic Scenes from Multiple Cameras. 3193-3203



Yuan Li, Zhi-Hao Lin, David A. Forsyth, Jia-Bin Huang, Shenlong Wang:
ClimateNeRF: Extreme Weather Synthesis in Neural Radiance Field. 3204-3215



Tiange Xiang, Adam Sun, Jiajun Wu, Ehsan Adeli , Li Fei-Fei:
Rendering Humans from Object-Occluded Monocular Videos. 3216-3227

-      Yuanbo Xiangli, Lining Xu, Xingang Pan, Nanxuan Zhao, Bo Dai, Dahua Lin:
AssetField: Assets Mining and Reconfiguration in Ground Feature Plane Representation. 3228-3238
-     Yingfei Liu, Junjie Yan, Fan Jia, Shuailin Li, Aqi Gao, Tiancai Wang, Xiangyu Zhang:
PETRv2: A Unified Framework for 3D Perception from Multi-Camera Images. 3239-3249
-     Takuhiro Kaneko:
MIMO-NeRF: Fast Neural Rendering with Multi-input Multi-output Neural Radiance Fields. 3250-3260
-     Zelin Gao, Weichen Dai, Yu Zhang:
Adaptive Positional Encoding for Bundle-Adjusting Neural Radiance Fields. 3261-3271
-     Yiming Wang, Qin Han, Marc Habermann, Kostas Daniilidis, Christian Theobalt, Lingjie Liu:
NeuS2: Fast Learning of Neural Implicit Surfaces for Multi-view Reconstruction. 3272-3283
-     Qitong Wang, Long Zhao, Liangzhe Yuan, Ting Liu, Xi Peng:
Learning from Semantic Alignment between Unpaired Multiviews for Egocentric Video Recognition. 3284-3294
-     Junpeng Jing, Jiankun Li, Pengfei Xiong, Jiangyu Liu, Shuaicheng Liu, Yichen Guo, Xin Deng, Mai Xu, Lai Jiang, Leonid Sigal:
Uncertainty Guided Adaptive Warping for Robust and Efficient Stereo Matching. 3295-3304
-     Martin Bråtelund, Felix Rydell:
Compatibility of Fundamental Matrices for Complete Viewing Graphs. 3305-3313
-     Pin Tang, Hai-Ming Xu, Chao Ma:
ProtoTransfer: Cross-Modal Prototype Transfer for Point Cloud Segmentation. 3314-3324
-     Jinqing Zhang, Yanan Zhang , Qingjie Liu, Yunhong Wang:
SA-BEV: Generating Semantic-Aware Bird's-Eye-View Feature for Multi-view 3D Object Detection. 3325-3334
-     Ziying Song, Haiyue Wei, Lin Bai, Lei Yang, Caiyan Jia:
GraphAlign: Enhancing Accurate Feature Alignment by Graph matching for Multi-Modal 3D Object Detection. 3335-3346
-     Mikhail Terekhov, Viktor Larsson:
Tangent Sampson Error: Fast Approximate Two-view Reprojection Error for Central Camera Models. 3347-3355
-     Gilles Puy, Alexandre Boulch, Renaud Marlet:
Using a Waffle Iron for Automotive Point Cloud Semantic Segmentation. 3356-3366
-     Levente Hajder, Lajos Lóczsi, Daniel Barath:
Fast Globally Optimal Surface Normal from an Affine Correspondence.



3367-3378



Marcel C. Bühler, Kripasindhu Sarkar, Tanmay Shah, Gengyan Li , Daoye Wang, Leonhard Helminger, Sergio Orts-Escalano, Dmitry Lagun, Otmar Hilliges, Thabo Beeler, Abhimitra Meka:

Preface: A Data-driven Volumetric Prior for Few-shot Ultra High-resolution Face Synthesis. 3379-3390



Brent Yi, Weijia Zeng, Sam Buchanan, Yi Ma:
Canonical Factors for Hybrid Neural Fields. 3391-3403



Haobo Jiang, Zheng Dang, Shuo Gu, Jin Xie, Mathieu Salzmann, Jian Yang:
Center-Based Decoupled Point Cloud Registration for 6D Object Pose Estimation. 3404-3414



Annika Hagemann, Moritz Knorr, Christoph Stiller:
Deep geometry-aware camera self-calibration from video. 3415-3425



Nathaniel Burgdorfer, Philippos Mordohai:
V-FUSE: Volumetric Depth Map Fusion with Long-Range Constraints. 3426-3435



Yuxiang Cai, Yifan Zhu, Haiwei Zhang, Bo Ren:
Consistent Depth Prediction for Transparent Object Reconstruction from RGB-D Camera. 3436-3445



Sungwon Hwang, Junha Hyung, Daejin Kim, Min-Jung Kim, Jaegul Choo:
FaceCLIPNeRF: Text-driven 3D Face Manipulation using Deformable Neural Radiance Fields. 3446-3456



Xiufeng Xie, Riccardo Gherardi, Zhihong Pan, Stephen Huang:
HollowNeRF: Pruning Hashgrid-Based NeRFs with Trainable Collision Mitigation. 3457-3467



Jae-Hyeok Lee, Dae-Shik Kim:
ICE-NeRF: Interactive Color Editing of NeRFs via Decomposition-Aware Weight Optimization. 3468-3478



Zhijian Huang, Sihao Lin, Guiyu Liu, Mukun Luo, Chaoqiang Ye, Hang Xu, Xiaojun Chang , Xiaodan Liang:
FULLER: Unified Multi-modality Multi-task 3D Perception via Multi-level Gradient Calibration. 3479-3488



Aarrushi Shandilya, Benjamin Attal, Christian Richardt, James Tompkin, Matthew O'Toole:
Neural Fields for Structured Lighting. 3489-3499



Tao Xie, Ke Wang, Siyi Lu, Yukun Zhang, Kun Dai, Xiaoyu Li, Jie Xu, Li Wang, Lijun Zhao, Xinyu Zhang, Ruifeng Li:
CO-Net: Learning Multiple Point Cloud Tasks at Once with A Cohesive Network. 3500-3510



Jiahui Zhang, Fangneng Zhan, Yingchen Yu, Kunhao Liu, Rongliang Wu, Xiaoqin Zhang, Ling Shao, Shijian Lu:
Pose-Free Neural Radiance Fields via Implicit Pose Regularization. 3511-3520

-      Xiao Pan, Zongxin Yang, Jianxin Ma, Chang Zhou, Yi Yang:
TransHuman: A Transformer-based Human Representation for Generalizable Neural Human Rendering. 3521-3532
-     Haoyu Wu, Alexandros Graikos, Dimitris Samaras:
S-VoLSDF: Sparse Multi-View Stereo Regularization of Neural Implicit Surfaces. 3533-3545
-     Chaoran Tian, Weihong Pan, Zimo Wang, Mao Mao, Guofeng Zhang, Hujun Bao, Ping Tan, Zhaopeng Cui:
DPS-Net: Deep Polarimetric Stereo Depth Estimation. 3546-3556
-     Changyong Shu, Jiajun Deng, Fisher Yu, Yifan Liu:
3DPPE: 3D Point Positional Encoding for Transformer-based Multi-Camera 3D Object Detection. 3557-3566
-     Qi Ma, Danda Pani Paudel , Ajad Chhatkuli, Luc Van Gool:
Deformable Neural Radiance Fields using RGB and Event Cameras. 3567-3577
-     Jingyang Zhang, Yao Yao, Shiwei Li, Jingbo Liu, Tian Fang, David McKinnon, Yanghai Tsin, Long Quan:
NeILF++: Inter-Reflectable Light Fields for Geometry and Material Estimation. 3578-3587
-     Chunlin Ren, Qingshan Xu, Shikun Zhang, Jiaqi Yang:
Hierarchical Prior Mining for Non-local Multi-View Stereo. 3588-3597
-     Shihao Wang, Yingfei Liu, Tiancai Wang, Ying Li, Xiangyu Zhang:
Exploring Object-Centric Temporal Modeling for Efficient Multi-View 3D Object Detection. 3598-3608
-     Sara Rojas, Jesus Zarzar, Juan C. Pérez, Artsiom Sanakoyeu, Ali K. Thabet, Albert Pumarola, Bernard Ghanem :
Re-ReND: Real-time Rendering of NeRFs across Devices. 3609-3618
-     Xiaoyang Huang, Yi Zhang, Kai Chen, Teng Li, Wenjun Zhang, Bingbing Ni:
Learning Shape Primitives via Implicit Convexity Regularization. 3619-3628
-     Ruihong Yin, Sezer Karaoglu, Theo Gevers:
Geometry-guided Feature Learning and Fusion for Indoor Scene Reconstruction. 3629-3638
-     Zhiwei Zhang, Zhizhong Zhang, Qian Yu, Ran Yi, Yuan Xie, Lizhuang Ma:
LiDAR-Camera Panoptic Segmentation via Geometry-Consistent and Semantic-Aware Alignment. 3639-3648
-     Wenjie Ding, Limeng Qiao, Xi Qiu, Chi Zhang:
PivotNet: Vectorized Pivot Learning for End-to-end HD Map Construction. 3649-3659
-     Ming Qian, Jincheng Xiong, Gui-Song Xia, Nan Xue:
Sat2Density: Faithful Density Learning from Satellite-Ground Image Pairs. 3660-3669
-     Xin Lai, Yuhui Yuan, Ruihang Chu, Yukang Chen, Han Hu, Jiaya Jia:
Mask-Attention-Free Transformer for 3D Instance Segmentation. 3670-3680

-      Xiaoyong Lu, Yaping Yan, Tong Wei, Songlin Du:
Scene-Aware Feature Matching. 3681-3690
-     Zhuoxiao Chen, Yadan Luo, Zheng Wang, Mahsa Baktashmotlagh, Zi Huang :
Revisiting Domain-Adaptive 3D Object Detection by Reliable, Diverse and Class-balanced Pseudo-Labeling. 3691-3703
-     Youmin Zhang, Fabio Tosi, Stefano Mattoccia , Matteo Poggi:
GO-SLAM: Global Optimization for Consistent 3D Instant Reconstruction. 3704-3714
-     Valter Piedade, Pedro Miraldo:
BANSAC: A dynamic BAyesian Network for adaptive SAmple Consensus. 3715-3724
-     Felix Rydell, Elim Shehu, Angélica Torres:
Theoretical and Numerical Analysis of 3D Reconstruction Using Point and Line Incidences. 3725-3734
-     Haozhe Lin, Zequn Chen, Jinzhi Zhang, Bing Bai, Yu Wang, Ruqi Huang, Lu Fang:
RealGraph: A Multiview Dataset for 4D Real-world Context Graph Generation. 3735-3745
-     Kaiqiang Xiong, Rui Peng, Zhe Zhang, Tianxing Feng, Jianbo Jiao , Feng Gao, Ronggang Wang:
CL-MVSNet: Unsupervised Multi-view Stereo with Dual-level Contrastive Learning. 3746-3757
-     Zhuofan Zong, Dongzhi Jiang, Guanglu Song, Zeyue Xue, Jingyong Su, Hongsheng Li, Yu Liu:
Temporal Enhanced Training of Multi-view 3D Object Detector via Historical Object Prediction. 3758-3767
-     Zitian Wang, Zehao Huang, Jiahui Fu, Naiyan Wang, Si Liu:
Object as Query: Lifting any 2D Object Detector to 3D Detection. 3768-3777
-     Ming Nie, Yujing Xue, Chunwei Wang, Chaoqiang Ye, Hang Xu, Xinge Zhu, Qingqiu Huang, Michael Bi Mi, Xinchao Wang, Li Zhang:
PARTNER: Level up the Polar Representation for LiDAR 3D Object Detection. 3778-3790
-     Chuxin Wang, Wenfei Yang, Tianzhu Zhang:
Not Every Side Is Equal: Localization Uncertainty Estimation for Semi-Supervised 3D Object Detection. 3791-3801
-     Yifan Zhang, Zhen Dong, Huanrui Yang, Ming Lu, Cheng-Ching Tseng, Yuan Du, Kurt Keutzer, Li Du, Shanghang Zhang:
QD-BEV : Quantization-aware View-guided Distillation for Multi-view 3D Object Detection. 3802-3812
-     Lvmin Zhang, Anyi Rao , Maneesh Agrawala :
Adding Conditional Control to Text-to-Image Diffusion Models. 3813-3824
-     Liwen Wu, Rui Zhu, Mustafa B. Yaldiz, Yinhao Zhu, Hong Cai, Janarbek Matai, Fatih Porikli, Tzu-Mao Li, Manmohan Chandraker, Ravi Ramamoorthi:



Factorized Inverse Path Tracing for Efficient and Accurate Material-Lighting Estimation. 3825-3835

- Jianren Wang, Sudeep Dasari, Mohan Kumar Srirama, Shubham Tulsiani, Abhinav Gupta:
Manipulate by Seeing: Creating Manipulation Controllers from Pre-Trained Representations. 3836-3845
- Chengliang Zhong, Yuhang Zheng, Yupeng Zheng, Hao Zhao, Li Yi, Xiaodong Mu, Ling Wang, Pengfei Li, Guyue Zhou, Chao Yang, Xinliang Zhang, Jian Zhao:
3D Implicit Transporter for Temporally Consistent Keypoint Discovery. 3846-3857
- Nathan Mankovich, Tolga Birdal:
Chordal Averaging on Flag Manifolds and Its Applications. 3858-3867
- Weikang Wan, Haoran Geng, Yun Liu, Zikang Shan, Yaodong Yang , Li Yi, He Wang:
UniDexGrasp++: Improving Dexterous Grasping Policy Learning via Geometry-aware Curriculum and Iterative Generalist-Specialist Learning. 3868-3879
- Zhiyu Huang, Haochen Liu, Chen Lv:
GameFormer: Game-theoretic Modeling and Learning of Transformer-based Interactive Prediction and Planning for Autonomous Driving. 3880-3890
- Gengshan Yang, Shuo Yang, John Z. Zhang, Zachary Manchester, Deva Ramanan:
PPR: Physically Plausible Reconstruction from Monocular Videos. 3891-3901
- Wenjia Wang, Yongtao Ge, Haiyi Mei, Zhongang Cai, Qingping Sun, Yanjun Wang, Chunhua Shen, Lei Yang, Taku Komura:
Zolly: Zoom Focal Length Correctly for Perspective-Distorted Human Mesh Reconstruction. 3902-3912
- Hyekang Park, Jongyoun Noh, Youngmin Oh, Donghyeon Baek, Bumsuk Ham:
ACLS: Adaptive and Conditional Label Smoothing for Network Calibration. 3913-3922
- Jun Luo, Matías Mendieta, Chen Chen, Shandong Wu:
PGFed: Personalize Each Client's Global Objective for Federated Learning. 3923-3933
- Angelina Wang, Olga Russakovsky:
Overwriting Pretrained Bias with Finetuning Data. 3934-3945
- Cheng Zhang, Xuanbai Chen, Siqi Chai, Chen Henry Wu, Dmitry Lagun, Thabo Beeler, Fernando De la Torre:
ITI-Gen: Inclusive Text-to-Image Generation. 3946-3957
- Robin Hesse, Simone Schaub-Meyer, Stefan Roth:
FunnyBirds: A Synthetic Vision Dataset for a Part-Based Analysis of Explainable AI Methods. 3958-3968
- Bo Dai, Linge Wang, Baoxiong Jia, Zeyu Zhang, Song-Chun Zhu, Chi Zhang, Yixin Zhu:

**X-VoE: Measuring eXplanatory Violation of Expectation in Physical Events.**

3969-3979



Irena Gao, Gabriel Ilharco, Scott M. Lundberg, Marco Túlio Ribeiro:
Adaptive Testing of Computer Vision Models. 3980-3991



Alexander Kirillov, Eric Mintun, Nikhila Ravi, Hanzi Mao, Chloé Rolland, Laura Gustafson, Tete Xiao, Spencer Whitehead, Alexander C. Berg, Wan-Yen Lo, Piotr Dollár, Ross B. Girshick:
Segment Anything. 3992-4003



Perrine Chassat, Juhyun Park, Nicolas J.-B. Brunel:
Shape Analysis of Euclidean Curves under Frenet-Serret Framework. 4004-4013



Shyam Nandan Rai, Fabio Cermelli, Dario Fontanel, Carlo Masone , Barbara Caputo:
Unmasking Anomalies in Road-Scene Segmentation. 4014-4023



Lu Qi, Jason Kuen, Tiancheng Shen, Jiuxiang Gu, Wenbo Li, Weidong Guo, Jiaya Jia, Zhe Lin, Ming-Hsuan Yang:
High Quality Entity Segmentation. 4024-4033



Haochen Wang, Xiaolong Jiang, Xu Tang, Yao Hu, Cilin Yan, Weidi Xie, Shuai Wang, Efstratios Gavves:
Towards Open-Vocabulary Video Instance Segmentation. 4034-4043



Yutao Hu, Qixiong Wang, Wenqi Shao, Enze Xie, Zhenguo Li, Jungong Han, Ping Luo:
Beyond One-to-One: Rethinking the Referring Image Segmentation. 4044-4054



Wenhao Tang, Sheng Huang, Xiaoxian Zhang, Fengtao Zhou, Yi Zhang, Bo Liu:
Multiple Instance Learning Framework with Masked Hard Instance Mining for Whole Slide Image Classification. 4055-4064



Colorado J. Reed, Ritwik Gupta, Shufan Li, Sarah Brockman, Christopher Funk, Brian Clipp, Kurt Keutzer, Salvatore Candido, Matt Uyttendaele, Trevor Darrell:
Scale-MAE: A Scale-Aware Masked Autoencoder for Multiscale Geospatial Representation Learning. 4065-4076



Pandeng Li, Chen-Wei Xie, Liming Zhao, Hongtao Xie, Jiannan Ge, Yun Zheng, Deli Zhao, Yongdong Zhang:
Progressive Spatio-Temporal Prototype Matching for Text-Video Retrieval. 4077-4087



Junwen He, Yifan Wang, Lijun Wang, Huchuan Lu, Bin Luo, Jun-Yan He, Jin-Peng Lan, Yifeng Geng, Xuansong Xie:
Towards Deeply Unified Depth-aware Panoptic Segmentation with Bi-directional Guidance Learning. 4088-4098



Liulei Li, Wenguan Wang, Yang Yi:
LogicSeg: Parsing Visual Semantics with Neural Logic Learning and Reasoning. 4099-4110

-      Kamal Gupta, Varun Jampani, Carlos Esteves, Abhinav Shrivastava, Ameesh Makadia, Noah Snavely, Abhishek Kar:
ASIC: Aligning Sparse in-the-wild Image Collections. 4111-4122
-     Yael Vinker, Yuval Alaluf, Daniel Cohen-Or, Ariel Shamir:
CLIPascene: Scene Sketching with Different Types and Levels of Abstraction. 4123-4133
-     Koutilya PNVR, Bharat Singh, Pallabi Ghosh, Behjat Siddiquie, David Jacobs:
LD-ZNet: A Latent Diffusion Approach for Text-Based Image Segmentation. 4134-4145
-     Tianshi Cao, Karsten Kreis, Sanja Fidler, Nicholas Sharp, Kangxue Yin:
TexFusion: Synthesizing 3D Textures with Text-Guided Image Diffusion Models. 4146-4158
-     Zhang Chen, Zhong Li, Liangchen Song, Lele Chen, Jingyi Yu, Junsong Yuan, Yi Xu:
NeuRBF: A Neural Fields Representation with Adaptive Radial Basis Functions. 4159-4171
-     William Peebles, Saining Xie:
Scalable Diffusion Models with Transformers. 4172-4182
-     Xin Yu, Peng Dai, Wenbo Li, Lan Ma, Zhengze Liu, Xiaojuan Qi:
Texture Generation on 3D Meshes with Point-UV Diffusion. 4183-4193
-     Eric R. Chan, Koki Nagano, Matthew A. Chan, Alexander W. Bergman, Jeong Joon Park, Axel Levy, Miika Aittala, Shalini De Mello, Tero Karras, Gordon Wetzstein:
Generative Novel View Synthesis with 3D-Aware Diffusion Models. 4194-4206
-     Enze Xie, Lewei Yao, Han Shi, Zhili Liu, Daquan Zhou, Zhaoqiang Liu, Jiawei Li, Zhenguo Li:
DiffFit: Unlocking Transferability of Large Diffusion Models via Simple Parameter-Efficient Fine-Tuning. 4207-4216
-     Kyle Sargent, Jing Yu Koh, Han Zhang, Huiwen Chang, Charles Herrmann, Pratul Srinivasan, Jiajun Wu, Deqing Sun:
VQ3D: Learning a 3D-Aware Generative Model on ImageNet. 4217-4227
-     Wenhong Ge, Tao Hu, Haoyu Zhao, Shu Liu, Ying-Cong Chen:
Ref-NeuS: Ambiguity-Reduced Neural Implicit Surface Learning for Multi-View Reconstruction with Reflection. 4228-4237
-     Kushagra Pandey, Stephan Mandt:
A Complete Recipe for Diffusion Generative Models. 4238-4249
-     Yiqi Zhong, Luming Liang, Ilya Zharkov, Ulrich Neumann:
MMVP: Motion-Matrix-based Video Prediction. 4250-4260
-     Tomer Stolik, Itai Lang, Shai Avidan:
SAGA: Spectral Adversarial Geometric Attack on 3D Meshes. 4261-4271
-     Qiufan Ji, Lin Wang , Cong Shi, Shengshan Hu, Yingying Chen, Lichao Sun:
Benchmarking and Analyzing Robust Point Cloud Recognition: Bag of Tricks for Defending Adversarial Examples. 4272-4281



- Naufal Suryanto, Yongsu Kim, Harashta Tatimma Larasati, Hyoeun Kang, Thi-Thu-Huong Le, Yoonyoung Hong, Hunmin Yang, Se-Yoon Oh, Howon Kim:
ACTIVE: Towards Highly Transferable 3D Physical Camouflage for Universal and Robust Vehicle Evasion. 4282-4291
- Peifei Zhu, Genki Osada, Hirokatsu Kataoka, Tsubasa Takahashi:
Frequency-aware GAN for Adversarial Manipulation Generation. 4292-4301
- Heeseon Kim, Minji Son, Minbeom Kim, Myung-Joon Kwon, Changick Kim:
Breaking Temporal Consistency: Generating Video Universal Adversarial Perturbations Using Image Models. 4302-4311
- Han Fang, Jiyi Zhang, Yupeng Qiu, Jiayang Liu, Ke Xu, Chengfang Fang, Ee-Chien Chang:
Tracing the Origin of Adversarial Attack for Forensic Investigation and Deterrence. 4312-4321
- Ziqi Zhou, Shengshan Hu, Ruizhi Zhao, Qian Wang, Leo Yu Zhang, Junhui Hou, Hai Jin:
Downstream-agnostic Adversarial Examples. 4322-4332
- Zhigang Su, Dawei Zhou, Nannan Wang, Decheng Liu, Zhen Wang, Xinbo Gao:
Hiding Visual Information via Obfuscating Adversarial Perturbations. 4333-4343
- Changjiang Li, Ren Pang, Zhaohan Xi, Tianyu Du, Shouling Ji, Yuan Yao, Ting Wang:
An Embarrassingly Simple Backdoor Attack on Self-supervised Learning. 4344-4355
- Kaixun Jiang, Zhaoyu Chen, Hao Huang, Jiafeng Wang, Dingkang Yang, Bo Li, Yan Wang, Wenqiang Zhang:
Efficient Decision-based Black-box Patch Attacks on Video Recognition. 4356-4366
- Satoshi Suzuki, Shin'ya Yamaguchi, Shoichiro Takeda, Sekitoshi Kanai, Naoki Makishima, Atsushi Ando, Ryo Masumura:
Adversarial Finetuning with Latent Representation Constraint to Mitigate Accuracy-Robustness Tradeoff. 4367-4378
- Qingwen Bu, Dong Huang, Heming Cui:
Towards Building More Robust Models with Frequency Bias. 4379-4388
- Ningfei Wang, Yunpeng Luo, Takami Sato, Kaidi Xu, Qi Alfred Chen:
Does Physical Adversarial Example Really Matter to Autonomous Driving? Towards System-Level Effect of Adversarial Object Evasion Attack. 4389-4400
- Kaijie Zhu, Xixu Hu, Jindong Wang, Xing Xie, Ge Yang:
Improving Generalization of Adversarial Training via Robust Critical Fine-Tuning. 4401-4411
- Xuannan Liu, Yaoyao Zhong, Yuhang Zhang, Lixiong Qin, Weihong Deng:
Enhancing Generalization of Universal Adversarial Perturbation through Gradient Aggregation. 4412-4421

-      Xingxing Wei, Yao Huang, Yitong Sun, Jie Yu:
Unified Adversarial Patch for Cross-modal Attacks in the Physical World.
4422-4431
-     Donghua Wang, Wen Yao, Tingsong Jiang, Chao Li, Xiaoqian Chen:
RFLA: A Stealthy Reflected Light Adversarial Attack in the Physical World.
4432-4442
-     Mingli Zhu, Shaokui Wei, Li Shen, Yanbo Fan, Baoyuan Wu:
Enhancing Fine-Tuning based Backdoor Defense with Sharpness-Aware Minimization. 4443-4454
-     Ka-Chun Shum, Hong-Wing Pang, Bin-Son Hua, Duc Thanh Nguyen, Sai-Kit Yeung:
Conditional 360-degree Image Synthesis for Immersive Indoor Scene Decoration. 4455-4465
-     Bin Chen, Jia-Li Yin, Shukai Chen, Bohao Chen, Ximeng Liu:
An Adaptive Model Ensemble Adversarial Attack for Boosting Adversarial Transferability. 4466-4475
-     Byung-Kwan Lee, Junho Kim, Yong Man Ro:
Mitigating Adversarial Vulnerability through Causal Parameter Estimation by Adversarial Double Machine Learning. 4476-4486
-     Yaguan Qian, Shuke He, Chenyu Zhao, Jiaqiang Sha, Wei Wang, Bin Wang:
LEA²: A Lightweight Ensemble Adversarial Attack via Non-overlapping Vulnerable Frequency Regions. 4487-4498
-     Yulin Jin, Xiaoyu Zhang, Jian Lou, Xu Ma, Zilong Wang, Xiaofeng Chen:
Explaining Adversarial Robustness of Neural Networks from Clustering Effect Perspective. 4499-4508
-     Ruyi Ding, Shijin Duan, Xiaolin Xu, Yunsi Fei:
VertexSerum: Poisoning Graph Neural Networks for Link Inference. 4509-4518
-     Thibault Maho, Seyed-Mohsen Moosavi-Dezfooli, Teddy Furon:
How to choose your best allies for a transferable attack? 4519-4528
-     Dongyoon Yang, Insung Kong, Yongdai Kim:
Enhancing Adversarial Robustness in Low-Label Regime via Adaptively Weighted Regularization and Knowledge Distillation. 4529-4538
-     Xinquan Chen, Xitong Gao, Juanjuan Zhao, Kejiang Ye, Cheng-Zhong Xu:
AdvDiffuser: Natural Adversarial Example Synthesis with Diffusion Models. 4539-4549
-     Tao Zhou, Qi Ye, Wenhan Luo, Kaihao Zhang, Zhiguo Shi, Jiming Chen:
F&F Attack: Adversarial Attack against Multiple Object Trackers by Inducing False Negatives and False Positives. 4550-4560
-     Lukas Struppek, Dominik Hintersdorf, Kristian Kersting:
Rickrolling the Artist: Injecting Backdoors into Text Encoders for Text-to-Image Synthesis. 4561-4573

-      Zhengzhi Lu, He Wang, Ziyi Chang, Guoan Yang, Hubert P. H. Shum: **Hard No-Box Adversarial Attack on Skeleton-Based Human Action Recognition with Skeleton-Motion-Informed Gradient.** 4574-4583
-     Xiaosen Wang, Zeliang Zhang, Jianping Zhang: **Structure Invariant Transformation for better Adversarial Transferability.** 4584-4596
-     Min Liu, Alberto L. Sangiovanni-Vincentelli, Xiangyu Yue: **Beating Backdoor Attack at Its Own Game.** 4597-4606
-     Wenshuo Ma, Yidong Li, Xiaofeng Jia, Wei Xu: **Transferable Adversarial Attack for Both Vision Transformers and Convolutional Networks via Momentum Integrated Gradients.** 4607-4616
-     Nabeel Hingun, Chawin Sitawarin, Jerry Li, David A. Wagner: **REAP: A Large-Scale Realistic Adversarial Patch Benchmark.** 4617-4628
-     Siquan Huang, Yijiang Li, Chong Chen, Leyu Shi, Ying Gao: **Multi-metrics adaptively identifies backdoors in Federated learning.** 4629-4639
-     Zhuoer Xu, Zhangxuan Gu, Jianping Zhang, Shiwen Cui, Changhua Meng, Weiqiang Wang: **Backpropagation Path Search On Adversarial Transferability.** 4640-4650
-     Teresa Yeo, Oguzhan Fatih Kar, Zahra Sodagar, Amir Zamir: **Rapid Network Adaptation: Learning to Adapt Neural Networks Using Test-Time Feedback.** 4651-4664
-     Jianshuo Dong, Han Qiu, Yiming Li, Tianwei Zhang, Yuanjie Li, Zeqi Lai, Chao Zhang, Shu-Tao Xia: **One-bit Flip is All You Need: When Bit-flip Attack Meets Model Training.** 4665-4675
-     Junfeng Guo, Ang Li, Lixu Wang, Cong Liu: **PolicyCleanse: Backdoor Detection and Mitigation for Competitive Reinforcement Learning.** 4676-4685
-     Shouwei Ruan, Yinpeng Dong, Hang Su, Jianteng Peng, Ning Chen, Xingxing Wei: **Towards Viewpoint-Invariant Visual Recognition via Adversarial Training.** 4686-4696
-     Mengnan Zhao, Lihe Zhang, Yuqiu Kong, Baocai Yin: **Fast Adversarial Training with Smooth Convergence.** 4697-4706
-     Virat Shejwalkar, Lingjuan Lyu, Amir Houmansadr: **The Perils of Learning From Unlabeled Data: Backdoor Attacks on Semi-supervised Learning.** 4707-4717
-     Hegui Zhu, Yuchen Ren , Xiaoyan Sui, Lianping Yang, Wuming Jiang: **Boosting Adversarial Transferability via Gradient Relevance Attack.** 4718-4727
-     Guanhao Gan, Yiming Li, Dongxian Wu, Shu-Tao Xia: **Towards Robust Model Watermark via Reducing Parametric Vulnerability.** 4728-4738

-      Yiran Liu, Xin Feng, Yunlong Wang, Wu Yang, Di Ming :
TRM-UAP: Enhancing the Transferability of Data-Free Universal Adversarial Perturbation via Truncated Ratio Maximization. 4739-4748
-     Guangnian Wan, Haitao Du, Xuejing Yuan, Jun Yang, Meiling Chen, Jie Xu:
Enhancing Privacy Preservation in Federated Learning via Learning Rate Perturbation. 4749-4758
-     Jie Zhang, Chen Chen, Weiming Zhuang, Lingjuan Lyu:
TARGET: Federated Class-Continual Learning via Exemplar-Free Distillation. 4759-4770
-     Sriram Yenamandra, Pratik Ramesh, Viraj Prabhu, Judy Hoffman:
FACTS: First Amplify Correlations and Then Slice to Discover Bias. 4771-4781
-     Yutong Wu, Xingshuo Han, Han Qiu, Tianwei Zhang:
Computation and Data Efficient Backdoor Attacks. 4782-4791
-     Yaopei Zeng, Lei Liu, Li Liu, Li Shen, Shaoguo Liu, Baoyuan Wu:
Global Balanced Experts for Federated Long-Tailed Learning. 4792-4802
-     Qucheng Peng, Ce Zheng, Chen Chen:
Source-free Domain Adaptive Human Pose Estimation. 4803-4813
-     Nicole Meister, Dora Zhao, Angelina Wang, Vikram V. Ramaswamy, Ruth Fong, Olga Russakovsky:
Gender Artifacts in Visual Datasets. 4814-4825
-     Haokun Chen, Ahmed Frikha, Denis Krompass, Jindong Gu, Volker Tresp:
FRAug: Tackling Federated Learning with Non-IID Features via Representation Augmentation. 4826-4836
-     Zahra Ghodsi, Mojtaba Javaheripour, Nojan Sheybani, Xinqiao Zhang, Ke Huang, Farinaz Koushanfar:
zPROBE: Zero Peak Robustness Checks for Federated Learning. 4837-4847
-     Myeongseob Ko, Ming Jin, Chenguang Wang, Ruoxi Jia:
Practical Membership Inference Attacks Against Large-Scale Multi-Modal Models: A Pilot Study. 4848-4858
-     Chen Yang, Meilu Zhu , Yifan Liu, Yixuan Yuan:
FedPD: Federated Open Set Recognition with Parameter Disentanglement. 4859-4868
-     Junxu Liu, Mingsheng Xue, Jian Lou, Xiaoyu Zhang, Li Xiong, Zhan Qin:
MUter: Machine Unlearning on Adversarially Trained Models. 4869-4879
-     William Thong, Przemyslaw Joniak, Alice Xiang:
Beyond Skin Tone: A Multidimensional Measure of Apparent Skin Color. 4880-4890
-     Jannik Brinkmann, Paul Swoboda, Christian Bartelt:
A Multidimensional Analysis of Social Biases in Vision Transformers. 4891-4900
-     Jiaxuan Li, Duc Minh Vo, Hideki Nakayama:
Partition-and-Debias: Agnostic Biases Mitigation via A Mixture of Biases-

-      **Specific Experts.** 4901-4911
Dongyao Zhu, Yanbo Fang, Bowen Lei, Yiqun Xie, Dongkuan Xu, Jie Zhang, Ruqi Zhang:
Rethinking Data Distillation: Do Not Overlook Calibration. 4912-4922
-     Rémi Nahon, Van-Tam Nguyen, Enzo Tartaglione:
Mining bias-target Alignment from Voronoi Cells. 4923-4932
-     Ming-Chang Chiu, Pin-Yu Chen, Xuezhe Ma:
Better May Not Be Fairer: A Study on Subgroup Discrepancy in Image Classification. 4933-4943
-     Hao Fang, Bin Chen, Xuan Wang, Zhi Wang, Shu-Tao Xia:
GIFD: A Generative Gradient Inversion Method with Feature Domain Optimization. 4944-4953
-     Hao Liang, Pietro Perona, Guha Balakrishnan:
Benchmarking Algorithmic Bias in Face Recognition: An Experimental Approach Using Synthetic Faces and Human Evaluation. 4954-4964
-     Guangyu Sun, Matías Mendieta, Jun Luo, Shandong Wu, Chen Chen:
FedPerfix: Towards Partial Model Personalization of Vision Transformers in Federated Learning. 4965-4975
-     Sungwon Han, Sungwon Park, Fangzhao Wu, Sundong Kim, Bin Zhu, Xing Xie, Meeyoung Cha:
Towards Attack-tolerant Federated Learning via Critical Parameter Analysis. 4976-4985
-     Ziheng Huang, Boheng Li, Yan Cai, Run Wang, Shangwei Guo, Liming Fang, Jing Chen, Lina Wang:
What can Discriminator do? Towards Box-free Ownership Verification of Generative Adversarial Networks. 4986-4996
-     Xiuwen Fang, Mang Ye, Xiyuan Yang:
Robust Heterogeneous Federated Learning under Data Corruption. 4997-5007
-     Yuhao Zhou, Mingjia Shi, Yuanxi Li, Yanan Sun, Qing Ye, Jiancheng Lv:
Communication-efficient Federated Learning with Single-Step Synthetic Features Compressor for Faster Convergence. 5008-5017
-     Jianqing Zhang, Yang Hua, Hao Wang, Tao Song, Zhengui Xue, Ruhui Ma, Jian Cao, Haibing Guan:
GPFL: Simultaneously Learning Global and Personalized Feature Information for Personalized Federated Learning. 5018-5028
-     Wenxuan Zeng, Meng Li, Wenjie Xiong , Tong Tong, Wen-Jie Lu, Jin Tan, Runsheng Wang, Ru Huang:
MPCViT: Searching for Accurate and Efficient MPC-Friendly Vision Transformer with Heterogeneous Attention. 5029-5040
-     Jan Hendrik Metzen, Robin Hutmacher, N. Grace Hua, Valentyn Boreiko, Dan Zhang:

**Identification of Systematic Errors of Image Classifiers on Rare Subgroups.**

5041-5050



Nadiya Shvai, Arcadi Llanza Carmona, Amir Nakib:

Adaptive Image Anonymization in the Context of Image Classification with Neural Networks. 5051-5060

Saeed Vahidian, Sreevatsank Kadaveru, Woonjoon Baek, Weijia Wang,

Vyacheslav Kungurtsev, Chen Chen, Mubarak Shah, Bill Lin:

When Do Curricula Work in Federated Learning? 5061-5071

Haotian Wang, Haoang Chi, Wenjing Yang, Zhipeng Lin, Mingyang Geng, Long

Lan, Jing Zhang, Dacheng Tao:

Domain Specified Optimization for Deployment Authorization. 5072-5082

Ming Li, Xiangyu Xu, Hehe Fan, Pan Zhou, Jun Liu, Jia-Wei Liu, Jiahe Li, Jussi

Keppo, Mike Zheng Shou, Shuicheng Yan:

STPrivacy: Spatio-Temporal Privacy-Preserving Action Recognition. 5083-

5092



Yuke Zhang, Dake Chen, Souvik Kundu, Chenghao Li, Peter A. Beerel :

SAL-ViT: Towards Latency Efficient Private Inference on ViT using Selective Attention Search with a Learnable Softmax Approximation. 5093-5102

Chi Zhang, Xiaoman Zhang, Ekanut Sotthiwat, Yanyu Xu, Ping Liu, Liangli Zhen, Yong Liu:

Generative Gradient Inversion via Over-Parameterized Networks in Federated Learning. 5103-5112

Abhipsa Basu, R. Venkatesh Babu, Danish Pruthi:

Inspecting the Geographical Representativeness of Images from Text-to-Image Models. 5113-5124

Yunqian Wen, Bo Liu , Jingyi Cao, Rong Xie, Li Song:

Divide and Conquer: a Two-Step Method for High Quality Face De-identification with Model Explainability. 5125-5134

Yizhe Li, Yu-Lin Tsai, Chia-Mu Yu, Pin-Yu Chen, Xuebin Ren:

Exploring the Benefits of Visual Prompting in Differential Privacy. 5135-5144

Lei Zhang, Zhibo Wang, Xiaowei Dong, Yunhe Feng, Xiaoyi Pang, Zhifei Zhang, Kui Ren:

Towards Fairness-aware Adversarial Network Pruning. 5145-5154

Hongwu Peng, Shaoyi Huang, Tong Zhou, Yukui Luo, Chenghong Wang, Zigeng Wang, Jiahui Zhao, Xi Xie, Ang Li, Tony Geng, Kaleel Mahmood, Wujie Wen, Xiaolin Xu, Caiwen Ding:

AutoReP: Automatic ReLU Replacement for Fast Private Network Inference. 5155-5165

Xingxuan Zhang, Renzhe Xu, Han Yu, Yancheng Dong, Pengfei Tian, Peng Cui:

Flatness-Aware Minimization for Domain Generalization. 5166-5179

Jingwei Sun, Ziyue Xu, Dong Yang, Vishwesh Nath, Wenqi Li, Can Zhao, Daguang Xu, Yiran Chen, Holger R. Roth:



Communication-Efficient Vertical Federated Learning with Limited Overlapping Samples. 5180-5189

- Gorjan Radevski, Dusan Grujicic, Matthew B. Blaschko, Marie-Francine Moens, Tinne Tuytelaars:

Multimodal Distillation for Egocentric Action Recognition. 5190-5201

- Peri Akiva, Jing Huang, Kevin J. Liang, Rama Kovvuri, Xingyu Chen, Matt Feiszli, Kristin J. Dana, Tal Hassner:

Self-Supervised Object Detection from Egocentric Videos. 5202-5214

- Lorenzo Mur-Labadia, Josechu J. Guerrero, Ruben Martinez-Cantin:

Multi-label affordance mapping from egocentric vision. 5215-5226

- Huiyu Wang, Mitesh Kumar Singh, Lorenzo Torresani:

Ego-Only: Egocentric Action Detection without Exocentric Transferring.

5227-5238

- Boxiao Pan, Bokui Shen, Davis Rempe, Despoina Paschalidou, Kaichun Mo, Yanchao Yang, Leonidas J. Guibas:

COPilot: Human-Environment Collision Prediction and Localization from Egocentric Videos. 5239-5249

- Yue Xu, Yong-Lu Li, Zhemin Huang, Michael Xu Liu, Cewu Lu, Yu-Wing Tai, Chi-Keung Tang:

EgoPCA: A New Framework for Egocentric Hand-Object Interaction Understanding. 5250-5261

- Shraman Pramanick, Yale Song, Sayan Nag, Kevin Qinghong Lin, Hardik Shah, Mike Zheng Shou, Rama Chellappa, Pengchuan Zhang:

EgoVLpv2: Egocentric Video-Language Pre-training with Fusion in the Backbone. 5262-5274

- Yiye Chen, Yunzhi Lin, Ruinian Xu, Patricio A. Vela:

WDiscOOD: Out-of-Distribution Detection via Whitened Linear Discriminant Analysis. 5275-5284

- Yandong Wen, Weiyang Liu, Yao Feng, Bhiksha Raj, Rita Singh, Adrian Weller, Michael J. Black, Bernhard Schölkopf:

Pairwise Similarity Learning is SimPLE. 5285-5295

- Zexi Li, Xinyi Shang, Rui He, Tao Lin, Chao Wu:

No Fear of Classifier Biases: Neural Collapse Inspired Federated Learning with Synthetic and Fixed Classifier. 5296-5306

- Jeffrey Gu, Kuan-Chieh Wang, Serena Yeung :

Generalizable Neural Fields as Partially Observed Neural Processes. 5307-5316

- Fabian Mentzer, Eirikur Agustsson, Michael Tschannen:

M2T: Masking Transformers Twice for Faster Decoding. 5317-5326

- Bill Psomas, Ioannis Kakogeorgiou, Konstantinos Karantzalos, Yannis Avrithis:

Keep It SimPool: Who Said Supervised Transformers Suffer from Attention Deficit? 5327-5337

-      Yuan Liu, Songyang Zhang, Jiacheng Chen, Zhaohui Yu, Kai Chen, Dahua Lin:
Improving Pixel-based MIM by Reducing Wasted Modeling Capability. 5338-5349
-     Kechun Liu, Yitong Jiang, Inchang Choi, Jinwei Gu:
Learning Image-Adaptive Codebooks for Class-Agnostic Image Restoration. 5350-5360
-     Ruchika Chavhan, Henry Gouk, Da Li, Timothy M. Hospedales:
Quality Diversity for Visual Pre-Training. 5361-5371
-     Chengkai Hou, Jieyu Zhang, Haonan Wang, Tianyi Zhou:
Subclass-balancing Contrastive Learning for Long-tailed Recognition. 5372-5384
-     Sotiris Anagnostidis, Aurélien Lucchi, Thomas Hofmann:
Mastering Spatial Graph Prediction of Road Networks. 5385-5395
-     Max van Spengler, Erwin Berkhouit, Pascal Mettes:
Poincaré ResNet. 5396-5405
-     Xiaotong Li, Zixuan Hu, Yixiao Ge, Ying Shan, Ling-Yu Duan:
Exploring Model Transferability through the Lens of Potential Energy. 5406-5415
-     Yixuan Wei, Han Hu, Zhenda Xie, Ze Liu, Zheng Zhang, Yue Cao, Jianmin Bao, Dong Chen, Baining Guo:
Improving CLIP Fine-tuning Performance. 5416-5426
-     Tianjiao Ding, Shengbang Tong, Kwan Ho Ryan Chan, Xili Dai, Yi Ma, Benjamin D. Haeffele:
Unsupervised Manifold Linearizing and Clustering. 5427-5438
-     Yeti Ziya Gürbüz, Ozan Sener, A. Aydin Alatan:
Generalized Sum Pooling for Metric Learning. 5439-5450
-     Ke Liu, Feng Liu, Haishuai Wang, Ning Ma, Jiajun Bu, Bo Han:
Partition Speeds Up Learning Implicit Neural Representations Based on Exponential-Increase Hypothesis. 5451-5460
-     Mannat Singh, Quentin Duval, Kalyan Vasudev Alwala, Haoqi Fan, Vaibhav Aggarwal, Aaron Adcock, Armand Joulin, Piotr Dollár, Christoph Feichtenhofer, Ross B. Girshick, Rohit Girdhar, Ishan Misra:
The effectiveness of MAE pre-pretraining for billion-scale pretraining. 5461-5471
-     Han Xiao, Wenzhao Zheng, Zheng Zhu, Jie Zhou, Jiwen Lu:
Token-Label Alignment for Vision Transformers. 5472-5481
-     Nishant Jain, Harkirat S. Behl, Yogesh Singh Rawat, Vibhav Vineet:
Efficiently Robustify Pre-Trained Models. 5482-5492
-     Tao Xie, Kun Dai, Siyi Lu, Ke Wang, Zhiqiang Jiang, Jinghan Gao, Dedong Liu, Jie Xu, Lijun Zhao, Ruifeng Li:
OFVL-MS: Once for Visual Localization across Multiple Indoor Scenes. 5493-5503

-  [Download](#) [Cite](#) [Share](#) Cheng Yan, Shiyu Zhang, Yang Liu, Guansong Pang, Wenjun Wang:
Feature Prediction Diffusion Model for Video Anomaly Detection. 5504-5514
-  [Download](#) [Cite](#) [Share](#) Chia-Hao Chen, Ying-Tian Liu, Zhifei Zhang, Yuan-Chen Guo, Song-Hai Zhang:
Joint Implicit Neural Representation for High-fidelity and Compact Vector Fonts. 5515-5525
-  [Download](#) [Cite](#) [Share](#) Zijian Wang, Yadan Luo, Liang Zheng, Zi Huang , Mahsa Baktashmotlagh:
How Far Pre-trained Models Are from Neural Collapse on the Target Dataset Informs their Transferability. 5526-5535
-  [Download](#) [Cite](#) [Share](#) Chengkun Wang, Wenzhao Zheng, Zheng Zhu, Jie Zhou, Jiwen Lu:
OPERA: Omni-Supervised Representation Learning with Hierarchical Supervisions. 5536-5547
-  [Download](#) [Cite](#) [Share](#) Kanchana Ranasinghe, Brandon McKinzie, Sachin Ravi, Yinfei Yang, Alexander Toshev, Jonathon Shlens:
Perceptual Grouping in Contrastive Vision-Language Models. 5548-5561
-  [Download](#) [Cite](#) [Share](#) Bingyin Zhao, Zhiding Yu, Shiyi Lan, Yutao Cheng, Anima Anandkumar, Yingjie Lao, José M. Álvarez:
Fully Attentional Networks with Self-emerging Token Labeling. 5562-5572
-  [Download](#) [Cite](#) [Share](#) Xudong Tian, Zhizhong Zhang, Xin Tan, Jun Liu, Chengjie Wang, Yanyun Qu, Guannan Jiang, Yuan Xie:
Instance and Category Supervision are Alternate Learners for Continual Learning. 5573-5582
-  [Download](#) [Cite](#) [Share](#) Hong Yan, Yang Liu, Yushen Wei, Zhen Li, Guanbin Li, Liang Lin:
SkeletonMAE: Graph-based Masked Autoencoder for Skeleton Sequence Pre-training. 5583-5595
-  [Download](#) [Cite](#) [Share](#) David Fan, Jue Wang, Shuai Liao, Yi Zhu, Vimal Bhat, Hector J. Santos-Villalobos, Rohith MV, Xinyu Li:
Motion-Guided Masking for Spatiotemporal Representation Learning. 5596-5606
-  [Download](#) [Cite](#) [Share](#) Enneng Yang, Li Shen, Zhenyi Wang, Shiwei Liu, Guibing Guo, Xingwei Wang:
Data Augmented Flatness-aware Gradient Projection for Continual Learning. 5607-5616
-  [Download](#) [Cite](#) [Share](#) Ziyi Wang, Xumin Yu, Yongming Rao, Jie Zhou, Jiwen Lu:
Take-A-Photo: 3D-to-2D Generative Pre-training of Point Cloud Models. 5617-5627
-  [Download](#) [Cite](#) [Share](#) Yefei He, Zhenyu Lou, Luoming Zhang, Jing Liu, Weijia Wu, Hong Zhou, Bohan Zhuang:
BiViT: Extremely Compressed Binary Vision Transformers. 5628-5640
-  [Download](#) [Cite](#) [Share](#) Sepehr Sameni, Simon Jenni, Paolo Favaro:
Spatio-Temporal Crop Aggregation for Video Representation Learning. 5641-5651
-  [Download](#) [Cite](#) [Share](#) Hanjae Kim, Jiyoung Lee, Seongheon Park, Kwanghoon Sohn:
Hierarchical Visual Primitive Experts for Compositional Zero-Shot Learning. 5652-5662

-  [Shengjiang Quan, Masahiro Hirano, Yuji Yamakawa:](#) **Semantic Information in Contrastive Learning.** 5663-5673
-  [Xuehan Bai, Yan Li, Yanhua Cheng, Wenjie Yang, Quan Chen, Han Li:](#) **Cross-Domain Product Representation Learning for Rich-Content E-Commerce.** 5674-5683
-  [Haoyang Cheng, Haitao Wen, Xiaoliang Zhang](#) , Heqian Qiu, Lanxiao Wang, Hongliang Li: **Contrastive Continuity on Augmentation Stability Rehearsal for Continual Self-Supervised Learning.** 5684-5694
-  [Mehmet Kerim Yucel](#) , Ramazan Gokberk Cinbis, Pinar Duygulu: **HybridAugment++: Unified Frequency Spectra Perturbations for Model Robustness.** 5695-5705
-  [Wenliang Zhao, Yongming Rao, Zuyan Liu, Benlin Liu, Jie Zhou, Jiwen Lu:](#) **Unleashing Text-to-Image Diffusion Models for Visual Perception.** 5706-5716
-  [Abhishek Aich, Samuel Schulter, Amit K. Roy-Chowdhury, Manmohan Chandraker, Yumin Suh:](#) **Efficient Controllable Multi-Task Architectures.** 5717-5728
-  [Ruihan Xu, Haokui Zhang, Wenze Hu, Shiliang Zhang, Xiaoyu Wang:](#) **ParCNetV2: Oversized Kernel with Enhanced Attention^{*}.** 5729-5739
-  [Zihao Sun, Yu Sun, Longxing Yang, Shun Lu, Jilin Mei, Wenxiao Zhao, Yu Hu:](#) **Unleashing the Power of Gradient Signal-to-Noise Ratio for Zero-Shot NAS.** 5740-5750
-  [Fudong Lin, Summer Crawford, Kaleb Guillot, Yihe Zhang, Yan Chen, Xu Yuan](#) , Li Chen, Shelby Williams, Robert Minvielle, Xiangming Xiao, Drew Gholson, Nicolas Ashwell, Tri Setiyono, Brenda Tubana, Lu Peng, Magdy A. Bayoumi, Nian-Feng Tzeng : **MMST-ViT: Climate Change-aware Crop Yield Prediction via Multi-Modal Spatial-Temporal Vision Transformer.** 5751-5761
-  [Pavan Kumar Anasosalu Vasu, James Gabriel, Jeff Zhu, Oncel Tuzel, Anurag Ranjan:](#) **FastViT: A Fast Hybrid Vision Transformer using Structural Reparameterization.** 5762-5772
-  [Sudong Cai:](#) **IIEU: Rethinking Neural Feature Activation from Decision-Making.** 5773-5783
-  [Nam Hyeon-Woo, Kim Yu-Ji, Byeongho Heo, Dongyoon Han, Seong Joon Oh, Tae-Hyun Oh:](#) **Scratching Visual Transformer's Back with Uniform Attention.** 5784-5795
-  [Xudong Wang, Li Lyна Zhang, Jiahang Xu, Quanlu Zhang, Yujing Wang, Yuqing Yang, Ningxin Zheng, Ting Cao, Mao Yang:](#) **SpaceEvo: Hardware-Friendly Search Space Design for Efficient INT8 Inference.** 5796-5805



Chen Tang, Li Lyra Zhang, Huiqiang Jiang, Jiahang Xu, Ting Cao, Quanlu Zhang, Yuqing Yang, Zhi Wang, Mao Yang:

ElasticViT: Conflict-aware Supernet Training for Deploying Fast Vision Transformer on Diverse Mobile Devices. 5806-5817



Jongbin Ryu, Dongyoon Han, Jongwoo Lim:

Gramian Attention Heads are Strong yet Efficient Vision Learners. 5818-5828



Yulin Wang , Yang Yue, Rui Lu, Tianjiao Liu, Zhao Zhong, Shiji Song, Gao Huang:
EfficientTrain: Exploring Generalized Curriculum Learning for Training Visual Backbones. 5829-5841



Jinhong Wang, Yi Cheng, Jintai Chen, Tingting Chen, Danny Chen, Jian Wu:
Ord2Seq: Regarding Ordinal Regression as Label Sequence Prediction. 5842-5852



Shipeng Bai, Jun Chen, Xintian Shen, Yixuan Qian, Yong Liu:
Unified Data-Free Compression: Pruning and Quantization without Fine-Tuning. 5853-5862



Runyi Yu, Zhennan Wang, Yinhuai Wang, Kehan Li, Chang Liu, Haoyi Duan,

Xiangyang Ji, Jie Chen:

LaPE: Layer-adaptive Position Embedding for Vision Transformers with Independent Layer Normalization. 5863-5873



Anurag Roy, Vinay Kumar Verma, Sravan Voonna, Kripabandhu Ghosh, Saptarshi Ghosh, Abir Das:

Exemplar-Free Continual Transformer with Convolutions. 5874-5884



Yongjie Chen, Hongmin Liu, Haoran Yin, Bin Fan:

Building Vision Transformers with Hierarchy Aware Feature Aggregation. 5885-5895



Mingyang Zhang, Xinyi Yu, Haodong Zhao, Linlin Ou:

ShiftNAS: Improving One-shot NAS via Probability Shift. 5896-5905



Akshaya Athwale, Arman Afrasiyabi, Justin Lagüe, Ichrak Shili, Ola Ahmad, Jean-François Lalonde:

DarSwin: Distortion Aware Radial Swin Transformer. 5906-5915



Xiaoxing Wang, Xiangxiang Chu, Yuda Fan, Zhixi Zhang, Bo Zhang, Xiaokang Yang, Junchi Yan:

ROME: Robustifying Memory-Efficient NAS via Topology Disentanglement and Gradient Accumulation. 5916-5926



Yixing Xu, Chao Li, Dong Li, Xiao Sheng, Fan Jiang, Lu Tian, Ashish Sirasao:

FDViT: Improve the Hierarchical Architecture of Vision Transformer. 5927-5937



Dongchen Han, Xuran Pan, Yizeng Han, Shiji Song, Gao Huang:

FLatten Transformer: Vision Transformer using Focused Linear Attention. 5938-5948



Xiangxiang Chu, Shun Lu, Xudong Li, Bo Zhang:

MixPath: A Unified Approach for One-shot Neural Architecture Search.



5949-5958

- Jingtao Wang, Zengjie Song, Yuxi Wang, Jun Xiao, Yuran Yang, Shuqi Mei, Zhaoxiang Zhang:
SSF: Accelerating Training of Spiking Neural Networks with Stabilized Spiking Flow. 5959-5968
- Yizeng Han, Dongchen Han, Zeyu Liu, Yulin Wang , Xuran Pan, Yifan Pu, Chao Deng, Junlan Feng, Shiji Song, Gao Huang:
Dynamic Perceiver for Efficient Visual Recognition. 5969-5979
- Sucheng Ren, Xingyi Yang, Songhua Liu, Xinchao Wang:
SG-Former: Self-guided Transformer with Evolving Token Reallocation. 5980-5991
- Weifeng Lin, Ziheng Wu, Jiayu Chen, Jun Huang, Lianwen Jin:
Scale-Aware Modulation Meet Transformer. 5992-6003
- Wenze Liu, Hao Lu, Hongtao Fu, Zhiguo Cao:
Learning to Upsample by Learning to Sample. 6004-6014
- Yansong Peng, Yueyi Zhang, Zhiwei Xiong, Xiaoyan Sun, Feng Wu:
GET: Group Event Transformer for Event-Based Vision. 6015-6025
- Zhipeng Huang, Zhizheng Zhang, Cuiling Lan, Zheng-Jun Zha, Yan Lu, Baining Guo:
Adaptive Frequency Filters As Efficient Global Token Mixers. 6026-6036
- Haokui Zhang, Wenze Hu, Xiaoyu Wang:
Fcaformer: Forward Cross Attention in Hybrid Vision Transformer. 6037-6046
- Yaolei Qi, Yuting He, Xiaoming Qi, Yuan Zhang, Guanyu Yang:
Dynamic Snake Convolution based on Topological Geometric Constraints for Tubular Structure Segmentation. 6047-6056
- Seyedalireza Khoshsirat, Chandra Kambhamettu:
Sentence Attention Blocks for Answer Grounding. 6057-6067
- Quang Hieu Vo, Linh-Tam Tran, Sung-Ho Bae, Lok-Won Kim, Choong Seon Hong:
MST-compression: Compressing and Accelerating Binary Neural Networks with Minimum Spanning Tree. 6068-6077
- Ilwi Yun, Chanyong Shin, Hyunku Lee, Hyuk-Jae Lee, Chae-Eun Rhee:
EGformer: Equirectangular Geometry-biased Transformer for 360 Depth Estimation. 6078-6089
- Guhnoo Yun, Juhan Yoo, Kijung Kim, Jeongho Lee, Dong Hwan Kim:
SPANet: Frequency-balancing Token Mixer using Spectral Pooling Aggregation Modulation. 6090-6101
- Jie Song, Zhengqi Xu, Sai Wu, Gang Chen, Mingli Song:
ModelGiF: Gradient Fields for Model Functional Distance. 6102-6112
- Gustavo Adolfo Vargas Hakim, David Osowiecki, Mehrdad Noori, Milad Cheraghalkhani, Ali Bahri, Ismail Ben Ayed, Christian Desrosiers:
ClusT3: Information Invariant Test-Time Training. 6113-6112

-      Borui Zhao, Renjie Song, Jiajun Liang:
Cumulative Spatial Knowledge Distillation for Vision Transformers. 6123-6132
-     Jong-Hyeon Baek, Daehyun Kim, Su-Min Choi, Hyo-Jun Lee, Hanul Kim, Yeong Jun Koh:
Luminance-aware Color Transform for Multiple Exposure Correction. 6133-6142
-     Qingyan Meng, Mingqing Xiao, Shen Yan, Yisen Wang, Zhouchen Lin, Zhi-Quan Luo:
Towards Memory- and Time-Efficient Backpropagation for Training Spiking Neural Networks. 6143-6153
-     Mateusz Michalkiewicz, Masoud Faraki, Xiang Yu, Manmohan Chandraker, Mahsa Baktashmotlagh:
Domain Generalization Guided by Gradient Signal to Noise Ratio of Parameters. 6154-6165
-     Borui Zhao, Quan Cui, Renjie Song, Jiajun Liang:
DOT: A Distillation-Oriented Trainer. 6166-6175
-     Yuhong Li, Jiajie Li, Cong Hao, Pan Li, Jinjun Xiong, Deming Chen:
Extensible and Efficient Proxy for Neural Architecture Search. 6176-6187
-     Utkarsh Singhal, Carlos Esteves, Ameesh Makadia, Stella X. Yu:
Learning to Transform for Generalizable Instance-wise Invariance. 6188-6198
-     Alexandre Kirchmeyer, Jia Deng:
Convolutional Networks with Oriented 1D Kernels. 6199-6209
-     Yanghao Wang, Zhongqi Yue, Xian-Sheng Hua, Hanwang Zhang:
Random Boxes Are Open-world Object Detectors. 6210-6220
-     Yuxin Fang, Shusheng Yang, Shijie Wang, Yixiao Ge, Ying Shan, Xinggang Wang:
Unleashing Vanilla Vision Transformer with Masked Image Modeling for Object Detection. 6221-6230
-     Qiming Xia, Jinhao Deng, Chenglu Wen, Hai Wu, Shaoshuai Shi, Xin Li, Cheng Wang:
Coln: Contrastive Instance Feature Mining for Outdoor 3D Object Detection with Very Limited Annotations. 6231-6240
-     Minying Zhang, Tianpeng Bu, Lulu Hu:
A Dynamic Dual-Processing Object Detection Framework Inspired by the Brain's Recognition Mechanism. 6241-6251
-     Yilong Lv, Min Li, Yujie He, Zhuzhen He, Shaopeng Li, Aitao Yang:
Anchor-Intermediate Detector: Decoupling and Coupling Bounding Boxes for Accurate Object Detection. 6252-6261
-     Declan McIntosh, Alexandra Branzan Albu:
Inter-Realization Channels: Unsupervised Anomaly Detection Beyond One-Class Classification. 6262-6272

-      Shuai Wang, Yao Teng, Limin Wang:
Deep Equilibrium Object Detection. 6273-6283
-     Jing Zhao, Li Sun, Qingli Li:
RecursiveDet: End-to-End Region-based Recursive Object Detection. 6284-6293
-     Xiang Yuan, Gong Cheng, Kebing Yan, Qinghua Zeng, Junwei Han:
Small Object Detection via Coarse-to-fine Proposal Generation and Imitation Learning. 6294-6304
-     Shenghao Fu, Junkai Yan, Yipeng Gao, Xiaohua Xie, Wei-Shi Zheng:
ASAG: Building Strong One-Decoder-Layer Sparse Detectors via Adaptive Sparse Anchor Generation. 6305-6315
-     Xiaofeng Mao, Yuefeng Chen, Yao Zhu, Da Chen, Hang Su, Rong Zhang, Hui Xue:
COCO-O: A Benchmark for Object Detectors under Natural Distribution Shifts. 6316-6327
-     Yuzhong Zhao, Qixiang Ye, Weijia Wu, Chunhua Shen, Fang Wan:
Generative Prompt Model for Weakly Supervised Object Localization. 6328-6338
-     Shanshan Lao, Guanglu Song, Boxiao Liu, Yu Liu, Yujiu Yang:
UniKD: Universal Knowledge Distillation for Mimicking Homogeneous or Heterogeneous Object Detectors. 6339-6349
-     Jaehyeok Bae, Jae-Han Lee, Seyun Kim:
PNI: Industrial Anomaly Detection using Position and Neighborhood Information. 6350-6360
-     Shanshan Lao, Guanglu Song, Boxiao Liu, Yu Liu, Yujiu Yang:
Masked Autoencoders Are Stronger Knowledge Distillers. 6361-6370
-     Ziyu Li, Jingming Guo, Tongtong Cao, Bingbing Liu, Wankou Yang:
GPA-3D: Geometry-aware Prototype Alignment for Unsupervised Domain Adaptive 3D Object Detection from Point Clouds. 6371-6380
-     Lingyu Xiao, Xiang Li, Sen Yang, Wankou Yang:
ADNet: Lane Shape Prediction via Anchor Decomposition. 6381-6390
-     Qipeng Liu, Luojun Lin, Zhifeng Shen, Zhifeng Yang:
Periodically Exchange Teacher-Student for Source-Free Object Detection. 6391-6401
-     Xinzhu Ma, Yongtao Wang, Yinmin Zhang, Zhiyi Xia, Yuan Meng, Zhihui Wang, Haojie Li, Wanli Ouyang:
Towards Fair and Comprehensive Comparisons for Image-Based 3D Object Detection. 6402-6412
-     Xianpeng Liu, Ce Zheng, Kelvin Cheng, Nan Xue, Guo-Jun Qi, Tianfu Wu:
Monocular 3D Object Detection with Bounding Box Denoising in 3D by Perceiver. 6413-6423
-     Hewei Guo, Liping Ren, Jingjing Fu, Yuwang Wang, Zhizheng Zhang, Cuiling Lan, Haoqian Wang, Xinwen Hou:



- Template-guided Hierarchical Feature Restoration for Anomaly Detection.** 6424-6435
Yuting Wang, Velibor Ilic, Jiatong Li, Branislav Kisacanin, Vladimir Pavlovic:
- ALWOD: Active Learning for Weakly-Supervised Object Detection.** 6436-6446
Hansol Kim, Youngjun Kwak, Minyoung Jung, Jinho Shin, Youngsung Kim, Changick Kim:
- ProtoFL: Unsupervised Federated Learning via Prototypical Distillation.** 6447-6456
Ting Lei, Fabian Caba, Qingchao Chen, Hailin Jin, Yuxin Peng, Yang Liu:
- Efficient Adaptive Human-Object Interaction Detection with Concept-guided Memory.** 6457-6467
Shilong Liu, Tianhe Ren, Jiayu Chen, Zhaoyang Zeng, Hao Zhang, Feng Li, Hongyang Li, Jun Huang, Hang Su, Jun Zhu, Lei Zhang:
- Detection Transformer with Stable Matching.** 6468-6477
Liangqi Li, Jiaxu Miao, Dahu Shi, Wenming Tan, Ye Ren, Yi Yang, Shiliang Pu:
- Distilling DETR with Visual-Linguistic Knowledge for Open-Vocabulary Object Detection.** 6478-6487
Tri Cao, Jiawen Zhu, Guansong Pang:
- Anomaly Detection under Distribution Shift.** 6488-6500
Aritra Bhowmik, Yu Wang, Nora Baka, Martin R. Oswald , Cees G. M. Snoek:
- Detecting Objects with Context-Likelihood Graphs and Graph Refinement.** 6501-6510
Yeonghwan Song, Seokwoo Jang, Dina Katabi, Jeany Son:
- Unsupervised Object Localization with Representer Point Selection.** 6511-6521
Yutong Lin, Yuhui Yuan, Zheng Zhang, Chen Li, Nanning Zheng, Han Hu:
- DETR Does Not Need Multi-Scale or Locality Design.** 6522-6531
Qiaoyi Su, Yuhong Chou, Yifan Hu, Jianing Li, Shijie Mei, Ziyang Zhang, Guoqi Li:
- Deep Directly-Trained Spiking Neural Networks for Object Detection.** 6532-6542
David Schinagl, Georg Krispel, Christian Fruhwirth-Reisinger, Horst Possegger, Horst Bischof :
- GACE: Geometry Aware Confidence Enhancement for Black-box 3D Object Detectors on LiDAR-Data.** 6543-6553
Yao Teng, Haisong Liu, Sheng Guo, Limin Wang:
- StageInteractor: Query-based Object Detector with Cross-stage Interaction.** 6554-6565
Yifan Pu, Yiru Wang, Zhuofan Xia, Yizeng Han, Yulin Wang , Weihao Gan, Zidong Wang, Shiji Song, Gao Huang:
- Adaptive Rotated Convolution for Rotated Object Detection.** 6566-6577
Manyuan Zhang, Guanglu Song, Yu Liu, Hongsheng Li:
- Decoupled DETR: Spatially Disentangling Localization and Classification for**

-      **Improved End-to-End Object Detection.** 6578-6587
Jiannan Wu, Yi Jiang, Bin Yan, Huchuan Lu, Zehuan Yuan, Ping Luo:
- Exploring Transformers for Open-world Instance Segmentation.** 6588-6598
-     Xiaojun Tang, Junsong Fan, Chuanchen Luo, Zhaoxiang Zhang, Man Zhang, Zongyuan Yang:
DDG-Net: Discriminability-Driven Graph Network for Weakly-supervised Temporal Action Localization. 6599-6609
-     Qiang Chen, Xiaokang Chen, Jian Wang, Shan Zhang, Kun Yao, Haocheng Feng, Junyu Han, Errui Ding, Gang Zeng, Jingdong Wang:
Group DETR: Fast DETR Training with Group-Wise One-to-Many Assignment. 6610-6619
-     Zhiwei Chen, Jinren Ding, Liujuan Cao, Yunhang Shen, Shengchuan Zhang, Guannan Jiang, Rongrong Ji:
Category-aware Allocation Transformer for Weakly Supervised Object Localization. 6620-6629
-     Zhuangzhuang Chen, Jin Zhang, Zhuonan Lai, Guanming Zhu, Zun Liu, Jie Chen, Jianqiang Li:
The Devil is in the Crack Orientation: A New Perspective for Crack Detection. 6630-6640
-     Yu Pei, Xian Zhao, Hao Li, Jingyuan Ma, Jingwei Zhang, Shiliang Pu:
Clusterformer: Cluster-based Transformer for 3D Object Detection in Point Clouds. 6641-6650
-     Dehua Zheng, Wenhui Dong, Hailin Hu, Xinghao Chen, Yunhe Wang:
Less is More: Focus Attention for Efficient DETR. 6651-6660
-     Hongyang Li, Hao Zhang, Zhaoyang Zeng, Shilong Liu, Feng Li, Tianhe Ren, Lei Zhang:
DFA3D: 3D Deformable Attention For 2D-to-3D Feature Lifting. 6661-6670
-     Ke Zhu, Minghao Fu, Jianxin Wu:
Multi-Label Self-Supervised Learning with Scene Images. 6671-6680
-     Mingqiao Ye, Lei Ke, Siyuan Li, Yu-Wing Tai, Chi-Keung Tang, Martin Danelljan, Fisher Yu:
Cascade-DETR: Delving into High-Quality Universal Object Detection. 6681-6691
-     Yanjing Li, Sheng Xu, Mingbao Lin, Jihao Yin, Baochang Zhang, Xianbin Cao:
Representation Disparity-aware Distillation for 3D Object Detection. 6692-6701
-     Khurram Azeem Hashmi, Goutham Kallemudi, Didier Stricker, Muhammad Zeshan Afzal:
FeatEnHancer: Enhancing Hierarchical Features for Object Detection and Beyond Under Low-Light Vision. 6702-6712
-     Tao Ma, Xuemeng Yang, Hongbin Zhou, Xin Li, Botian Shi, Junjie Liu, Yuchen Yang, Zhizheng Liu, Liang He, Yu Qiao, Yikang Li, Hongsheng Li:

**DetZero: Rethinking Offboard 3D Object Detection with Long-term Sequential Point Clouds.** 6713-6724

Zhuofan Zong, Guanglu Song, Yu Liu:
DETRs with Collaborative Hybrid Assignments Training. 6725-6735



Jiong Wang, Huiming Zhang, Haiwen Hong, Xuan Jin, Yuan He, Hui Xue, Zhou Zhao:
Open-Vocabulary Object Detection With an Open Corpus. 6736-6746



Saksham Suri, Sai Saketh Rambhatla, Rama Chellappa, Abhinav Shrivastava:
SparseDet: Improving Sparsely Annotated Object Detection with Pseudo-positive Mining. 6747-6758



Xinyi Zhang, Naiqi Li, Jiawei Li, Tao Dai, Yong Jiang, Shu-Tao Xia:
Unsupervised Surface Anomaly Detection with Diffusion Probabilistic Model. 6759-6768



Haiyang Wang, Hao Tang, Shaoshuai Shi, Aoxue Li, Zhenguo Li, Bernt Schiele, Liwei Wang:
UniTR: A Unified and Efficient Multi-Modal Transformer for Bird's-Eye-View Representation. 6769-6779



Xincheng Yao, Ruoqi Li, Zefeng Qian, Yan Luo, Chongyang Zhang:
Focus the Discrepancy: Intra- and Inter-Correlation Learning for Image Anomaly Detection. 6780-6790



Junkai Xu, Liang Peng, Haoran Chen, Hao Li, Wei Qian, Ke Li, Wenxiao Wang, Deng Cai:
MonoNeRD: NeRF-like Representations for Monocular 3D Object Detection. 6791-6801



Feng Liu, Xiaosong Zhang, Zhiliang Peng, Zonghao Guo, Fang Wan, Xiangyang Ji, Qixiang Ye:
Integrally Migrating Pre-trained Transformer Encoder-decoders for Visual Object Detection. 6802-6811



Ziye Chen, Yu Liu, Mingming Gong, Bo Du, Guoqi Qian, Kate Smith-Miles:
Generating Dynamic Kernels via Transformers for Lane Detection. 6812-6821



Lu Zhang, Chenbo Zhang, Jiajia Zhao, Jihong Guan, Shuigeng Zhou:
Meta-ZSDETR: Zero-shot DETR with Meta-learning. 6822-6831



Di Wu, Pengfei Chen, Xuehui Yu, Guorong Li, Zhenjun Han, Jianbin Jiao:
Spatial Self-Distillation for Object Detection with Inaccurate Bounding Boxes. 6832-6842



Ming Li, Jie Wu, Xionghui Wang, Chen Chen, Jie Qin, Xuefeng Xiao, Rui Wang, Min Zheng, Xin Pan:
AlignDet: Aligning Pre-training and Fine-tuning in Object Detection. 6843-6853



Zhengzhong Tu, Peyman Milanfar, Hossein Talebi:
MULLER: Multilayer Laplacian Resizer for Vision. 6854-6864

-      Guodong Wang, Yunhong Wang, Jie Qin, Dongming Zhang, Xiuguo Bao, Di Huang:
Unilaterally Aggregated Contrastive Learning with Hierarchical Augmentation for Anomaly Detection. 6865-6874
-     Jiahao Chang, Shuo Wang, Hai-Ming Xu, Zehui Chen, Chenhongyi Yang, Feng Zhao:
DETRDistill: A Universal Knowledge Distillation Framework for DETR-families. 6875-6885
-     Kuan-Chih Huang, Ming-Hsuan Yang, Yi-Hsuan Tsai:
Delving into Motion-Aware Matching for Monocular 3D Object Tracking. 6886-6895
-     Zhiqi Li, Zhiding Yu, Wenhui Wang, Anima Anandkumar, Tong Lu, José M. Álvarez:
FB-BEV: BEV Representation from Forward-Backward View Transformations. 6896-6905
-     Zehui Chen, Zhenyu Li, Shuo Wang, Dengpan Fu, Feng Zhao:
Learning with Noisy Data for Semi-Supervised 3D Object Detection. 6906-6916
-     Na Dong, Yongqiang Zhang, Mingli Ding, Gim Hee Lee:
Boosting Long-tailed Object Detection via Step-wise Learning on Smooth-tail Data. 6917-6926
-     Xin Liu, Fatemeh Karimi Nejadasl, Jan C. van Gemert, Olaf Booij, Silvia L. Pintea:
Objects do not disappear: Video object detection by single-frame object location anticipation. 6927-6938
-     Long Zhao, Liangzhe Yuan, Boqing Gong, Yin Cui, Florian Schroff, Ming-Hsuan Yang, Hartwig Adam, Ting Liu:
Unified Visual Relationship Detection with Vision and Language Models. 6939-6950
-     Didi Zhu, Yinchuan Li, Junkun Yuan, Zexi Li, Kun Kuang, Chao Wu:
Universal Domain Adaptation via Compressive Attention Matching. 6951-6962
-     Wenzhang Zhou, Heng Fan, Tiejian Luo, Libo Zhang:
Unsupervised Domain Adaptive Detection with Network Stability Analysis. 6963-6972
-     Tao Tu, Shun-Po Chuang, Yu-Lun Liu, Cheng Sun, Ke Zhang, Donna Roy, Cheng-Hao Kuo, Min Sun:
ImGeoNet: Image-induced Geometry-aware Voxel Representation for Multi-view 3D Object Detection. 6973-6984
-     Yufei Yin, Jiajun Deng, Wengang Zhou, Li Li, Houqiang Li:
Cyclic-Bootstrap Labeling for Weakly Supervised Object Detection. 6985-6995
-     Zhenhuan Liu, Liang Li, Jiayu Xiao, Zheng-Jun Zha, Qingming Huang:
Text-Driven Generative Domain Adaptation with Spectral Consistency Regularization. 6996-7006

-      Daniel Silver, Aditya Ranjan, Tirthak Patel, Harshitta Gandhi, William Cutler, Devesh Tiwari:
MosaiQ: Quantum Generative Adversarial Networks for Image Generation on NISQ Computers. 7007-7016
-     Ruihan Gao, Wenzhen Yuan, Jun-Yan Zhu:
Controllable Visual-Tactile Synthesis. 7017-7029
-     Hadas Orgad, Bahjat Kawar, Yonatan Belinkov:
Editing Implicit Assumptions in Text-to-Image Diffusion Models. 7030-7038
-     David Svitov, Dmitrii Gudkov, Renat Bashirov, Victor Lempitsky:
DINAR: Diffusion Inpainting of Neural Textures for One-Shot Human Avatars. 7039-7049
-     Vadim Sushko, Ruyu Wang, Juergen Gall:
Smoothness Similarity Regularization for Few-Shot GAN Adaptation. 7050-7059
-     Chanyue Wu, Dong Wang, Yunpeng Bai, Hanyu Mao, Ying Li, Qiang Shen:
HSR-Diff: Hyperspectral Image Super-Resolution via Conditional Diffusion Models. 7060-7070
-     Jason J. Yu, Fereshteh Forghani, Konstantinos G. Derpanis, Marcus A. Brubaker:
Long-Term Photometric Consistent Novel View Synthesis with Diffusion Models. 7071-7081
-     Lijiang Li, Huixia Li, Xiawu Zheng, Jie Wu, Xuefeng Xiao, Rui Wang, Min Zheng, Xin Pan, Fei Chao, Rongrong Ji:
AutoDiffusion: Training-Free Optimization of Time Steps and Architectures for Automated Diffusion Model Acceleration. 7082-7091
-     Nannan Li, Kevin J. Shih, Bryan A. Plummer:
Collecting The Puzzle Pieces: Disentangled Self-Driven Human Pose Transfer by Permuting Textures. 7092-7103
-     Chen Naveh:
Multi-Directional Subspace Editing in Style-Space. 7104-7114
-     Stella Bounareli, Christos Tzelepis , Vasileios Argyriou, Ioannis Patras, Georgios Tzimiropoulos:
HyperReenact: One-Shot Reenactment via Jointly Learning to Refine and Retarget Faces. 7115-7125
-     Taegyeong Lee, Jeonghun Kang, Hyeonyu Kim, Taehwan Kim:
Generating Realistic Images from In-the-wild Sounds. 7126-7136
-     Sherwin Bahmani, Jeong Joon Park, Despoina Paschalidou, Xinguang Yan, Gordon Wetzstein, Leonidas J. Guibas, Andrea Tagliasacchi:
CC3D: Layout-Conditioned Generation of Compositional 3D Scenes. 7137-7147
-     Rishabh Jain, Mayur Hemani, Duygu Ceylan, Krishna Kumar Singh, Jingwan Lu, Mausoom Sarkar, Balaji Krishnamurthy:
UMFuse: Unified Multi View Fusion for Human Editing applications. 7148-7157

-  [Sheng-Yu Wang, Alexei A. Efros, Jun-Yan Zhu, Richard Zhang: Evaluating Data Attribution for Text-to-Image Models.](#) 7158-7169
-  [Shengxi Li, Jialu Zhang, Yifei Li, Mai Xu, Xin Deng, Li Li: Neural Characteristic Function Learning for Conditional Image Generation.](#) 7170-7180
-  [Liyuan Ma, Tingwei Gao, Haitian Jiang, Haibin Shen, Kejie Huang: WaveIPT: Joint Attention and Flow Alignment in the Wavelet domain for Pose Transfer.](#) 7181-7191
-  [Junyi Zhang, Jiaqi Guo, Shizhao Sun, Jian-Guang Lou, Dongmei Zhang: LayoutDiffusion: Improving Graphic Layout Generation by Discrete Diffusion Probabilistic Models.](#) 7192-7202
-  [Fei Gao, Yifan Zhu, Chang Jiang, Nannan Wang: Human-Inspired Facial Sketch Synthesis with Dynamic Adaptation.](#) 7203-7213
-  [Savas Özkan, Mete Özay, Tom Robinson: Conceptual and Hierarchical Latent Space Decomposition for Face Editing.](#) 7214-7223
-  [Seogkyu Jeon, Bei Liu, Pilhyeon Lee, Kibeom Hong, Jianlong Fu, Hyeran Byun: Improving Diversity in Zero-Shot GAN Adaptation with Semantic Variations.](#) 7224-7233
-  [Minjung Shin, Yunji Seo, Jeongmin Bae, Young Sun Choi, Hyunsu Kim, Hyeran Byun, Youngjung Uh: BallGAN: 3D-aware Image Synthesis with a Spherical Background.](#) 7234-7245
-  [Bram Wallace, Akash Gokul, Stefano Ermon, Nikhil Naik: End-to-End Diffusion Latent Optimization Improves Classifier Guidance.](#) 7246-7256
-  [Li Siyao, Tianpei Gu, Weiye Xiao, Henghui Ding, Ziwei Liu, Chen Change Loy: Deep Geometrized Cartoon Line Inbetweening.](#) 7257-7266
-  [Jianglin Fu, Shikai Li, Yuming Jiang, Kwan-Yee Lin, Wayne Wu, Ziwei Liu: UnitedHuman: Harnessing Multi-Source Data for High-Resolution Human Generation.](#) 7267-7277
-  [Yang Zhao, Tingbo Hou, Yu-Chuan Su, Xuhui Jia, Yandong Li, Matthias Grundmann: Towards Authentic Face Restoration with Iterative Diffusion Models and Beyond.](#) 7278-7288
-  [Ligong Han, Yinxiao Li, Han Zhang, Peyman Milanfar, Dimitris N. Metaxas, Feng Yang: SVDiff: Compact Parameter Space for Diffusion Fine-Tuning.](#) 7289-7300
-  [Andranik Sargsyan, Shant Navasardyan, Xingqian Xu, Humphrey Shi: MI-GAN: A Simple Baseline for Image Inpainting on Mobile Devices.](#) 7301-7311

-      Patrick Esser, Johnathan Chiu, Parmida Atighehchian, Jonathan Granskog, Anastasis Germanidis:
Structure and Content-Guided Video Synthesis with Diffusion Models. 7312-7322
-     Yuxin Jiang, Liming Jiang, Shuai Yang, Chen Change Loy:
Scenimefy: Learning to Craft Anime Scene via Semi-Supervised Image-to-Image Translation. 7323-7333
-     Shiyue Cao, Yueqin Yin, Lianghua Huang, Yu Liu, Xin Zhao , Deli Zhao, Kaiqi Huang:
Efficient-VQGAN: Towards High-Resolution Image Generation with Efficient Vision Transformers. 7334-7343
-     Chen Henry Wu, Fernando De la Torre:
A Latent Space of Stochastic Diffusion Models for Zero-Shot Image Editing and Guidance. 7344-7353
-     Amandeep Kumar, Ankan Kumar Bhunia, Sanath Narayan, Hisham Cholakkal, Rao Muhammad Anwer, Salman H. Khan, Ming-Hsuan Yang, Fahad Shahbaz Khan:
Generative Multiplane Neural Radiance for 3D-Aware Image Generation. 7354-7364
-     Lang Nie, Chunyu Lin, Kang Liao, Shuaicheng Liu, Yao Zhao:
Parallax-Tolerant Unsupervised Deep Image Stitching. 7365-7374
-     Desai Xie, Ping Hu, Xin Sun, Sören Pirk, Jianming Zhang, Radomír Mech, Arie E. Kaufman:
GAIT: Generating Aesthetic Indoor Tours with Deep Reinforcement Learning. 7375-7385
-     Mohammad Reza Karimi Dastjerdi, Jonathan Eisenmann, Yannick Hold-Geoffroy, Jean-François Lalonde:
EverLight: Indoor-Outdoor Editable HDR Lighting Estimation. 7386-7395
-     Wenkai Dong, Song Xue, Xiaoyue Duan, Shumin Han:
Prompt Tuning Inversion for Text-Driven Image Editing Using Diffusion Models. 7396-7406
-     Tiankai Hang, Shuyang Gu, Chen Li, Jianmin Bao, Dong Chen, Han Hu, Xin Geng, Baining Guo:
Efficient Diffusion Training via Min-SNR Weighting Strategy. 7407-7417
-     Jinheng Xie, Yuexiang Li, Yawen Huang, Haozhe Liu, Wentian Zhang, Yefeng Zheng, Mike Zheng Shou:
BoxDiff: Text-to-Image Synthesis with Training-Free Box-Constrained Diffusion. 7418-7427
-     Susung Hong, Gyuseong Lee, Wooseok Jang, Seungryong Kim:
Improving Sample Quality of Diffusion Models Using Self-Attention Guidance. 7428-7437
-     Luozhou Wang, Shuai Yang, Shu Liu, Ying-Cong Chen:
Not All Steps are Created Equal: Selective Diffusion Distillation for Image Manipulation. 7438-7447

-  [Download](#) [Cite](#) [Share](#) Li Niu, Junyan Cao, Wenyan Cong, Liqing Zhang:
Deep Image Harmonization with Learnable Augmentation. 7448-7457
-  [Download](#) [Cite](#) [Share](#) Xin Yang, Xiaogang Xu, Yingcong Chen:
Out-of-domain GAN inversion via Invertibility Decomposition for Photo-Realistic Human Face Manipulation. 7458-7467
-  [Download](#) [Cite](#) [Share](#) Wing Yin Yu, Lai-Man Po, Ray C. C. Cheung, Yuzhi Zhao, Yu Xue, Kun Li:
Bidirectionally Deformable Motion Modulation For Video-based Human Pose Transfer. 7468-7478
-  [Download](#) [Cite](#) [Share](#) Chieh-Yun Chen, Yi-Chung Chen, Hong-Han Shuai, Wen-Huang Cheng:
Size Does Matter: Size-aware Virtual Try-on via Clothing-oriented Transformation Try-on Network. 7479-7488
-  [Download](#) [Cite](#) [Share](#) Moayed Haji Ali, Andrew Bond, Levent Karacan, Tolga Birdal, Erkut Erdem, Duygu Ceylan, Aykut Erdem:
VidStyleODE: Disentangled Video Editing via StyleGAN and NeuralODEs. 7489-7500
-  [Download](#) [Cite](#) [Share](#) Xintian Shen, Jiangning Zhang, Jun Chen, Shipeng Bai, Yue Han, Yabiao Wang, Chengjie Wang, Yong Liu:
Learning Global-aware Kernel for Image Harmonization. 7501-7510
-  [Download](#) [Cite](#) [Share](#) Songwei Ge, Taesung Park, Jun-Yan Zhu, Jia-Bin Huang:
Expressive Text-to-Image Generation with Rich Text. 7511-7522
-  [Download](#) [Cite](#) [Share](#) Chongshan Lu, Fukun Yin, Xin Chen, Wen Liu, Tao Chen, Gang Yu, Jiayuan Fan:
A Large-Scale Outdoor Multi-modal Dataset and Benchmark for Novel View Synthesis and Implicit Scene Reconstruction. 7523-7533
-  [Download](#) [Cite](#) [Share](#) Jiahe Li, Jiawei Zhang, Xiao Bai, Jun Zhou, Lin Gu:
Efficient Region-Aware Neural Radiance Fields for High-Fidelity Talking Portrait Synthesis. 7534-7544
-  [Download](#) [Cite](#) [Share](#) Lingzhi Zhang, Zhengjie Xu, Connelly Barnes, Yuqian Zhou, Qing Liu, He Zhang, Sohrab Amirghodsi, Zhe Lin, Eli Shechtman, Jianbo Shi:
Perceptual Artifacts Localization for Image Synthesis Tasks. 7545-7556
-  [Download](#) [Cite](#) [Share](#) Minho Park, Jooyeol Yun, Seunghwan Choi, Jaegul Choo:
Learning to Generate Semantic Layouts for Higher Text-Image Correspondence in Text-to-Image Synthesis. 7557-7566
-  [Download](#) [Cite](#) [Share](#) Zipeng Xu, Enver Sangineto, Nicu Sebe :
StylerDALLE: Language-Guided Style Transfer Using a Vector-Quantized Tokenizer of a Large-Scale Generative Model. 7567-7577
-  [Download](#) [Cite](#) [Share](#) Chaeyeon Chung, Yeojeong Park, Seunghwan Choi, Munkhsoyol Ganbat, Jaegul Choo:
Shortcut-V2V: Compression Framework for Video-to-Video Translation based on Temporal Redundancy Reduction. 7578-7588
-  [Download](#) [Cite](#) [Share](#) Jay Zhangjie Wu, Yixiao Ge, Xiantao Wang, Stan Weixian Lei, Yuchao Gu, Yufei Shi, Wynne Hsu, Ying Shan, Xiaohu Qie, Mike Zheng Shou:
Tune-A-Video: One-Shot Tuning of Image Diffusion Models for Text-to-Video Generation. 7589-7599

-  [Download](#) [Cite](#) [Share](#) Kaede Shiohara, Xingchao Yang, Takafumi Taketomi:
BlendFace: Re-designing Identity Encoders for Face-Swapping. 7600-7610
-  [Download](#) [Cite](#) [Share](#) Zhentao Yu, Zixin Yin, Deyu Zhou, Duomin Wang, Finn Wong, Baoyuan Wang:
Talking Head Generation with Probabilistic Audio-to-Visual Diffusion Priors. 7611-7621
-  [Download](#) [Cite](#) [Share](#) Jiapeng Zhu, Ceyuan Yang, Yujun Shen, Zifan Shi, Bo Dai, Deli Zhao, Qifeng Chen:
LinkGAN: Linking GAN Latents to Pixels for Controllable Image Synthesis. 7622-7632
-  [Download](#) [Cite](#) [Share](#) Ziyi Li, Qinye Zhou, Xiaoyun Zhang, Ya Zhang , Yanfeng Wang, Weidi Xie:
Open-vocabulary Object Segmentation with Diffusion Models. 7633-7642
-  [Download](#) [Cite](#) [Share](#) Zhizhong Wang, Lei Zhao, Wei Xing:
StyleDiffusion: Controllable Disentangled Style Transfer via Diffusion Models. 7643-7655
-  [Download](#) [Cite](#) [Share](#) Yuan Gong, Yong Zhang, Xiaodong Cun, Fei Yin, Yanbo Fan, Xuan Wang, Baoyuan Wu, Yujiu Yang:
ToonTalker: Cross-Domain Face Reenactment. 7656-7666
-  [Download](#) [Cite](#) [Share](#) Yunji Kim, Jiyoung Lee, Jin-Hwa Kim, Jung-Woo Ha, Jun-Yan Zhu:
Dense Text-to-Image Generation with Attention Modulation. 7667-7677
-  [Download](#) [Cite](#) [Share](#) Yue Song, Jichao Zhang, Nicu Sebe , Wei Wang:
Householder Projector for Unsupervised Latent Semantics Discovery. 7678-7688
-  [Download](#) [Cite](#) [Share](#) Li Niu, Linfeng Tan, Xinhao Tao, Junyan Cao, Fengjun Guo, Teng Long, Liqing Zhang:
Deep Image Harmonization with Globally Guided Feature Transformation and Relation Distillation. 7689-7698
-  [Download](#) [Cite](#) [Share](#) Ceyuan Yang, Yujun Shen, Zhiyi Zhang, Yinghao Xu, Jiapeng Zhu, Zhirong Wu, Bolei Zhou:
One-Shot Generative Domain Adaptation. 7699-7708
-  [Download](#) [Cite](#) [Share](#) Cheng-Hung Chan, Cheng-Yang Yuan, Cheng Sun, Hwann-Tzong Chen:
Hashing Neural Video Decomposition with Multiplicative Residuals in Space-Time. 7709-7719
-  [Download](#) [Cite](#) [Share](#) Xinqian Xu, Zhangyang Wang, Eric J. Zhang, Kai Wang, Humphrey Shi:
Versatile Diffusion: Text, Images and Variations All in One Diffusion Model. 7720-7731
-  [Download](#) [Cite](#) [Share](#) Qiucheng Wu, Yujian Liu, Handong Zhao, Trung Bui, Zhe Lin, Yang Zhang, Shiyu Chang:
Harnessing the Spatial-Temporal Attention of Diffusion Models for High-Fidelity Text-to-Image Synthesis. 7732-7742
-  [Download](#) [Cite](#) [Share](#) Arda Senocak, Hyeonggon Ryu, Junsik Kim, Tae-Hyun Oh, Hanspeter Pfister, Joon Son Chung:
Sound Source Localization is All about Cross-Modal Alignment. 7743-7753
-  [Download](#) [Cite](#) [Share](#) Shentong Mo, Weiguo Pian, Yapeng Tian:
Class-Incremental Grouping Network for Continual Audio-Visual Learning.

-      7754-7764
Weiguo Pian, Shentong Mo, Yunhui Guo, Yapeng Tian:
Audio-Visual Class-Incremental Learning. 7765-7777
-     Jeongsoo Choi, Joanna Hong, Yong Man Ro:
DiffV2S: Diffusion-based Video-to-Speech Synthesis with Vision-guided Speaker Embedding. 7778-7787
-     Yujin Jeong, Wonjeong Ryoo, Seunghyun Lee, Dabin Seo, Wonmin Byeon, Sangpil Kim, Jinkyu Kim:
The Power of Sound (TPoS): Audio Reactive Video Generation with Stable Diffusion. 7788-7798
-     Urwa Muaz, Wondong Jang, Rohun Tripathi, Santhosh Mani, Wenbin Ouyang, Ravi Teja Gadde, Baris Gecer, Sergio Elizondo, Reza Madad, Naveen Nair:
SIDGAN: High-Resolution Dubbed Video Generation via Shift-Invariant Learning. 7799-7808
-     Zhe Niu, Brian Mak:
On the Audio-visual Synchronization for Lip-to-Speech Synthesis. 7809-7818
-     Mingfei Chen, Kun Su, Eli Shlizerman:
Be Everywhere - Hear Everything (BEE): Audio Scene Reconstruction by Sparse Audio-Visual Samples. 7819-7828
-     Heeseung Yun, Joonil Na, Gunhee Kim:
Dense 2D-3D Indoor Prediction with Sound via Aligned Cross-Modal Distillation. 7829-7838
-     Jie Hong, Zeeshan Hayder, Junlin Han, Pengfei Fang, Mehrtash Harandi, Lars Petersson:
Hyperbolic Audio-visual Zero-shot Learning. 7839-7849
-     Sanjoy Chowdhury, Sreyan Ghosh, Subhrajyoti Dasgupta, Anton Ratnarajah, Utkarsh Tyagi, Dinesh Manocha:
AdVerb: Visually Guided Audio Dereverberation. 7850-7862
-     Ziyang Chen, Shengyi Qian, Andrew Owens:
Sound Localization from Motion: Jointly Learning Sound Direction and Camera Rotation. 7863-7874
-     Lukas Höllerin, Ang Cao, Andrew Owens, Justin Johnson, Matthias Nießner:
Text2Room: Extracting Textured 3D Meshes from 2D Text-to-Image Models. 7875-7886
-     Noah Stier, Baptiste Angles, Liang Yang, Yajie Yan, Alex Colburn, Ming Chuang:
LivePose: Online 3D Reconstruction from Monocular Video with Dynamic Camera Poses. 7887-7896
-     Shuwei Shao, Zhongcai Pei, Weihai Chen, Xingming Wu, Zhengguo Li:
NDDepth: Normal-Distance Assisted Monocular Depth Estimation. 7897-7906
-     Yueru Luo, Chaoda Zheng, Xu Yan, Tang Kun, Chao Zheng, Shuguang Cui, Zhen Li:

**LATR: 3D Lane Detection from Monocular Images with Transformer.** 7907-7918

- Xiaosong Jia, Yulu Gao, Li Chen, Junchi Yan, Patrick Langechuan Liu, Hongyang Li:
DriveAdapter: Breaking the Coupling Barrier of Perception and Planning in End-to-End Autonomous Driving. 7919-7929

- Sergey Prokudin, Qianli Ma, Maxime Raafat, Julien Valentin, Siyu Tang :
Dynamic Point Fields. 7930-7942

- Haiwen Feng, Peter Kulits, Shichen Liu, Michael J. Black, Victoria Fernández Abrevaya:
Generalizing Neural Human Fitting to Unseen Poses With Articulated SE(3) Equivariance. 7943-7954

- Siwei Zhang, Qianli Ma, Yan Zhang, Sadegh Aliakbarian, Darren Cosker, Siyu Tang :
Probabilistic Human Mesh Recovery in 3D Scenes from Egocentric Views. 7955-7966

- Shashank Tripathi, Agniv Chatterjee, Jean-Claude Passy, Hongwei Yi, Dimitrios Tzionas, Michael J. Black:
DECO: Dense Estimation of 3D Human-Scene Contact In The Wild. 7967-7979

- Pengfei Ren, Chao Wen, Xiaozheng Zheng, Zhou Xue , Haifeng Sun, Qi Qi, Jingyu Wang, Jianxin Liao:
Decoupled Iterative Refinement Framework for Interacting Hands Reconstruction from a Single RGB Image. 7980-7991

- Mattias P. Heinrich, Alexander Bigalke, Christoph Großbröhmer, Lasse Hansen:
Chasing clouds: Differentiable volumetric rasterisation of point clouds as a highly efficient and accurate loss for large-scale deformable 3D registration. 7992-8002

- Rizhao Cai, Yawen Cui, Zhi Li, Zitong Yu, Haoliang Li, Yongjian Hu, Alex C. Kot:
Rehearsal-Free Domain Continual Face Anti-Spoofing: Generalize More and Forget Less. 8003-8014

- Ling Gao , Hang Su, Daniel Gehrig , Marco Cannici, Davide Scaramuzza, Laurent Kneip:
A 5-Point Minimal Solver for Event Camera Relative Motion Estimation. 8015-8025

- Juan Carlos Dibene, Zhixiang Min, Enrique Dunn:
General Planar Motion from a Pair of 3D Correspondences. 8026-8036

- Christophe Bolduc, Justine Giroux, Marc Hébert, Claude Demers, Jean-François Lalonde:
Beyond the Pixel: a Photometrically Calibrated HDR Dataset for Luminance and Color Prediction. 8037-8047

- Zixiang Zhao, Haowen Bai, Yuanzhi Zhu, Jiangshe Zhang, Shuang Xu, Yulun Zhang, Kai Zhang, Deyu Meng, Radu Timofte, Luc Van Gool:
DDFM: Denoising Diffusion Model for Multi-Modality Image Fusion. 8048-8059

-      Zhexin Liang, Chongyi Li, Shangchen Zhou, Ruicheng Feng, Chen Change Loy:
Iterative Prompt Learning for Unsupervised Backlit Image Enhancement. 8060-8069
-     Rundong Luo, Wenjing Wang, Wenhan Yang, Jiaying Liu:
Similarity Min-Max: Zero-Shot Day-Night Domain Adaptation. 8070-8080
-     Jinyuan Liu, Zhu Liu, Guanyao Wu, Long Ma, Risheng Liu, Wei Zhong, Zhongxuan Luo, Xin Fan:
Multi-interactive Feature Learning and a Full-time Multi-modality Benchmark for Image Fusion and Segmentation. 8081-8090
-     Jeremy Klotz, Mohit Gupta, Aswin C. Sankaranarayanan:
Computational 3D Imaging with Position Sensors. 8091-8100
-     Mian Wei, Sotiris Nousias, Rahul Gulve, David B. Lindell, Kiriakos N. Kutulakos:
Passive Ultra-Wideband Single-Photon Imaging. 8101-8112
-     Federica Arrigoni, Tomás Pajdla, Andrea Fusiello:
Viewing Graph Solvability in Practice. 8113-8121
-     Yaqing Ding, Chiang-Heng Chien, Viktor Larsson, Karl Åström, Benjamin B. Kimia:
Minimal Solutions to Generalized Three-View Relative Pose Problem. 8122-8130
-     Varun Sundar, Andrei Ardelean, Tristan Swedish, Claudio Bruschini, Edoardo Charbon, Mohit Gupta:
SoDaCam: Software-defined Cameras via Single-Photon Imaging. 8131-8142
-     Stefano Gasperini, Nils Morbitzer, HyunJun Jung, Nassir Navab, Federico Tombari:
Robust Monocular Depth Estimation under Challenging Conditions. 8143-8152
-     Tianhang Wang, Guang Chen, Kai Chen, Zhengfa Liu, Bo Zhang, Alois Knoll, Changjun Jiang:
UMC: A Unified Bandwidth-efficient and Multi-resolution based Collaborative Perception Framework. 8153-8162
-     Shan Wang, Yanhao Zhang, Akhil Perincherry, Ankit Vora, Hongdong Li:
View Consistent Purification for Accurate Cross-View Localization. 8163-8172
-     Ruochen Jiao, Xiangguo Liu, Takami Sato, Qi Alfred Chen, Qi Zhu:
Semi-supervised Semantics-guided Adversarial Training for Robust Trajectory Prediction. 8173-8183
-     Junyuan Deng, Qi Wu, Xieyuanli Chen, Songpengcheng Xia, Zhen Sun, Guoqing Liu, Wenxian Yu, Ling Pei:
NeRF-LOAM: Neural Implicit Representation for Large-Scale Incremental LiDAR Odometry and Mapping. 8184-8193
-     Xiyue Zhu, Vlas Zyrianov, Zhijian Liu, Shenlong Wang:
MapPrior: Bird's-Eye View Map Layout Estimation with Generative Models. 8194-8205

-      Bernhard Jaeger, Kashyap Chitta, Andreas Geiger:
Hidden Biases of End-to-End Driving Models. 8206-8215
-     Ronghao Dang, Liuyi Wang, Zongtao He, Shuai Su, Jiagui Tang, Chengju Liu, Qijun Chen:
Search for or Navigate to? Dual Adaptive Thinking for Object Navigation. 8216-8225
-     Yiyao Zhu, Di Luan, Shaojie Shen:
BiFF: Bi-level Future Fusion with Polyline-based Coordinate for Interactive Trajectory Prediction. 8226-8237
-     Sivabalan Manivasagam, Ioan Andrei Bârsan, Jingkang Wang, Ze Yang, Raquel Urtasun:
Towards Zero Domain Gap: A Comprehensive Study of Realistic LiDAR Simulation for Autonomy Testing. 8238-8248
-     Tuo Feng, Wenguan Wang, Xiaohan Wang, Yi Yang, Qinghua Zheng:
Clustering based Point Cloud Representation Learning for 3D Analysis. 8249-8260
-     Görkay Aydemir, Adil Kaan Akan, Fatma Güney:
ADAPT: Efficient Multi-Agent Trajectory Prediction with Adaptation. 8261-8271
-     Yibo Liu, Kelly Zhu, Guile Wu, Yuan Ren, Bingbing Liu, Yang Liu, Jinjun Shan:
MV-DeepSDF: Implicit Modeling with Multi-Sweep Point Clouds for 3D Vehicle Reconstruction in Autonomous Driving. 8272-8282
-     Kunyang Lin, Peihao Chen, Diwei Huang, Thomas H. Li, Mingkui Tan, Chuang Gan:
Learning Vision-and-Language Navigation from YouTube Videos. 8283-8292
-     Liang Zhang, Nathaniel Xu, Pengfei Yang, Gaojie Jin, Cheng-Chao Huang, Lijun Zhang:
TrajPAC: Towards Robustness Verification of Pedestrian Trajectory Prediction Models. 8293-8305
-     Bo Jiang, Shaoyu Chen, Qing Xu, Bencheng Liao, Jiajie Chen, Helong Zhou, Qian Zhang, Wenyu Liu, Chang Huang, Xinggang Wang:
VAD: Vectorized Scene Representation for Efficient Autonomous Driving. 8306-8316
-     Hao Chen, Jiaze Wang, Kun Shao, Furui Liu, Jianye Hao, Chenyong Guan, Guangyong Chen, Pheng-Ann Heng:
Traj-MAE: Masked Autoencoders for Trajectory Prediction. 8317-8328
-     Chengtang Yao, Lidong Yu, Yuwei Wu, Yunde Jia:
Sparse Point Guided 3D Lane Detection. 8329-8338
-     Dingyuan Zhang , Dingkang Liang, Zhikang Zou, Jingyu Li, Xiaoqing Ye, Zhe Liu, Xiao Tan, Xiang Bai:
A Simple Vision Transformer for Weakly Semi-supervised 3D Object Detection. 8339-8349

-      Mozhgan Pourkeshavarz, Changhe Chen, Amir Rasouli:
Learn TAROT with MENTOR: A Meta-Learned Self-supervised Approach for Trajectory Prediction. 8350-8359
-     Yilun Chen, Zhiding Yu, Yukang Chen, Shiyi Lan, Anima Anandkumar, Jiaya Jia, José M. Álvarez:
FocalFormer3D : Focusing on Hard Instance for 3D Object Detection. 8360-8371
-     Wenwen Tong, Chonghao Sima, Tai Wang, Li Chen, Silei Wu, Hanming Deng, Yi Gu, Lewei Lu, Ping Luo, Dahua Lin, Hongyang Li:
Scene as Occupancy. 8372-8381
-     Jeffrey Yunfan Liu, Yun Chen, Ze Yang, Jingkang Wang, Sivabalan Manivasagam, Raquel Urtasun:
Real-Time Neural Rasterization for Large Scenes. 8382-8393
-     Amir Belder, Refael Vivanti, Ayellet Tal:
A Game of Bundle Adjustment - Learning Efficient Convergence. 8394-8403
-     Mao Ye, Gregory P. Meyer, Yuning Chai, Qiang Liu:
Efficient Transformer-based 3D Object Detection with Dynamic Token Halting. 8404-8416
-     Jiuming Liu, Guangming Wang, Zhe Liu, Chaokang Jiang, Marc Pollefeys, Hesheng Wang:
RegFormer: An Efficient Projection-Aware Transformer Network for Large-Scale Point Cloud Registration. 8417-8426
-     Yan Xia, Mariia Gladkova, Rui Wang, Qianyun Li, Uwe Still, João F. Henriques, Daniel Cremers:
CASSPR: Cross Attention Single Scan Place Recognition. 8427-8438
-     Dongkwon Jin, Dahun Kim, Chang-Su Kim:
Recursive Video Lane Detection. 8439-8448
-     Jiayu Yang, Enze Xie, Miaomiao Liu, José M. Álvarez:
Parametric Depth Based Feature Representation Learning for Object Detection and Segmentation in Bird's-Eye View. 8449-8458
-     Hongge Chen, Zhao Chen, Gregory P. Meyer, Dennis Park, Carl Vondrick, Ashish Shrivastava, Yuning Chai:
SHIFT3D: Synthesizing Hard Inputs For Tricking 3D Detectors. 8459-8469
-     Maosheng Ye, Jiamiao Xu, Xunnong Xu, Tengfei Wang, Tongyi Cao, Qifeng Chen:
Bootstrap Motion Forecasting With Self-Consistent Constraints. 8470-8480
-     Tzofi Klinghoffer, Jonah Philion, Wenzheng Chen, Or Litany, Zan Gojcic, Jungseock Joo, Ramesh Raskar, Sanja Fidler, José M. Álvarez:
Towards Viewpoint Robustness in Bird's Eye View Segmentation. 8481-8490
-     Sehwan Choi, Jungho Kim, Junyong Yun, Jun Won Choi:
R-Pred: Two-Stage Motion Prediction Via Tube-Query Attention-Based Trajectory Refinement. 8491-8501
-     Zhijie Yan, Pengfei Li, Zheng Fu, Shaocong Xu, Yongliang Shi, Xiaoxue Chen, Yuhang Zheng, Yang Li, Tianyu Liu, Chuxuan Li, Nairui Luo, Xu Gao, Yilun Chen,



Zuoxu Wang, Yifeng Shi, Pengfei Huang, Zhengxiao Han, Jirui Yuan, Jiangtao Gong, Guyue Zhou, Hang Zhao, Hao Zhao:

INT2: Interactive Trajectory Prediction at Intersections. 8502-8513



Hongyu Zhou, Zheng Ge, Zeming Li, Xiangyu Zhang:

MatrixVT: Efficient Multi-Camera to BEV Transformation for 3D Perception. 8514-8523



Pengfei Zhu, Mengshi Qi, Xia Li, Weijian Li, Huadong Ma:

Unsupervised Self-Driving Attention Prediction via Uncertainty Mining and Knowledge Embedding. 8524-8534



Xuechao Chen, Shuangjie Xu, Xiaoyi Zou, Tongyi Cao, Dit-Yan Yeung, Lu Fang:
SVQNet: Sparse Voxel-Adjacent Query Network for 4D Spatio-Temporal LiDAR Semantic Segmentation. 8535-8544



Ari Seff, Brian Cera, Dian Chen, Mason Ng, Aurick Zhou, Nigamaa Nayakanti, Khaled S. Refaat, Rami Al-Rfou, Benjamin Sapp:
MotionLM: Multi-Agent Motion Forecasting as Language Modeling. 8545-8556



Yigit Baran Can, Alexander Liniger, Danda Pani Paudel , Luc Van Gool:
Improving Online Lane Graph Extraction by Object-Lane Clustering. 8557-8567



Mahyar Najibi, Jingwei Ji, Yin Zhou, Charles R. Qi, Xinchen Yan, Scott Ettinger, Dragomir Anguelov:
Unsupervised 3D Perception with 2D Vision-Language Distillation for Autonomous Driving. 8568-8578



Wencheng Han, Junbo Yin, Jianbing Shen:
Self-Supervised Monocular Depth Estimation by Direction-aware Cumulative Convolution Network. 8579-8589



Nakul Agarwal, Yi-Ting Chen:
Ordered Atomic Activity for Fine-grained Interactive Traffic Scenario Understanding. 8590-8602



Zeyu Wang, Dingwen Li, Chenxu Luo, Cihang Xie, Xiaodong Yang:
DistillBEV: Boosting Multi-Camera 3D Object Detection with Cross-Modal Knowledge Distillation. 8603-8612



Thomas E. Huang, Yifan Liu, Luc Van Gool, Fisher Yu:
Video Task Decathlon: Unifying Image and Video Tasks in Autonomous Driving. 8613-8623



Ziyang Xie, Ziqi Pang, Yu-Xiong Wang:
MV-Map: Offboard HD Map Generation with Multi-view Consistency. 8624-8634



Guile Wu, Tongtong Cao, Bingbing Liu, Xingxin Chen, Yuan Ren:
Towards Universal LiDAR-Based 3D Object Detection by Multi-Domain Knowledge Transfer. 8635-8644



Jie Cheng , Xiaodong Mei, Ming Liu:
Forecast-MAE: Self-supervised Pre-training for Motion Forecasting with

**Masked Autoencoders.** 8645-8655

Zequn Qin, Jingyu Chen, Chao Chen, Xiaozhi Chen, Xi Li:
UniFusion: Unified Multi-view Fusion Transformer for Spatial-Temporal Representation in Bird's-Eye-View. 8656-8665



Lun Luo, Shuhang Zheng, Yixuan Li, Yongzhi Fan, Beinan Yu, Si-Yuan Cao, Junwei Li, Hui-Liang Shen:
BEVPlace: Learning LiDAR-based Place Recognition using Bird's Eye View Images. 8666-8675



Binglu Wang, Lei Zhang, Zhaozhong Wang, Yongqiang Zhao, Tianfei Zhou:
Core: Cooperative Reconstruction for Multi-Agent Perception. 8676-8686



Chongjian Ge, Junsong Chen, Enze Xie, Zhongdao Wang, Lanqing Hong, Huchuan Lu, Zhenguo Li, Ping Luo:
MetaBEV: Solving Sensor Failures for 3D Detection and Map Segmentation. 8687-8697



Zhu Yu, Zehua Sheng, Zili Zhou, Lun Luo, Si-Yuan Cao, Hong Gu, Huaqi Zhang, Hui-Liang Shen:
Aggregating Feature Point Cloud for Depth Completion. 8698-8709



Haoyuan Li, Haoye Dong, Hanchao Jia, Dong Huang, Michael C. Kampffmeyer, Liang Lin, Xiaodan Liang:
Coordinate Transformer: Achieving Single-stage Multi-person Mesh Recovery from Videos. 8710-8719



Rajeev Yasarla, Hong Cai, Jisoo Jeong, Yunxiao Shi, Risheek Garrepalli, Fatih Porikli:
MAMo: Leveraging Memory and Attention for Monocular Video Depth Estimation. 8720-8730



Jongsung Lee, Gyeongsu Cho, Jeongin Park, Kyongjun Kim, Seongoh Lee, Jung-Hee Kim, Seong-Gyun Jeong, Kyungdon Joo:
SlaBins: Fisheye Depth Estimation using Slanted Bins on Road Environments. 8731-8740



Renke Wang, Guimin Que, Shuo Chen, Xiang Li, Jun Li, Jian Yang:
Creative Birds: Self-Supervised Single-View 3D Style Transfer. 8741-8750



Haotian Bai, Yiqi Lin, Yize Chen, Lin Wang :
Dynamic PlenOctree for Adaptive Sampling Refinement in Explicit NeRF. 8751-8761



Yuguang Li, Kai Wang, Hui Li, Seon-Min Rhee, Seungju Han, Jihye Kim, Min Yang, Ran Yang, Feng Zhu:
CORE: Co-planarity Regularized Monocular Geometry Estimation with Weak Supervision. 8762-8771



Foivos Paraperas Papantoniou, Alexandros Lattas, Stylianos Moschoglou, Stefanos Zafeiriou:
Relightify: Relightable 3D Faces from a Single Image via Diffusion Models. 8772-8783



Bruce X. B. Yu, Zhi Zhang, Yongxu Liu, Sheng-Hua Zhong, Yan Liu, Chang Wen Chen:

GLA-GCN: Global-local Adaptive Graph Convolutional Network for 3D Human Pose Estimation from Monocular Video. 8784-8795



Junho Kim, Eun Sun Lee, Young Min Kim:

Calibrating Panoramic Depth Estimation for Practical Localization and Mapping. 8796-8806



Christopher Wewer, Eddy Ilg, Bernt Schiele, Jan Eric Lenssen:

SimNP: Learning Self-Similarity Priors Between Neural Points. 8807-8818



Dongyue Chen, Tingxuan Huang, Zhimin Song, Shizhuo Deng, Tong Jia:

AGG-Net: Attention Guided Gated-convolutional Network for Depth Image Completion. 8819-8828



Stanislaw Szymanowicz, Christian Rupprecht, Andrea Vedaldi:

Viewset Diffusion: (0-)Image-Conditioned 3D Generative Models from 2D Data. 8829-8839



Haotian Dong, Enhui Ma, Lubo Wang, Miaohui Wang, Wuyuan Xie, Qing Guo, Ping Li, Lingyu Liang, Kairui Yang, Di Lin:

CVSformer: Cross-View Synthesis Transformer for Semantic Scene Completion. 8840-8849



Yan Di, Chenyangguang Zhang, Ruida Zhang, Fabian Manhardt, Yongzhi Su, Jason R. Rambach, Didier Stricker, Xiangyang Ji, Federico Tombari:

U-RED: Unsupervised 3D Shape Retrieval and Deformation for Partial Point Clouds. 8850-8861



Zhaoxuan Zhang, Bo Dong, Tong Li, Felix Heide, Pieter Peers, Baocai Yin, Xin Yang:

Single Depth-image 3D Reflection Symmetry and Shape Prediction. 8862-8872



Kieran Saunders, George Vogiatzis, Luis J. Manso:

Self-supervised Monocular Depth Estimation: Let's Talk About The Weather. 8873-8883



Alexey Bokhovkin, Shubham Tulsiani, Angela Dai:

Mesh2Tex: Generating Mesh Textures from Image Queries. 8884-8894



Zijie Wu, Yaonan Wang, Mingtao Feng, He Xie, Ajmal Mian:

Sketch and Text Guided Diffusion Model for Colored Point Cloud Generation. 8895-8905



Xiaoyang Lyu, Peng Dai, Zizhang Li, Dongyu Yan , Yi Lin, Yifan Peng, Xiaojuan Qi:

Learning A Room with the Occ-SDF Hybrid: Signed Distance Function Mingled with Occupancy Aids Scene Representation. 8906-8916



Chi Zhang, Wei Yin, Gang Yu, Zhibin Wang, Tao Chen, Bin Fu, Joey Tianyi Zhou, Chunhua Shen:

Robust Geometry-Preserving Depth Estimation Using Differentiable Rendering. 8917-8927

-      Jianglong Ye, Naiyan Wang, Xiaolong Wang:
FeatureNeRF: Learning Generalizable NeRFs by Distilling Foundation Models. 8928-8939
-     Yangyi Huang, Hongwei Yi, Weiyang Liu, Haofan Wang, Boxi Wu, Wenxiao Wang, Binbin Lin, Debing Zhang, Deng Cai:
One-shot Implicit Animatable Avatars with Model-based Priors. 8940-8951
-     Xinya Chen, Jiaxin Huang, Yanrui Bin, Lu Yu, Yiyi Liao:
VeRi3D: Generative Vertex-based Radiance Fields for 3D Controllable Human Image Synthesis. 8952-8963
-     Yutao Jiang, Yang Zhou, Yuan Liang, Wenxi Liu, Jianbo Jiao , Yuhui Quan, Shengfeng He :
Diffuse3D: Wide-Angle 3D Photography via Bilateral Diffusion. 8964-8974
-     Zheng Dang, Mathieu Salzmann:
AutoSynth: Learning to Generate 3D Training Data for Object Point Cloud Registration. 8975-8985
-     Yufei Zhang, Hanjing Wang, Jeffrey O. Kephart, Qiang Ji:
Body Knowledge and Uncertainty Modeling for Monocular 3D Human Body Reconstruction. 8986-8998
-     Tianke Zhang, Xiangeng Chu, Yunfei Liu, Lijian Lin, Zhendong Yang, Zhengzhuo Xu, Chengkun Cao, Fei Yu, Changyin Zhou, Chun Yuan, Yu Li:
Accurate 3D Face Reconstruction with Facial Component Tokens. 8999-9008
-     Wei Yin, Chi Zhang, Hao Chen, Zhipeng Cai, Gang Yu, Kaixuan Wang, Xiaozhi Chen, Chunhua Shen:
Metric3D: Towards Zero-shot Metric 3D Prediction from A Single Image. 9009-9019
-     Binghui Zuo, Zimeng Zhao, Wenqian Sun, Wei Xie, Zhou Xue , Yangang Wang:
Reconstructing Interacting Hands with Interaction Prior from Monocular Images. 9020-9030
-     Guangcong Wang, Zhaoxi Chen, Chen Change Loy, Ziwei Liu:
SparseNeRF: Distilling Depth Ranking for Few-shot Novel View Synthesis. 9031-9042
-     Yiran Yang, Dongshuo Yin, Xuee Rong, Xian Sun, Wenhui Diao, Xinming Li:
Beyond the limitation of monocular 3D detector via knowledge distillation. 9043-9052
-     Zenghao Chai, Tianke Zhang, Tianyu He, Xu Tan, Tadas Baltrusaitis, HsiangTao Wu, Runnan Li, Sheng Zhao, Chun Yuan, Jiang Bian:
HiFace: High-Fidelity 3D Face Reconstruction by Learning Static and Dynamic Details. 9053-9064
-     Jiacong Xu, Yi Zhang, Jiawei Peng, Wufei Ma, Artur Jesslen, Pengliang Ji, Qixin Hu, Jiehua Zhang, Qihao Liu, Jiahao Wang, Wei Ji, Chen Wang, Xiaoding Yuan, Prakhar Kaushik, Guofeng Zhang, Jie Liu, Yushan Xie, Yawen Cui, Alan L. Yuille, Adam Kortylewski:
Animal3D: A Comprehensive Dataset of 3D Animal Pose and Shape. 9065-9075

-      Jiahao Li, Zongxin Yang, Xiaohan Wang, Jianxin Ma, Chang Zhou, Yi Yang:
JOTR: 3D Joint Contrastive Learning with Transformers for Occluded Human Mesh Recovery. 9076-9087
-     Xueteng Yang, Yihao Luo, Yuliang Xiu, Wei Wang, Hao Xu, Zhaoxin Fan:
D-IF: Uncertainty-aware Human Digitization via Implicit Distribution Field. 9088-9098
-     Xuepeng Shi, Georgi Dikov, Gerhard Reitmayr, Tae-Kyun Kim, Mohsen Ghafoorian:
3D Distillation: Improving Self-Supervised Monocular Depth Estimation on Reflective Surfaces. 9099-9109
-     Junzhe Zhang, Yushi Lan, Shuai Yang, Fangzhou Hong, Quan Wang, Chai Kiat Yeo, Ziwei Liu, Chen Change Loy:
DeformToon3d: Deformable Neural Radiance Fields for 3D Toonification. 9110-9120
-     Renrui Zhang, Han Qiu, Tai Wang, Ziyu Guo, Ziteng Cui, Yu Qiao, Hongsheng Li, Peng Gao:
MonoDETR: Depth-guided Transformer for Monocular 3D Object Detection. 9121-9132
-     Jun Hoong Chan, Bohan Yu, Heng Guo, Jieji Ren, Zongqing Lu, Boxin Shi:
ReLeaPS : Reinforcement Learning-based Illumination Planning for Generalized Photometric Stereo. 9133-9141
-     Vaibhav Vavilala, David A. Forsyth:
Convex Decomposition of Indoor Scenes. 9142-9152
-     Muhammad Zubair Irshad, Sergey Zakharov, Katherine Liu, Vitor Guizilini, Thomas Kollar, Adrien Gaidon, Zsolt Kira, Rares Ambrus:
NeO 360: Neural Fields for Sparse View Synthesis of Outdoor Scenes. 9153-9164
-     Yuanbo Yang, Yifei Yang, Hanlei Guo, Rong Xiong, Yue Wang, Yiyi Liao:
UrbanGIRAFFE: Representing Urban Scenes as Compositional Generative Neural Feature Fields. 9165-9176
-     Yuxiang Lan, Yachao Zhang, Xu Ma, Yanyun Qu, Yun Fu:
Efficient Converted Spiking Neural Network for 3D and 2D Classification. 9177-9186
-     Lin Geng Foo, Jia Gong, Hossein Rahmani, Jun Liu:
Distribution-Aligned Diffusion for Human Mesh Recovery. 9187-9198
-     Vitor Guizilini, Igor Vasiljevic, Dian Chen, Rares Ambrus, Adrien Gaidon:
Towards Zero-Shot Scale-Aware Monocular Depth Estimation. 9199-9209
-     Alex Costanzino , Pierluigi Zama Ramirez , Matteo Poggi, Fabio Tosi, Stefano Mattoccia , Luigi Di Stefano:
Learning Depth Estimation for Transparent and Mirror Surfaces. 9210-9221
-     Xiang Zhang, Zeyuan Chen, Fangyin Wei, Zhuowen Tu:
Uni-3D: A Universal Model for Panoptic 3D Scene Reconstruction. 9222-9232

-      Ling Luo, Pinaki Nath Chowdhury, Tao Xiang, Yi-Zhe Song, Yulia Gryaditskaya: **3D VR Sketch Guided 3D Shape Prototyping and Exploration.** 9233-9242
-     Mingqi Shao, Chongkun Xia, Zhendong Yang, Junnan Huang, Xueqian Wang: **Transparent Shape from a Single View Polarization Image.** 9243-9252
-     Zhangyang Xiong, Di Kang, Derong Jin, Weikai Chen, Linchao Bao, Shuguang Cui, Xiaoguang Han:
Get3DHuman: Lifting StyleGAN-Human into a 3D Generative Model using Pixel-aligned Reconstruction Priors. 9253-9263
-     Ruoshi Liu, Rundi Wu, Basile Van Hoorick, Pavel Tokmakov, Sergey Zakharov, Carl Vondrick:
Zero-1-to-3: Zero-shot One Image to 3D Object. 9264-9275
-     Guangkai Xu, Wei Yin, Hao Chen, Chunhua Shen, Kai Cheng, Feng Zhao:
FrozenRecon: Pose-free 3D Scene Reconstruction with Frozen Depth Models. 9276-9286
-     Mohammad Samiul Arshad, William J. Beksi:
LIST: Learning Implicitly from Spatial Transformers for Single-View 3D Reconstruction. 9287-9296
-     Ta Ying Cheng, Matheus Gadelha, Sören Pirk, Thibault Groueix, Radomír Mech, Andrew Markham, Niki Trigoni:
3DMiner: Discovering Shapes from Large-Scale Unannotated Image Datasets. 9297-9307
-     Wei Xie, Zimeng Zhao, Shiying Li, Binghui Zuo, Yangang Wang:
Nonrigid Object Contact Estimation With Regional Unwrapping Transformer. 9308-9317
-     Shoukang Hu, Fangzhou Hong, Liang Pan, Haiyi Mei, Lei Yang, Ziwei Liu:
SHERF: Generalizable Human NeRF from a Single Image. 9318-9330
-     Nan Jiang, Tengyu Liu, Zhexuan Cao, Jieming Cui, Zhiyuan Zhang, Yixin Chen, He Wang, Yixin Zhu, Siyuan Huang:
Full-Body Articulated Human-Object Interaction. 9331-9342
-     Jingjia Shi, Shuaifeng Zhi, Kai Xu:
PlaneRecTR: Unified Query Learning for 3D Plane Recovery from a Single View. 9343-9352
-     Anh-Quan Cao, Raoul de Charette:
SceneRF: Self-Supervised Monocular 3D Scene Reconstruction with Radiance Fields. 9353-9364
-     Yi Zhang, Pengliang Ji, Angtian Wang, Jieru Mei, Adam Kortylewski, Alan L. Yuille:
3D-Aware Neural Body Fitting for Occlusion Robust 3D Human Pose Estimation. 9365-9376
-     Zhengming Zhou, Qiulei Dong:
Two-in-One Depth: Bridging the Gap Between Monocular and Binocular Self-supervised Depth Estimation. 9377-9387
-     Yufei Wang, Bo Li, Ge Zhang, Qi Liu , Tao Gao, Yuchao Dai:
LRRU: Long-short Range Recurrent Updating Networks for Depth

-      Completion. 9388-9398
-     Yunpeng Zhang, Zheng Zhu, Dalong Du:
OccFormer: Dual-path Transformer for Vision-based 3D Semantic Occupancy Prediction. 9399-9409
-     Kailin Li, Lixin Yang, Haoyu Zhen, Zenan Lin, Xinyu Zhan, Licheng Zhong, Jian Xu, Kejian Wu, Cewu Lu:
Chord: Category-level Hand-held Object Reconstruction via Shape Deformation. 9410-9420
-     Jiawei Yao, Chuming Li, Keqiang Sun, Yingjie Cai, Hao Li, Wanli Ouyang, Hongsheng Li:
NDC-Scene: Boost Monocular 3D Semantic Scene Completion in Normalized Device Coordinates Space. 9421-9431
-     Yiran Wang, Min Shi, Jiaqi Li, Zihao Huang, Zhiguo Cao, Jianming Zhang, Ke Xian, Guosheng Lin:
Neural Video Depth Stabilizer. 9432-9442
-     Feishi Wang, Jieji Ren, Heng Guo, Mingjun Ren, Boxin Shi:
DiLiGenT-Π: Photometric Stereo for Planar Surfaces with Rich Details - Benchmark Dataset and Beyond. 9443-9453
-     Mathis Petrovich, Michael J. Black, Güл Varol:
TMR: Text-to-Motion Retrieval Using Contrastive 3D Human Motion Synthesis. 9454-9463
-     Shuai Li, Sisi Zhuang, Wenfeng Song, Xinyu Zhang, Hejia Chen, Aimin Hao:
Sequential Texts Driven Cohesive Motions Synthesis with Natural Transitions. 9464-9474
-     Chenxin Xu, Robby T. Tan, Yuhong Tan, Siheng Chen, Xinchao Wang, Yanfeng Wang:
Auxiliary Tasks Benefit 3D Skeleton-based Human Motion Prediction. 9475-9486
-     Changxing Deng, Ao Luo, Haibin Huang, Shaodan Ma, Jiangyu Liu, Shuaicheng Liu:
Explicit Motion Disentangling for Efficient Optical Flow Estimation. 9487-9496
-     Gianluca Mancusi, Aniello Panariello, Angelo Porrello, Matteo Fabbri, Simone Calderara, Rita Cucchiara:
TrackFlow: Multi-Object Tracking with Normalizing Flows. 9497-9509
-     Ling-Hao Chen, Jiawei Zhang, Yewen Li, Yiren Pang, Xiaobo Xia, Tongliang Liu:
HumanMAC: Masked Motion Completion for Human Motion Prediction. 9510-9521
-     Jiazen Liu, Xirong Li:
Geometrized Transformer for Self-Supervised Homography Estimation. 9522-9531
-     Shuai Yuan , Shuzhi Yu, Hannah Kim, Carlo Tomasi:
SemARFlow: Injecting Semantics into Unsupervised Optical Flow

**Estimation for Autonomous Driving.** 9532-9543

Konstantin Pakulev, Alexander Vakhitov, Gonzalo Ferrer:

NeSS-ST: Detecting Good and Stable Keypoints with a Neural Stability Score and the Shi-Tomasi detector. 9544-9554

Yidong Cai, Jie Liu, Jie Tang, Gangshan Wu:

Robust Object Modeling for Visual Tracking. 9555-9566

Julian Tanke, Linguang Zhang, Amy Zhao, Chengcheng Tang, Yujun Cai, Lezi Wang, Po-Chen Wu, Juergen Gall, Cem Keskin:

Social Diffusion: Long-term Multiple Human Motion Anticipation. 9567-9577

Ben Kang, Xin Chen, Dong Wang, Houwen Peng, Huchuan Lu:

Exploring Lightweight Hierarchical Vision Transformers for Efficient Visual Tracking. 9578-9587

Sadegh Aliakbarian, Fatemeh Sadat Saleh, David Collier, Pashmina Cameron, Darren Cosker:

HMD-NeMo: Online 3D Avatar Motion Generation From Sparse Observations. 9588-9597

Rui Li, Shenglong Zhou, Dong Liu:

Learning Fine-Grained Features for Pixel-wise Video Correspondences.

9598-9607



Ao Luo, Fan Yang, Xin Li, Lang Nie, Chunyu Lin, Haoqiang Fan, Shuaicheng Liu:

GAFlow: Incorporating Gaussian Attention into Optical Flow. 9608-9617

Miao Fan, Mingrui Chen, Chen Hu, Shuchang Zhou:

Occ²Net: Robust Image Matching Based on 3D Occupancy Estimation for Occluded Regions. 9618-9628

Jiye Lee, Hanbyul Joo:

Locomotion-Action-Manipulation: Synthesizing Human-Scene Interactions in Complex 3D Environments. 9629-9640

Liushuai Shi, Le Wang, Sanping Zhou, Gang Hua:

Trajectory Unified Transformer for Pedestrian Trajectory Prediction. 9641-9650

Haotian Liu, Guang Chen, Sanqing Qu, Yanping Zhang, Zhijun Li, Alois Knoll, Changjun Jiang:

TMA: Temporal Motion Aggregation for Event-based Optical Flow. 9651-9660

Federico Paredes-Vallés, Kirk Y. W. Scheper, Christophe De Wagter, Guido C. H. E. de Croon:

Taming Contrast Maximization for Learning Sequential, Low-latency, Event-based Optical Flow. 9661-9671

Rémi Pautrat, Iago Suárez, Yifan Yu, Marc Pollefeys, Viktor Larsson:

GlueStick: Robust Image Matching by Sticking Points and Lines Together.

9672-9682



Mattia Segù, Bernt Schiele, Fisher Yu:

DARTH: Holistic Test-time Adaptation for Multiple Object Tracking. 9683-



9693

- Emanuele Santellani, Christian Sormann, Mattia Rossi, Andreas Kuhn, Friedrich Fraundorfer:
S-TREK: Sequential Translation and Rotation Equivariant Keypoints for local feature extraction. 9694-9703
- Yuanyou Xu, Zongxin Yang, Yi Yang:
Integrating Boxes and Masks: A Multi-Object Framework for Unified Visual Tracking and Segmentation. 9704-9717
- Fabien Delattre, David Dirnfeld, Phat Nguyen, Stephen Scarano, Michael J. Jones, Pedro Miraldo, Erik G. Learned-Miller:
Robust Frame-to-Frame Camera Rotation Estimation in Crowded Scenes. 9718-9728
- Yonghao Dong, Le Wang, Sanping Zhou, Gang Hua:
Sparse Instance Conditioned Multimodal Trajectory Prediction. 9729-9738
- Jianyuan Wang, Christian Rupprecht, David Novotný:
PoseDiffusion: Solving Pose Estimation via Diffusion-aided Bundle Adjustment. 9739-9749
- Shuxiao Ding , Eike Rehder, Lukas Schneider, Marius Cordts, Juergen Gall:
3DMOTFormer: Graph Transformer for Online 3D Multi-Object Tracking. 9750-9760
- Takahiro Maeda, Norimichi Ukita:
Fast Inference and Update of Probabilistic Density Estimation on Trajectory Prediction. 9761-9771
- Hai Jiang , Haipeng Li, Songchen Han, Haoqiang Fan, Bing Zeng, Shuaicheng Liu:
Supervised Homography Learning with Realistic Dataset Generation. 9772-9781
- Qingyao Xu, Weibo Mao, Jingze Gong, Chenxin Xu, Siheng Chen, Weidi Xie, Ya Zhang, Yanfeng Wang:
Joint-Relation Transformer for Multi-Person Motion Prediction. 9782-9792
- Wachirawit Ponghiran, Chamika Mihiranga Liyanagedera, Kaushik Roy:
Event-based Temporally Dense Optical Flow Estimation with Sequential Learning. 9793-9802
- Brandon Y. Feng , Hadi Alzayer, Michael Rubinstein, William T. Freeman, Jia-Bin Huang:
3D Motion Magnification: Visualizing Subtle Motions with Time-Varying Radiance Fields. 9803-9812
- Xinglong Luo, Kunming Luo, Ao Luo, Zhengning Wang, Ping Tan, Shuaicheng Liu:
Learning Optical Flow from Event Camera with Rendered Dataset. 9813-9823
- Hung Tran, Vuong Le, Svetha Venkatesh, Truyen Tran:
Persistent-Transient Duality: A Multi-mechanism Approach for Modeling Human-Object Interaction. 9824-9833

-      Weilong Yan, Robby T. Tan, Bing Zeng, Shuaicheng Liu:
Deep Homography Mixture for Single Image Rolling Shutter Correction.
9834-9843
-     Xueqian Li, Jianqiao Zheng, Francesco Ferroni, Jhony Kaesemeyer Pontes, Simon Lucey:
Fast Neural Scene Flow. 9844-9856
-     Chang Nie, Guangming Wang, Zhe Liu, Luca Cavalli, Marc Pollefeys, Hesheng Wang:
RLSAC: Reinforcement Learning enhanced Sample Consensus for End-to-End Robust Estimation. 9857-9866
-     Ruopeng Gao, Limin Wang:
MeMOTR: Long-Term Memory-Augmented Transformer for Multi-Object Tracking. 9867-9876
-     Tian-Xing Xu, Yuan-Chen Guo, Yu-Kun Lai, Song-Hai Zhang:
MBPTrack: Improving 3D Point Cloud Tracking with Memory networks and Box Priors. 9877-9886
-     Yutao Cui, Chenkai Zeng, Xiaoyu Zhao, Yichun Yang, Gangshan Wu, Limin Wang:
SportsMOT: A Large Multi-Object Tracking Dataset in Multiple Sports Scenes. 9887-9897
-     Rui Li, Baopeng Zhang, Jun Liu, Wei Liu, Jian Zhao, Zhu Teng:
Heterogeneous Diversity Driven Active Learning for Multi-Object Tracking. 9898-9907
-     Kehong Gong, Dongze Lian, Heng Chang, Chuan Guo, Zihang Jiang, Xinxin Zuo , Michael Bi Mi, Xinchao Wang:
TM2D: Bimodality Driven 3D Dance Generation via Music-Text Integration. 9908-9918
-     Teli Ma, Mengmeng Wang, Jimin Xiao, Huifeng Wu, Yong Liu:
Synchronize Feature Extracting and Matching: A Single Branch Framework for 3D Object Tracking. 9919-9929
-     Yiheng Liu, Junta Wu, Yi Fu:
Collaborative Tracking Learning for Frame-Rate-Insensitive Multi-Object Tracking. 9930-9939
-     Xin Li, Yuqing Huang , Zhenyu He, Yaowei Wang, Huchuan Lu, Ming-Hsuan Yang:
CiteTracker: Correlating Image and Text for Visual Tracking. 9940-9949
-     Nikos Athanasiou, Mathis Petrovich, Michael J. Black, Gü̈l Varol:
SINC: Spatial Composition of 3D Human Motions for Simultaneous Action Generation. 9950-9961
-     Kai Liu, Sheng Jin, Zhihang Fu, Ze Chen, Rongxin Jiang, Jieping Ye:
Uncertainty-aware Unsupervised Multi-Object Tracking. 9962-9971
-     Bowen Li, Ziyuan Huang, Junjie Ye, Yiming Li, Sebastian A. Scherer, Hang Zhao, Changhong Fu:



- PVT++: A Simple End-to-End Latency-Aware Visual Tracking Framework.**
9972-9982
- Inhwan Bae, Jean Oh, Hae-Gon Jeon:
EigenTrajectory: Low-Rank Descriptors for Multi-Modal Trajectory Forecasting. 9983-9995
- Zhexiong Wan, Yuxin Mao, Jing Zhang, Yuchao Dai:
RPEFlow: Multimodal Fusion of RGB-PointCloud-Event for Joint Optical Flow and Scene Flow Estimation. 9996-10006
- Wencan Cheng, Jong Hwan Ko:
Multi-Scale Bidirectional Recurrent Network with Hybrid Correlation for Point Cloud Based Scene Flow Estimation. 10007-10016
- Cheng-Che Cheng, Min-Xuan Qiu, Chen-Kuo Chiang, Shang-Hong Lai:
ReST: A Reconfigurable Spatial-Temporal Graph Model for Multi-Camera Multi-Object Tracking. 10017-10026
- Carl Doersch, Yi Yang, Mel Vecerík, Dilara Gokay, Ankush Gupta, Yusuf Aytar, João Carreira, Andrew Zisserman:
TAPIR: Tracking Any Point with per-frame Initialization and temporal Refinement. 10027-10038
- Yun Wang , Cheng Chi, Min Lin, Xin Yang:
IHNet: Iterative Hierarchical Network Guided by High-Resolution Estimated Information for Scene Flow Estimation. 10039-10048
- Evonne Ng, Sanjay Subramanian, Dan Klein, Angjoo Kanazawa, Trevor Darrell, Shiry Ginosar:
Can Language Models Learn to Listen? 10049-10059
- Lei Lai, Zhongkai Shangguan, Jimuyang Zhang, Eshed Ohn-Bar:
XVO: Generalized Visual Odometry via Cross-Modal Self-Training. 10060-10071
- Jenny Schmalfuss, Lukas Mehl, Andrés Bruhn:
Distracting Downpour: Adversarial Weather Attacks for Motion Estimation. 10072-10082
- Dawei Yang, Jianfeng He, Yinчao Ma, Qianjin Yu, Tianzhu Zhang:
Foreground-Background Distribution Modeling Transformer for Visual Object Tracking. 10083-10093
- Reza Ghoddoosian, Ish Dwivedi, Nakul Agarwal, Behzad Dariush:
Weakly-Supervised Action Segmentation and Unseen Error Detection in Anomalous Instructional Videos. 10094-10104
- Daochang Liu, Qiyue Li, Anh-Dung Dinh, Tingting Jiang, Mubarak Shah, Chang Xu:
Diffusion Action Segmentation. 10105-10115
- Muhammad Adi Nugroho, Sangmin Woo, Sumin Lee, Changick Kim:
Audio-Visual Glance Network for Efficient Video Recognition. 10116-10125
- Kun Xia, Le Wang, Sanping Zhou, Gang Hua, Wei Tang:
Learning from Noisy Pseudo Labels for Semi-Supervised Temporal Action Localization. 10126-10135

-  [Yifei Chen, Dapeng Chen, Ruijin Liu, Hao Li, Wei Peng:](#) **Video Action Recognition with Attentive Semantic Units.** 10136-10146
-  [Yunyao Mao, Jiajun Deng, Wengang Zhou, Yao Fang, Wanli Ouyang, Houqiang Li:](#) **Masked Motion Predictors are Strong 3D Action Representation Learners.** 10147-10157
-  [Kranthi Kumar Rachavarapu, A. N. Rajagopalan:](#) **Boosting Positive Segments for Weakly-Supervised Audio-Visual Video Parsing.** 10158-10168
-  [Guiqin Wang, Peng Zhao, Cong Zhao, Shusen Yang, Jie Cheng, Liziwei Leng, Jianxing Liao, Qinghai Guo:](#) **Weakly-Supervised Action Localization by Hierarchically-structured Latent Attention Modeling.** 10169-10179
-  [Juntae Lee, Mihir Jain, Sungrack Yun:](#) **Few-Shot Common Action Localization via Cross-Attentional Fusion of Context and Temporal Dynamics.** 10180-10189
-  [Chihiro Nakatani, Hiroaki Kawashima, Norimichi Ukita:](#) **Interaction-aware Joint Attention Estimation Using People Attributes.** 10190-10199
-  [Ronghui Li, Junfan Zhao, Yachao Zhang, Mingyang Su, Zeping Ren, Han Zhang, Yansong Tang, Xiu Li:](#) **FineDance: A Fine-grained Choreography Dataset for 3D Full Body Dance Generation.** 10200-10209
-  [Yuanhao Zhai, Ziyi Liu !\[\]\(cc6909d841d43c4b6544c60064314dbc_img.jpg\), Zhenyu Wu, Yi Wu, Chunluan Zhou, David S. Doermann, Junsong Yuan, Gang Hua:](#) **SOAR: Scene-debiasing Open-set Action Recognition.** 10210-10220
-  [Jungho Lee, Minhyeok Lee, Suhwan Cho, Sungmin Woo, Sungjun Jang, Sangyoun Lee:](#) **Leveraging Spatio-Temporal Dependency for Skeleton-Based Action Recognition.** 10221-10230
-  [Sangwon Kim, Dasom Ahn, ByoungChul Ko:](#) **Cross-Modal Learning with 3D Deformable Attention for Action Recognition.** 10231-10241
-  [Wangmeng Xiang, Chao Li, Yuxuan Zhou, Biao Wang, Lei Zhang:](#) **Generative Action Description Prompts for Skeleton-based Action Recognition.** 10242-10251
-  [Jihwan Kim, Miso Lee, Jae-Pil Heo:](#) **Self-Feedback DETR for Temporal Action Detection.** 10252-10262
-  [Zhiheng Li, Wenjia Geng, Muheng Li, Lei Chen, Yansong Tang, Jiwen Lu, Jie Zhou:](#) **Skip-Plan: Procedure Planning in Instructional Videos via Condensed Action Space Learning.** 10263-10272
-  [Giacomo Zara, Alessandro Conti !\[\]\(9279b08fecf641ac9de61c9be2ccebe9_img.jpg\), Subhankar Roy, Stéphane Lathuilière, Paolo Rota, Elisa Ricci:](#)



The Unreasonable Effectiveness of Large Language-Vision Models for Source-free Video Domain Adaptation. 10273-10283



Alessandro Flaborea, Luca Collorone, Guido Maria D'Amely di Melendugno, Stefano D'Arrigo, Bardh Prenkaj, Fabio Galasso:

Multimodal Motion Conditioned Diffusion Model for Skeleton-based Video Anomaly Detection. 10284-10295



Chenrui Shi, Che Sun, Yuwei Wu, Yunde Jia:

Video Anomaly Detection via Sequentially Learning Multiple Pretext Tasks. 10296-10306



Joungbin An, Hyolim Kang, Su Ho Han, Ming-Hsuan Yang, Seon Joo Kim: **MiniROAD: Minimal RNN Framework for Online Action Detection.** 10307-10316



Emad Bahrami Rad, Gianpiero Francesca, Juergen Gall:

How Much Temporal Long-Term Context is Needed for Action Segmentation? 10317-10327



Sauradip Nag, Xiatian Zhu, Jiankang Deng, Yi-Zhe Song, Tao Xiang:

DiffTAD: Temporal Action Detection with Proposal Denoising Diffusion. 10328-10340



Anshul Shah, Benjamin Lundell, Harpreet Sawhney, Rama Chellappa:

STEPs: Self-Supervised Key Step Extraction and Localization from Unlabeled Procedural Videos. 10341-10353



Lei Chen, Zhan Tong, Yibing Song, Gangshan Wu, Limin Wang:

Efficient Video Action Detection with Token Dropout and Context Refinement. 10354-10365



Jingwen Guo, Hong Liu, Shitong Sun, Tianyu Guo, Min Zhang, Chenyang Si:

FSAR: Federated Skeleton-based Action Recognition with Adaptive Topology Structure and Knowledge Distillation. 10366-10376



Frederic Z. Zhang, Yuhui Yuan, Dylan Campbell, Zhuoyao Zhong, Stephen Gould: **Exploring Predicate Visual Context in Detecting of Human-Object Interactions.** 10377-10387



Shuqiang Cao, Weixin Luo, Bairui Wang, Wei Zhang, Lin Ma:

E2E-LOAD: End-to-End Long-form Online Action Detection. 10388-10398



Qinying Liu, Zilei Wang, Shenghai Rong, Junjie Li, Yixin Zhang:

Revisiting Foreground and Background Separation in Weakly-supervised Temporal Action Localization: A Clustering-based Approach. 10399-10409



Jungho Lee, Minhyeok Lee, Dogyoon Lee, Sangyoun Lee:

Hierarchically Decomposed Graph Convolutional Networks for Skeleton-Based Action Recognition. 10410-10419



Numair Khan, Lei Xiao, Douglas Lanman:

Tiled Multiplane Images for Practical 3D Photography. 10420-10430



Shantanu Gupta, Mohit Gupta:

Eulerian Single-Photon Vision. 10431-10442

-      Shangchen Zhou, Chongyi Li, Kelvin C. K. Chan, Chen Change Loy:
ProPainter: Improving Propagation and Transformer for Video Inpainting. 10443-10452
-     Jinyang Tai:
Global Perception Based Autoregressive Neural Processes. 10453-10463
-     Jiaming Liu, Rushil Anirudh, Jayaraman J. Thiagarajan, Stewart He, K. Aditya Mohan, Ulugbek S. Kamilov, Hyojin Kim:
DOLCE: A Model-Based Probabilistic Diffusion Framework for Limited-Angle CT Reconstruction. 10464-10474
-     Chao Wang, Ana Serrano, Xingang Pan, Bin Chen, Karol Myszkowski, Hans-Peter Seidel, Christian Theobalt, Thomas Leimkühler:
GlowGAN: Unsupervised Learning of HDR Images from LDR Images in the Wild. 10475-10485
-     Berthy T. Feng, Jamie Smith, Michael Rubinstein, Huiwen Chang, Katherine L. Bouman, William T. Freeman:
Score-Based Diffusion Models as Principled Priors for Inverse Imaging. 10486-10497
-     Yuki Fujimura, Takahiro Kushida, Takuya Funatomi, Yasuhiro Mukaigawa:
NLOS-NeuS: Non-line-of-sight Neural Implicit Surface. 10498-10507
-     Ting Jiang, Chuan Wang, Xinpeng Li, Ru Li, Haoqiang Fan, Shuaicheng Liu:
MEFLUT: Unsupervised 1D Lookup Tables for Multi-exposure Image Fusion. 10508-10517
-     Wenjie Wei, Malu Zhang, Hong Qu, Ammar Belatreche, Jian Zhang, Hong Chen:
Temporal-Coded Spiking Neural Networks with Dynamic Firing Threshold: Learning with Event-Driven Backpropagation. 10518-10528
-     Yanhua Yu, Siyuan Shen, Zi Wang , Binbin Huang, Yuehan Wang, Xingyue Peng, Suan Xia, Ping Liu, Ruiqian Li, Shiying Li:
Enhancing Non-line-of-sight Imaging via Learnable Inverse Kernel and Attention Mechanisms. 10529-10539
-     Tao Lv, Hao Ye, Quan Yuan, Zhan Shi, Yibo Wang, Shuming Wang, Xun Cao:
Aperture Diffraction for Compact Snapshot Spectral Imaging. 10540-10550
-     JoonKyu Park, Sanghyun Son, Kyoung Mu Lee:
Content-Aware Local GAN for Photo-Realistic Super-Resolution. 10551-10560
-     Berk Iskender , Marc Louis Klasky, Yoram Bresler:
RED-PSM: Regularization by Denoising of Partially Separable Models for Dynamic Imaging. 10561-10570
-     Goutam Bhat, Michaël Gharbi, Jiawen Chen, Luc Van Gool, Zhihao Xia:
Self-Supervised Burst Super-Resolution. 10571-10580
-     Jinxiu Liang, Yixin Yang, Boyu Li, Peiqi Duan, Yong Xu, Boxin Shi:
Coherent Event Guided Low-Light Video Enhancement. 10581-10591
-     Sacha Jungerman, Atul Ingle, Mohit Gupta:
Panoramas from Photons. 10592-10602

-      Anqi Yang, Eunhee Kang, Hyong-Euk Lee, Aswin C. Sankaranarayanan: **Designing Phase Masks for Under-Display Cameras.** 10603-10611
-     Ping Wang , Lishun Wang, Xin Yuan: **Deep Optics for Video Snapshot Compressive Imaging.** 10612-10622
-     Sachin Shah, Sakshum Kulshrestha, Christopher A. Metzler: **TiDy-PSFs: Computational Imaging with Time-Averaged Dynamic Point-Spread-Functions.** 10623-10633
-     Mingde Yao, Jie Huang, Xin Jin, Ruikang Xu, Shenglong Zhou, Man Zhou, Zhiwei Xiong: **Generalized Lightness Adaptation with Channel Selective Normalization.** 10634-10645
-     Delin Qu, Yizhen Lao, Zhigang Wang, Dong Wang, Bin Zhao, Xuelong Li: **Towards Nonlinear-Motion-Aware and Occlusion-Robust Rolling Shutter Correction.** 10646-10654
-     Saurabh Yadav, Koteswar Rao Jeripothula : **FCCNs: Fully Complex-valued Convolutional Networks using Complex-valued Color Model and Loss Function.** 10655-10664
-     Yan Yang, Liyuan Pan, Liu Liu: **Event Camera Data Pre-training.** 10665-10675
-     Suhyeon Lee, Hyungjin Chung, Minyoung Park, Jonghyuk Park, Wi-Sun Ryu, Jong Chul Ye: **Improving 3D Imaging with Pre-Trained Perpendicular 2D Diffusion Models.** 10676-10686
-     Mengwei Ren, Mauricio Delbracio, Hossein Talebi, Guido Gerig, Peyman Milanfar: **Multiscale Structure Guided Diffusion for Image Deblurring.** 10687-10699
-     Xiang Zhang, Lei Yu, Wen Yang, Jianzhuang Liu, Gui-Song Xia: **Generalizing Event-Based Motion Deblurring in Real-World Scenarios.** 10700-10710
-     Seongmin Hong, Inbum Park, Se Young Chun: **On the Robustness of Normalizing Flows for Inverse Problems in Imaging.** 10711-10721
-     Felipe Gutierrez-Barragan, Fangzhou Mu, Andrei Ardelean, Atul Ingle, Claudio Bruschini, Edoardo Charbon, Yin Li, Mohit Gupta, Andreas Velten: **Learned Compressive Representations for Single-Photon 3D Imaging.** 10722-10732
-     Enze Ye, Yuhang Wang, Hong Zhang, Yiqin Gao, Huan Wang, He Sun: **Recovering a Molecule's 3D Dynamics from Liquid-phase Electron Microscopy Movies.** 10733-10743
-     Muyao Niu, Zhihang Zhong, Yinqiang Zheng: **NIR-assisted Video Enhancement via Unpaired 24-hour Data.** 10744-10754
-     Dorian Chan, Mark Sheinin, Matthew O'Toole: **SpinCam: High-Speed Imaging via a Rotating Point-Spread Function.** 10755-10765

-      Kang Liao, Lang Nie, Chunyu Lin, Zishuo Zheng, Yao Zhao:
RecRecNet: Rectangling Rectified Wide-Angle Images by Thin-Plate Spline Model and DoF-based Curriculum Learning. 10766-10775
-     Shuchen Weng, Peixuan Zhang, Zheng Chang, Xinlong Wang, Si Li, Boxin Shi:
Affective Image Filter: Reflecting Emotions from Text to Images. 10776-10785
-     Feng Zhang, Bin Xu, Zhiqiang Li, Xinran Liu, Qingbo Lu, Changxin Gao, Nong Sang:
Towards General Low-Light Raw Noise Synthesis and Modeling. 10786-10796
-     Jin Wang, Wenming Weng, Yueyi Zhang, Zhiwei Xiong:
Unsupervised Video Deraining with An Event Camera. 10797-10806
-     Cong Wang , Yu-Ping Wang, Dinesh Manocha:
LoLep: Single-View View Synthesis with Locally-Learned Planes and Self-Attention Occlusion Inference. 10807-10817
-     Xiaoyu Huang, Dhruv Batra, Akshara Rai, Andrew Szot:
Skill Transformer: A Monolithic Policy for Mobile Manipulation. 10818-10828
-     Klemen Kotar, Aaron Walsman, Roozbeh Mottaghi:
ENTL: Embodied Navigation Trajectory Learner. 10829-10838
-     Hanqing Wang, Wei Liang, Luc Van Gool, Wenguan Wang:
Dreamwalker: Mental Planning for Continuous Vision-Language Navigation. 10839-10849
-     Kunal Pratap Singh, Jordi Salvador, Luca Weihs, Aniruddha Kembhavi:
Scene Graph Contrastive Learning for Embodied Navigation. 10850-10860
-     Zhengyi Luo, Jinkun Cao, Alexander Winkler, Kris Kitani, Weipeng Xu:
Perpetual Humanoid Control for Real-time Simulated Avatars. 10861-10870
-     Yuhang Yang, Wei Zhai, Hongchen Luo, Yang Cao, Jiebo Luo, Zheng-Jun Zha:
Grounding 3D Object Affordance from 2D Interactions in Images. 10871-10881
-     Jacob Krantz, Théophile Gervet, Karmesh Yadav, Austin S. Wang, Chris Paxton, Roozbeh Mottaghi, Dhruv Batra, Jitendra Malik, Stefan Lee, Devendra Singh Chaplot:
Navigating to Objects Specified by Images. 10882-10891
-     Albert J. Zhai, Shenlong Wang:
PEANUT: Predicting and Navigating to Unseen Targets. 10892-10901
-     Byeonghwi Kim, Jinyeon Kim, Yuyeong Kim, Cheolhong Min, Jonghyun Choi:
Context-Aware Planning and Environment-Aware Memory for Instruction Following Embodied Agents. 10902-10912
-     Ruihai Wu, Chuanruo Ning, Hao Dong:
Learning Foresightful Dense Visual Affordance for Deformable Object Manipulation. 10913-10922



Enrico Cancelli, Tommaso Campari, Luciano Serafini, Angel X. Chang, Lamberto Ballan :

Exploiting Proximity-Aware Tasks for Embodied Social Navigation. 10923-10933



Rui Liu, Xiaohan Wang, Wenguan Wang, Yi Yang:

Bird's-Eye-View Scene Graph for Vision-Language Navigation. 10934-10946



Zike Yan, Haoxiang Yang, Hongbin Zha:

Active Neural Mapping. 10947-10958



Jinyu Chen, Wenguan Wang, Si Liu, Hongsheng Li, Yi Yang:

Omnidirectional Information Gathering for Knowledge Transfer-based Audio-Visual Navigation. 10959-10969



Pierre Marza, Laëtitia Matignon, Olivier Simonin, Christian Wolf:

Multi-Object Navigation with dynamically learned neural implicit representations. 10970-10981



Conghui Hu, Can Zhang, Gim Hee Lee:

Unsupervised Feature Representation Learning for Domain-generalized Cross-domain Image Retrieval. 10982-10991



Dmitry Baranchuk, Matthijs Douze, Yash Upadhyay, I. Zeki Yalniz:

DeDrift: Robust Similarity Search under Content Drift. 10992-11001



Shihao Shao, Kaifeng Chen, Arjun Karpur, Qinghua Cui, André Araújo, Bingyi Cao:

Global Features are All You Need for Image Retrieval and Reranking. 11002-11012



Bailin Yang, Haoqiang Sun, Frederick W. B. Li

Zheng Chen, Jianlu Cai, Chao Song:

HSE: Hybrid Species Embedding for Deep Metric Learning. 11013-11023



Chang Zou, Zeqi Chen, Zhichao Cui, Yuehu Liu, Chi Zhang:

Discrepant and Multi-instance Proxies for Unsupervised Person Re-identification. 11024-11034



Bin Yang, Jun Chen, Mang Ye:

Towards Grand Unified Representation Learning for Unsupervised Visible-Infrared Person Re-Identification. 11035-11045



Gabriele Moreno Berton, Gabriele Trivigno, Barbara Caputo, Carlo Masone :

EigenPlaces: Training Viewpoint Robust Models for Visual Place Recognition. 11046-11056



Kaiqu Liang, Samuel Albanie:

Simple Baselines for Interactive Video Retrieval with Questions and Answers. 11057-11067



Xin Chen, Bin Wang, Yongsheng Gao:

Fan-Beam Binarization Difference Projection (FB-BDP): A Novel Local Object Descriptor for Fine-Grained Leaf Image Retrieval. 11068-11077



Chull Hwan Song, Taebaek Hwang, Jooyoung Yoon, Shunghyun Choi, Yeong Hyeon Gu:



- Conditional Cross Attention Network for Multi-Space Embedding without Entanglement in Only a SINGLE Network.** 11078-11087
Jianbing Wu, Hong Liu, Yuxin Su, Wei Shi, Hao Tang:
- Learning Concordant Attention via Target-aware Alignment for Visible-Infrared Person Re-identification.** 11088-11097
Shafiq Ahmad, Pietro Morerio, Alessio Del Bue:
- Person Re-Identification without Identification via Event Anonymization.** 11098-11107
Gabriele Trivigno, Gabriele Moreno Berton, Juan Aragon, Barbara Caputo, Carlo Masone :
- Divide&Classify: Fine-Grained Classification for City-Wide Visual Place Recognition.** 11108-11118
Albert Mohwald, Tomás Jenícek, Ondrej Chum:
- Dark Side Augmentation: Generating Diverse Night Examples for Metric Learning.** 11119-11129
Peiyan Guan, Renjing Pei, Bin Shao, Jianzhuang Liu, Weimian Li, Jiaxi Gu, Hang Xu, Songcen Xu, Youliang Yan, Edmund Y. Lam:
- PIDRo: Parallel Isomeric Attention with Dynamic Routing for Text-Video Retrieval.** 11130-11139
Zhiyin Shao, Xinyu Zhang, Changxing Ding, Jian Wang, Jingdong Wang:
- Unified Pre-training with Pseudo Texts for Text-To-Image Person Re-identification.** 11140-11150
Hao Yu, Xu Cheng, Wei Peng, Weihao Liu, Guoying Zhao:
- Modality Unifying Network for Visible-Infrared Person Re-Identification.** 11151-11161
Peng Xu, Xiatian Zhu:
- DeepChange: A Long-Term Person Re-Identification Benchmark with Clothes Change.** 11162-11171
Ziyang Luo, Pu Zhao, Can Xu, Xiubo Geng, Tao Shen, Chongyang Tao, Jing Ma, Qingwei Lin, Dixin Jiang:
- LexLIP: Lexicon-Bottlenecked Language-Image Pre-Training for Large-Scale Image-Text Sparse Retrieval.** 11172-11183
Jiangming Shi, Yachao Zhang, Xiangbo Yin, Yuan Xie, Zhizhong Zhang, Jianping Fan, Zhongchao Shi, Yanyun Qu:
- Dual Pseudo-Labels Interactive Self-Training for Semi-Supervised Visible-Infrared Person Re-Identification.** 11184-11194
Yifei Zhou, Zilu Li, Abhinav Shrivastava, Hengshuang Zhao, Antonio Torralba, Tai-Peng Tian, Ser-Nam Lim:
- BT²: Backward-compatible Training with Basis Transformation.** 11195-11204
Xinlong Yang, Haixin Wang, Jinan Sun, Shikun Zhang, Chong Chen, Xian-Sheng Hua, Xiao Luo :
- Prototypical Mixing and Retrieval-based Refinement for Label Noise-resistant Image Retrieval.** 11205-11215

-      Zhongyan Zhang, Lei Wang, Luping Zhou, Piotr Koniusz: **Learning Spatial-context-aware Global Visual Feature Representation for Instance Image Retrieval.** 11216-11225
-     Yunquan Zhu, Xinkai Gao, Bo Ke, Ruizhi Qiao, Xing Sun: **Coarse-to-Fine: Learning Compact Discriminative Representation for Single-Stage Image Retrieval.** 11226-11235
-     Xingye Fang, Yang Yang, Ying Fu: **Visible-Infrared Person Re-Identification via Semantic Alignment and Affinity Inference.** 11236-11245
-     Hao Ni, Yuke Li, Lianli Gao, Heng Tao Shen, Jingkuan Song: **Part-Aware Transformer for Generalizable Person Re-identification.** 11246-11255
-     Nikolaos-Antonios Ypsilantis, Kaifeng Chen, Bingyi Cao, Mário Lipovský, Pelin Dogan-Schönberger, Grzegorz Makosa, Boris Bluntschli, Mojtaba Seyedhosseini, Ondrej Chum, André Araújo: **Towards Universal Image Embeddings: A Large-Scale Dataset and Challenge for Generic Image Representations.** 11256-11267
-     Jianfeng Dong, Minsong Zhang, Zheng Zhang, Xianke Chen, Daizong Liu, Xiaoye Qu, Xun Wang, Baolong Liu: **Dual Learning with Dynamic Knowledge Distillation for Partially Relevant Video Retrieval.** 11268-11278
-     Kang Ma, Ying Fu, Dezhi Zheng, Yunjie Peng, Chunshui Cao, Yongzhen Huang: **Fine-grained Unsupervised Domain Adaptation for Gait Recognition.** 11279-11288
-     Anwesan Pal, Sahil Wadhwa, Ayush Jaiswal, Xu Zhang, Yue Wu, Rakesh Chada, Pradeep Natarajan, Henrik I. Christensen : **FashionNTM: Multi-turn Fashion Image Retrieval via Cascaded Memory.** 11289-11300
-     Tianrui Guan, Aswath Muthuselvam, Montana Hoover, Xijun Wang, Jing Liang, Adarsh Jagan Sathyamoorthy, Damon Conover, Dinesh Manocha: **CrossLoc3D: Aerial-Ground Cross-Source 3D Place Recognition.** 11301-11310
-     Yixuan Zhou, Yi Qu, Xing Xu, Hengtao Shen: **ImbSAM: A Closer Look at Sharpness-Aware Minimization in Class-Imbalanced Recognition.** 11311-11321
-     Juwon Seo, Ji-Su Kang, Gyeong-Moon Park: **LFS-GAN: Lifelong Few-Shot Image Generation.** 11322-11332
-     Yuyang Liu, Yang Cong, Dipam Goswami, Xialei Liu, Joost van de Weijer: **Augmented Box Replay: Overcoming Foreground Shift for Incremental Object Detection.** 11333-11343
-     David Brüggemann, Christos Sakaridis , Tim Brödermann, Luc Van Gool: **Contrastive Model Adaptation for Cross-Condition Robustness in Semantic Segmentation.** 11344-11353

-      Yixin Zhang, Zilei Wang, Junjie Li, Jiafan Zhuang, Zihan Lin:
Towards Effective Instance Discrimination Contrastive Loss for Unsupervised Domain Adaptation. 11354-11365
-     Sheng Cheng, Tejas Gokhale, Yezhou Yang:
Adversarial Bayesian Augmentation for Single-Source Domain Generalization. 11366-11376
-     Fan Lyu, Qing Sun, Fanhua Shang, Liang Wan, Wei Feng:
Measuring Asymmetric Gradient Discrepancy in Parallel Continual Learning. 11377-11386
-     Changlong Gao, Chengxu Liu, Yujie Dun, Xueming Qian:
CSDA: Learning Category-Scale Joint Feature for Domain Adaptive Object Detection. 11387-11396
-     Kenneth Borup, Cheng Perng Phoo, Bharath Hariharan:
Distilling from Similar Tasks for Transfer Learning on a Budget. 11397-11407
-     Wonguk Cho, Jinha Park, Taesup Kim:
Complementary Domain Adaptation and Generalization for Unsupervised Continual Domain Shift Learning. 11408-11418
-     Geon Lee, Sanghoon Lee, Dohyung Kim, Younghoon Shin, Yongsang Yoon, Bumsub Ham:
Camera-Driven Representation Learning for Unsupervised Domain Adaptive Person Re-identification. 11419-11428
-     Muhammad Gul Zain Ali Khan, Muhammad Ferjad Naeem, Luc Van Gool, Didier Stricker, Federico Tombari, Muhammad Zeshan Afzal:
Introducing Language Guidance in Prompt-based Continual Learning. 11429-11439
-     Huiwen Xu, U Kang:
Fast and Accurate Transferability Measurement by Evaluating Intra-class Feature Variance. 11440-11448
-     Qiankun Gao, Chen Zhao, Yifan Sun, Teng Xi, Gang Zhang, Bernard Ghanem , Jian Zhang:
A Unified Continual Learning Framework with General Parameter-Efficient Tuning. 11449-11459
-     Nicola K. Dinsdale, Mark Jenkinson, Ana I. L. Namburete:
SFHarmony: Source Free Domain Adaptation for Distributed Neuroimaging Analysis. 11460-11471
-     Vivek Chavan, Paul Koch, Marian Schlüter, Clemens Briese:
Towards Realistic Evaluation of Industrial Continual Learning Scenarios with an Emphasis on Energy Consumption and Computational Footprint. 11472-11484
-     Kaihong Wang, Donghyun Kim, Rogério Feris, Margrit Betke:
CDAC: Cross-domain Attention Consistency in Transformer for Domain Adaptive Semantic Segmentation. 11485-11495

-      Joonhyung Park, Hyunjin Seo, Eunho Yang:
PC-Adapter: Topology-Aware Adapter for Efficient Domain Adaption on Point Clouds with Rectified Pseudo-label. 11496-11506
-     Ji Zhang, Lianli Gao, Xu Luo, Hengtao Shen, Jingkuan Song:
DETA: Denoised Task Adaptation for Few-Shot Learning. 11507-11517
-     Chaoqi Chen, Luyao Tang, Leitian Tao, Hong-Yu Zhou, Yue Huang, Xiaoguang Han, Yizhou Yu:
Activate and Reject: Towards Safe Domain Generalization under Category Shift. 11518-11529
-     Xiran Wang, Jian Zhang, Lei Qi, Yinghuan Shi:
Generalizable Decision Boundaries: Dualistic Meta-Learning for Open Set Domain Generalization. 11530-11539
-     Wenxuan Zhang, Paul Janson, Kai Yi , Ivan Skorokhodov , Mohamed Elhoseiny :
Continual Zero-Shot Learning through Semantically Guided Generative Random Walks. 11540-11551
-     Yuwei Yang, Munawar Hayat, Zhao Jin, Hongyuan Zhu, Yinjie Lei:
Zero-Shot Point Cloud Segmentation by Semantic-Visual Aware Synthesis. 11552-11562
-     Qihao Zhao, Chen Jiang, Wei Hu, Fan Zhang, Jun Liu:
MDCS: More Diverse Experts with Consistency Self-distillation for Long-tailed Recognition. 11563-11574
-     Vimal K. B., Saketh Bachu, Tanmay Garg, Niveditha Lakshmi Narasimhan, Raghavan Konuru, Vineeth N. Balasubramanian:
Building a Winning Team: Selecting Source Model Ensembles using a Submodular Transferability Estimation Approach. 11575-11586
-     Yizhe Xiong, Hui Chen, Zijia Lin, Sicheng Zhao, Guiguang Ding:
Confidence-based Visual Dispersal for Few-shot Unsupervised Domain Adaptation. 11587-11597
-     Miaoyu Li, Yachao Zhang, Xu Ma, Yanyun Qu, Yun Fu:
BEV-DG: Cross-Modal Learning under Bird's-Eye View for Domain Generalization of 3D Semantic Segmentation. 11598-11608
-     Sarinda Samarasinghe, Mamshad Nayeem Rizve, Navid Kardan, Mubarak Shah:
CDFSL-V: Cross-Domain Few-Shot Learning for Videos. 11609-11618
-     Samitha Herath, Basura Fernando, Ehsan Abbasnejad, Munawar Hayat, Shahram Khadivi, Mehrtash Harandi, Hamid Rezatofighi, Gholamreza Haffari:
Energy-based Self-Training and Normalization for Unsupervised Domain Adaptation. 11619-11628
-     Kecheng Zheng, Wei Wu, Ruili Feng, Kai Zhu, Jiawei Liu, Deli Zhao, Zheng-Jun Zha, Wei Chen, Yujun Shen:
Regularized Mask Tuning: Uncovering Hidden Knowledge in Pre-trained Vision-Language Models. 11629-11639

-      Tamasha Malepathirana , Damith A. Senanayake, Saman K. Halgamuge: **NAPA-VQ: Neighborhood Aware Prototype Augmentation with Vector Quantization for Continual Learning.** 11640-11650
-     Zeyi Huang, Andy Zhou, Zijian Lin, Mu Cai, Haohan Wang, Yong Jae Lee: **A Sentence Speaks a Thousand Images: Domain Generalization through Distilling CLIP with Language Guidance.** 11651-11661
-     Yutong Feng, Biao Gong, Jianwen Jiang, Yiliang Lv, Yujun Shen, Deli Zhao, Jingren Zhou: **ViM: Vision Middleware for Unified Downstream Transferring.** 11662-11673
-     Parantak Singh, You Li, Ankur Sikarwar, Weixian Lei, Difei Gao, Morgan B. Talbot, Ying Sun, Mike Zheng Shou, Gabriel Kreiman, Mengmi Zhang: **Learning to Learn: How to Continuously Teach Humans and Machines.** 11674-11685
-     Jinjing Zhu, Yunhao Luo, Xu Zheng, Hao Wang, Lin Wang : **A Good Student is Cooperative and Reliable: CNN-Transformer Collaborative Learning for Semantic Segmentation.** 11686-11696
-     Jun-Yeong Moon, Keon-Hee Park, Jung Uk Kim, Gyeong-Moon Park: **Online Class Incremental Learning on Stochastic Blurry Task Boundary via Mask and Visual Prompt Tuning.** 11697-11707
-     Jiahua Dong, Wenqi Liang, Yang Cong, Gan Sun: **Heterogeneous Forgetting Compensation for Class-Incremental Learning.** 11708-11717
-     Seunghee Koh, Hyounguk Shon, Janghyeon Lee, Hyeong Gwon Hong, Junmo Kim: **Disposable Transfer Learning for Selective Source Task Unlearning.** 11718-11726
-     Byung Hyun Lee, Okchul Jung, Jonghyun Choi, Se Young Chun: **Online Continual Learning on Hierarchical Label Expansion.** 11727-11736
-     Jingyi Zhang, Jiaxing Huang, Xueying Jiang, Shijian Lu: **Black-box Unsupervised Domain Adaptation with Bi-directional Atkinson-Shiffrin Memory.** 11737-11748
-     Yingfan Tao, Jingna Sun, Hao Yang, Li Chen, Xu Wang, Wenming Yang, Daniel K. Du, Min Zheng: **Local and Global Logit Adjustments for Long-Tailed Learning.** 11749-11758
-     Adrian Bulat, Ricardo Guerrero, Brais Martínez, Georgios Tzimiropoulos: **FS-DETR: Few-Shot DEtection TRansformer with prompting and without re-training.** 11759-11768
-     Mingze Gao, Qilong Wang, Zhenyi Lin, Pengfei Zhu, Qinghua Hu, Jingbo Zhou: **Tuning Pre-trained Model via Moment Probing.** 11769-11779
-     Hao Cheng, Siyuan Yang, Joey Tianyi Zhou, Lanqing Guo, Bihan Wen : **Frequency Guidance Matters in Few-Shot Learning.** 11780-11790
-     Haoyu He, Jianfei Cai, Jing Zhang, Dacheng Tao, Bohan Zhuang: **Sensitivity-Aware Visual Parameter-Efficient Fine-Tuning.** 11791-11801

-      Yushu Li, Xun Xu, Yongyi Su, Kui Jia:
On the Robustness of Open-World Test-Time Training: Self-Training with Dynamic Prototype Expansion. 11802-11812
-     Dahuin Jung, Dongyoon Han, Jihwan Bang, Hwanjun Song:
Generating Instance-level Prompts for Rehearsal-free Continual Learning. 11813-11823
-     Zelin Zang, Lei Shang, Senqiao Yang, Fei Wang, Baigui Sun, Xuansong Xie, Stan Z. Li:
Boosting Novel Category Discovery Over Domains with Soft Contrastive Learning and All in One Classifier. 11824-11833
-     Zhiqi Kang, Enrico Fini, Moin Nabi, Elisa Ricci, Kartek Alahari:
A soft nearest-neighbor framework for continual semi-supervised learning. 11834-11843
-     Jiewen Yang, Xinpeng Ding, Ziyang Zheng, Xiaowei Xu, Xiaomeng Li:
GraphEcho: Graph-Driven Unsupervised Domain Adaptation for Echocardiogram Video Segmentation. 11844-11853
-     Dídac Surís, Sachit Menon, Carl Vondrick:
ViperGPT: Visual Inference via Python Execution for Reasoning. 11854-11864
-     Junyang Wang, Yuanhong Xu, Juhua Hu, Ming Yan, Jitao Sang, Qi Qian:
Improved Visual Fine-tuning with Natural Language Supervision. 11865-11875
-     Zihan Lin, Zilei Wang, Yixin Zhang:
Preparing the Future for Continual Semantic Segmentation. 11876-11886
-     Min Zhang, Junkun Yuan, Yue He, Wenbin Li, Zhengyu Chen, Kun Kuang:
MAP: Towards Balanced Generalization of IID and OOD through Model-Agnostic Adapters. 11887-11897
-     Yixuan Pei, Zhiwu Qing, Shiwei Zhang, Xiang Wang, Yingya Zhang, Deli Zhao, Xueming Qian:
Space-time Prompting for Video Class-incremental Learning. 11898-11908
-     Haiyang Yu, Xiaocong Wang, Bin Li, Xiangyang Xue:
Chinese Text Recognition with A Pre-Trained CLIP-Like Model Through Image-IDS Aligning. 11909-11918
-     Samuel Schulter, Vijay Kumar B. G, Yumin Suh, Konstantinos M. Dafnis, Zhixing Zhang, Shiyu Zhao, Dimitris N. Metaxas:
OmniLabel: A Challenging Benchmark for Language-Based Object Detection. 11919-11928
-     Jiapeng Li, Ping Wei, Wenjuan Han, Lifeng Fan:
IntentQA: Context-aware Video Intent Reasoning. 11929-11940
-     Xiaohua Zhai, Basil Mustafa, Alexander Kolesnikov, Lucas Beyer:
Sigmoid Loss for Language Image Pre-Training. 11941-11952
-     Aleksandar Shtedritski, Christian Rupprecht, Andrea Vedaldi:
What does CLIP know about a red circle? Visual prompt engineering for

**VLMs.** 11953-11963

Tan Wang, Kevin Lin, Linjie Li, Chung-Ching Lin, Zhengyuan Yang, Hanwang Zhang, Zicheng Liu, Lijuan Wang:

Equivariant Similarity for Vision-Language Foundation Models. 11964-11974



Zun Wang, Jialu Li, Yicong Hong, Yi Wang, Qi Wu, Mohit Bansal, Stephen Gould, Hao Tan, Yu Qiao:

Scaling Data Generation in Vision-and-Language Navigation. 11975-11986



Shenghan Su, Lin Gu, Yue Yang, Zenghui Zhang, Tatsuya Harada:

Name Your Colour For the Task: Artificially Discover Colour Naming via Colour Quantisation Transformer. 11987-11997



Hongxiang Li, Meng Cao, Xuxin Cheng, Yaowei Li, Zhihong Zhu, Yuexian Zou:

G2L: Semantically Aligned and Uniform Video Grounding via Geodesic and Game Theory. 11998-12008



Yibo Cui, Liang Xie, Yakun Zhang, Meishan Zhang , Ye Yan, Erwei Yin:

Grounded Entity-Landmark Adaptive Pre-training for Vision-and-Language Navigation. 12009-12019



Sarah Ibrahim, Xiaohang Sun, Pichao Wang, Amanmeet Garg, Ashutosh Sanan, Mohamed Omar:

Audio-Enhanced Text-to-Video Retrieval using Text-Conditioned Feature Alignment. 12020-12030



Hexiang Hu, Yi Luan, Yang Chen, Urvashi Khandelwal, Mandar Joshi, Kenton Lee, Kristina Toutanova, Ming-Wei Chang:

Open-domain Visual Entity Recognition: Towards Recognizing Millions of Wikipedia Entities. 12031-12041



Xin Feng, Yifeng Xu, Guangming Lu, Wenjie Pei:

Hierarchical Contrastive Learning for Pattern-Generalizable Image Corruption Detection. 12042-12051



Yuchun Miao, Lefei Zhang, Liangpei Zhang, Dacheng Tao:

DDS2M: Self-Supervised Denoising Diffusion Spatio-Spectral Model for Hyperspectral Image Restoration. 12052-12062



Yun Guo, Xueyao Xiao, Yi Chang, Shumin Deng, Luxin Yan:

From Sky to the Ground: A Large-scale Benchmark and Simple Baseline Towards Real Rain Removal. 12063-12073



Zhiheng Fu, Longguang Wang, Lian Xu, Zhiyong Wang, Hamid Laga, Yulan Guo, Farid Boussaïd, Mohammed Bennamoun:

VAPCNet: Viewpoint-Aware 3D Point Cloud Completion. 12074-12084



Guangyang Wu, Xiaohong Liu, Kunming Luo, Xi Liu, Qingqing Zheng, Shuaicheng Liu, Xinyang Jiang, Guangtao Zhai, Wenyi Wang:

AccFlow: Backward Accumulation for Long-Range Optical Flow. 12085-12094



Chenjie Cao, Yanwei Fu:

Improving Transformer-based Image Matching by Cascaded Capturing Spatially Informative Keypoints. 12095-12105

-      Yunlong Liu, Tao Huang, Weisheng Dong, Fangfang Wu, Xin Li, Guangming Shi: **Low-Light Image Enhancement with Multi-stage Residue Quantization and Brightness-aware Attention.** 12106-12115
-     Yizhong Pan, Xiao Liu, Xiangyu Liao, Yuanzhouhan Cao, Chao Ren: **Random Sub-Samples Generation for Self-Supervised Real Image Denoising.** 12116-12125
-     Wenqi Ouyang, Yi Dong, Xiaoyang Kang, Peiran Ren, Xin Xu, Xuansong Xie: **RSFNet: A White-Box Image Retouching Approach using Region-Specific Color Filters.** 12126-12135
-     Ajay Jaiswal, Xingguang Zhang, Stanley H. Chan, Zhangyang Wang: **Physics-Driven Turbulence Image Restoration with Stochastic Refinement.** 12136-12147
-     Weiran Gou, Ziyao Yi, Yan Xiang, Shaoqing Li, Zibin Liu, Dehui Kong , Ke Xu: **SYENet: A Simple Yet Effective Network for Multiple Low-Level Vision Tasks with Real-time Performance on Mobile Device.** 12148-12161
-     Yeong Il Jang, Keuntek Lee, Gu Yong Park, Seyun Kim, Nam Ik Cho: **Self-supervised Image Denoising with Downsampled Invariance Loss and Conditional Blind-Spot Network.** 12162-12171
-     Wenyu Li, Yan Xu , Yang Yang, Haoran Ji, Yue Lang: **Variational Degeneration to Structural Refinement: A Unified Framework for Superimposed Image Decomposition.** 12172-12182
-     Guandu Liu, Yukang Ding, Mading Li, Ming Sun, Xing Wen, Bin Wang: **Reconstructed Convolution Module Based Look-Up Tables for Efficient Image Super-Resolution.** 12183-12192
-     Jiaying Lin, Rynson W. H. Lau: **Self-supervised Pre-training for Mirror Detection.** 12193-12202
-     Bingna Xu, Yong Guo, Luoqian Jiang, Mianjie Yu, Jian Chen: **Downscaled Representation Matters: Improving Image Rescaling with Collaborative Downscaled Images.** 12203-12213
-     Nisha Varghese, Ashish Kumar, A. N. Rajagopalan: **Self-supervised Monocular Underwater Depth Recovery, Image Restoration, and a Real-sea Video Dataset.** 12214-12224
-     Xiang Ji, Zhixiang Wang, Zhihang Zhong, Yinqiang Zheng: **Rethinking Video Frame Interpolation from Shutter Mode Induced Degradation.** 12225-12234
-     Xiang Ji, Zhixiang Wang, Shin'ichi Satoh, Yinqiang Zheng: **Single Image Deblurring with Row-dependent Blur Magnitude.** 12235-12246
-     Hao Chen, Chenyuan Qu, Yu Zhang, Chen Chen, Jianbo Jiao : **Multi-view Self-supervised Disentanglement for General Image Denoising.** 12247-12257
-     Jungwoo Kim, Min H. Kim: **Joint Demosaicing and Deghosting of Time-Varying Exposures for Single-Shot HDR Imaging.** 12258-12267

-      Xunpeng Yi, Han Xu, Hao Zhang, Linfeng Tang, Jiayi Ma:
Diff-Retinex: Rethinking Low-light Image Enhancement with A Generative Diffusion Model. 12268-12277
-     Zheng Chen, Yulun Zhang, Jinjin Gu, Linghe Kong, Xiaokang Yang, Fisher Yu:
Dual Aggregation Transformer for Image Super-Resolution. 12278-12287
-     Jun-Sang Yoo, Hongjae Lee, Seung-Won Jung:
Video Object Segmentation-aware Video Frame Interpolation. 12288-12299
-     Yunhao Zou, Chenggang Yan, Ying Fu:
RawHDR: High Dynamic Range Image Reconstruction from a Single Raw Image. 12300-12310
-     Jiangxin Dong, Jinshan Pan, Zhongbao Yang, Jinhui Tang:
Multi-scale Residual Low-Pass Filter Network for Image Deblurring. 12311-12320
-     Yuhui Dai, Junkang Zhang, Faming Fang, Guixu Zhang:
Indoor Depth Recovery Based on Deep Unfolding with Non-Local Prior. 12321-12330
-     Hongyang Zhou, Xiaobin Zhu, Jianqing Zhu, Zheng Han, Shi-Xue Zhang, Jingyan Qin, Xu-Cheng Yin:
Learning Correction Filter via Degradation-Adaptive Regression for Blind Single Image Super-Resolution. 12331-12341
-     Zhengyu Liang, Yingqian Wang, Longguang Wang, Jungang Yang, Shilin Zhou, Yulan Guo:
Learning Non-Local Spatial-Angular Correlation for Light Field Image Super-Resolution. 12342-12352
-     Changfeng Yu, Shiming Chen, Yi Chang, Yibing Song, Luxin Yan:
Both Diverse and Realism Matter: Physical Attribute and Style Alignment for Rainy Image Generation. 12353-12363
-     Man Zhou, Jie Huang, Naishan Zheng, Chongyi Li:
Learned Image Reasoning Prior Penetrates Deep Unfolding Network for Panchromatic and Multi-Spectral Image Fusion. 12364-12373
-     Yilin Liu, Jiang Li, Yunkui Pang, Dong Nie, Pew-Thian Yap:
The Devil is in the Upsampling: Architectural Decisions Made Simpler for Denoising with Deep Image Prior. 12374-12383
-     Hao Feng, Wendi Wang, Jiajun Deng, Wengang Zhou, Li Li, Houqiang Li:
SimFIR: A Simple Framework for Fisheye Image Rectification with Self-supervised Representation Learning. 12384-12393
-     Qi Zhu, Man Zhou, Naishan Zheng, Chongyi Li, Jie Huang, Feng Zhao:
Exploring Temporal Frequency Spectrum in Deep Video Deblurring. 12394-12403
-     Yufei Wang, Yi Yu, Wenhan Yang, Lanqing Guo, Lap-Pui Chau, Alex C. Kot, Bihan Wen :
ExposureDiffusion: Learning to Expose for Low-light Image Enhancement. 12404-12414

-      Zinuo Li, Xuhang Chen , Chi-Man Pun, Xiaodong Cun:
High-Resolution Document Shadow Removal via A Large-Scale Real-World Dataset and A Frequency-Aware Shadow Erasing Net. 12415-12424
-     Bin Duan, Ming Zhong, Yan Yan:
Towards Saner Deep Image Registration. 12425-12434
-     Xiaoyu Shi, Zhaoyang Huang, Weikang Bian, Dasong Li, Manyuan Zhang, Ka Chun Cheung, Simon See, Hongwei Qin, Jifeng Dai, Hongsheng Li:
VideoFlow: Exploiting Temporal Cues for Multi-frame Optical Flow Estimation. 12435-12446
-  Lv Tang, Xinfeng Zhang, Gai Zhang, Xiaoqi Ma:
Scene Matters: Model-based Deep Video Compression. 12447-12457
-  Hoonhee Cho, Yuhwan Jeong, Taewoo Kim, Kuk-Jin Yoon:
Non-Coaxial Event-guided Motion Deblurring with Spatial Alignment. 12458-12469
-  Yuanhao Cai, Hao Bian, Jing Lin, Haoqian Wang, Radu Timofte, Yulun Zhang:
Retinexformer: One-stage Retinex-based Transformer for Low-light Image Enhancement. 12470-12479
-  Ao Li , Le Zhang , Yun Liu, Ce Zhu:
Feature Modulation Transformer: Cross-Refinement of Global Representation via High-Frequency Prior for Image Super-Resolution. 12480-12490
-  Dongxu Zhao, Daniel Lichy, Pierre-Nicolas Perrin, Jan-Michael Frahm, Soumyadip Sengupta:
MVPSNet: Fast Generalizable Multi-view Photometric Stereo. 12491-12502
-  Chengxu Liu, Xuan Wang , Shuai Li, Yuzhi Wang, Xueming Qian:
FSI: Frequency and Spatial Interactive Learning for Image Restoration in Under-Display Cameras. 12503-12512
-  Zixiang Zhao, Jiangshe Zhang, Xiang Gu, Chengli Tan, Shuang Xu, Yulun Zhang, Radu Timofte, Luc Van Gool:
Spherical Space Feature Decomposition for Guided Depth Map Super-Resolution. 12513-12524
-  Naishan Zheng, Man Zhou, Yanmeng Dong, Xiangyu Rui, Jie Huang, Chongyi Li, Feng Zhao:
Empowering Low-Light Image Enhancer through Customized Learnable Priors. 12525-12535
-  Ke Xu, Gerhard Petrus Hancke, Rynson W. H. Lau:
Learning Image Harmonization in the Linear Color Space. 12536-12545
-  Binbin Song, Xiangyu Chen, Shuning Xu, Jiantao Zhou:
Under-Display Camera Image Restoration with Scattering Effect. 12546-12555
-  Jiamian Wang, Huan Wang, Yulun Zhang, Yun Fu, Zhiqiang Tao:
Iterative Soft Shrinkage Learning for Efficient Image Super-Resolution. 12556-12565

-      Yuhui Quan, Xin Yao, Hui Ji:
Single Image Defocus Deblurring via Implicit Neural Inverse Kernels. 12566-12576
-     Chunming He, Kai Li, Guoxia Xu, Yulun Zhang, Runze Hu, Zhenhua Guo, Xiu Li:
Degradation-Resistant Unfolding Network for Heterogeneous Image Fusion. 12577-12587
-     Donghwan Seo, Abhijith Punnappurath, Luxi Zhao, Abdelrahman Abdelhamed, SaiKiran Tedla, Sanguk Park, Jihwan Choe, Michael S. Brown:
Graphics2RAW: Mapping Computer Graphics Images to Sensor RAW Images. 12588-12597
-     Haoyuan Wang, Xiaogang Xu, Ke Xu, Rynson W. H. Lau:
Lighting up NeRF via Unsupervised Decomposition and Enhancement. 12598-12607
-     Xin Lin, Chao Ren, Xiao Liu, Jie Huang, Yinjie Lei:
Unsupervised Image Denoising in Real-World Scenarios via Self-Collaboration Parallel Generative Adversarial Branches. 12608-12618
-     Tian Ye, Sixiang Chen, Jinbin Bai, Jun Shi, Chenghao Xue, Jingxia Jiang, Junjie Yin, Erkang Chen, Yun Liu:
Adverse Weather Removal with Codebook Priors. 12619-12630
-     Yuhui Quan, Haoran Huang, Shengfeng He , Ruotao Xu:
Deep Video Demoiréing via Compact Invertible Dyadic Decomposition. 12631-12640
-     Han Yang, Tianyu Wang, Xiaowei Hu, Chi-Wing Fu:
SILT: Shadow-aware Iterative Label Tuning for Learning to Detect Shadows from Noisy Labels. 12641-12652
-     Shangrong Yang, Chunyu Lin, Kang Liao, Yao Zhao:
Innovating Real Fisheye Image Correction with Dual Diffusion Architecture. 12653-12662
-     Jiayu Sun, Ke Xu, Youwei Pang, Lihe Zhang, Huchuan Lu, Gerhard P. Hancke, Rynson W. H. Lau:
Adaptive Illumination Mapping for Shadow Detection in Raw Images. 12663-12672
-     Xiaodong Yang, Zhuang Ma, Zhiyu Ji, Zhe Ren:
GEDepth: Ground Embedding for Monocular Depth Estimation. 12673-12681
-     Aiping Zhang, Wenqi Ren, Yi Liu, Xiaochun Cao:
Lightweight Image Super-Resolution with Superpixel Token Interaction. 12682-12691
-     Siming Zheng, Xin Yuan:
Unfolding Framework with Prior of Convolution-Transformer Mixture and Uncertainty Estimation for Video Snapshot Compressive Imaging. 12692-12703
-     Haechang Lee, Dongwon Park, Wongi Jeong, Kijeong Kim, Hyunwoo Je, Dongil Ryu, Se Young Chun:



Efficient Unified Demosaicing for Bayer and Non-Bayer Patterned Image Sensors. 12704-12713



Haesoo Chung, Nam Ik Cho:

LAN-HDR: Luminance-based Alignment Network for High Dynamic Range Video Reconstruction. 12714-12723



Li Niu, Xing Zhao, Bo Zhang, Liqing Zhang:

Fine-grained Visible Watermark Removal. 12724-12733



Yupeng Zhou, Zhen Li, Chun-Le Guo, Song Bai, Ming-Ming Cheng, Qibin Hou:
SRFormer: Permuted Self-Attention for Single Image Super-Resolution. 12734-12745



Xiang Li, Jiangxin Dong, Jinhui Tang, Jinshan Pan:

DLGSANet: Lightweight Dynamic Local and Global Self-Attention Network for Image Super-Resolution. 12746-12755



Yuwei Qiu, Kaihao Zhang, Chenxi Wang, Wenhan Luo, Hongdong Li, Zhi Jin:
MB-TaylorFormer: Multi-branch Efficient Transformer Expanded by Taylor Formula for Image Dehazing. 12756-12767



Fei Li, Linfeng Zhang, Zikun Liu, Juan Lei, Zhenbo Li:

Multi-Frequency Representation Enhancement with Privilege Information for Video Super-Resolution. 12768-12779



Jongmin Park, Jooyoung Lee, Munchurl Kim:

COMPASS: High-Efficiency Deep Image Compression with Arbitrary-scale Spatial Scalability. 12780-12789



Steven Tel, Zongwei Wu, Yulun Zhang, Barthélémy Heyrman, Cédric Demonceaux, Radu Timofte, Dominique Ginhac:

Alignment-free HDR Deghosting with Semantics Consistent Transformer. 12790-12799



Nikola Zubic, Daniel Gehrig , Mathias Gehrig, Davide Scaramuzza:

From Chaos Comes Order: Ordering Event Representations for Object Recognition and Detection. 12800-12810



Gang Fu, Qing Zhang, Lei Zhu, Chunxia Xiao, Ping Li:

Towards High-Quality Specular Highlight Removal by Leveraging Large-Scale Synthetic Data. 12811-12819



Masakazu Yoshimura, Junji Otsuka, Atsushi Irie, Takeshi Ohashi:

DynamicISP: Dynamically Controlled Image Signal Processor for Image Recognition. 12820-12830



Huiyuan Fu, Wenkai Zheng, Xicong Wang, Jiaxuan Wang, Heng Zhang, Huadong Ma:

Dancing in the Dark: A Benchmark towards General Low-light Video Enhancement. 12831-12840



Sheng Shen, Huanjing Yue, Jingyu Yang:

Dec-Adapter: Exploring Efficient Decoder-Side Adapter for Bridging Screen Content and Natural Image Compression. 12841-12850

-      Zidong Cao, Hao Ai, Yan-Pei Cao, Ying Shan, Xiaohu Qie, Lin Wang : **OmniZoomer: Learning to Move and Zoom in on Sphere at High-Resolution.** 12851-12861
-     Xuanhua He, Keyu Yan, Rui Li, Chengjun Xie, Jie Zhang, Man Zhou: **Pyramid Dual Domain Injection Network for Pan-sharpening.** 12862-12871
-     Shuzhou Yang, Moxuan Ding, Yanmin Wu, Zihan Li, Jian Zhang: **Implicit Neural Representation for Cooperative Low-light Image Enhancement.** 12872-12881
-     Egor I. Ershov, Vasily Tesalin, Ivan Ermakov, Michael S. Brown: **Physically-plausible illumination distribution estimation.** 12882-12890
-     Jun Cheng, Tao Liu, Shan Tan: **Score Priors Guided Deep Variational Inference for Unsupervised Real-World Single Image Denoising.** 12891-12902
-     Eunhye Lee, Jinsu Yoo, Yunjeong Yang, Sungyong Baik, Tae Hyun Kim: **Semantic-Aware Dynamic Parameter for Video Inpainting Transformer.** 12903-12912
-     Miaoyu Li, Ying Fu, Ji Liu, Yulun Zhang: **Pixel Adaptive Deep Unfolding Transformer for Hyperspectral Image Reconstruction.** 12913-12922
-     Yuyan Zhou, Dong Liang, Songcan Chen, Sheng-Jun Huang, Shuo Yang, Chongyi Li: **Improving Lens Flare Removal with General-Purpose Pipeline and Multiple Light Sources Recovery.** 12923-12933
-     Mengyao Li , Liquan Shen, Peng Ye, Guorui Feng, Zheyin Wang: **RFD-ECNet: Extreme Underwater Image Compression with Reference to Feature Dictionary.** 12934-12943
-     Su-Kai Chen, Hung-Lin Yen, Yu-Lun Liu, Min-Hung Chen, Hou-Ning Hu, Wen-Hsiao Peng, Yen-Yu Lin: **Learning Continuous Exposure Value Representations for Single-Image HDR Reconstruction.** 12944-12954
-     Yuning Cui, Wenqi Ren, Xiaochun Cao, Alois Knoll: **Focal Network for Image Restoration.** 12955-12965
-     Weiying Zheng, Cheng Xu, Xuemiao Xu, Wenxi Liu, Shengfeng He : **CIRI: Curricular Inactivation for Residue-aware One-shot Video Inpainting.** 12966-12976
-     Xiaoyu Liu, Ming Liu, Junyi Li, Shuai Liu, Xiaotao Wang, Lei Lei, Wangmeng Zuo: **Beyond Image Borders: Learning Feature Extrapolation for Unbounded Image Composition.** 12977-12986
-     Zhicun Yin, Ming Liu, Xiaoming Li, Hui Yang, Longan Xiao, Wangmeng Zuo: **MetaF2N: Blind Image Super-Resolution by Learning Efficient Model Adaptation from Faces.** 12987-12998
-     Lanqing Guo, Chong Wang, Wenhan Yang, Yufei Wang, Bihan Wen : **Boundary-Aware Divide and Conquer: A Diffusion-based Solution for**

**Unsupervised Shadow Removal.** 12999-13008

Xiaoguang Li, Qing Guo, Rabab Abdelfattah, Di Lin, Wei Feng, Ivor W. Tsang , Song Wang:

Leveraging Inpainting for Single-Image Shadow Removal. 13009-13018

Zeqiang Lai, Chenggang Yan, Ying Fu:

Hybrid Spectral Denoising Transformer with Guided Attention. 13019-13029

SaiKiran Tedla, Beixuan Yang, Michael S. Brown:

Examining Autoexposure for Challenging Scenes. 13030-13039

Wei Shang, Dongwei Ren, Chaoyu Feng, Xiaotao Wang, Lei Lei, Wangmeng Zuo:

Self-supervised Learning to Bring Dual Reversed Rolling Shutter Images Alive. 13040-13048

Bin Xia, Yulun Zhang, Shiyin Wang, Yitong Wang, Xinglong Wu, Yapeng Tian, Wenming Yang, Luc Van Gool:

DiffIR: Efficient Diffusion Model for Image Restoration. 13049-13059

Sixiang Chen, Tian Ye, Jinbin Bai, Erkang Chen, Jun Shi, Lei Zhu:

Sparse Sampling Transformer with Uncertainty-Driven Ranking for Unified Removal of Raindrops and Rain Streaks. 13060-13071

Lin Zhang, Xin Li, Dongliang He, Fu Li, Errui Ding, Zhaoxiang Zhang:

LMR: A Large-Scale Multi-Reference Dataset for Reference-based Super-Resolution. 13072-13081

Yinglong Wang, Zhen Liu, Jianzhuang Liu, Songcen Xu, Shuaicheng Liu:

Low-Light Image Enhancement with Illumination-Aware Gamma Correction and Complete Image Modelling Network. 13082-13091

Qiming Hu, Xiaojie Guo:

Single Image Reflection Separation via Component Synergy. 13092-13101

Fan Zhang, Shaodi You, Yu Li, Ying Fu:

Learning Rain Location Prior for Nighttime Deraining. 13102-13111

Myungsuk Choi, Hana Lee, Hyong-Euk Lee:

Exploring Positional Characteristics of Dual-Pixel Data for Camera Autofocus. 13112-13122

Keunsoo Ko , Chang-Su Kim:

Continuously Masked Transformer for Image Inpainting. 13123-13132

Zixi Tuo, Huan Yang, Jianlong Fu, Yujie Dun, Xueming Qian:

Learning Data-Driven Vector-Quantized Degradation Model for Animation Video Super-Resolution. 13133-13143

Long Sun, Jiangxin Dong, Jinhui Tang, Jinshan Pan:

Spatially-Adaptive Feature Modulation for Efficient Image Super-Resolution. 13144-13153

Yijun Yang, Angelica I. Avilés-Rivero, Huazhu Fu, Ye Liu, Weiming Wang, Lei Zhu:

Video Adverse-Weather-Component Suppression Network via Weather Messenger and Adversarial Backpropagation. 13154-13164

-      Haoyu Chen , Jingjing Ren, Jinjin Gu, Hongtao Wu, Xuequan Lu, Haoming Cai, Lei Zhu:
Snow Removal in Video: A New Dataset and A Novel Method. 13165-13176
-     Xiaoming Zhang , Tianrui Li, Xiaole Zhao :
Boosting Single Image Super-Resolution via Partial Channel Shifting. 13177-13186
-     Pengxu Wei, Yujing Sun, Xingbei Guo, Chang Liu, Guanbin Li, Jie Chen, Xiangyang Ji, Liang Lin:
Towards Real-World Burst Image Super-Resolution: Benchmark and Method. 13187-13196
-     Xin Luo, Yunan Zhu, Shunxin Xu, Dong Liu:
On the Effectiveness of Spectral Discriminators for Perceptual Quality Improvement. 13197-13206
-     Yunshan Qi, Lin Zhu, Yu Zhang, Jia Li:
E²NeRF: Event Enhanced Neural Radiance Fields from Blurry Images. 13208-13218
-     Yunhao Zou, Chenggang Yan, Ying Fu:
Iterative Denoiser and Noise Estimator for Self-Supervised Image Denoising. 13219-13228
-     Xin Jin, Jia-Wen Xiao, Linghao Han, Chunle Guo, Ruixun Zhang, Xiaolei Liu, Chongyi Li:
Lighting Every Darkness in Two Pairs : A Calibration-Free Pipeline for RAW Denoising. 13229-13238
-     Yuhui Quan, Huan Teng, Ruotao Xu, Jun Huang, Hui Ji:
Fingerprinting Deep Image Restoration Models. 13239-13249
-     Yukuan Min, Aming Wu, Cheng Deng:
Environment-Invariant Curriculum Relation Learning for Fine-Grained Scene Graph Generation. 13250-13261
-     Xuan Wei, Zhidan Ran, Xiaobo Lu:
DCPB: Deformable Convolution based on the Poincaré Ball for Top-view Fisheye Cameras. 13262-13271
-     Peng Tu, Xu Xie, Guo Ai, Yuexiang Li, Yawen Huang, Yefeng Zheng:
FemtoDet: An Object Detection Baseline for Energy Versus Performance Tradeoffs. 13272-13281
-     Hemanth Saratchandran, Shin-Fang Chng, Sameera Ramasinghe, Lachlan E. MacDonald, Simon Lucey:
Curvature-Aware Training for Coordinate Networks. 13282-13292
-     Dror Aiger, André Araújo, Simon Lynen:
Yes, we CANN: Constrained Approximate Nearest Neighbors for local feature-based visual localization. 13293-13303
-     Chen Li, Edward G. Jones, Steve Furber:
Unleashing the Potential of Spiking Neural Networks with Dynamic Confidence. 13304-13314

-      Gaku Nakano:
Minimal Solutions to Uncalibrated Two-view Geometry with Known Epipoles. 13315-13324
-     Yilong Chen , Zhixiong Nan, Tao Xiang:
FBLNet: FeedBack Loop Network for Driver Attention Prediction. 13325-13334
-     Aming Wu, Da Chen, Cheng Deng:
Deep Feature Deblurring Diffusion for Detecting Out-of-Distribution Objects. 13335-13345
-     Dawit Mureja Argaw, Joon-Young Lee, Markus Woodson, In So Kweon, Fabian Caba Heilbron:
Long-range Multimodal Pretraining for Movie Understanding. 13346-13357
-     Wenjie Yang, Yiyi Chen, Yan Li, Yanhua Cheng, Xudong Liu, Quan Chen, Han Li:
Cross-view Semantic Alignment for Livestreaming Product Recognition. 13358-13367
-     Mingfei Han, Yali Wang, Zhihui Li, Lina Yao, Xiaojun Chang , Yu Qiao:
HTML: Hybrid Temporal-scale Multimodal Learning Framework for Referring Video Object Segmentation. 13368-13377
-     Ming Wang, Xianda Guo, Beibei Lin, Tian Yang, Zheng Zhu, Lincheng Li, Shunli Zhang, Xin Yu:
DyGait: Exploiting Dynamic Representations for High-performance Gait Recognition. 13378-13387
-     Chaorui Deng, Da Chen , Qi Wu:
Identity-Consistent Aggregation for Video Object Detection. 13388-13398
-     Yuecong Xu, Jianfei Yang, Yunjiao Zhou, Zhenghua Chen, Min Wu, Xiaoli Li:
Augmenting and Aligning Snippets for Few-Shot Video Domain Adaptation. 13399-13410
-     Jiayi Shao, Xiaohan Wang, Ruijie Quan, Junjun Zheng, Jiang Yang, Yi Yang:
Action Sensitivity Learning for Temporal Action Localization. 13411-13423
-     Song Tang, Chuang Li, Pu Zhang, Rongnian Tang:
SwinLSTM: Improving Spatiotemporal Prediction Accuracy using Swin Transformer and LSTM. 13424-13433
-     Lingyi Hong, Wenchao Chen, Zhongying Liu, Wei Zhang, Pinxue Guo, Zhaoyu Chen, Wenqiang Zhang:
LVOS: A Benchmark for Long-term Video Object Segmentation. 13434-13446
-     Bingkun Huang, Zhiyu Zhao, Guozhen Zhang, Yu Qiao , Limin Wang:
MGMAE: Motion Guided Masking for Video Masked Autoencoding. 13447-13458
-     Nicolas Aziere, Sinisa Todorovic:
Markov Game Video Augmentation for Action Segmentation. 13459-13468
-     Théo Ladune, Pierrick Philippe, Félix Henry, Gordon Clare, Thomas Leguay:
COOL-CHIC: Coordinate-based Low Complexity Hierarchical Image Codec. 13469-13476

-      Adrian Bulat, Enrique Sanchez, Brais Martínez, Georgios Tzimiropoulos:
ReGen: A good Generative zero-shot video classifier should be Rewarded. 13477-13487
-     Muhammad Kashif Ali, Dongjin Kim, Tae Hyun Kim:
Task Agnostic Restoration of Natural Video Dynamics. 13488-13498
-     Or Hirschorn, Shai Avidan:
Normalizing Flows for Human Pose Anomaly Detection. 13499-13508
-     Zixuan Zhao, Dongqi Wang, Xu Zhao:
Movement Enhancement toward Multi-Scale Video Feature Representation for Temporal Action Detection. 13509-13518
-     An-Lan Wang, Kun-Yu Lin, Jia-Run Du, Jingke Meng, Wei-Shi Zheng:
Event-Guided Procedure Planning from Instructional Videos with Text Supervision. 13519-13529
-     Sunjae Yoon, Gwanhyeong Koo, Dahyun Kim, Chang D. Yoo:
SCANet: Scene Complexity Aware Network for Weakly-Supervised Video Moment Retrieval. 13530-13540
-     Guanxiong Sun, Chi Wang, Zhaoyu Zhang, Jiankang Deng, Stefanos Zafeiriou, Yang Hua:
Spatio-temporal Prompting Network for Robust Video Feature Extraction. 13541-13551
-     Joseph Fiorese, Ishan Rajendrakumar Dave, Mubarak Shah:
TeD-SPAD: Temporal Distinctiveness for Self-supervised Privacy-preservation for video Anomaly Detection. 13552-13563
-     Yuan Tian, Guo Lu, Guangtao Zhai, Zhiyong Gao:
Non-Semantics Suppressed Mask Learning for Unsupervised Video Semantic Compression. 13564-13576
-     Shen Yan, Xuehan Xiong, Arsha Nagrani, Anurag Arnab, Zhonghao Wang, Weina Ge, David Ross, Cordelia Schmid:
UnLoc: A Unified Framework for Video Localization Tasks. 13577-13587
-     Haodong Duan, Mingze Xu, Bing Shuai, Davide Modolo, Zhuowen Tu, Joseph Tighe, Alessandro Bergamo:
SkeleTR: Towards Skeleton-based Action Recognition in the Wild. 13588-13598
-     Tengda Han, Max Bain, Arsha Nagrani, Gülden Varol, Weidi Xie, Andrew Zisserman:
AutoAD II: The Sequel - Who, When, and What in Movie Audio Description. 13599-13609
-     Chiara Plizzari, Toby Perrett, Barbara Caputo, Dima Damen:
What can a cook in Italy teach a mechanic in India? Action Recognition Generalisation Over Scenarios and Locations. 13610-13620
-     Wayner Barrios, Mattia Soldan , Alberto Mario Ceballos-Arroyo, Fabian Caba Heilbron, Bernard Ghanem :
Localizing Moments in Long Video Via Multimodal Guidance. 13621-13632

-      Di Yang , Yaohui Wang, Antitza Dantcheva, Quan Kong, Lorenzo Garattoni, Gianpiero Francesca, François Brémond:
LAC - Latent Action Composition for Skeleton-based Action Segmentation.
13633-13644
-     Yangyang Xu, Shengfeng He , Kwan-Yee K. Wong, Ping Luo:
RIGID: Recurrent GAN Inversion and Editing of Real Face Videos. 13645-13655
-     Wentao Bao, Lele Chen, Libing Zeng, Zhong Li, Yi Xu, Junsong Yuan, Yu Kong:
Uncertainty-aware State Space Transformer for Egocentric 3D Hand Trajectory Forecasting. 13656-13665
-     Wenhao Wu, Yuxin Song, Zhun Sun, Jingdong Wang, Chang Xu, Wanli Ouyang:
What Can Simple Arithmetic Operations Do for Temporal Modeling? 13666-13676
-     Bo Fang, Wenhao Wu, Chang Liu, Yu Zhou, Yuxin Song, Weiping Wang, Xiangbo Shu, Xiangyang Ji, Jingdong Wang:
UATVR: Uncertainty-Adaptive Text-Video Retrieval. 13677-13687
-     Hanjun Li, Xiujun Shu, Sunan He, Ruizhi Qiao, Wei Wen, Taian Guo, Bei Gan, Xing Sun:
D3G: Exploring Gaussian Prior for Temporal Sentence Grounding with Glance Annotation. 13688-13700
-     Ke Fan, Zechen Bai, Tianjun Xiao, Dominik Zietlow, Max Horn, Zixu Zhao, Carl-Johann Simon-Gabriel, Mike Zheng Shou, Francesco Locatello, Bernt Schiele, Thomas Brox, Zheng Zhang, Yanwei Fu, Tong He:
Unsupervised Open-Vocabulary Object Localization in Videos. 13701-13709
-     Bin Shao, Jianzhuang Liu, Renjing Pei, Songcen Xu, Peng Dai, Juwei Lu, Weimian Li, Youliang Yan:
HiVLP: Hierarchical Interactive Video-Language Pre-Training. 13710-13720
-     Yulin Pan, Xiangteng He, Biao Gong, Yiliang Lv, Yujun Shen, Yuxin Peng, Deli Zhao:
Scanning Only Once: An End-to-end Framework for Fast Temporal Grounding in Long Videos. 13721-13731
-     Syed Talal Wasim, Muhammad Uzair Khattak, Muzammal Naseer, Salman Khan, Mubarak Shah, Fahad Shahbaz Khan:
Video-FocalNets: Spatio-Temporal Focal Modulation for Video Action Recognition. 13732-13743
-     Yasser Abdelaziz Dahou Djilali, Sanath Narayan, Haithem Boussaid, Ebtesam Almazrouei, Mérouane Debbah:
Lip2Vec: Efficient and Robust Visual Speech Recognition via Latent-to-Latent Visual to Audio Representation Mapping. 13744-13755
-     Georg Heigold, Daniel Keysers, Matthias Minderer, Mario Lucic, Alexey A. Gritsenko, Fisher Yu, Alex Bewley, Thomas Kipf:
Video OWL-ViT: Temporally-consistent open-world localization in video. 13756-13765

-      Fida Mohammad Thoker, Hazel Doughty, Cees G. M. Snoek:
Tubelet-Contrastive Self-Supervision for Video-Efficient Generalization.
13766-13777
-     Jiahao Wang, Guo Chen, Yifei Huang, Limin Wang, Tong Lu:
Memory-and-Anticipation Transformer for Online Action Understanding.
13778-13789
-     Borui Jiang, Yang Jin, Zhentao Tan, Yadong Mu:
Video Action Segmentation via Contextually Refined Temporal Keypoints.
13790-13799
-     Jinhyun Jang, Jungin Park, Jin Kim, Hyeongjun Kwon, Kwanghoon Sohn:
Knowing Where to Focus: Event-aware Transformer for Video Grounding.
13800-13810
-     Yingping Liang, Jiaming Liu, Debing Zhang, Ying Fu:
MPI-Flow: Learning Realistic Optical Flow with Multiplane Images. 13811-13822
-     Yicong Li, Junbin Xiao, Chun Feng, Xiang Wang, Tat-Seng Chua:
Discovering Spatio-Temporal Rationales for Video Question Answering.
13823-13832
-     Qiangqiang Wu, Tianyu Yang, Wei Wu, Antoni B. Chan:
Scalable Video Object Segmentation with Simplified Framework. 13833-13843
-     Yikai Wang, Yinpeng Dong, Fuchun Sun, Xiao Yang:
Root Pose Decomposition Towards Generic Non-rigid 3D Reconstruction with Monocular Videos. 13844-13854
-     Chuhan Zhang, Ankush Gupta, Andrew Zisserman:
Helping Hands: An Object-Aware Ego-Centric Video Recognition Model.
13855-13866
-     Yisheng Zhu, Hu Han, Zhengtao Yu, Guangcan Liu:
Modeling the Relative Visual Tempo for Self-supervised Skeleton-based Action Recognition. 13867-13876
-     Xiangtai Li, Haobo Yuan, Wenwei Zhang, Guangliang Cheng, Jiangmiao Pang, Chen Change Loy:
Tube-Link: A Flexible Cross Tube Framework for Universal Video Segmentation. 13877-13887
-     Zhiwu Qing, Shiwei Zhang, Ziyuan Huang, Yingya Zhang, Changxin Gao, Deli Zhao, Nong Sang:
Disentangling Spatial and Temporal Learning for Efficient Image-to-Video Transfer Learning. 13888-13898
-     Guangyi Chen, Xiao Liu, Guangrun Wang, Kun Zhang, Philip H. S. Torr, Xiao-Ping Zhang, Yansong Tang:
Tem-adapter: Adapting Image-Text Pretraining for Video Question Answer.
13899-13909

-      Qiao Wu, Jiaqi Yang, Kun Sun, Chu'ai Zhang, Yanning Zhang, Mathieu Salzmann: **MixCycle: Mixup Assisted Semi-Supervised 3D Single Object Tracking with Cycle Consistency.** 13910-13920
-     Jun Zhou, Kai Chen, Linlin Xu, Qi Dou, Jing Qin: **Deep Fusion Transformer Network with Weighted Vector-Wise Keypoints Voting for Robust 6D Object Pose Estimation.** 13921-13931
-     Jianhui Liu, Yukang Chen, Xiaoqing Ye, Xiaojuan Qi: **IST-Net: Prior-free Category-level Pose Estimation with Implicit Space Transformation.** 13932-13942
-     Shuiwang Li, Yangxiang Yang, Dan Zeng, Xucheng Wang: **Adaptive and Background-Aware Vision Transformer for Real-Time UAV Tracking.** 13943-13954
-     Jiehong Lin, Zewei Wei, Yabin Zhang, Kui Jia: **VI-Net: Boosting Category-level 6D Object Pose Estimation via Learning Decoupled Rotations on the Spherical Representations.** 13955-13965
-     Ding Ma, Xiangqian Wu: **Tracking by Natural Language Specification with Long Short-term Context Decoupling.** 13966-13975
-     Ruyi Lian, Haibin Ling: **CheckerPose: Progressive Dense Keypoint Localization for Object Pose Estimation with Graph Neural Network.** 13976-13987
-     Long Wang, Shen Yan, Jianan Zhen, Yu Liu, Maojun Zhang, Guofeng Zhang, Xiaowei Zhou: **Deep Active Contours for Real-time 6-DoF Object Tracking.** 13988-13998
-     Heng Zhao, Shenxing Wei, Dahu Shi, Wenming Tan, Zheyang Li, Ye Ren, Xing Wei, Yi Yang, Shiliang Pu: **Learning Symmetry-Aware Geometry Correspondences for 6D Object Pose Estimation.** 13999-14008
-     Ruiqi Wang, Xinggang Wang, Te Li, Rong Yang, Minhong Wan, Wen-Yu Liu: **Query6DoF: Learning Sparse Queries as Implicit Shape Prior for Category-Level 6DoF Pose Estimation.** 14009-14018
-     Boyan Wan, Yifei Shi, Kai Xu: **SOCS: Semantically-aware Object Coordinate Space for Category-Level 6D Object Pose Estimation under Large Shape Variations.** 14019-14028
-     Yang Hai, Rui Song, Jiaojiao Li, David Ferstl, Yinlin Hu: **Pseudo Flow Consistency for Self-Supervised 6D Object Pose Estimation.** 14029-14039
-     Denys Rozumnyi, Jirí Matas, Marc Pollefeys, Vittorio Ferrari, Martin R. Oswald : **Tracking by 3D Model Estimation of Unknown Objects in Videos.** 14040-14050
-     Chen Lin, Andrew J. Hanson, Sonya M. Hanson: **Algebraically rigorous quaternion framework for the neural network pose estimation problem.** 14051-14060

-      Fulin Liu, Yinlin Hu, Mathieu Salzmann:
Linear-Covariance Loss for End-to-End Learning of 6D Pose Estimation.
14061-14071
-     Rémi Pautrat, Shaohui Liu, Petr Hruba, Marc Pollefeys, Daniel Barath:
Vanishing Point Estimation in Uncalibrated Images with Prior Gravity Direction. 14072-14081
-     Minhao Li, Zheng Qin, Zhirui Gao, Renjiao Yi, Chenyang Zhu, Yulan Guo, Kai Xu:
2D3D-MATR: 2D-3D Matching Transformer for Detection-free Registration between Images and Point Clouds. 14082-14092
-     Simian Luo, Xuelin Qian, Yanwei Fu, Yinda Zhang, Ying Tai, Zhenyu Zhang, Chengjie Wang, Xiangyang Xue:
Learning Versatile 3D Shape Generation with Improved Auto-regressive Models. 14093-14103
-     Zhaoqi Su , Liangxiao Hu, Siyou Lin, Hongwen Zhang, Shengping Zhang, Justus Thies, Yebin Liu:
CaPhy: Capturing Physical Properties for Animatable Human Avatars.
14104-14114
-     Yaohua Zha, Jinpeng Wang, Tao Dai, Bin Chen, Zhi Wang, Shu-Tao Xia:
Instance-aware Dynamic Prompt Tuning for Pre-trained Point Cloud Models. 14115-14124
-     Jingen Jiang, Mingyang Zhao, Shiqing Xin, Yanchao Yang, Hanxiao Wang, Xiaohong Jia, Dong-Ming Yan:
Structure-Aware Surface Reconstruction via Primitive Assembly. 14125-14134
-     Emmanuel Hartman, Emery Pierson, Martin Bauer, Nicolas Charon, Mohamed Daoudi :
BaRe-ESA: A Riemannian Framework for Unregistered Human Body Shapes.
14135-14145
-     Shan He, Haonan He, Shuo Yang, Xiaoyan Wu, Pengcheng Xia, Bing Yin, Cong Liu, Lirong Dai, Chang Xu:
Speech4Mesh: Speech-Assisted Monocular 3D Facial Reconstruction for Speech-Driven 3D Facial Animation. 14146-14156
-     Jihun Kim, Hyekjun Kwon, Yunseo Yang, Kuk-Jin Yoon:
Learning Point Cloud Completion without Complete Point Clouds: A Pose-Aware Approach. 14157-14167
-     Siyu Ren , Junhui Hou , Xiaodong Chen, Ying He, Wenping Wang:
GeoUDF: Surface Reconstruction from 3D Point Clouds via Geometry-guided Distance Representation. 14168-14178
-     Arjun Mani, Ishaan Preetam Chandratreya, Elliot Creager, Carl Vondrick, Richard S. Zemel:
SurfsUp: Learning Fluid Simulation for Novel Surfaces. 14179-14189
-     Di Liu, Xiang Yu, Meng Ye, Qilong Zhangli, Zhuowei Li, Zhixing Zhang, Dimitris N. Metaxas:



DeFormer: Integrating Transformers with Deformable Models for 3D Shape Abstraction from a Single Image. 14190-14200

Meng Ye, Dong Yang, Mikael Kanski, Leon Axel, Dimitris N. Metaxas:
Neural Deformable Models for 3D Bi-Ventricular Heart Shape Reconstruction and Modeling from 2D Sparse Cardiac Magnetic Resonance Imaging. 14201-14210



George Kiyohiro Nakayama, Mikaela Angelina Uy, Jiahui Huang, Shi-Min Hu, Ke Li, Leonidas J. Guibas:
DiffFacto: Controllable Part-Based 3D Point Cloud Generation with Cross Diffusion. 14211-14221



Baowen Zhang, Jiahe Li, Xiaoming Deng, Yinda Zhang, Cuixia Ma, Hongan Wang:
Self-supervised Learning of Implicit Shape Representation with Dense Correspondence for Deformable Objects. 14222-14232



Tiago Novello, Vinícius da Silva , Guilherme G. Schardong , Luiz Schirmer, Hélio Lopes, Luiz Velho:
Neural Implicit Surface Evolution. 14233-14243



Zisheng Chen, Hongbin Xu, Weitao Chen, Zhipeng Zhou, Haihong Xiao , Baigui Sun, Xuansong Xie, Wenxiong Kang:
PointDC: Unsupervised Semantic Segmentation of 3D Point Clouds via Cross-modal Distillation and Super-Voxel Clustering. 14244-14253



Ziya Erkoç, Fangchang Ma, Qi Shan, Matthias Nießner, Angela Dai:
HyperDiffusion: Generating Implicit Neural Fields with Weight-Space Diffusion. 14254-14264



Ruihai Wu, Chenrui Tie, Yushi Du, Yan Zhao, Hao Dong:
Leveraging SE(3) Equivariance for Learning 3D Geometric Shape Assembly. 14265-14274



Qingyao Shuai, Chi Zhang, Kaizhi Yang, Xuejin Chen:
DPF-Net: Combining Explicit Shape Priors in Deformable Primitive Field for Unsupervised Structural Reconstruction of 3D Objects. 14275-14283



Jie Wang, Lihe Ding, Tingfa Xu, Shaocong Dong, Xinli Xu, Long Bai, Jianan Li:
Sample-adaptive Augmentation for Point Cloud Recognition Against Real-world Corruptions. 14284-14293



Yunbo Tao, Daizong Liu, Pan Zhou, Yulai Xie, Wei Du, Wei Hu:
3DHacker: Spectrum-based Decision Boundary Generation for Hard-label 3D Point Cloud Attack. 14294-14304



Ruikai Cui, Shi Qiu, Saeed Anwar, Jiawei Liu, Chaoyue Xing, Jing Zhang, Nick Barnes:
P2C: Self-Supervised Point Cloud Completion from Single Partial Clouds. 14305-14314



Yidi Shao, Chen Change Loy, Bo Dai:
Towards Multi-Layered 3D Garments Animation. 14315-14324



Ruixiang Jiang, Can Wang, Jingbo Zhang, Menglei Chai, Mingming He, Dongdong Chen, Jing Liao:



AvatarCraft: Transforming Text into Neural Human Avatars with Parameterized Shape and Pose Control. 14325-14336

- Hyeonseop Song, Seokhun Choi, Hoseok Do, Chul Lee, Taehyeong Kim:
Blending-NeRF: Text-Driven Localized Editing in Neural Radiance Fields. 14337-14347

- Suyi Chen, Hao Xu, Ru Li, Guanghui Liu, Chi-Wing Fu, Shuaicheng Liu:
SIRA-PCR: Sim-to-Real Adaptation for 3D Point Cloud Registration. 14348-14359

- Ruowei Wang, Yu Liu, Pei Su, Jianwei Zhang, Qijun Zhao:
3D Semantic Subspace Traverser: Empowering 3D Generative Model with Shape Editing Capability. 14360-14371

- Chen Zhang, Ganzhangqin Yuan, Wenbing Tao:
DMNet: Delaunay Meshing Network for 3D Shape Representation. 14372-14382

- Cheng-Yao Hong, Yu-Ying Chou, Tyng-Luh Liu:
Attention Discriminant Sampling for Point Clouds. 14383-14394

- Juil Koo, Seungwoo Yoo, Minh Hieu Nguyen, Minhyuk Sung:
SALAD: Part-Level Latent Diffusion for 3D Shape Generation and Manipulation. 14395-14405

- Jiaze Sun, Zhixiang Chen, Tae-Kyun Kim:
MAPConNet: Self-supervised 3D Pose Transfer with Mesh and Point Contrastive Learning. 14406-14416

- Xuanyu Yi, Jiajun Deng, Qianru Sun, Xian-Sheng Hua, Joo-Hwee Lim, Hanwang Zhang:
Invariant Training 2D-3D Joint Hard Samples for Few-Shot Point Cloud Recognition. 14417-14428

- Meir Yossef Levi, Guy Gilboa:
EPiC: Ensemble of Partial Point Clouds for Robust Classification. 14429-14438

- Siyou Lin, Boyao Zhou, Zerong Zheng, Hongwen Zhang, Yebin Liu:
Leveraging Intrinsic Properties for Non-Rigid Garment Alignment. 14439-14450

- Mingze Sun, Shiwei Mao, Puhua Jiang, Maks Ovsjanikov, Ruqi Huang:
Spatially and Spectrally Consistent Deep Functional Maps. 14451-14461

- Zhe Zhu , Honghua Chen, Xing He, Weiming Wang, Jing Qin, Mingqiang Wei:
SVDFormer: Complementing Point Cloud via Self-view Augmentation and Self-structure Dual-generator. 14462-14472

- Jiepeng Wang, Congyi Zhang, Peng Wang, Xin Li, Peter J. Cobb, Christian Theobalt, Wenping Wang:
Batch-based Model Registration for Fast 3D Sherd Reconstruction. 14473-14483

- Siming Yan, Zhenpei Yang, Haoxiang Li, Chen Song, Li Guan, Hao Kang, Gang Hua, Qixing Huang:



Implicit Autoencoder for Point-Cloud Self-Supervised Representation Learning. 14484-14496

- [List](#) [Download](#) [Search](#) [Share](#) Ren-Wu Li, Ling-Xiao Zhang, Chunpeng Li, Yu-Kun Lai, Lin Gao:
E3Sym: Leveraging E(3) Invariance for Unsupervised 3D Planar Reflective Symmetry Detection. 14497-14507

- [List](#) [Download](#) [Search](#) [Share](#) Omer Gralnik, Guy Gafni, Ariel Shamir:
Semantify: Simplifying the Control of 3D Morphable Models using CLIP. 14508-14518

- [List](#) [Download](#) [Search](#) [Share](#) Nissim Maruani, Roman Klokov, Maks Ovsjanikov, Pierre Alliez, Mathieu Desbrun:
VoroMesh: Learning Watertight Surface Meshes with Voronoi Diagrams. 14519-14528

- [List](#) [Download](#) [Search](#) [Share](#) Qi Zuo, Yafei Song, Jianfang Li, Lin Liu, Liefeng Bo:
DG3D: Generating High Quality 3D Textured Shapes by Learning to Discriminate Multi-Modal Diffusion-Renderings. 14529-14538

- [List](#) [Download](#) [Search](#) [Share](#) Abril Corona-Figueroa, Sam Bond-Taylor, Neelanjan Bhowmik, Yona Falinie A. Gaus, Toby P. Breckon, Hubert P. H. Shum, Chris G. Willcocks:
Unaligned 2D to 3D Translation with Conditional Vector-Quantized Code Diffusion using Transformers. 14539-14548

- [List](#) [Download](#) [Search](#) [Share](#) Fangzhou Lin, Yun Yue, Songlin Hou, Xuechu Yu, Yajun Xu, Kazunori D. Yamada, Ziming Zhang:
Hyperbolic Chamfer Distance for Point Cloud Completion. 14549-14560

- [List](#) [Download](#) [Search](#) [Share](#) Aryan Mikaeili, Or Perel, Mehdi Safaee, Daniel Cohen-Or, Ali Mahdavi-Amiri:
SKED: Sketch-guided Text-based 3D Editing. 14561-14573

- [List](#) [Download](#) [Search](#) [Share](#) Francesca Babiloni, Matteo Maggioni, Thomas Tanay, Jiankang Deng, Ales Leonardis, Stefanos Zafeiriou:
Adaptive Spiral Layers for Efficient 3D Representation Learning on Meshes. 14574-14585

- [List](#) [Download](#) [Search](#) [Share](#) Manuel Kaufmann, Jie Song, Chen Guo, Kaiyue Shen, Tianjian Jiang, Chengcheng Tang, Juan José Zarate, Otmar Hilliges:
EMDB: The Electromagnetic Database of Global 3D Human Pose and Shape in the Wild. 14586-14597

- [List](#) [Download](#) [Search](#) [Share](#) Yufu Wang, Kostas Daniilidis:
ReFit: Recurrent Fitting Network for 3D Human Recovery. 14598-14608

- [List](#) [Download](#) [Search](#) [Share](#) Wenhao Chai, Zhongyu Jiang, Jenq-Neng Hwang, Gaoang Wang:
Global Adaptation meets Local Generalization: Unsupervised Domain Adaptation for 3D Human Pose Estimation. 14609-14619

- [List](#) [Download](#) [Search](#) [Share](#) Tze Ho Elden Tse, Franziska Mueller, Zhengyang Shen, Danhang Tang, Thabo Beeler, Mingsong Dou, Yinda Zhang, Sasa Petrovic, Hyung Jin Chang, Jonathan Taylor, Bardia Doosti:
Spectral Graphomer: Spectral Graph-based Transformer for Egocentric Two-Hand Reconstruction using Multi-View Color Images. 14620-14631

-      Xiaozheng Zheng, Zhuo Su, Chao Wen, Zhou Xue  Xiaojie Jin: **Realistic Full-Body Tracking from Sparse Observations via Joint-Level Modeling.** 14632-14642
-     Mu Zhou, Lucas Stoffl, Mackenzie Weygandt Mathis, Alexander Mathis: **Rethinking pose estimation in crowds: overcoming the detection information bottleneck and ambiguity.** 14643-14653
-     Yucheng Xing  , Xin Wang  : **HDG-ODE: A Hierarchical Continuous-Time Model for Human Pose Forecasting.** 14654-14666
-     Juntao Jian, Xiuping Liu, Manyi Li, Ruizhen Hu, Jian Liu: **AffordPose: A Large-scale Dataset of Hand-Object Interactions with Affordance-driven Hand Pose.** 14667-14678
-     Mingyi Shi, Sebastian Starke, Yuting Ye, Taku Komura, Jungdam Won: **PhaseMP: Robust 3D Pose Estimation via Phase-conditioned Human Motion Prior.** 14679-14691
-     Kaifeng Zhao, Yan Zhang, Shaofei Wang, Thabo Beeler, Siyu Tang  : **Synthesizing Diverse Human Motions in 3D Indoor Scenes.** 14692-14703
-     Rohan Choudhury, Kris M. Kitani, László A. Jeni: **TEMPO: Efficient Multi-View Pose Estimation, Tracking, and Forecasting.** 14704-14714
-     Wenkang Shan, Zhenhua Liu, Xinfeng Zhang, Zhao Wang, Kai Han, Shanshe Wang, Siwei Ma, Wen Gao: **Diffusion-Based 3D Human Pose Estimation with Multi-Hypothesis Aggregation.** 14715-14725
-     Sungchan Park, Eunyi You, Inhoe Lee, Joonseok Lee: **Towards Robust and Smooth 3D Multi-Person Pose Estimation from Monocular Videos in the Wild.** 14726-14736
-     Shubham Goel, Georgios Pavlakos, Jathushan Rajasegaran, Angjoo Kanazawa, Jitendra Malik: **Humans in 4D: Reconstructing and Tracking Humans with Transformers.** 14737-14748
-     Shih-Yang Su, Timur M. Bagautdinov, Helge Rhodin: **NPC: Neural Point Characters from Video.** 14749-14759
-     Hanyang Kong, Kehong Gong, Dongze Lian, Michael Bi Mi, Xinchao Wang: **Priority-Centric Human Motion Generation in Discrete Latent Space.** 14760-14770
-     Taeksoo Kim, Shunsuke Saito, Hanbyul Joo: **NCHO: Unsupervised Learning for Neural 3D Composition of Humans and Objects.** 14771-14782
-     Hyeongjin Nam, Daniel Sungho Jung, Yeonguk Oh, Kyoung Mu Lee: **Cyclic Test-Time Adaptation on Monocular Video for 3D Human Mesh Reconstruction.** 14783-14793

-      Rongyu Chen, Linlin Yang, Angela Yao:
MHEntropy: Entropy Meets Multiple Hypotheses for Pose and Shape Recovery. 14794-14803
-     Boyuan Jiang, Lei Hu, Shihong Xia:
Probabilistic Triangulation for Uncalibrated Multi-View 3D Human Pose Estimation. 14804-14814
-     Runyang Feng, Yixing Gao, Tze Ho Elden Tse, Xueqing Ma, Hyung Jin Chang:
DiffPose: SpatioTemporal Diffusion Model for Video-Based Human Pose Estimation. 14815-14826
-     Buzhen Huang, Jingyi Ju, Zhihao Li, Yangang Wang:
Reconstructing Groups of People with Hypergraph Relational Reasoning. 14827-14837
-     Yuran Sun, Alan William Dougherty, Zhuoying Zhang, Yi-King Choi, Chuan Wu:
MixSynthFormer: A Transformer Encoder-like Structure with Mixed Synthetic Self-attention for Efficient Human Pose Estimation. 14838-14847
-     Zhiying Leng, Shun-Cheng Wu, Mahdi Saleh, Antonio Montanaro, Hao Yu, Yin Wang, Nassir Navab, Xiaohui Liang , Federico Tombari:
Dynamic Hyperbolic Attention Network for Fine Hand-object Reconstruction. 14848-14858
-     Yiming Zhao, Denys Rozumnyi, Jie Song, Otmar Hilliges, Marc Pollefeys, Martin R. Oswald :
Human from Blur: Human Pose Tracking from Blurry Images. 14859-14869
-     Zijian Dong, Xu Chen, Jinlong Yang, Michael J. Black, Otmar Hilliges, Andreas Geiger:
AG3D: Learning to Generate 3D Avatars from 2D Image Collections. 14870-14881
-     Sirui Xu, Zhengyuan Li, Yu-Xiong Wang, Liang-Yan Gui:
InterDiff: Generating 3D Human-Object Interactions with Physics-Informed Diffusion. 14882-14894
-     ChangHee Yang, Kyeongbo Kong, Sung-Jun Min, Dongyoong Wee, Ho-Deok Jang, Geonho Cha, Suk-Ju Kang:
SEFD: Learning to Distill Complex Pose and Occlusion. 14895-14906
-     Dongkai Wang, Shiliang Zhang:
3D Human Mesh Recovery with Sequentially Global Rotation Estimation. 14907-14916
-     Yingxuan You, Hong Liu, Ti Wang, Wenhao Li, Runwei Ding, Xia Li:
Co-Evolution of Pose and Mesh for 3D Human Body Estimation from Video. 14917-14927
-     Zhisheng Huang, Yujin Chen, Di Kang, Jinlu Zhang, Zhigang Tu:
PHRIT: Parametric Hand Representation with Implicit Template. 14928-14938
-     Kai Zhai, Qiang Nie, Bo Ouyang, Xiang Li, Shanlin Yang:
HopFIR: Hop-wise GraphFormer with Intragroup Joint Refinement for 3D

**Human Pose Estimation.** 14939-14949

Dripta S. Raychaudhuri, Calvin-Khang Ta, Arindam Dutta, Rohit Lal, Amit K. Roy-Chowdhury:

Prior-guided Source-free Domain Adaptation for Human Pose Estimation.
14950-14960

Lu Dai, Liqian Ma, Shenhan Qian, Hao Liu, Ziwei Liu, Hui Xiong:

Cloth2Body: Generating 3D Human Body Mesh from 2D Clothing. 14961-14971

Ginger Delmas, Philippe Weinzaepfel, Francesc Moreno-Noguer, Grégory Rogez:

PoseFix: Correcting 3D Human Poses with Natural Language. 14972-14982

Huan Liu, Qiang Chen, Zichang Tan, Jiang-Jiang Liu, Jian Wang, Xiangbo Su, Xiaolong Li, Kun Yao, Junyu Han, Errui Ding, Yao Zhao, Jingdong Wang:

Group Pose: A Simple Baseline for End-to-End Multi-person Pose Estimation. 14983-14992

Samaneh Azadi, Akbar Shah, Thomas Hayes, Devi Parikh, Sonal Gupta:

Make-An-Animation: Large-Scale Text-conditional 3D Human Motion Generation. 14993-15002

Yuxuan Xue, Bharat Lal Bhatnagar, Riccardo Marin, Nikolaos Sarafianos, Yuanlu Xu, Gerard Pons-Moll, Tony Tung:

NSF: Neural Surface Fields for Human Modeling from Monocular Depth.
15004-15014

Huajin Pi, Sida Peng, Minghui Yang, Xiaowei Zhou, Hujun Bao:

Hierarchical Generation of Human-Object Interactions with Diffusion Probabilistic Models. 15015-15027

Hojun Jang, Minkwan Kim, Jinseok Bae, Young Min Kim:

Dynamic Mesh Recovery from Partial Point Cloud Sequence. 15028-15038

Wentao Zhu, Xiaoxuan Ma, Zhaoyang Liu, Libin Liu, Wayne Wu, Yizhou Wang:

MotionBERT: A Unified Perspective on Learning Human Motion Representations. 15039-15053

Wentian Qu, Zhaopeng Cui, Yinda Zhang, Chenyu Meng, Cuixia Ma, Xiaoming Deng, Hongan Wang:

Novel-view Synthesis and Pose Estimation for Hand-Object Interaction from Sparse Views. 15054-15065

Shujie Zhang, Tianyue Zheng, Zhe Chen, Jingzhi Hu, Abdelwahed Khamis, Jiajun Liu, Jun Luo:

OCHID-Fi: Occlusion-Robust Hand Pose Estimation in 3D via RF-Vision.
15066-15075

Jie Yang, Ailing Zeng, Feng Li, Shilong Liu, Ruimao Zhang, Lei Zhang:

Neural Interactive Keypoint Detection. 15076-15086

Lennart Bramlage, Michelle Karg, Cristóbal Curio:

Plausible Uncertainties for Human Pose Regression. 15087-15096

Zhiyang Dou, Qingxuan Wu, Cheng Lin, Zeyu Cao, Qiangqiang Wu, Weilin Wan, Taku Komura, Wenping Wang:

**TORE: Token Reduction for Efficient Human Mesh Recovery with Transformer.** 15097-15109

- Jinnan Chen, Chen Li, Gim Hee Lee:
Weakly-supervised 3D Pose Transfer with Keypoints. 15110-15119

- Ahmed Abdelreheem , Ivan Skorokhodov , Maks Ovsjanikov, Peter Wonka:
SATR: Zero-Shot Semantic Segmentation of 3D Shapes. 15120-15133

- Hu Xu, Saining Xie, Po-Yao Huang, Licheng Yu, Russell Howes, Gargi Ghosh, Luke Zettlemoyer, Christoph Feichtenhofer:
CiT: Curation in Training for Effective Vision-Language Data. 15134-15143

- Muhammad Uzair Khattak, Syed Talal Wasim, Muzammal Naseer, Salman Khan, Ming-Hsuan Yang, Fahad Shahbaz Khan:
Self-regulating Prompts: Foundational Model Adaptation without Forgetting. 15144-15154

- Effrosyni Mavroudi, Triantafyllos Afouras, Lorenzo Torresani:
Learning to Ground Instructional Articles in Videos through Narrations. 15155-15167

- Shuhei Kurita, Naoki Katsura, Eri Onami:
RefEgo: Referring Expression Comprehension Dataset from First-Person Perception of Ego4D. 15168-15178

- Yiming Zhang, ZeMing Gong, Angel X. Chang:
Multi3DRefer: Grounding Text Description to Multiple 3D Objects. 15179

- Mohammad Mahdi Derakhshani, Enrique Sanchez, Adrian Bulat, Victor Guilherme Turrisi da Costa, Cees G. M. Snoek, Georgios Tzimiropoulos, Brais Martínez:
Bayesian Prompt Learning for Image-Language Model Generalization. 15191-15200

- Arushi Goel, Basura Fernando, Frank Keller, Hakan Bilen :
Who are you referring to? Coreference resolution in image narrations. 15201-15212

- Simon Kornblith, Lala Li, Zirui Wang, Thao Nguyen:
Guiding image captioning models toward more specific captions. 15213-15223

- Jihyung Kil, Soravit Changpinyo, Xi Chen, Hexiang Hu, Sebastian Goodman, Wei-Lun Chao, Radu Soricut:
PreSTU: Pre-Training for Scene-Text Understanding. 15224-15234

- Wang Lin, Tao Jin, Ye Wang, Wenwen Pan, Linjun Li, Xize Cheng, Zhou Zhao:
Exploring Group Video Captioning with Efficient Relational Approximation. 15235-15244

- Eric Slyman, Minsuk Kahng, Stefan Lee:
VLSlice: Interactive Vision-and-Language Slice Discovery. 15245-15255

- Dhruvesh Patel, Hamid Eghbalzadeh, Nitin Kamra, Michael Louis Iuzzolino, Unnat Jain, Ruta Desai:



- Pretrained Language Models as Visual Planners for Human Assistance.**
15256-15268
- Chongyan Chen, Samreen Anjum, Danna Gurari:
VQA Therapy: Exploring Answer Differences by Visually Grounding Answers. 15269-15279
- Cuican Yu, Guansong Lu, Yihan Zeng, Jian Sun, Xiaodan Liang, Huibin Li, Zongben Xu, Songcen Xu, Wei Zhang, Hang Xu:
Towards High-Fidelity Text-Guided 3D Face Generation and Manipulation Using only Images. 15280-15291
- Alberto Baldrati, Lorenzo Agnolucci , Marco Bertini, Alberto Del Bimbo:
Zero-Shot Composed Image Retrieval with Textual Inversion. 15292-15301
- Miaoge Li, Dongsheng Wang, Xinyang Liu, Zequn Zeng, Ruiying Lu, Bo Chen, Mingyuan Zhou:
PatchCT: Aligning Patch Set and Label Set with Conditional Transport for Multi-Label Image Classification. 15302-15312
- Minsu Kim, Jeong Hun Yeo, Jeongsoo Choi, Yong Man Ro:
Lip Reading for Low-resource Languages by Learning and Combining General Speech Knowledge and Language-specific Knowledge. 15313-15325
- Zoey Guo, Yiwen Tang, Ray Zhang, Dong Wang, Zhigang Wang, Bin Zhao, Xuelong Li:
ViewRefer: Grasp the Multi-view Knowledge for 3D Visual Grounding.
15326-15337
- Shubo Liu, Hongsheng Zhang, Yuankai Qi , Peng Wang, Yanning Zhang, Qi Wu:
AerialVLN: Vision-and-Language Navigation for UAVs. 15338-15348
- Matthew Trager, Pramuditha Perera, Luca Zancato, Alessandro Achille, Parminder Bhatia, Stefano Soatto:
Linear Spaces of Meanings: Compositional Structures in Vision-Language Models. 15349-15358
- Qinghao Ye, Guohai Xu, Ming Yan, Haiyang Xu, Qi Qian, Ji Zhang, Fei Huang:
HiTeA: Hierarchical Temporal-Aware Video-Language Pre-training. 15359-15370
- Rishi Hazra, Brian Chen, Akshara Rai, Nitin Kamra, Ruta Desai:
EgoTV: Egocentric Task Verification from Natural Language Task Descriptions. 15371-15383
- Yi-Syuan Chen, Yun-Zhu Song, Cheng Yu Yeo, Bei Liu, Jianlong Fu, Hong-Han Shuai:
SINC: Self-Supervised In-Context Learning for Vision-Language Tasks.
15384-15396
- Yanyuan Qiao, Zheng Yu, Qi Wu:
VLN-PETL: Parameter-Efficient Transfer Learning for Vision-and-Language Navigation. 15397-15406
- Peize Sun, Shoufa Chen, Chenchen Zhu, Fanyi Xiao, Ping Luo, Saining Xie, Zhicheng Yan:

**Going Denser with Open-Vocabulary Part Segmentation.** 15407-15419

Jiajin Tang, Ge Zheng, Sibei Yang:

Temporal Collection and Distribution for Referring Video Object Segmentation. 15420-15430

Huan Li, Ping Wei, Zeyu Ma, Nanning Zheng:

Inverse Compositional Learning for Weakly-supervised Relation Grounding. 15431-15441

Cheng-En Wu, Yu Tian, Haichao Yu, Heng Wang, Pedro Morgado, Yu Hen Hu, Linjie Yang:

Why Is Prompt Tuning for Vision-Language Models Robust to Noisy Labels? 15442-15451

Seungju Han, Jack Hessel, Nouha Dziri, Yejin Choi, Youngjae Yu:

Champagne: Learning Real-world Conversation from Large-Scale Web Videos. 15452-15463

Jiashuo Fan, Yaoyuan Liang, Leyao Liu, Shao-Lun Huang, Lei Zhang:

RCA-NOC: Relative Contrastive Alignment for Novel Object Captioning. 15464-15474

Ximeng Sun, Pengchuan Zhang, Peizhao Zhang, Hardik Shah, Kate Saenko, Xide Xia:

DIME-FM : DIstilling Multimodal and Efficient Foundation Models. 15475-15487

Yassine Ouali, Adrian Bulat, Brais Martínez, Georgios Tzimiropoulos:

Black Box Few-Shot Adaptation for Vision-Language models. 15488-15500

Dongwon Kim, Namyup Kim, Cuiling Lan, Suha Kwak:

Shatter and Gather: Learning Referring Image Segmentation with Text Supervision. 15501-15511

Yaojie Shen, Xin Gu, Kai Xu, Heng Fan, Longyin Wen, Libo Zhang:

Accurate and Fast Compressed Video Captioning. 15512-15521

Heng Zhang, Daqing Liu, Zezhong Lv, Bing Su, Dacheng Tao:

Exploring Temporal Concurrency for Video-Language Representation Learning. 15522-15532

Liliane Momeni, Mathilde Caron, Arsha Nagrani, Andrew Zisserman, Cordelia Schmid:

Verbs in Action: Improving verb understanding in video-language models. 15533-15545

Huijie Yao, Wengang Zhou, Hao Feng, Hezhen Hu, Hao Zhou, Houqiang Li:

Sign Language Translation with Iterative Prototype. 15546-15555

Dahun Kim, Anelia Angelova, Weicheng Kuo:

Contrastive Feature Masking Open-Vocabulary Vision Transformer. 15556-15566

Yuwei Zhang, Chih-Hui Ho, Nuno Vasconcelos:

Toward Unsupervised Realistic Visual Question Answering. 15567-15578

-      Zihan Wang, Xiangyang Li, Jiahao Yang, Yeqi Liu, Shuqiang Jiang:
GridMM: Grid Memory Map for Vision-and-Language Navigation. 15579-15590
-     Le Zhuo, Zhaokai Wang, Baisen Wang, Yue Liao, Chenxi Bao, Stanley Peng, Songhao Han, Aixi Zhang, Fei Fang, Si Liu:
Video Background Music Generation: Dataset, Method and Evaluation. 15591-15601
-     Chaorui Deng, Qi Chen, Pengda Qin, Da Chen, Qi Wu:
Prompt Switch: Efficient CLIP Adaptation for Text-Video Retrieval. 15602-15612
-     Beier Zhu, Yulei Niu, Yucheng Han, Yue Wu, Hanwang Zhang:
Prompt-aligned Gradient for Prompt Tuning. 15613-15623
-     Baoshuo Kan, Teng Wang, Wenpeng Lu, Xiantong Zhen, Weili Guan, Feng Zheng:
Knowledge-Aware Prompt Tuning for Generalizable Vision-Language Models. 15624-15634
-     Zongyang Ma, Ziqi Zhang, Yuxin Chen, Zhongang Qi, Yingmin Luo, Zekun Li, Chunfeng Yuan, Bing Li, Xiaohu Qie, Ying Shan, Weiming Hu:
Order-Prompted Tag Sequence Generation for Video Tagging. 15635-15644
-     Sarah M. Pratt, Ian Covert, Rosanne Liu, Ali Farhadi:
What does a platypus look like? Generating customized prompts for zero-shot image classification. 15645-15655
-     Junhyeong Cho, Gilhyun Nam, Sungyeon Kim, Hunmin Yang, Suha Kwak:
PromptStyler: Prompt-driven Style Generation for Source-free Domain Generalization. 15656-15666
-     Runhui Huang, Jianhua Han, Guansong Lu, Xiaodan Liang, Yihan Zeng, Wei Zhang, Hang Xu:
DiffDis: Empowering Generative Diffusion Model with Cross-Modal Discrimination Capability. 15667-15677
-     Cheng Shi, Sibei Yang:
EdaDet: Open-Vocabulary Object Detection Using Early Dense Alignment. 15678-15688
-     Xize Cheng, Tao Jin, Rongjie Huang, Linjun Li, Wang Lin, Zehan Wang, Ye Wang, Huadai Liu, Aoxiong Yin, Zhou Zhao:
MixSpeech: Cross-Modality Self-Learning with Audio-Visual Stream Mixup for Visual Speech Translation and Recognition. 15689-15699
-     Karsten Roth, Jae-Myung Kim, A. Sophia Koepke, Oriol Vinyals, Cordelia Schmid, Zeynep Akata:
Waffling around for Performance: Visual Classification with Random Words and Broad Concepts. 15700-15711
-     Yanyuan Qiao, Yuankai Qi , Zheng Yu, Jing Liu, Qi Wu:
March in Chat: Interactive Prompting for Remote Embodied Referring Expression. 15712-15721

-      Jaime Spencer, Simon Hadfield, Chris Russell, Richard Bowden:
Kick Back & Relax: Learning to Reconstruct the World by Watching SlowTV.
15722-15733
-     Wuyang Li, Xiaoqing Guo, Yixuan Yuan:
Novel Scenes & Classes: Towards Adaptive Open-set Object Detection.
15734-15744
-     Aditya Ganeshan, R. Kenny Jones, Daniel Ritchie:
Improving Unsupervised Visual Program Inference with Code Rewriting Families. 15745-15755
-     Weilai Xiang, Hongyu Yang, Di Huang, Yunhong Wang:
Denoising Diffusion Autoencoders are Unified Self-supervised Learners.
15756-15766
-     Pengwan Yang, Cees G. M. Snoek, Yuki M. Asano:
Self-Ordering Point Clouds. 15767-15776
-     Sai Saketh Rambhatla, Ishan Misra, Rama Chellappa, Abhinav Shrivastava:
MOST: Multiple Object localization with Self-supervised Transformers for object discovery. 15777-15788
-     Sookwan Han, Hanbyul Joo:
CHORUS: Learning Canonicalized 3D Human-Object Spatial Relations from Unbounded Synthesized Images. 15789-15800
-     Zhaopeng Dou, Zhongdao Wang, Yali Li, Shengjin Wang:
Identity-Seeking Self-Supervised Representation Learning for Generalizable Person Re-identification. 15801-15812
-     Yankai Jiang, Mingze Sun, Heng Guo, Xiaoyu Bai, Ke Yan, Le Lu, Minfeng Xu:
Anatomical Invariance Modeling and Semantic Alignment for Self-supervised Learning in 3D Medical Image Analysis. 15813-15823
-     Zekun Li, Lei Qi, Yinghuan Shi, Yang Gao:
IOMatch: Simplifying Open-Set Semi-Supervised Learning with Joint Inliers and Outliers Utilization. 15824-15833
-     Guan Gui, Zhen Zhao , Lei Qi, Luping Zhou, Lei Wang, Yinghuan Shi:
Enhancing Sample Utilization through Sample Adaptive Augmentation in Semi-Supervised Learning. 15834-15843
-     Manyi Zhang, Xuyang Zhao, Jun Yao, Chun Yuan, Weiran Huang:
When Noisy Labels Meet Long Tail Dilemmas: A Representation Calibration Method. 15844-15854
-     Yifan Yang, Shuhai Zhang, Zixiong Huang, Yubing Zhang, Mingkui Tan:
Cross-Ray Neural Radiance Fields for Novel-view Synthesis from Unconstrained Image Collections. 15855-15865
-     Zhihong Pan, Riccardo Gherardi, Xiufeng Xie, Stephen Huang:
Effective Real Image Editing with Accelerated Iterative Diffusion Inversion.
15866-15875
-     Siming Fan, Jingtian Piao, Chen Qian, Hongsheng Li, Kwan-Yee Lin:
Simulating Fluids in Real-World Still Images. 15876-15885

-      Chenyang Qi, Xiaodong Cun, Yong Zhang, Chenyang Lei, Xintao Wang, Ying Shan, Qifeng Chen:
FateZero: Fusing Attentions for Zero-shot Text-based Video Editing. 15886-15896
-     Yuxiang Wei, Yabo Zhang, Zhilong Ji, Jinfeng Bai, Lei Zhang, Wangmeng Zuo:
ELITE: Encoding Visual Concepts into Textual Embeddings for Customized Text-to-Image Generation. 15897-15907
-     Levon Khachatryan, Andranik Mojsisyan, Vahram Tadevosyan, Roberto Henschel, Zhangyang Wang, Shant Navasardyan, Humphrey Shi:
Text2Video-Zero: Text-to-Image Diffusion Models are Zero-Shot Video Generators. 15908-15918
-     Byungjun Kim, Patrick Kwon, Kwangho Lee, Myunggi Lee, Sookwan Han, Daesik Kim, Hanbyul Joo:
Chupa: Carving 3D Clothed Humans from Skinned Shape Priors using 2D Diffusion Probabilistic Models. 15919-15930
-     Karl Holmquist, Bastian Wandt:
DiffPose: Multi-hypothesis Human Pose Estimation using Diffusion Models. 15931-15941
-     Xuan Ju, Ailing Zeng, Chenchen Zhao, Jianan Wang, Lei Zhang, Qiang Xu:
HumanSD: A Native Skeleton-Guided Diffusion Model for Human Image Generation. 15942-15952
-     Mikihiro Tanaka, Kent Fujiwara:
Role-aware Interaction Generation from Textual Description. 15953-15963
-     Ye Yuan, Jiaming Song, Umar Iqbal, Arash Vahdat, Jan Kautz:
PhysDiff: Physics-Guided Human Motion Diffusion Model. 15964-15975
-     Xiang Guo, Jiadai Sun, Yuchao Dai, Guanying Chen, Xiaoqing Ye, Xiao Tan, Errui Ding, Yumeng Zhang, Jingdong Wang:
Forward Flow for Novel View Synthesis of Dynamic Scenes. 15976-15987
-     Jiachuan Wang, Shimin Di , Lei Chen, Charles Wang Wai Ng:
Noise2Info: Noisy Image to Information of Noise for Self-Supervised Image Denoising. 15988-15997
-     Eyal Gomel, Tal Shaharabany, Lior Wolf:
Box-based Refinement for Weakly Supervised and Unsupervised Localization Tasks. 15998-16008
-     Yijiang Li, Xinjiang Wang, Lihe Yang, Litong Feng, Wayne Zhang, Ying Gao:
Diverse Cotraining Makes Strong Semi-Supervised Segmentor. 16009-16021
-     Yue Fan, Anna Kukleva, Dengxin Dai, Bernt Schiele:
SSB: Simple but Strong Baseline for Boosting Performance of Open-Set Semi-Supervised Learning. 16022-16032
-     Suqin Yuan, Lei Feng, Tongliang Liu:
Late Stopping: Avoiding Confidently Learning from Mislabeled Examples. 16033-16042

- 
- Di Huang, Sida Peng, Tong He, Honghui Yang, Xiaowei Zhou, Wanli Ouyang: **Ponder: Point Cloud Pre-training via Neural Rendering.** 16043-16052
- 
- Kaiyou Song, Shan Zhang, Zimeng Luo, Tong Wang, Jin Xie: **Semantics-Consistent Feature Search for Self-Supervised Visual Representation Learning.** 16053-16062
- 
- Yuwei Yang, Hai Li, Yiran Chen: **Stable and Causal Inference for Discriminative Self-supervised Deep Visual Representations.** 16063-16074
- 
- Yue Duan, Zhen Zhao, Lei Qi, Luping Zhou, Lei Wang, Yinghuan Shi: **Towards Semi-supervised Learning with Non-random Missing Labels.** 16075-16085
- 
- Jing Wu, Jennifer A. Hobbs, Naira Hovakimyan: **Hallucination Improves the Performance of Unsupervised Visual Representation Learning.** 16086-16097
- 
- Mariana-luliana Georgescu, Eduardo Fonseca, Radu Tudor Ionescu, Mario Lucic, Cordelia Schmid, Anurag Arnab: **Audiovisual Masked Autoencoders.** 16098-16108
- 
- Zhengfeng Lai, Noranart Vesdapunt, Ning Zhou, Jun Wu, Cong Phuoc Huynh, Xuelu Li, Kah Kuen Fu, Chen-Nee Chuah: **PADCLIP: Pseudo-labeling with Adaptive Debiasing in CLIP for Unsupervised Domain Adaptation.** 16109-16119
- 
- Fanbin Lu, Xufeng Yao, Chi-Wing Fu, Jiaya Jia: **Removing Anomalies as Noises for Industrial Defect Localization.** 16120-16129
- 
- Aojun Zhou, Yang Li, Zipeng Qin, Jianbo Liu, Junting Pan, Renrui Zhang, Rui Zhao, Peng Gao, Hongsheng Li: **SparseMAE: Sparse Training Meets Masked Autoencoders.** 16130-16140
- 
- Lihe Yang, Zhen Zhao, Lei Qi, Yu Qiao , Yinghuan Shi, Hengshuang Zhao: **Shrinking Class Space for Enhanced Certainty in Semi-Supervised Learning.** 16141-16150
- 
- Chen Liang, Wenguan Wang, Jiaxu Miao, Yi Yang: **Logic-induced Diagnostic Reasoning for Semi-supervised Semantic Segmentation.** 16151-16162
- 
- Chaoqiang Zhao, Matteo Poggi, Fabio Tosi, Lei Zhou, Qiyu Sun, Yang Tang, Stefano Mattoccia : **GasMono: Geometry-Aided Self-Supervised Monocular Depth Estimation for Indoor Scenes.** 16163-16174
- 
- Yao Wei, Yanchao Sun, Ruijie Zheng, Sai Vemprala, Rogerio Bonatti, Shuhang Chen, Ratnesh Madaan, Zhongjie Ba, Ashish Kapoor, Shuang Ma: **Is Imitation All You Need? Generalized Decision-Making with Dual-Phase Training.** 16175-16185
- 
- Aaditya Singh, Kartik Sarangmath, Prithvijit Chattopadhyay, Judy Hoffman: **Benchmarking Low-Shot Robustness to Natural Distribution Shifts.** 16186-

-       16196
Imanol G. Estepa, Ignacio Sarasúa, Bhalaji Nagarajan, Petia Radeva:
All4One: Symbiotic Neighbour Contrastive Learning via Self-Attention and Redundancy Reduction. 16197-16207
-      Yiwen Huang, Yixuan Sun, Chenghang Lai, Qing Xu, Xiaomei Wang, Xuli Shen, Weifeng Ge:
Weakly Supervised Learning of Semantic Correspondence through Cascaded Online Correspondence Refinement. 16208-16217
-      Sha Meng, Dian Shao, Jiacheng Guo, Shan Gao:
Tracking without Label: Unsupervised Multiple Object Tracking via Contrastive Similarity Learning. 16218-16227
-      Vivien Cabannes, Léon Bottou, Yann LeCun, Randall Balestriero:
Active Self-Supervised Learning: A Few Low-Cost Relationships Are All You Need. 16228-16237
-      Chen Wei, Karttikeya Mangalam, Po-Yao Huang, Yanghao Li, Haoqi Fan, Hu Xu, Huiyu Wang, Cihang Xie, Alan L. Yuille, Christoph Feichtenhofer:
Diffusion Models as Masked Autoencoders. 16238-16248
-      Mitchell Keren Taraday, Chaim Baskin:
Enhanced Meta Label Correction for Coping with Label Corruption. 16249-16258
-      Huimin Wu, Chenyang Lei, Xiao Sun, Peng-Shuai Wang, Qifeng Chen, Kwang-Ting Cheng, Stephen Lin, Zhirong Wu:
Randomized Quantization: A Generic Augmentation for Data Agnostic Self-supervised Learning. 16259-16270
-      Long Tian, Jingyi Feng, Xiaoqiang Chai, Wenchao Chen, Liming Wang, Xiyang Liu, Bo Chen:
Prototypes-oriented Transductive Few-shot Learning with Conditional Transport. 16271-16280
-      Yuanyi Zhong, Haoran Tang, Jun-Kun Chen, Yu-Xiong Wang:
Contrastive Learning Relies More on Spatial Inductive Bias Than Supervised Learning: An Empirical Study. 16281-16290
-      Jie Hu, Chen Chen, Liujuan Cao, Shengchuan Zhang, Annan Shu, Guannan Jiang, Rongrong Ji:
Pseudo-label Alignment for Semi-supervised Instance Segmentation. 16291-16301
-      Shuo Li, Yue He, Weiming Zhang, Wei Zhang, Xiao Tan, Junyu Han, Errui Ding, Jingdong Wang:
CFCG: Semi-Supervised Semantic Segmentation via Cross-Fusion and Contour Guidance Supervision. 16302-16312
-      Junqiang Huang, Zichao Guo:
Pixel-Wise Contrastive Distillation. 16313-16323
-      Qiankun Ma, Jiayao Gao, Bo Zhan, Yunpeng Guo, Jiliu Zhou, Yan Wang:
Rethinking Safe Semi-supervised Learning: Transferring the Open-set

**Problem to A Close-set One.** 16324-16333

Jungsoo Lee, Debasmit Das, Jaegul Choo, Sungha Choi:

Towards Open-Set Test-Time Adaptation Utilizing the Wisdom of Crowds in Entropy Minimization. 16334



Jiaming Li, Xiangru Lin, Wei Zhang, Xiao Tan, Yingying Li, Junyu Han, Errui Ding, Jingdong Wang, Guanbin Li:

Gradient-based Sampling for Class Imbalanced Semi-supervised Object Detection. 16344-16354



Zhihao Gu, Liang Liu, Xu Chen, Ran Yi, Jiangning Zhang, Yabiao Wang, Chengjie Wang, Annan Shu, Guannan Jiang, Lizhuang Ma:

Remembering Normality: Memory-guided Knowledge Distillation for Unsupervised Anomaly Detection. 16355-16363



Pan Du, Suyun Zhao, Zisen Sheng, Cuiping Li, Hong Chen:

Semi-Supervised Learning via Weight-aware Distillation under Class Distribution Mismatch. 16364-16374



Sunghyun Park, Seunghan Yang, Jaegul Choo, Sungrack Yun:

Label Shift Adapter for Test-Time Adaptation under Covariate and Label Shifts. 16375-16385



Mingkai Zheng, Shan You, Lang Huang, Chen Luo, Fei Wang, Chen Qian, Chang Xu:

SimMatchV2: Semi-Supervised Learning with Graph Consistency. 16386-16396



JoonHo Lee, Jae Oh Woo, Hankyu Moon, Kwonho Lee:

Unsupervised Accuracy Estimation of Deep Visual Models using Domain-Adaptive Adversarial Perturbation without Source Samples. 16397-16406



Nina Shvetsova, Felix Petersen, Anna Kukleva, Bernt Schiele, Hilde Kuehne:

Learning by Sorting: Self-supervised Learning with Group Ordering Constraints. 16407-16417



Yasar Abbas Ur Rehman, Yan Gao, Pedro Porto Buarque de Gusmão, Mina Alibeigi, Jiajun Shen, Nicholas D. Lane:

L-DAWA: Layer-wise Divergence Aware Weight Aggregation in Federated Self-Supervised Visual Representation Learning. 16418-16427



Peiyan Gu, Chuyu Zhang, Ruijie Xu, Xuming He:

Class-relation Knowledge Distillation for Novel Class Discovery. 16428-16437



Hiroki Nakamura, Masashi Okada, Tadahiro Taniguchi:

Representation Uncertainty in Self-Supervised Learning as Variational Inference. 16438-16447



Ahmed Hatem, Yiming Qian, Yang Wang:

Point-TTA: Test-Time Adaptation for Point Cloud Registration Using Multitask Meta-Auxiliary Learning. 16448-16458



Tim Lebailly, Thomas Stegmüller, Behzad Bozorgtabar, Jean-Philippe Thiran, Tinne Tuytelaars:



Adaptive Similarity Bootstrapping for Self-Distillation based Representation Learning. 16459-16468

Xiaoxiao Sheng, Zhiqiang Shen, Gang Xiao, Longguang Wang, Yulan Guo, Hehe Fan:

Point Contrastive Prediction with Semantic Clustering for Self-Supervised Learning on Point Cloud Videos. 16469-16478

Fangfei Lin, Bing Bai, Yiwen Guo, Hao Chen, Yazhou Ren, Zenglin Xu: **MHCN: A Hyperbolic Neural Network Model for Multi-view Hierarchical Clustering.** 16479-16489

Mohammadreza Salehi, Efstratios Gavves, Cees G. M. Snoek, Yuki M. Asano: **Time Does Tell: Self-Supervised Time-Tuning of Dense Image Representations.** 16490-16501

Marc Botet Colomer, Pier Luigi Dovesi, Theodoros Panagiotakopoulos, Joao Frederico Carvalho, Linus Härenstam-Nielsen, Hossein Azizpour, Hedvig Kjellström, Daniel Cremers, Matteo Poggi: **To Adapt or Not to Adapt? Real-Time Adaptation for Semantic Segmentation.** 16502-16513

SoonCheol Noh, DongEon Jeong, Jee-Hyong Lee: **Simple and Effective Out-of-Distribution Detection via Cosine-based Softmax Loss.** 16514-16523

Takanori Asanomi, Shinnosuke Matsuo, Daiki Suehiro, Ryoma Bise: **MixBag: Bag-Level Data Augmentation for Learning from Label Proportions.** 16524-16533

Zhiqiang Shen, Xiaoxiao Sheng, Hehe Fan, Longguang Wang, Yulan Guo, Qiong Liu, Hao Wen, Xi Zhou: **Masked Spatio-Temporal Structure Prediction for Self-supervised Learning on Point Cloud Videos.** 16534-16543

Xin Wen, Bingchen Zhao, Xiaojuan Qi: **Parametric Classification for Generalized Category Discovery: A Baseline Study.** 16544-16554

Zixu Zhao, Jiaze Wang, Max Horn, Yizhuo Ding, Tong He, Zechen Bai, Dominik Zietlow, Carl-Johann Simon-Gabriel, Bing Shuai, Zhuowen Tu, Thomas Brox, Bernt Schiele, Yanwei Fu, Francesco Locatello, Zheng Zhang, Tianjun Xiao: **Object-Centric Multiple Object Tracking.** 16555-16565

Yan Fang, Feng Zhu, Bowen Cheng, Luoqi Liu, Yao Zhao, Yunchao Wei: **Locating Noise is Halfway Denoising for Semi-Supervised Segmentation.** 16566-16576

Bingchen Zhao, Xin Wen, Kai Han: **Learning Semi-supervised Gaussian Mixture Models for Generalized Category Discovery.** 16577-16587

Dominik A. Kloepfer, Dylan Campbell, João F. Henriques: **LoCUS: Learning Multiscale 3D-consistent Features from Posed Images.** 16588-16598

-      Qi Qian:
Stable Cluster Discrimination for Deep Clustering. 16599-16608
-     Teng Long, Nanne van Noord :
Cross-modal Scalable Hyperbolic Hierarchical Clustering. 16609-16618
-     Shichao Dong, Ruibo Li, Jiacheng Wei, Fayao Liu, Guosheng Lin:
Collaborative Propagation on Multiple Instance Graphs for 3D Instance Segmentation with Single-point Supervision. 16619-16628
-     Rui Qian, Shuangrui Ding, Xian Liu, Dahua Lin:
Semantics Meets Temporal Correspondence: Self-supervised Object-centric Learning in Videos. 16629-16641
-     Hyungmin Kim, Sungho Suh, Daehwan Kim, Daun Jeong , Hansang Cho, Junmo Kim:
Proxy Anchor-based Unsupervised Learning for Continuous Generalized Category Discovery. 16642-16651
-     Daiqing Li, Huan Ling, Amlan Kar, David Acuna, Seung Wook Kim, Karsten Kreis, Antonio Torralba, Sanja Fidler:
DreamTeacher: Pretraining Image Backbones with Deep Generative Models. 16652-16662
-     Muhammad Jehanzeb Mirza, Inkyu Shin, Wei Lin, Andreas Schriebl, Kunyang Sun, Jaesung Choe, Mateusz Kozinski, Horst Possegger, In So Kweon, Kuk-Jin Yoon, Horst Bischof:
MATE: Masked Autoencoders are Online 3D Test-Time Learners. 16663-16672
-     Huaxi Huang, Hui Kang, Sheng Liu, Olivier Salvado, Thierry Rakotoarivelo, Dadong Wang, Tongliang Liu:
PADDLES: Phase-Amplitude Spectrum Disentangled Early Stopping for Learning with Noisy Labels. 16673-16684
-     Chen Li, Xiaoling Hu, Shahira Abousamra, Chao Chen:
Calibrating Uncertainty for Semi-Supervised Crowd Counting. 16685-16695
-     Subhadeep Roy, Shankhanil Mitra, Soma Biswas, Rajiv Soundararajan:
Test Time Adaptation for Blind Image Quality Assessment. 16696-16705
-     Jie Chen, Hua Mao, Wai Lok Woo, Xi Peng:
Deep Multiview Clustering by Contrasting Cluster Assignments. 16706-16715
-     Stefano Zorzi, Friedrich Fraundorfer:
Re: PolyWorld - A Graph Neural Network for Polygonal Scene Parsing. 16716-16725
-     Favyen Bastani, Piper Wolters, Ritwik Gupta, Joe Ferdinando, Aniruddha Kembhavi:
SatlasPretrain: A Large-Scale Dataset for Remote Sensing Image Understanding. 16726-16736
-     Runmin Dong, Lichao Mou, Mengxuan Chen, Weijia Li, Xin-Yi Tong, Shuai Yuan, Lixian Zhang, Juepeng Zheng, Xiao Xiang Zhu, Haohuan Fu:



- Large-Scale Land Cover Mapping with Fine-Grained Classes via Class-Aware Semi-Supervised Semantic Segmentation.** 16737-16747
Yuxuan Li, Qibin Hou, Zhaohui Zheng, Ming-Ming Cheng, Jian Yang, Xiang Li:
- Large Selective Kernel Network for Remote Sensing Object Detection.** 16748-16759
Lei Wang, Min Dai, Jianan He, Jingwei Huang:
- Towards Geospatial Foundation Models via Continual Pretraining.** 16760-16770
Lei Wang, Min Dai, Jianan He, Jingwei Huang:
- Regularized Primitive Graph Learning for Unified Vector Mapping.** 16771-16780
Hengwei Zhao , Xinyu Wang, Jingtao Li, Yanfei Zhong:
- Class Prior-Free Positive-Unlabeled Learning with Taylor Variational Loss for Hyperspectral Remote Sensing Imagery.** 16781-16790
Maximilian Bernhard, Niklas Strauß, Matthias Schubert:
- MapFormer: Boosting Change Detection by Using Pre-change Information.** 16791-16800
Fabian Deuser, Konrad Habel, Norbert Oswald:
- Sample4Geo: Hard Negative Sampling For Cross-View Geo-Localisation.** 16801-16810
Gang Yang, Xiangyong Cao, Wenzhe Xiao, Man Zhou, Aiping Liu, Xun Chen, Deyu Meng:
- PanFlowNet: A Flow-Based Deep Network for Pan-sharpening.** 16811-16821
Yinhe Liu, Sunan Shi, Junjue Wang, Yanfei Zhong:
- Seeing Beyond the Patch: Scale-Adaptive Semantic Segmentation of High-resolution Remote Sensing Imagery based on Reinforcement Learning.** 16822-16832
Lvfang Tao, Wei Gao, Ge Li, Chenhao Zhang:
- AdaNIC: Towards Practical Neural Image Compression via Dynamic Transform Routing.** 16833-16842
Yanyu Li, Ju Hu, Yang Wen, Georgios Evangelidis, Kamyar Salahi, Yanzhi Wang, Sergey Tulyakov, Jian Ren:
- Rethinking Vision Transformers for MobileNet Size and Speed.** 16843-16854
Chensheng Peng, Guangming Wang, Xian Wan Lo, Xinrui Wu, Chenfeng Xu, Masayoshi Tomizuka, Wei Zhan, Hesheng Wang:
- DELFLOW: Dense Efficient Learning of Scene Flow for Large-Scale Point Clouds.** 16855-16864
Matthew Dutson, Yin Li, Mohit Gupta:
- Eventful Transformers: Leveraging Temporal Redundancy in Vision Transformers.** 16865-16877
Man Yao, Jiakui Hu, Guangshe Zhao, Yaoyuan Wang, Ziyang Zhang, Bo Xu, Guoqi Li:
- Inherent Redundancy in Spiking Neural Networks.** 16878-16888

-      Hayoung Yun, Hanjoo Cho:
Achievement-based Training Progress Balancing for Multi-Task Learning.
16889-16898
-     Shuangrui Ding, Peisen Zhao, Xiaopeng Zhang, Rui Qian, Hongkai Xiong, Qi Tian:
Prune Spatio-temporal Tokens by Semantic-aware Temporal Accumulation.
16899-16910
-     Yunqiang Li, Jan C. van Gemert, Torsten Hoefer, Bert Moons, Evangelos Eleftheriou, Bram-Ernst Verhoef:
Differentiable Transportation Pruning. 16911-16921
-     Alberto Ancilotto, Francesco Paissan, Elisabetta Farella:
XiNet: Efficient Neural Networks for tinyML. 16922-16931
-     Natalia Frumkin, Dibakar Gope, Diana Marculescu :
Jumping through Local Minima: Quantization in the Loss Landscape of Vision Transformers. 16932-16942
-     Ian Colbert, Alessandro Pappalardo, Jakoba Petri-Koenig:
A2Q: Accumulator-Aware Quantization with Guaranteed Overflow Avoidance. 16943-16952
-     Rui Chen, Qiyu Wan, Pavana Prakash , Lan Zhang, Xu Yuan, Yanmin Gong, Xin Fu, Miao Pan:
Workie-Talkie: Accelerating Federated Learning by Overlapping Computing and Communications via Contrastive Regularization. 16953-16963
-     Xinlin Li, Bang Liu, Rui Heng Yang, Vanessa Courville, Chao Xing, Vahid Partovi Nia:
DenseShift : Towards Accurate and Efficient Low-Bit Power-of-Two Quantization. 16964-16974
-     Parsa Nooralinejad, Ali Abbasi, Soroush Abbasi Koohpayegani, Kossar Pourahmadi Meibodi, Rana Muhammad Shahroz Khan, Soheil Kolouri, Hamed Pirsiavash:
PRANC: Pseudo RAndom Networks for Compacting deep models. 16975-16985
-     Fartash Faghri, Hadi Pouransari, Sachin Mehta, Mehrdad Farajtabar, Ali Farhadi, Mohammad Rastegari, Oncel Tuzel:
Reinforce Data, Multiply Impact: Improved Model Accuracy and Robustness with Dataset Reinforcement. 16986-16997
-     Thomas Heitzinger, Martin Kampel:
A Fast Unified System for 3D Object Detection and Tracking. 16998-17008
-     Xiao-Ming Wu, Dian Zheng, Zuhao Liu, Wei-Shi Zheng:
Estimator Meets Equilibrium Perspective: A Rectified Straight Through Estimator for Binary Neural Networks Training. 17009-17018
-     Zhikai Li, Qingyi Gu:
I-ViT: Integer-only Quantization for Efficient Vision Transformer Inference.
17019-17029

-      Peijie Dong, Lujun Li, Zimian Wei, Xin Niu, Zhiliang Tian, Hengyue Pan:
EMQ: Evolving Training-free Proxies for Automated Mixed Precision Quantization. 17030-17040
-     Yae Jee Cho, Gauri Joshi, Dimitrios Dimitriadis:
Local or Global: Selective Knowledge Assimilation for Federated Learning with Limited Labels. 17041-17050
-     Ahmad Sajedi, Samir Khaki, Ehsan Amjadian, Lucy Z. Liu, Yuri A. Lawryshyn, Konstantinos N. Plataniotis:
DataDAM: Efficient Dataset Distillation with Attention Matching. 17051-17061
-     Yonatan Dukler, Benjamin Bowman, Alessandro Achille, Aditya Golatkar, Ashwin Swaminathan, Stefano Soatto:
SAFE: Machine Unlearning With Shard Graphs. 17062-17072
-     Davide Abati, Haitam Ben Yahia, Markus Nagel, Amirhossein Habibian:
ResQ: Residual Quantization for Video Perception. 17073-17083
-     Sara Shoouri, Mingyu Yang, Zichen Fan, Hun-Seok Kim:
Efficient Computation Sharing for Multi-Task Visual Scene Understanding. 17084-17095
-     Arman Karimian, Roberto Tron:
Essential Matrix Estimation using Convex Relaxations in Orthogonal Space. 17096-17106
-     Cheng Fu, Hanxian Huang, Zixuan Jiang, Yun Ni, Lifeng Nai, Gang Wu, Liqun Cheng, Yanqi Zhou, Sheng Li, Andrew Li, Jishen Zhao:
TripLe: Revisiting Pretrained Model Reuse and Progressive Learning for Efficient Vision Transformer Scaling and Searching. 17107-17117
-     Mengzhao Chen, Wenqi Shao, Peng Xu, Mingbao Lin, Kaipeng Zhang, Fei Chao, Rongrong Ji, Yu Qiao, Ping Luo:
DiffRate : Differentiable Compression Rate for Efficient Vision Transformers. 17118-17128
-     Longrong Yang, Xianpan Zhou, Xuewei Li, Liang Qiao, Zheyang Li, Ziwei Yang, Gaoang Wang, Xi Li:
Bridging Cross-task Protocol Inconsistency for Distillation in Dense Object Detection. 17129-17138
-     Zhendong Yang, Ailing Zeng, Zhe Li, Tianke Zhang, Chun Yuan, Yu Li:
From Knowledge Distillation to Self-Knowledge Distillation: A Unified Approach with Normalized Loss and Customized Soft Labels. 17139-17148
-     Damien Robert , Hugo Raguet, Loïc Landrieu:
Efficient 3D Semantic Segmentation with Superpoint Transformer. 17149-17158
-     Daquan Zhou, Kai Wang, Jianyang Gu , Xiangyu Peng, Dongze Lian, Yifan Zhang, Yang You, Jiashi Feng:
Dataset Quantization. 17159-17170

-      Shibo Jie, Haoqing Wang, Zhi-Hong Deng:
Revisiting the Parameter Efficiency of Adapters from the Perspective of Precision Redundancy. 17171-17180
-     Zhikai Li, Junrui Xiao, Lianwei Yang, Qingyi Gu:
RepQ-ViT: Scale Reparameterization for Post-Training Quantization of Vision Transformers. 17181-17190
-     Ruoyu Feng, Yixin Gao, Xin Jin, Runsen Feng, Zhibo Chen:
Semantically Structured Image Compression via Irregular Group-Based Decoupling. 17191-17201
-     Song Park, Sanghyuk Chun, Byeongho Heo, Wonjae Kim, Sangdoo Yun:
SeiT: Storage-Efficient Vision Training with Tokens Using 1% of Pixel Storage. 17202-17213
-     Mengzhao Chen, Mingbao Lin, Zihang Lin, Yuxin Zhang, Fei Chao, Rongrong Ji:
SMMix: Self-Motivated Image Mixing for Vision Transformers. 17214-17224
-     Penghui Yang, Ming-Kun Xie, Chen-Chen Zong, Lei Feng, Gang Niu, Masashi Sugiyama, Sheng-Jun Huang:
Multi-Label Knowledge Distillation. 17225-17234
-     Yuxi Ren, Jie Wu, Peng Zhang, Manlin Zhang, Xuefeng Xiao, Qian He, Rui Wang, Min Zheng, Xin Pan:
UGC: Unified GAN Compression for Efficient Image-to-Image Translation. 17235-17245
-     Mathias Parger, Chengcheng Tang, Thomas Neff, Christopher D. Twigg, Cem Keskin, Robert Wang, Markus Steinberger:
MotionDeltaCNN: Sparse CNN Inference of Frame Differences in Moving Camera Videos with Spherical Buffers and Padded Convolutions. 17246-17255
-     Han Cai, Junyan Li, Muyan Hu, Chuang Gan, Song Han:
EfficientViT: Lightweight Multi-Scale Attention for High-Resolution Dense Prediction. 17256-17267
-     Yanqing Liu, Jianyang Gu, Kai Wang, Zheng Zhu, Wei Jiang, Yang You:
DREAM: Efficient Dataset Distillation by Representative Matching. 17268-17278
-     Changhun Lee, Hyungjun Kim, Eunhyeok Park, Jae-Joon Kim:
INSTA-BNN: Binary Neural Network with INSTAnce-aware Threshold. 17279-17288
-     Zanlin Ni, Yulin Wang, Jiangwei Yu, Haojun Jiang, Yue Cao, Gao Huang:
Deep Incubation: Training Large Models by Divide-and-Conquering. 17289-17299
-     Tianlong Chen, Xuxi Chen, Xianzhi Du, Abdullah Rashwan, Fan Yang, Huizhong Chen, Zhangyang Wang, Yeqing Li:
AdaMV-MoE: Adaptive Multi-Task Vision Mixture-of-Experts. 17300-17311
-     Ting-An Chen, De-Nian Yang, Ming-Syan Chen:
Overcoming Forgetting Catastrophe in Quantization-Aware Training. 17312-



17321



Guoxuan Xia, Christos-Savvas Bouganis:
Window-Based Early-Exit Cascades for Uncertainty Estimation: When Deep Ensembles are More Efficient than Single Models. 17322-17334



Junyong Choi, Hyeon Cho, Seokhwa Cheung, Wonjun Hwang :
ORC: Network Group-based Knowledge Distillation using Online Role Change. 17335-17344



Yufei Guo, Xiaode Liu, Yuanpei Chen, Liwen Zhang, Weihang Peng, Yuhan Zhang, Xuhui Huang, Zhe Ma:
RMP-Loss: Regularizing Membrane Potential Distribution for Spiking Neural Networks. 17345-17355



Shangqian Gao, Zeyu Zhang, Yanfu Zhang, Feihu Huang, Heng Huang:
Structural Alignment for Network Pruning through Partial Regularization. 17356-17366



Lujun Li, Peijie Dong, Zimian Wei, Ya Yang:
Automated Knowledge Distillation via Monte Carlo Tree Search. 17367-17378



Abdelrahman M. Shaker, Muhammad Maaz, Hanoona Abdul Rasheed, Salman H. Khan, Ming-Hsuan Yang, Fahad Shahbaz Khan:
SwiftFormer: Efficient Additive Attention for Transformer-based Real-time Mobile Vision Applications. 17379-17390



Yuzhang Shang, Bingxin Xu, Gaowen Liu, Ramana Rao Kompella, Yan Yan:
Causal-DFQ: Causality Guided Data-free Network Quantization. 17391-17400



Kaixin Xu, Zhe Wang, Xue Geng, Min Wu, Xiaoli Li, Weisi Lin:
Efficient Joint Optimization of Layer-Adaptive Weight Pruning in Deep Neural Networks. 17401-17411



Song Guo, Lei Zhang, Xiawu Zheng, Yan Wang, Yuchao Li, Fei Chao, Chenglin Wu, Shengchuan Zhang, Rongrong Ji:
Automatic Network Pruning via Hilbert-Schmidt Independence Criterion Lasso under Information Bottleneck Principle. 17412-17423



Jiali Tang, Shuo Chen, Gang Niu, Masashi Sugiyama, Chen Gong:
Distribution Shift Matters for Knowledge Distillation with Webly Collected Images. 17424-17434



Zheng Fang, Xiaoyang Wang, Haocheng Li, Jie Liu, Qiugui Hu, Jimin Xiao:
FastRecon: Few-shot Industrial Anomaly Detection via Fast Feature Reconstruction. 17435-17444



Cheng Han, Qifan Wang, Yiming Cui, Zhiwen Cao, Wenguan Wang, Siyuan Qi, Dongfang Liu:
E²VPT: An Effective and Efficient Approach for Visual Prompt Tuning. 17445-17456



Zunnan Xu, Zhihong Chen, Yong Zhang, Yibing Song, Xiang Wan, Guanbin Li:
Bridging Vision and Language Encoders: Parameter-Efficient Tuning for

-      Referring Image Segmentation. 17457-17466
Sharath Girish, Abhinav Shrivastava, Kamal Gupta:
SHACIRA: Scalable HAsh-grid Compression for Implicit Neural Representations. 17467-17478
-     Wanli Chen, Xufeng Yao, Xinyun Zhang, Bei Yu:
Efficient Deep Space Filling Curve. 17479-17488
-     Xiuyu Li, Yijiang Liu, Long Lian, Huanrui Yang, Zhen Dong, Daniel Kang, Shanghang Zhang, Kurt Keutzer:
Q-Diffusion: Quantizing Diffusion Models. 17489-17499
-     Yumeng Shi, Shihao Bai, Xiuying Wei, Ruihao Gong, Jianlei Yang:
Lossy and Lossless (L^2) Post-training Model Size Compression. 17500-17510
-     Yong Guo, David Stutz, Bernt Schiele:
Robustifying Token Attention for Vision Transformers. 17511-17522
-     Quankai Gao, Qiangeng Xu, Hao Su, Ulrich Neumann, Zexiang Xu:
Strivec: Sparse Tri-Vector Radiance Fields. 17523-17533
-     Francesco Pittaluga, Bingbing Zhuang:
LDP-Feat: Image Features with Local Differential Privacy. 17534-17544
-     Yichen Xie, Chenfeng Xu, Marie-Julie Rakotosaona, Patrick Rim, Federico Tombari, Kurt Keutzer, Masayoshi Tomizuka, Wei Zhan:
SparseFusion: Fusing Multi-Modal Sparse Representations for Multi-Sensor 3D Object Detection. 17545-17556
-     Ankit Dhiman, R. Srinath, Harsh Rangwani, Rishabh Parihar, Lokesh R. Boregowda, Srinath Sridhar, R. Venkatesh Babu:
Strata-NeRF : Neural Radiance Fields for Stratified Scenes. 17557-17568
-     Youngseok Kim, Juyeb Shin, Sanmin Kim, In-Jae Lee, Jun Won Choi, Dongsuk Kum:
CRN: Camera Radar Net for Accurate, Robust, Efficient 3D Perception. 17569-17580
-     Philipp Lindenberger, Paul-Edouard Sarlin, Marc Pollefeys:
LightGlue: Local Feature Matching at Light Speed. 17581-17592
-     Dongwoo Lee, Jeongtaek Oh, Jaesung Rim, Sunghyun Cho, Kyoung Mu Lee:
ExBluRF: Efficient Radiance Fields for Extreme Motion Blurred Images. 17593-17602
-     Tong Wei, Yash Patel, Alexander Shekhovtsov, Jirí Matas, Daniel Barath:
Generalized Differentiable RANSAC. 17603-17614
-     Xinyi Ye, Weiyue Zhao, Tianqi Liu, Zihao Huang, Zhiguo Cao, Xin Li:
Constraining Depth Map Geometry for Multi-View Stereo: A Dual-Depth Approach with Saddle-shaped Depth Cells. 17615-17624
-     Chonghyuk Song, Gengshan Yang, Kangle Deng, Jun-Yan Zhu, Deva Ramanan:
Total-Recon: Deformable Scene Reconstruction for Embodied View Synthesis. 17625-17636

-      Xiangyu Wang, Jingsen Zhu, Qi Ye, Yuchi Huo, Yunlong Ran, Zhihua Zhong, Jiming Chen:
Seal-3D: Interactive Pixel-Level Editing for Neural Radiance Fields. 17637-17647
-     Mingzhi Yuan, Kexue Fu, Zhihao Li, Yucong Meng, Manning Wang:
PointMBF: A Multi-scale Bidirectional Fusion Network for Unsupervised RGB-D Point Cloud Registration. 17648-17659
-     Haiyang Ying, Baowei Jiang, Jinzhi Zhang, Di Xu, Tao Yu, Qionghai Dai, Lu Fang:
PARF: Primitive-Aware Radiance Fusion for Indoor Scene Novel View Synthesis. 17660-17670
-     Guangyan Chen, Meiling Wang, Li Yuan, Yi Yang, Yufeng Yue:
Rethinking Point Cloud Registration as Masking and Reconstruction. 17671-17681
-     Tianchen Zhao, Xuefei Ning, Ke Hong, Zhongyuan Qiu, Pu Lu, Yali Zhao, Linfeng Zhang, Lipu Zhou, Guohao Dai, Huazhong Yang, Yu Wang:
Ada3D : Exploiting the Spatial Redundancy with Adaptive Inference for Efficient 3D Object Detection. 17682-17692
-     Jiaxiang Tang, Hang Zhou, Xiaokang Chen, Tianshu Hu, Errui Ding, Jingdong Wang, Gang Zeng:
Delicate Textured Mesh Recovery from NeRF via Adaptive Surface Refinement. 17693-17703
-     Ziyue Feng, Liang Yang, Pengsheng Guo, Bing Li:
CVRecon: Rethinking 3D Geometric Feature Learning For Neural Reconstruction. 17704-17714
-     Zizhang Li, Xiaoyang Lyu, Yuanyuan Ding, Mengmeng Wang, Yiyi Liao, Yong Liu:
RICO: Regularizing the Unobservable for Indoor Compositional Reconstruction. 17715-17725
-     Dongting Hu, Zhenkai Zhang, Tingbo Hou, Tongliang Liu, Huan Fu, Mingming Gong:
Multiscale Representation for Real-Time Anti-Aliasing Neural Rendering. 17726-17737
-     Jieming Lou, Weide Liu, Zhuo Chen, Fayao Liu, Jun Cheng:
ELFNet: Evidential Local-global Fusion for Stereo Matching. 17738-17747
-     Tuan Duc Ngo, Binh-Son Hua, Khoi Nguyen:
GaPro: Box-Supervised 3D Point Cloud Instance Segmentation Using Gaussian Processes as Pseudo Labelers. 17748-17757
-     Andrea Porfiri Dal Cin, Giacomo Boracchi, Luca Magri:
Multi-body Depth and Camera Pose Estimation from Multiple Views. 17758-17768
-     Ashkan Mirzaei, Tristan Aumentado-Armstrong, Marcus A. Brubaker, Jonathan Kelly, Alex Levinstein, Konstantinos G. Derpanis, Igor Gilitschenski:
Reference-guided Controllable Inpainting of Neural Radiance Fields. 17769-17779

-      Peng Xiang, Xin Wen, Yu-Shen Liu, Hui Zhang, Yi Fang, Zhizhong Han:
Retro-FPN: Retrospective Feature Pyramid Network for Point Cloud Semantic Segmentation. 17780-17792
-     Jihao Liu, Tai Wang, Boxiao Liu, Qihang Zhang, Yu Liu, Hongsheng Li:
GeoMIM: Towards Better 3D Knowledge Transfer via Masked Image Modeling for Multi-view 3D Understanding. 17793-17803
-     Xiaofeng Wang, Zheng Zhu, Wenbo Xu, Yunpeng Zhang, Yi Wei, Xu Chi, Yun Ye, Dalong Du, Jiwen Lu, Xingang Wang:
OpenOccupancy: A Large Scale Benchmark for Surrounding Semantic Occupancy Perception. 17804-17813
-     Nikola Popovic, Danda Pani Paudel , Luc Van Gool:
Surface Normal Clustering for Implicit Representation of Manhattan Scenes. 17814-17824
-     Jaesung Choe, Christopher B. Choy, Jaesik Park, In So Kweon, Anima Anandkumar:
Spacetime Surface Regularization for Neural Dynamic Scene Reconstruction. 17825-17835
-     Junho Kim, Changwoon Choi, Hojun Jang, Young Min Kim:
LDL: Line Distance Functions for Panoramic Localization. 17836-17846
-     Yiheng Zhang, Zhaofan Qiu, Yingwei Pan, Ting Yao, Tao Mei:
Learning Neural Implicit Surfaces with Object-Aware Radiance Fields. 17847-17856
-     Fengrui Tian, Shaoyi Du, Yueqi Duan:
MonoNeRF: Learning a Generalizable Dynamic Radiance Field from Monocular Videos. 17857-17867
-     Ming-Fang Chang, Akash Sharma, Michael Kaess, Simon Lucey:
Neural Radiance Fields with LiDAR Maps. 17868-17877
-     Baixin Xu, Jiarui Zhang, Kwan-Yee Lin, Chen Qian, Ying He:
Deformable Model-Driven Neural Rendering for High-Fidelity 3D Reconstruction of Human Heads Under Low-View Settings. 17878-17888
-     Vitor Guizilini, Igor Vasiljevic, Jiading Fang, Rares Ambrus, Sergey Zakharov, Vincent Sitzmann, Adrien Gaidon:
DeLiRa: Self-Supervised Depth, Light, and Radiance Fields. 17889-17899
-     Jonathan Lorraine, Kevin Xie, Xiaohui Zeng, Chen-Hsuan Lin, Towaki Takikawa, Nicholas Sharp, Tsung-Yi Lin, Ming-Yu Liu, Sanja Fidler, James Lucas:
ATT3D: Amortized Text-to-3D Object Synthesis. 17900-17910
-     Andrea Ramazzina, Mario Bijelic, Stefanie Walz, Alessandro Sanvito, Dominik Scheuble, Felix Heide:
ScatterNeRF: Seeing Through Fog with Physically-Based Inverse Neural Rendering. 17911-17922
-     Philippe Weinzaepfel, Thomas Lucas, Vincent Leroy, Yohann Cabon, Vaibhav Arora, Romain Brégier, Gabriela Csurka, Leonid Antsfeld, Boris Chidlovskii, Jérôme Revaud:



CroCo v2: Improved Cross-view Completion Pre-training for Stereo Matching and Optical Flow. 17923-17934

Shuzhe Wang, Juho Kannala, Marc Pollefeys, Daniel Barath:
Guiding Local Feature Matching with Surface Curvature. 17935-17945

Baao Xie, Bohan Li, Zequn Zhang, Junting Dong, Xin Jin, Jingyu Yang, Wenjun Zeng:
NaviNeRF: NeRF-based 3D Representation Disentanglement by Latent Semantic Navigation. 17946-17956

Le Hui, Linghua Tang, Yuchao Dai, Jin Xie, Jian Yang:
Efficient LiDAR Point Cloud Oversegmentation Network. 17957-17966

Stephan Alaniz, Massimiliano Mancini , Zeynep Akata:
Iterative Superquadric Recomposition of 3D Objects from Multiple Views. 17967-17977

Zeke Xie, Xindi Yang, Yujie Yang, Qi Sun, Yixiang Jiang, Haoran Wang, Yunfeng Cai, Mingming Sun:
S3IM: Stochastic Structural SIMilarity and Its Unreasonable Effectiveness for Neural Fields. 17978-17988

Akshay Mundra, Mallikarjun B. R., Jiayi Wang, Marc Habermann, Christian Theobalt, Mohamed Elgharib:
LiveHand: Real-time and Photorealistic Neural Hand Rendering. 17989-17999

Cheng Sun, Guangyan Cai , Zhengqin Li, Kai Yan, Cheng Zhang, Carl S. Marshall, Jia-Bin Huang, Shuang Zhao, Zhao Dong:
Neural-PBIR Reconstruction of Shape, Material, and Illumination. 18000-18010

Sanmin Kim, Youngseok Kim, In-Jae Lee, Dongsuk Kum:
Predict to Detect: Prediction-guided 3D Object Detection using Sequential Images. 18011-18020

Qi Cai, Yingwei Pan, Ting Yao, Chong-Wah Ngo, Tao Mei:
ObjectFusion: Multi-modal 3D Object Detection with Object-Centric Fusion. 18021-18030

Jules Sanchez, Jean-Emmanuel Deschaud , François Goulette:
Domain generalization of 3D semantic segmentation in autonomous driving. 18031-18041

Tianqi Liu, Xinyi Ye, Weiyue Zhao, Zhiyu Pan, Min Shi, Zhiguo Cao:
When Epipolar Constraint Meets Non-local Operators in Multi-View Stereo. 18042-18051

Zongyi Xu, Bo Yuan, Shanshan Zhao, Qianni Zhang, Xinbo Gao:
Hierarchical Point-based Active Learning for Semi-supervised Point Cloud Semantic Segmentation. 18052-18062

Dave Zhenyu Chen, Ronghang Hu, Xinlei Chen, Matthias Nießner, Angel X. Chang:



UniT3D: A Unified Transformer for 3D Dense Captioning and Visual Grounding. 18063-18073

Frederik Warburg, Ethan Weber, Matthew Tancik, Aleksander Holynski, Angjoo Kanazawa:

Nerfbusters: Removing Ghostly Artifacts from Casually Captured NeRFs. 18074-18084

Fangyin Wei, Thomas A. Funkhouser, Szymon Rusinkiewicz:
Clutter Detection and Removal in 3D Scenes with View-Consistent Inpainting. 18085-18095

Inyong Koo, Inyoung Lee, Se-Ho Kim, Heeseon Kim, Woo-Jin Jeon, Changick Kim:
PG-RCNN: Semantic Surface Point Generation for 3D Object Detection. 18096-18105

Maoteng Zheng, Nengcheng Chen, Junfeng Zhu, Xiaoru Zeng, Huanbin Qiu, Yuyao Jiang, Xingyue Lu, Hao Qu:
Distributed bundle adjustment with block-based sparse matrix compression for super large scale datasets. 18106-18116

Tong Wei, Jirí Matas, Daniel Barath:
Adaptive Reordering Sampler with Neurally Guided MAGSAC. 18117-18127

Linfei Pan, Johannes L. Schönberger, Viktor Larsson, Marc Pollefeys:
Privacy Preserving Localization via Coordinate Permutations. 18128-18137

Jihong Ju, Ching Wei Tseng, Oleksandr Bailo, Georgi Dikov, Mohsen Ghafoorian:
DG-Recon: Depth-Guided Neural 3D Scene Reconstruction. 18138-18148

Muyu Xu, Fangneng Zhan, Jiahui Zhang, Yingchen Yu, Xiaoqin Zhang, Christian Theobalt, Ling Shao, Shijian Lu:
WaveNeRF: Wavelet-based Generalizable Neural Radiance Fields. 18149-18158

Ziming Chen, Yifeng Shi, Jinrang Jia:
TransIFF: An Instance-Level Feature Fusion Framework for Vehicle-Infrastructure Cooperative 3D Detection with Transformers. 18159-18168

Quan Liu, Hongzi Zhu, Yunsong Zhou, Hongyang Li, Shan Chang, Minyi Guo:
Density-invariant Features for Distant Point Cloud Registration. 18169-18179

Zhenwei Zhu , Liying Yang, Ning Li, Chaohao Jiang, Yanyan Liang:
UMIFormer: Mining the Correlations between Similar Tokens for Multi-View 3D Reconstruction. 18180-18189

Shengyu Huang, Zan Gojcic, Zian Wang, Francis Williams, Yoni Kasten, Sanja Fidler, Konrad Schindler, Or Litany:
Neural LiDAR Fields for Novel View Synthesis. 18190-18200

Yuxin Wang, Wayne Wu, Dan Xu:
Learning Unified Decompositional and Compositional NeRF for Editable Novel View Synthesis. 18201-18210

Liying Yang, Zhenwei Zhu , Xuxin Lin, Jian Nong, Yanyan Liang:
Long-Range Grouping Transformer for Multi-View 3D Reconstruction.

-      18211-18221
Junjie Yan, Yingfei Liu, Jianjian Sun, Fan Jia, Shuailin Li, Tiancai Wang, Xiangyu Zhang:
Cross Modal Transformer: Towards Fast and Robust 3D Object Detection.
18222-18232
-     Yadan Luo, Zhuoxiao Chen, Zhen Fang, Zheng Zhang, Mahsa Baktashmotagh, Zi Huang :
Kecor: Kernel Coding Rate Maximization for Active 3D Object Detection.
18233-18244
-     Luoyuan Xu, Tao Guan, Yuesong Wang, Wenkai Liu, Zhaojie Zeng, Junle Wang, Wei Yang:
C2F2NeUS: Cascade Cost Frustum Fusion for High Fidelity and Generalizable Neural Surface Reconstruction. 18245-18255
-     Yanwei Li, Zhiding Yu, Jonah Philion, Anima Anandkumar, Sanja Fidler, Jiaya Jia, Jose Alvarez:
End-to-end 3D Tracking with Decoupled Queries. 18256-18265
-     Zezhou Cheng, Carlos Esteves, Varun Jampani, Abhishek Kar, Subhransu Maji, Ameesh Makadia:
LU-NeRF: Scene and Pose Estimation by Synchronizing Local Unposed NeRFs. 18266-18275
-     Chao Chen, Yu-Shen Liu, Zhizhong Han:
GridPull: Towards Scalability in Learning Implicit Representations from 3D Point Clouds. 18276-18288
-     Weng Fei Low, Gim Hee Lee:
Robust e-NeRF: NeRF from Sparse & Noisy Events under Non-Uniform Motion. 18289-18300
-     Jiaxi Zeng, Chengtang Yao, Lidong Yu, Yuwei Wu, Yunde Jia:
Parameterized Cost Volume for Stereo Matching. 18301-18311
-     Sijia Jiang, Jing Hua, Zhizhong Han:
Coordinate Quantized Neural Implicit Representations for Multi-view Reconstruction. 18312-18323
-     Yiming Xie, Huaiyu Jiang, Georgia Gkioxari, Julian Straub:
Pixel-Aligned Recurrent Queries for Multi-View 3D Object Detection. 18324-18334
-     Wentao Jiang, Hao Xiang, Xinyu Cai, Runsheng Xu, Jiaqi Ma, Yikang Li, Gim Hee Lee, Si Liu:
Optimizing the Placement of Roadside LiDARs for Autonomous Driving. 18335-18344
-     Jiteng Mu, Shen Sang, Nuno Vasconcelos, Xiaolong Wang:
ActorsNeRF: Animatable Few-shot Human Rendering with Generalizable NeRFs. 18345-18355
-     Yifan Zhan, Shohei Nobuhara, Ko Nishino, Yinqiang Zheng:
NeRFrac: Neural Radiance Fields through Refractive Surface. 18356-18366



- Lizhao Liu, Zhuangwei Zhuang, Shangxin Huang, Xunlong Xiao, Tianhang Xiang, Cen Chen, Jingdong Wang, Mingkui Tan:
CPCM: Contextual Point Cloud Modeling for Weakly-supervised Point Cloud Semantic Segmentation. 18367-18376
- Noah Stier, Anurag Ranjan, Alex Colburn, Yajie Yan, Liang Yang, Fangchang Ma, Baptiste Angles:
FineRecon: Depth-aware Feed-forward Network for Detailed 3D Reconstruction. 18377-18386
- Erik Sandström, Yue Li, Luc Van Gool, Martin R. Oswald :
Point-SLAM: Dense Neural Point Cloud-based SLAM. 18387-18398
- Nermin Samet, Oriane Siméoni, Gilles Puy, Georgy Ponomatkin, Renaud Marlet, Vincent Lepetit:
You Never Get a Second Chance To Make a Good First Impression: Seeding Active Learning for 3D Semantic Segmentation. 18399-18411
- Jonas Kulhanek, Torsten Sattler:
Tetra-NeRF: Representing Neural Radiance Fields Using Tetrahedra. 18412-18423
- Luca Bartolomei , Matteo Poggi, Fabio Tosi, Andrea Conti, Stefano Mattoccia :
Active Stereo Without Pattern Projector. 18424-18436
- Jia-Wei Liu, Yan-Pei Cao, Tianyuan Yang, Zhongcong Xu, Jussi Keppo, Ying Shan, Xiaohu Qie, Mike Zheng Shou:
HOSNeRF: Dynamic Human-Object-Scene Neural Radiance Fields from a Single Video. 18437-18448
- Wentao Hu, Jia Zheng , Zixin Zhang, Xiaojun Yuan, Jian Yin, Zihan Zhou:
PlankAssembly: Robust 3D Reconstruction from Three Orthographic Views with Learnt Shape Programs. 18449-18459
- Yushuang Wu, Xiao Li, Jinglu Wang, Xiaoguang Han, Shuguang Cui, Yan Lu:
Efficient View Synthesis with Neural Radiance Distribution Field. 18460-18469
- Jiahao Lu, Jiacheng Deng, Chuxin Wang, Jianfeng He, Tianzhu Zhang:
Query Refinement Transformer for 3D Instance Segmentation. 18470-18480
- Xuesong Chen, Shaoshuai Shi, Chao Zhang, Benjin Zhu, Qiang Wang, Ka Chun Cheung, Simon See, Hongsheng Li:
TrajectoryFormer: 3D Object Tracking Transformer with Predictive Trajectory Hypotheses. 18481-18490
- Ruilong Li, Hang Gao, Matthew Tancik, Angjoo Kanazawa:
NerfAcc: Efficient Sampling Accelerates NeRFs. 18491-18500
- Zongcheng Li, Xiaoxiao Long, Yusen Wang, Tuo Cao, Wenping Wang, Fei Luo, Chunxia Xiao:
NeTO: Neural Reconstruction of Transparent Objects with Self-Occlusion Aware Refraction-Tracing. 18501-18511
- Dave Zhenyu Chen, Yawar Siddiqui, Hsin-Ying Lee, Sergey Tulyakov, Matthias Nießner:

**Text2Tex: Text-driven Texture Synthesis via Diffusion Models.** 18512-18522

Ziqi Wang, Fei Luo, Xiaoxiao Long, Wenxiao Zhang, Chunxia Xiao:
Learning Long-range Information with Dual-Scale Transformers for Indoor Scene Completion. 18523-18533



Haisong Liu, Yao Teng, Tao Lu, Haiguang Wang, Limin Wang:
SparseBEV: High-Performance Sparse 3D Object Detection from Multi-Camera Videos. 18534-18544



Peihao Li, Shaohui Wang, Chen Yang, Bingbing Liu, Weichao Qiu, Haoqian Wang:
NeRF-MS: Neural Radiance Fields with Multi-Sequence. 18545-18554



Ze Yang, Ruibo Li, Evan Ling, Chi Zhang, Yiming Wang, Dezhao Huang, Keng Teck Ma, Minhoe Hur, Guosheng Lin:
Label-Guided Knowledge Distillation for Continual Semantic Segmentation on 2D Images and 3D Point Clouds. 18555-18566



Mohsen Gholami, Mohammad Akbari, Xinglu Wang, Behnam Kamranian, Yong Zhang:
ETran: Energy-Based Transferability Estimation. 18567-18576



Mohammad Fahes, Tuan-Hung Vu, Andrei Bursuc, Patrick Pérez, Raoul de Charette:
PØDA: Prompt-driven Zero-shot Domain Adaptation. 18577-18587



Tao Sun, Cheng Lu, Haibin Ling:
Local Context-Aware Active Domain Adaptation. 18588-18597



Tianlun Zheng, Zhineng Chen, Bingchen Huang, Wei Zhang, Yu-Gang Jiang:
MRN: Multiplexed Routing Network for Incremental Multilingual Text Recognition. 18598-18607



Songhua Liu, Xinchao Wang:
Few-Shot Dataset Distillation via Translative Pre-Training. 18608-18618



Fei Ye, Adrian G. Bors:
Wasserstein Expansible Variational Autoencoder for Discriminative and Generative Continual Learning. 18619-18629



Tian Yu Liu, Stefano Soatto:
Tangent Model Composition for Ensembling and Continual Fine-tuning. 18630-18640



Xu Zheng, Tianbo Pan, Yunhao Luo, Lin Wang:
Look at the Neighbor: Distortion-aware Unsupervised Domain Adaptation for Panoramic Semantic Segmentation. 18641-18652



Lihua Zhou, Mao Ye, Xiatian Zhu, Siying Xiao, Xuqian Fan, Ferrante Neri:
Homeomorphism Alignment for Unsupervised Domain Adaptation. 18653-18664



Songlin Dong, Haoyu Luo, Yuhang He, Xing Wei, Jie Cheng, Yihong Gong:
Knowledge Restore and Transfer for Multi-Label Class-Incremental Learning. 18665-18674

-      Dayuan Jian, Mohammad Rostami:
Unsupervised Domain Adaptation for Training Event-Based Networks Using Contrastive Learning and Uncorrelated Conditioning. 18675-18685
-     Edoardo Cetin, Antonio Carta, Oya Çeliktutan:
A Simple Recipe to Meta-Learn Forward and Backward Transfer. 18686-18696
-     Xiuwei Chen, Xiaobin Chang:
Dynamic Residual Classifier for Class Incremental Learning. 18697-18706
-     Yunqiao Yang, Long-Kai Huang, Ying Wei:
Concept-wise Fine-tuning Matters in Preventing Negative Transfer. 18707-18717
-     Yujie Wei, Jiaxin Ye, Zhizhong Huang, Junping Zhang, Hongming Shan:
Online Prototype Learning for Online Continual Learning. 18718-18728
-     Liqiang He, Wei Wang, Albert Chen, Min Sun, Cheng-Hao Kuo, Sinisa Todorovic:
Bidirectional Alignment for Domain Adaptive Detection with Transformers. 18729-18739
-     Wenxuan Ma, Shuang Li, Jinming Zhang, Chi Harold Liu, Jingxuan Kang, Yulin Wang , Gao Huang:
Borrowing Knowledge From Pre-trained Language Model: A New Data-efficient Visual Learning Paradigm. 18740-18751
-     Yunhao Ge, Yuecheng Li, Shuo Ni, Jiaping Zhao, Ming-Hsuan Yang, Laurent Itti:
CLR: Channel-wise Lightweight Reprogramming for Continual Learning. 18752-18762
-     Haozhi Cao, Yuecong Xu, Jianfei Yang, Pengyu Yin, Shenghai Yuan, Lihua Xie:
Multi-Modal Continual Test-Time Adaptation for 3D Semantic Segmentation. 18763-18773
-     Aristeidis Panos, Yuriko Kobe, Daniel Olmeda Reino, Rahaf Aljundi, Richard E. Turner:
First Session Adaptation: A Strong Replay-Free Baseline for Class-Incremental Learning. 18774-18784
-     Debabrata Pal, Deeptej More, Sai Bhargav, Dipesh Tamboli, Vaneet Aggarwal, Biplab Banerjee:
Domain Adaptive Few-Shot Open-Set Learning. 18785-18794
-     Wenyu Zhang, Li Shen, Chuan-Sheng Foo:
Rethinking the Role of Pre-Trained Networks in Source-Free Domain Adaptation. 18795-18805
-     Hasan Abed Al Kader Hammoud, Ameya Prabhu, Ser-Nam Lim, Philip H. S. Torr, Adel Bibi, Bernard Ghanem :
Rapid Adaptation in Online Continual Learning: Are We Evaluating It Right? 18806-18815
-     Nian Liu, Kepan Nan, Wangbo Zhao, Yuanwei Liu, Xiwen Yao, Salman Khan, Hisham Cholakkal, Rao Muhammad Anwer, Junwei Han, Fahad Shahbaz Khan:



- Multi-grained Temporal Prototype Learning for Few-shot Video Object Segmentation.** 18816-18825
- Nikola Đukic, Alan Lukezic, Vitjan Zavrtanik, Matej Kristan: **A Low-Shot Object Counting Network With Iterative Prototype Adaptation.** 18826-18835
- Zhiqiang Gao, Kaizhu Huang, Rui Zhang, Dawei Liu, Jieming Ma: **Towards Better Robustness against Common Corruptions for Unsupervised Domain Adaptation.** 18836-18847
- Mengxue Kang, Jinpeng Zhang, Jinming Zhang, Xiashuang Wang, Yang Chen, Zhe Ma, Xuhui Huang: **Alleviating Catastrophic Forgetting of Incremental Object Detection via Within-Class and Between-Class Knowledge Distillation.** 18848-18858
- Fusheng Hao, Fengxiang He, Liu Liu, Fuxiang Wu, Dacheng Tao, Jun Cheng: **Class-Aware Patch Embedding Adaptation for Few-Shot Image Classification.** 18859-18869
- Mengmeng Jing, Xiantong Zhen, Jingjing Li, Cees G. M. Snoek: **Order-preserving Consistency Regularization for Domain Adaptation and Generalization.** 18870-18881
- Sunandini Sanyal, Ashish Ramayee Asokan, Suyaansh Bhambri, Akshay R. Kulkarni, Jogendra Nath Kundu, R. Venkatesh Babu: **Domain-Specificity Inducing Transformers for Source-Free Domain Adaptation.** 18882-18891
- Xingyi Yang, Xinchao Wang: **Diffusion Model as Representation Learner.** 18892-18903
- Jinhao Du, Shan Zhang, Qiang Chen, Haifeng Le, Yanpeng Sun, Yao Ni, Jian Wang, Bin He, Jingdong Wang: **σ -Adaptive Decoupled Prototype for Few-Shot Object Detection.** 18904-18914
- Hyundong Jin, Gyeong-Hyeon Kim, Chanho Ahn, Eunwoo Kim: **Growing a Brain with Sparsity-Inducing Generation for Continual Learning.** 18915-18924
- Jian Zhang, Lei Qi, Yinghuan Shi, Yang Gao: **DomainAdaptor: A Novel Approach to Test-time Adaptation.** 18925-18935
- Shaoyu Zhang, Chen Chen, Silong Peng: **Reconciling Object-Level and Global-Level Objectives for Long-Tail Detection.** 18936-18946
- Xueying Jiang, Jiaxing Huang, Sheng Jin, Shijian Lu: **Domain Generalization via Balancing Training Difficulty and Model Capability.** 18947-18957
- Sobhan Hemati, Guojun Zhang, Amir Hossein Estiri, Xi Chen: **Understanding Hessian Alignment for Domain Generalization.** 18958-18968
- Deblina Bhattacharjee, Sabine Süsstrunk, Mathieu Salzmann: **Vision Transformer Adapters for Generalizable Multitask Learning.** 18969-

-      18980 Xinyue Huo, Lingxi Xie, Wengang Zhou, Houqiang Li, Qi Tian:
Focus on Your Target: A Dual Teacher-Student Framework for Domain-adaptive Semantic Segmentation. 18981-18992
-     Zijing Zhao , Sitong Wei, Qingchao Chen, Dehui Li, Yifan Yang, Yuxin Peng, Yang Liu:
Masked Retraining Teacher-Student Framework for Domain Adaptive Object Detection. 18993-19003
-     Lanqing Hu, Meina Kan, Shiguang Shan, Xilin Chen:
DandelionNet: Domain Composition with Instance Adaptive Classification for Domain Generalization. 19004-19013
-     Sanghun Jung, Jungsoo Lee, Nanhee Kim, Amirreza Shaban, Byron Boots, Jaegul Choo:
CAFA: Class-Aware Feature Alignment for Test-Time Adaptation. 19014-19025
-     Anders Christensen, Massimiliano Mancini , A. Sophia Koepke, Ole Winther, Zeynep Akata:
Image-free Classifier Injection for Zero-Shot Classification. 19026-19035
-     Quanzhang Wang, Renzhen Wang, Yichen Wu , Xixi Jia, Deyu Meng:
CBA: Improving Online Continual Learning via Continual Bias Adaptor. 19036-19046
-     Jiang-Tian Zhai, Xialei Liu, Andrew D. Bagdanov, Ke Li, Ming-Ming Cheng:
Masked Autoencoders are Efficient Class Incremental Learners. 19047-19056
-     Jintao Guo, Lei Qi, Yinghuan Shi:
DomainDrop: Suppressing Domain-Sensitive Channels for Domain Generalization. 19057-19067
-     Zangwei Zheng, Mingyuan Ma, Kai Wang, Ziheng Qin, Xiangyu Yue, Yang You:
Preventing Zero-Shot Transfer Degradation in Continual Learning of Vision-Language Models. 19068-19079
-     Bingchen Zhao, Oisin Mac Aodha:
Incremental Generalized Category Discovery. 19080-19090
-     Gengwei Zhang, Liyuan Wang, Guoliang Kang, Ling Chen, Yunchao Wei:
SLCA: Slow Learner with Classifier Alignment for Continual Learning on a Pre-trained Model. 19091-19101
-     Fu-En Yang, Chien-Yi Wang, Yu-Chiang Frank Wang:
Efficient Model Personalization in Federated Learning via Client-Specific Prompt Generation. 19102-19111
-     Zenan Huang, Haobo Wang, Junbo Zhao, Nenggan Zheng:
iDAG: Invariant DAG Searching for Domain Generalization. 19112-19122
-     Sabbir Ahmed , Abdullah Al Arafat , Mamshad Nayeem Rizve, Rahim Hossain, Zhishan Guo, Adnan Siraj Rakin:
SSDA: Secure Source-Free Domain Adaptation. 19123-19133

-      Dong Zhao, Shuang Wang, Qi Zang, Dou Quan, Xiutiao Ye, Rui Yang, Licheng Jiao: **Learning Pseudo-Relations for Cross-domain Semantic Segmentation.** 19134-19146
-     Kai Zhu, Kecheng Zheng, Ruili Feng, Deli Zhao, Yang Cao, Zheng-Jun Zha: **Self-Organizing Pathway Expansion for Non-Exemplar Class-Incremental Learning.** 19147-19156
-     Ba Hung Ngo, Yeon Jeong Chae, Jung Eun Kwon, Jae Hyeon Park, Sung In Cho: **Improved Knowledge Transfer for Semi-supervised Domain Adaptation via Trico Training Strategy.** 19157-19166
-     Ziqi Gu, Chunyan Xu, Jian Yang, Zhen Cui: **Few-shot Continual Infomax Learning.** 19167-19176
-     Suman Saha, Lukas Hoyer, Anton Obukhov, Dengxin Dai, Luc Van Gool: **EDAPS: Enhanced Domain-Adaptive Panoptic Segmentation.** 19177-19188
-     Jay Zhangjie Wu, David Junhao Zhang, Wynne Hsu, Mengmi Zhang, Mike Zheng Shou: **Label-Efficient Online Continual Object Detection in Streaming Video.** 19189-19198
-     Kai Huang, Feigege Wang, Ye Xi, Yutao Gao: **Prototypical Kernel Learning and Open-set Foreground Perception for Generalized Few-shot Semantic Segmentation.** 19199-19208
-     Seonghyeon Moon, Samuel S. Sohn, Honglu Zhou, Sejong Yoon, Vladimir Pavlovic, Muhammad Haris Khan, Mubbasis Kapadia: **MSI: Maximize Support-Set Information for Few-Shot Segmentation.** 19209-19219
-     Xiaohua Chen, Yucan Zhou, Dayan Wu, Chule Yang, Bo Li, Qinghua Hu, Weiping Wang: **AREA: Adaptive Reweighting via Effective Area for Long-Tailed Classification.** 19220-19230
-     Prithvijit Chattopadhyay, Kartik Sarangmath, Vivek Vijaykumar, Judy Hoffman: **Pasta: Proportional Amplitude Spectrum Training Augmentation for Syn-to-Real Domain Generalization.** 19231-19243
-     Haifeng Xia, Kai Li, Zhengming Ding: **Personalized Semantics Excitation for Federated Image Classification.** 19244-19253
-     Haifeng Xia, Kai Li, Martin Renqiang Min, Zhengming Ding: **Few-Shot Video Classification via Representation Fusion and Promotion Learning.** 19254-19263
-     Stefano Gasperini, Alvaro Marcos-Ramiro, Michael Schmidt, Nassir Navab, Benjamin Busam, Federico Tombari: **Segmenting Known Objects and Unseen Unknowns without Prior Knowledge.** 19264-19275
-     Yuli Zou, Weijian Deng, Liang Zheng: **Adaptive Calibrator Ensemble: Navigating Test Set Difficulty in Out-of-**

**Distribution Scenarios.** 19276-19285

Jintian Ji, Songhe Feng:

Anchor Structure Regularization Induced Multi-view Subspace Clustering via Enhanced Tensor Rank Minimization. 19286-19295

Xinheng Wu, Jie Lu, Zhen Fang, Guangquan Zhang:

Meta OOD Learning For Continuously Adaptive OOD Detection. 19296-19307

Jiexi Yan, Zhihui Yin, Erkun Yang, Yanhua Yang, Heng Huang:

Learning with Diversity: Self-Expanded Equalization for Better Generalized Deep Metric Learning. 19308-19317

Xinghao Wu, Xuefeng Liu, Jianwei Niu, Guogang Zhu, Shaojie Tang

Bold but Cautious: Unlocking the Potential of Personalized Federated Learning through Cautiously Aggressive Collaboration. 19318-19327

Erdong Hu, Yuxin Tang, Anastasios Kyrillidis, Chris Jermaine:

Federated Learning Over Images: Vertical Decompositions and Pre-Trained Backbones Are Difficult to Beat. 19328-19339

Andong Deng, Xingjian Li, Di Hu, Tianyang Wang, Haoyi Xiong, Cheng-Zhong Xu:

Towards Inadequately Pre-trained Models in Transfer Learning. 19340-19351

Tuong Do, Binh X. Nguyen, Vuong Pham, Toan Tran, Erman Tjiputra, Quang D. Tran, Anh Nguyen:

Reducing Training Time in Cross-Silo Federated Learning using Multigraph Topology. 19352-19362

Yufei Guo, Yuhang Zhang, Yuanpei Chen, Weihang Peng, Xiaode Liu, Liwen Zhang, Xuhui Huang, Zhe Ma:

Membrane Potential Batch Normalization for Spiking Neural Networks. 19363-19373

Xiaoyuan Guan, Zhouwu Liu, Wei-Shi Zheng, Yuren Zhou, Ruixuan Wang:

Revisit PCA-based technique for Out-of-Distribution Detection. 19374-19382

Zhibin Dong, Siwei Wang, Jiaqi Jin, Xinwang Liu, En Zhu:

Cross-view Topology Based Consistent and Complementary Information for Deep Multi-view Clustering. 19383-19394

Jianqi Ma, Zhetong Liang, Wangmeng Xiang, Xi Yang, Lei Zhang:

A Benchmark for Chinese-English Scene Text Image Super-resolution. 19395-19404

Cheng Da, Chuwei Luo, Qi Zheng, Cong Yao:

Vision Grid Transformer for Document Layout Analysis. 19405-19415

Tongkun Guan, Wei Shen, Xue Yang, Qi Feng, Zekun Jiang, Xiaokang Yang:

Self-supervised Character-to-Character Distillation for Text Recognition. 19416-19427

Jiabang He, Lei Wang, Yi Hu, Ning Liu, Hui Liu, Xing Xu, Heng Tao Shen:

ICL-D3IE: In-Context Learning with Diverse Demonstrations Updating for Document Information Extraction. 19428-19437

-      Mingxin Huang , Jiaxin Zhang, Dezhi Peng, Hao Lu, Can Huang, Yuliang Liu, Xiang Bai, Lianwen Jin:
ESTextSpotter: Towards Better Scene Text Spotting with Explicit Synergy in Transformer. 19438-19448
-     Wei Pan, Anna Zhu, Xinyu Zhou, Brian Kenji Iwana, Shilin Li:
Few shot font generation via transferring similarity guided global style and quantization local style. 19449-19459
-     Haoyu Cao, Changcun Bao, Chaohu Liu, Huang Chen, Kun Yin, Hao Liu, Yinsong Liu, Degiang Jiang, Xing Sun:
Attention Where It Matters: Rethinking Visual Document Understanding with Selective Region Concentration. 19460-19470
-     Jordy Van Landeghem, Rafal Powalski, Rubén Tito, Dawid Jurkiewicz, Matthew B. Blaschko, Lukasz Borchmann, Mickaël Coustaty, Sien Moens, Michal Pietruszka, Bertrand Anckaert, Tomasz Stanislawek, Paweł Józiak, Ernest Valveny:
Document Understanding Dataset and Evaluation (DUDE). 19471-19483
-     Changxu Cheng, Peng Wang, Cheng Da, Qi Zheng, Cong Yao:
LISTER: Neighbor Decoding for Length-Insensitive Scene Text Recognition. 19484-19494
-     Lucas Morin, Martin Danelljan, Maria Isabel Agea, Ahmed S. Nassar, Valéry Weber, Ingmar Meijer, Peter W. J. Staar, Fisher Yu:
MolGrapher: Graph-based Visual Recognition of Chemical Structures. 19495-19504
-     Daehee Kim, Yoonsik Kim, Donghyun Kim, Yumin Lim, Geewook Kim, Taeho Kil:
SCOB: Universal Text Understanding via Character-wise Supervised Contrastive Learning with Online Text Rendering for Bridging Domain Gap. 19505-19516
-     Heng Li, Xiangping Wu, Qingcai Chen, Qianjin Xiang:
Foreground and Text-lines Aware Document Image Rectification. 19517-19526
-     Haofu Liao, Aruni RoyChowdhury, Weijian Li, Ankan Bansal, Yuting Zhang, Zhuowen Tu, Ravi Kumar Satzoda, R. Manmatha, Vijay Mahadevan:
DocTr: Document Transformer for Structured Information Extraction in Documents. 19527-19537
-     Yang Fu, Shibei Meng, Saihui Hou, Xuecai Hu, Yongzhen Huang:
GPGait: Generalized Pose-based Gait Recognition. 19538-19547
-     Lei Shen, Jianlong Jin, Ruixin Zhang, Huaen Li, Kai Zhao, Yingyi Zhang, Jingyun Zhang, Shouhong Ding, Yang Zhao, Wei Jia:
RPG-Palm: Realistic Pseudo-data Generation for Palmprint Recognition. 19548-19559
-     Feng Liu, Minchul Kim , ZiAng Gu, Anil Jain, Xiaoming Liu:
Learning Clothing and Pose Invariant 3D Shape Representation for Long-Term Person Re-Identification. 19560-19569
-     Hongji Guo, Qiang Ji:
Physics-Augmented Autoencoder for 3D Skeleton-Based Gait Recognition.

-      19570-19581
Lei Wang , Bo Liu, Fangfang Liang, Bincheng Wang:
Hierarchical Spatio-Temporal Representation Learning for Gait Recognition. 19582-19592
-     Fadi Boutros, Jonas Henry Grebe, Arjan Kuijper, Naser Damer:
IDiff-Face: Synthetic-based Face Recognition through Fuzzy Identity-Conditioned Diffusion Models. 19593-19604
-     Hatef Otroshi-Shahreza, Sébastien Marcel:
Template Inversion Attack against Face Recognition Systems using 3D Face Reconstruction. 19605-19615
-     Yuxi Mi, Yuge Huang, Jiazen Ji, Minyi Zhao, Jiaxiang Wu, Xingkun Xu, Shouhong Ding, Shuigeng Zhou:
Privacy-Preserving Face Recognition Using Random Frequency Components. 19616-19627
-     Koushik Srivatsan, Muzammal Naseer, Karthik Nandakumar:
FLIP: Cross-domain Face Anti-spoofing with Language Guidance. 19628-19639
-     Jonathan T. Barron, Ben Mildenhall, Dor Verbin, Pratul P. Srinivasan, Peter Hedman:
Zip-NeRF: Anti-Aliased Grid-Based Neural Radiance Fields. 19640-19648
-     Feng Wang, Sinan Tan, Xinghang Li, Zeyue Tian, Yafei Song, Huaping Liu:
Mixed Neural Voxels for Fast Multi-view Video Synthesis. 19649-19659
-     Yufei Ye, Poorvi Hebbar, Abhinav Gupta, Shubham Tulsiani:
Diffusion-Guided Reconstruction of Everyday Hand-Object Interaction Clips. 19660-19671
-     Justin Kerr, Chung Min Kim, Ken Goldberg, Angjoo Kanazawa, Matthew Tancik:
LERF: Language Embedded Radiance Fields. 19672-19682
-     Ayaan Haque, Matthew Tancik, Alexei A. Efros, Aleksander Holynski, Angjoo Kanazawa:
Instruct-NeRF2NeRF: Editing 3D Scenes with Instructions. 19683-19693
-     Jonathan Ventura, Zuzana Kukelova, Torsten Sattler, Dániel Baráth:
P1AC: Revisiting Absolute Pose From a Single Affine Correspondence. 19694-19704
-     Vanessa Sklyarova, Jenya Chelishev, Andreea Dogaru, Igor Medvedev, Victor Lempitsky, Egor Zakharov:
Neural Haircut: Prior-Guided Strand-Based Hair Reconstruction. 19705-19716
-     Wenbo Hu, Yuling Wang, Lin Ma, Bangbang Yang, Lin Gao, Xiao Liu, Yuewen Ma :
Tri-MipRF: Tri-Mip Representation for Efficient Anti-Aliasing Neural Radiance Fields. 19717-19726
-     Amirreza Shaban, Joonho Lee, Sanghun Jung, Xiangyun Meng, Byron Boots:
LiDAR-UDA: Self-ensembling Through Time for Unsupervised LiDAR Domain

**Adaptation.** 19727-19737

Qianqian Wang, Yen-Yu Chang, Ruojin Cai, Zhengqi Li, Bharath Hariharan, Aleksander Holynski, Noah Snavely:

Tracking Everything Everywhere All at Once. 19738-19749

Rawal Khirodkar, Aayush Bansal, Lingni Ma, Richard A. Newcombe, Minh Vo, Kris Kitani:

EgoHumans: An Egocentric 3D Multi-Human Benchmark. 19750-19762

Lue Fan, Yuxue Yang, Yiming Mao, Feng Wang, Yuntao Chen, Naiyan Wang, Zhaoxiang Zhang:

Once Detected, Never Lost: Surpassing Human Performance in Offline LiDAR based 3D Object Detection. 19763-19772

Shoufa Chen, Peize Sun, Yibing Song, Ping Luo:

DiffusionDet: Diffusion Model for Object Detection. 19773-19786

Jiaqi Wang, Pan Zhang, Tao Chu, Yuhang Cao, Yujie Zhou, Tong Wu, Bin Wang , Conghui He, Dahua Lin:

V3Det: Vast Vocabulary Visual Detection Dataset. 19787-19797

Yang Zheng, Adam W. Harley, Bokui Shen, Gordon Wetzstein, Leonidas J. Guibas: **PointOdyssey: A Large-Scale Synthetic Dataset for Long-Term Point Tracking.** 19798-19808



Hoonhee Cho, Hyeonseong Kim, Yujeong Chae, Kuk-Jin Yoon:

Label-Free Event-based Object Recognition via Joint Learning with Image Reconstruction from Events. 19809-19820

Yan Han, Peihao Wang, Souvik Kundu, Ying Ding, Zhangyang Wang:

Vision HGNN: An Image is More than a Graph of Nodes. 19821-19831

Shuning Chang, Pichao Wang, Hao Luo, Fan Wang, Mike Zheng Shou:

Revisiting Vision Transformer from the View of Path Ensemble. 19832-19842

Jia Ning, Chen Li, Zheng Zhang, Chunyu Wang, Zigang Geng, Qi Dai, Kun He, Han Hu:

All in Tokens: Unifying Output Space of Visual Tasks via Soft Token. 19843-19853

Haoxin Li, Yuan Liu, Hanwang Zhang, Boyang Li:

Mitigating and Evaluating Static Bias of Action Representations in the Background and the Foreground. 19854-19866

Haosen Shi, Shen Ren, Tianwei Zhang, Sinno Jialin Pan:

Deep Multitask Learning with Progressive Parameter Sharing. 19867-19878

Shuyuan Tu, Qi Dai, Zuxuan Wu, Zhi-Qi Cheng, Han Hu, Yu-Gang Jiang:

Implicit Temporal Modeling with Learnable Alignment for Video Recognition. 19879-19890

Kunchang Li, Yali Wang, Yizhuo Li, Yi Wang, Yinan He, Limin Wang, Yu Qiao:

Unmasked Teacher: Towards Training-Efficient Video Foundation Models. 19891-19903

-      Lu Yang, Liulei Li, Xueshi Xin, Yifan Sun, Qing Song, Wenguan Wang: **Large-Scale Person Detection and Localization using Overhead Fisheye Cameras.** 19904-19914
-     Silvia L. Pintea, Yancong Lin, Jouke Dijkstra, Jan C. van Gemert: **A step towards understanding why classification helps regression.** 19915-19924
-     Wei Cheng, Ruixiang Chen, Siming Fan, Wanqi Yin, Keyu Chen, Zhongang Cai, Jingbo Wang, Yang Gao, Zhengming Yu, Zhengyu Lin, Daxuan Ren, Lei Yang, Ziwei Liu, Chen Change Loy, Chen Qian, Wayne Wu, Dahua Lin, Bo Dai, Kwan-Yee Lin: **DNA-Rendering: A Diverse Neural Actor Repository for High-Fidelity Human-centric Rendering.** 19925-19936
-     Lingdong Kong, Youquan Liu, Xin Li, Runnan Chen, Wenwei Zhang, Jiawei Ren, Liang Pan, Kai Chen, Ziwei Liu: **Robo3D: Towards Robust and Reliable 3D Perception against Corruptions.** 19937-19949
-     Oren Barkan, Tal Reiss, Jonathan Weill, Ori Katz, Roy Hirsch, Itzik Malkiel, Noam Koenigstein: **Efficient Discovery and Effective Evaluation of Visual Perceptual Similarity: A Benchmark and Beyond.** 19950-19961
-     Clarence Lee, M. Ganesh Kumar, Cheston Tan: **DetermiNet: A Large-Scale Diagnostic Dataset for Complex Visually-Grounded Referencing using Determiners.** 19962-19971
-     Yonglu Li, Yue Xu, Xinyu Xu, Xiaohan Mao, Yuan Yao, Siqi Liu, Cewu Lu: **Beyond Object Recognition: A New Benchmark towards Object Concept Learning.** 19972-19983
-     Eslam Mohamed Bakr, Pengzhan Sun, Xiaoqian Shen , Faizan Farooq Khan, Li Erran Li, Mohamed Elhoseiny : **HRS-Bench: Holistic, Reliable and Scalable Benchmark for Text-to-Image Models.** 19984-19996
-     Risa Shinoda, Ryo Hayamizu, Kodai Nakashima, Nakamasa Inoue, Rio Yokota, Hirokatsu Kataoka: **SegRCDB: Semantic Segmentation via Formula-Driven Supervised Learning.** 19997-20006
-     Dan Liu, Jin Hou, Shaoli Huang, Jing Liu, Yuxin He, Bochuan Zheng, Jifeng Ning, Jindong Zhang: **LoTE-Animal: A Long Time-span Dataset for Endangered Animal Behavior Understanding.** 20007-20018
-     Ruisheng Wang, Shangfeng Huang, Hongxin Yang: **Building3D: An Urban-Scale Dataset and Benchmarks for Learning Roof Structures from Point Clouds.** 20019-20029
-     Dong Won Lee, Chaitanya Ahuja, Paul Pu Liang, Sanika Natu, Louis-Philippe Morency: **Lecture Presentations Multimodal Dataset: Towards Understanding Multimodality in Educational Videos.** 20030-20041

-      Dogyun Park, Suhyun Kim:
Probabilistic Precision and Recall Towards Reliable Evaluation of Generative Models. 20042-20052
-     Chenchen Zhu, Fanyi Xiao, Andres Alvarado, Yasmine Babaei, Jiabo Hu, Hichem El-Mohri, Sean Chang Culatana, Roshan Sumbaly, Zhicheng Yan:
EgoObjects: A Large-Scale Egocentric Dataset for Fine-Grained Object Understanding. 20053-20063
-     Ru Peng, Qiuyang Duan, Haobo Wang, Jiachen Ma, Yanbo Jiang, Yongjun Tu, Xiu Jiang, Junbo Zhao:
CAME: Contrastive Automated Model Evaluation. 20064-20075
-     Xiaqing Pan, Nicholas Charron, Yongqian Yang, Scott Peters, Thomas Whelan, Chen Kong, Omkar M. Parkhi, Richard A. Newcombe, Carl Yuheng Ren:
Aria Digital Twin: A New Benchmark Dataset for Egocentric 3D Machine Perception. 20076-20086
-     Haoning Wu, Erli Zhang, Liang Liao, Chaofeng Chen, Jingwen Hou, Annan Wang, Wenxiu Sun, Qiong Yan, Weisi Lin:
Exploring Video Quality Assessment on User Generated Contents from Aesthetic and Technical Perspectives. 20087-20097
-     Paola Cascante-Bonilla, Khaled Shehada, James Seale Smith, Sivan Doveh, Donghyun Kim, Rameswar Panda, GüL Varol, Aude Oliva, Vicente Ordonez, Rogério Feris, Leonid Karlinsky:
Going Beyond Nouns With Vision & Language Models Using Synthetic Data. 20098-20108
-     Yue Zhu, Nermin Samet, David Picard:
H3WB: Human3.6M 3D WholeBody Dataset and Benchmark. 20109-20120
-     Mina Alibeigi, William Ljungbergh, Adam Tonderski, Georg Hess, Adam Lilja, Carl Lindström, Daria Motorniuk, Junsheng Fu, Jenny Widahl, Christoffer Petersson:
Zenseact Open Dataset: A large-scale and diverse multimodal dataset for autonomous driving. 20121-20131
-     Kevis-Kokitsi Maninis, Stefan Popov, Matthias Nießner, Vittorio Ferrari:
CAD-Estate: Large-scale CAD Model Annotation in RGB Videos. 20132-20142
-     Dongyoon Han, Junsuk Choe, Seonghyeok Chun, John Joon Young Chung, Minsuk Chang, Sangdoo Yun, Jean Y. Song, Seong Joon Oh:
Neglected Free Lunch - Learning Image Classifiers Using Annotation Byproducts. 20143-20155
-     Kian Eng Ong, Xun Long Ng, Yanchao Li, Wenjie Ai, Kuangyi Zhao, Si Yong Yeo, Jun Liu:
Chaotic World: A Large and Challenging Benchmark for Human Behavior Understanding in Chaotic Events. 20156-20166
-     Henghui Ding, Chang Liu, Shuting He, Xudong Jiang , Philip H. S. Torr, Song Bai:
MOSE: A New Dataset for Video Object Segmentation in Complex Scenes. 20167-20177
-     Yannic Neuhaus, Maximilian Augustin, Valentyn Boreiko, Matthias Hein:
Spurious Features Everywhere - Large-Scale Detection of Harmful Spurious



Features in ImageNet. 20178-20189

Nirat Saini, Hanyu Wang, Archana Swaminathan, Vinoj Jayasundara, Bo He, Kamal Gupta, Abhinav Shrivastava:

Chop & Learn: Recognizing and Generating Object-State Compositions.

20190-20201



Huiyang Shao, Qianqian Xu, Peisong Wen, Peifeng Gao, Zhiyong Yang, Qingming Huang:

Building Bridge Across the Time: Disruption and Restoration of Murals In the Wild. 20202-20212



Xin Wang, Taein Kwon , Mahdi Rad, Bowen Pan, Ishani Chakraborty, Sean Andrist, Dan Bohus, Ashley Feniello, Bugra Tekin, Felipe Vieira Frujeri, Neel Joshi, Marc Pollefeys:

HoloAssist: an Egocentric Human Interaction Dataset for Interactive AI Assistants in the Real World. 20213-20224



Zhitao Yang, Zhongang Cai, Haiyi Mei, Shuai Liu, Zhaoxi Chen, Weiye Xiao, Yukun Wei, Zhongfei Qing, Chen Wei, Bo Dai, Wayne Wu, Chen Qian, Dahua Lin, Ziwei Liu, Lei Yang:

SynBody: Synthetic Dataset with Layered Human Models for 3D Human Perception and Modeling. 20225-20235



Runjia Li, Shuyang Sun, Mohamed Elhoseiny , Philip H. S. Torr:

OxfordTVG-HIC: Can Machine Make Humorous Captions from Images?

20236-20246



Lojze Zust, Janez Pers, Matej Kristan:

LaRS: A Diverse Panoptic Maritime Obstacle Detection Dataset and Benchmark. 20247-20257



Erica Weng, Hana Hoshino, Deva Ramanan, Kris Kitani:

Joint Metrics Matter: A Better Standard for Trajectory Forecasting. 20258-

20269



Yiqian Wu, Jing Zhang, Hongbo Fu, Xiaogang Jin:

LPFF: A Portrait Dataset for Face Generators Across Large Poses. 20270-

20280



Roman Shapovalov, Yanir Kleiman, Ignacio Rocco, David Novotný, Andrea

Vedaldi, Changan Chen, Filippos Kokkinos, Benjamin Graham, Natalia Neverova:

Replay: Multi-modal Multi-view Acted Videos for Casual Holography. 20281-

20291



Yiteng Xu, Peishan Cong, Yichen Yao, Runnan Chen, Yuenan Hou, Xinge Zhu,

Xuming He, Jingyi Yu, Yuexin Ma:

Human-centric Scene Understanding for 3D Large-scale Scenarios. 20292-

20302



Ryo Nakamura, Hirokatsu Kataoka, Sora Takashima, Edgar Josafat Martinez-Noriega, Rio Yokota, Nakamasa Inoue:

Pre-training Vision Transformers with Very Limited Synthesized Images.

20303-20312

-      Laura Gustafson, Chloé Rolland, Nikhila Ravi, Quentin Duval, Aaron Adcock, Cheng-Yang Fu, Melissa Hall, Candace Ross:
Facet: Fairness in Computer Vision Evaluation Benchmark. 20313-20325
-     Jingyuan Yang, Qirui Huang, Tingting Ding, Dani Lischinski, Daniel Cohen-Or, Hui Huang:
EmoSet: A Large-scale Visual Emotion Dataset with Rich Attributes. 20326-20337
-     Lijun Li, Linrui Tian, Xindi Zhang, Qi Wang, Bang Zhang, Liefeng Bo, Mengyuan Liu, Chen Chen:
RenderIH: A Large-scale Synthetic Dataset for 3D Interacting Hand Pose Estimation. 20338-20348
-     Yushi Hu, Benlin Liu, Jungo Kasai, Yizhong Wang, Mari Ostendorf, Ranjay Krishna, Noah A. Smith:
TIFA: Accurate and Interpretable Text-to-Image Faithfulness Evaluation with Question Answering. 20349-20360
-     Sruthi Sudhakar, Jon Hanzelka, Josh Bobillot, Tanmay Randhavane, Neel Joshi, Vibhav Vineet:
Exploring the Sim2Real Gap using Digital Twins. 20361-20370
-     Bingyang Zhou, Haoyu Zhou, Tianhai Liang, Qiaojun Yu, Siheng Zhao, Yuwei Zeng, Jun Lv, Siyuan Luo, Qiancai Wang, Xinyuan Yu, Haonan Chen, Cewu Lu, Lin Shao:
ClothesNet: An Information-Rich 3D Garment Model Repository with Simulated Clothes Environment. 20371-20381
-     Jiangwei Yu, Xiang Li, Xinran Zhao, Hongming Zhang, Yu-Xiong Wang:
Video State-Changing Object Segmentation. 20382-20391
-     Xinran Liu, Xiaoqiong Liu, Ziruo Yi, Xin Zhou, Thanh Le, Libo Zhang, Yan Huang, Qing Yang, Heng Fan:
PlanarTrack: A Large-scale Challenging Benchmark for Planar Object Tracking. 20392-20401
-     Dingkang Yang, Shuai Huang, Zhi Xu, Zhenpeng Li, Shunli Wang, Mingcheng Li, Yuzheng Wang, Yang Liu, Kun Yang, Zhaoyu Chen, Yan Wang, Jing Liu, Peixuan Zhang, Peng Zhai, Lihua Zhang:
AIDE: A Vision-Driven Multi-View, Multi-Modal, Multi-Tasking Dataset for Assistive Driving Perception. 20402-20413
-     Yan Luo, Min Shi, Yu Tian, Tobias Elze, Mengyu Wang:
Harvard Glaucoma Detection and Progression: A Multimodal Multitask Dataset and Generalization-Reinforced Semi-Supervised Learning. 20414-20425
-     Ran Gong, Jiangyong Huang, Yizhou Zhao, Haoran Geng, Xiaofeng Gao, Qingyang Wu, Wensi Ai, Ziheng Zhou, Demetri Terzopoulos, Song-Chun Zhu, Baoxiong Jia, Siyuan Huang:
ARNOLD: A Benchmark for Language-Grounded Task Learning With Continuous States in Realistic 3D Scenes. 20426-20438

-      Faizan Farooq Khan, Xiang Li, Andrew J. Temple, Mohamed Elhoseiny : **FishNet: A Large-scale Dataset and Benchmark for Fish Recognition, Detection, and Functional Trait Prediction.** 20439-20449
-     Guoyuan An, Woo Jae Kim, Saelyne Yang, Rong Li, Yuchi Huo, Sung-Eui Yoon: **Towards Content-based Pixel Retrieval in Revisited Oxford and Paris.** 20450-20461
-     Andong Deng, Taojiannan Yang, Chen Chen: **A Large-scale Study of Spatiotemporal Representation Learning with a New Benchmark on Action Recognition.** 20462-20474
-     Zilin Fang, Andrey Ignatov, Eduard Zamfir, Radu Timofte: **SQAD: Automatic Smartphone Camera Quality Assessment and Benchmarking.** 20475-20485
-     Qing Jiang, Jiapeng Wang, Dezheng Peng, Chongyu Liu, Lianwen Jin: **Revisiting Scene Text Recognition: A Data Perspective.** 20486-20497
-     Ryuichiro Hataya, Han Bao, Hiromi Arai: **Will Large-scale Generative Models Corrupt Future Datasets?** 20498-20508
-     Huajian Huang, Yinzhe Xu, Yingshu Chen, Sai-Kit Yeung: **360VOT: A New Benchmark Dataset for Omnidirectional Visual Object Tracking.** 20509-20519
-     Shu Nakamura, Yasutomo Kawanishi, Shohei Nobuhara, Ko Nishino: **DeePoint: Visual Pointing Recognition and Direction Estimation.** 20520-20530
-     Zhihua Li, Lijun Yin: **Contactless Pulse Estimation Leveraging Pseudo Labels and Self-Supervision.** 20531-20540
-     Hongxia Xie, Ming-Xian Lee, Tzu-Jui Chen, Hung-Jen Chen, Hou-I Liu, Hong-Han Shuai, Wen-Huang Cheng: **Most Important Person-guided Dual-branch Cross-Patch Attention for Group Affect Recognition.** 20541-20551
-     Shaowei Liu, Yang Zhou, Jimei Yang, Saurabh Gupta, Shenlong Wang: **ContactGen: Generative Contact Modeling for Grasp Generation.** 20552-20563
-     Balamurugan Thambiraja, Ihsanul Habibie, Sadegh Aliakbarian, Darren Cosker, Christian Theobalt, Justus Thies: **Imitator: Personalized Speech-driven 3D Facial Animation.** 20564-20574
-     Yihua Cheng, Feng Lu: **DVGaze: Dual-View Gaze Estimation.** 20575-20584
-     Jun Dan, Yang Liu, Haoyu Xie, Jiankang Deng, Haoran Xie, Xuansong Xie, Baigui Sun: **TransFace: Calibrating Transformer Training for Face Recognition from a Data-Centric Perspective.** 20585-20596
-     Yuchen Liu, Yabo Chen, Mengran Gou, Chun-Ting Huang, Yaoming Wang, Wenrui Dai, Hongkai Xiong:

**Towards Unsupervised Domain Generalization for Face Anti-Spoofing.**

20597-20607



Xiaohang Ren, Xingyu Chen, Pengfei Yao, Heung-Yeung Shum, Baoyuan Wang:
Reinforced Disentanglement for Face Swapping without Skip Connection.
20608-20618



Peiqi Jiao, Yuecong Min, Yanan Li, Xiaotao Wang, Lei Lei, Xilin Chen:
CoSign: Exploring Co-occurrence Signals in Skeleton-based Continuous Sign Language Recognition. 20619-20629



Ziqiao Peng, Haoyu Wu, Zhenbo Song, Hao Xu, Xiangyu Zhu, Jun He, Hongyan Liu, Zhaoxin Fan:
EmoTalk: Speech-Driven Emotional Disentanglement for 3D Face Animation. 20630-20640



Zhiyu Wu, Jinshi Cui:
LA-Net: Landmark-Aware Learning for Reliable Facial Expression Recognition under Label Noise. 20641-20650



Kai Yang, Hong Shang, Tianyang Shi, Xinghan Chen, Jingkai Zhou, Zhongqian Sun, Wei Yang:
ASM: Adaptive Skinning Model for High-Quality 3D Face Modeling. 20651-20660



Fu-Zhao Ou, Baoliang Chen, Chongyi Li, Shiqi Wang, Sam Kwong:
Troubleshooting Ethnic Quality Bias with Curriculum Domain Adaptation for Face Image Quality Assessment. 20661-20672



Jiancan Zhou, Xi Jia, Qiufu Li, Linlin Shen, Jinming Duan:
UniFace: Unified Cross-Entropy Loss for Deep Face Recognition. 20673-20682



Taeryung Lee, Yeonguk Oh, Kyoung Mu Lee:
Human Part-wise 3D Motion Context Learning for Sign Language Recognition. 20683-20693



Xiang Zhang, Taoyue Wang, Xiaotian Li, Huiyuan Yang, Lijun Yin:
Weakly-Supervised Text-driven Contrastive Learning for Facial Behavior Understanding. 20694-20705



Xiaozheng Zheng, Chao Wen, Zhou Xue, Pengfei Ren, Jingyu Wang:
HaMuCo: Hand Pose Estimation via Multiview Collaborative Self-Supervised Learning. 20706-20716



Xiaotian Li, Taoyue Wang, Geran Zhao, Xiang Zhang, Xi Kang, Lijun Yin:
ReactioNet: Learning High-order Facial Behavior from Universal Stimulus-Reaction by Dyadic Relation Reasoning. 20717-20728



Shuai Shen, Wanhua Li, Xiaobing Wang, Dafeng Zhang, Zhezhu Jin, Jie Zhou, Jiwen Lu:
CLIP-Cluster: CLIP-Guided Attribute Hallucination for Face Clustering.
20729-20738



Jingbo Wang, Ye Yuan, Zhengyi Luo, Kevin Xie, Dahua Lin, Umar Iqbal, Sanja Fidler, Sameh Khamis:

**Learning Human Dynamics in Autonomous Driving Scenarios.** 20739-20749

Yihao Zhi, Xiaodong Cun, Xuelin Chen, Xi Shen, Wen Guo, Shaoli Huang, Shenghua Gao:

LivelySpeaker: Towards Semantic-Aware Co-Speech Gesture Generation.
20750-20760



Ying Guo , Cheng Zhen, Pengfei Yan:

Controllable Guide-Space for Generalizable Face Forgery Detection. 20761-20770



Zhenfeng Fan, Zhiheng Zhang, Shuang Yang, Chongyang Zhong, Min Cao, Shihong Xia:

Unpaired Multi-domain Attribute Translation of 3D Facial Shapes with a Square and Symmetric Geometric Map. 20771-20781



Luchuan Song, Guojun Yin, Zhenchao Jin, Xiaoyi Dong, Chenliang Xu:

Emotional Listener Portrait: Realistic Listener Motion Simulation in Conversation. 20782-20792



Nithin Gopalakrishnan Nair, Anoop Cherian, Suhas Lohit, Ye Wang, Toshiaki Koike-Akino, Vishal M. Patel, Tim K. Marks:

Steered Diffusion: A Generalized Framework for Plug-and-Play Conditional Image Synthesis. 20793-20803



Jiali Ma, Zhongqi Yue, Tomoyuki Kagaya, Tomoki Suzuki, Jayashree Karlekar, Sugiri Pranata, Hanwang Zhang:

Invariant Feature Regularization for Fair Face Recognition. 20804-20813



Benjia Zhou, Zhigang Chen, Albert Clapés, Jun Wan, Yanyan Liang, Sergio Escalera , Zhen Lei, Du Zhang:

Gloss-free Sign Language Translation: Improving from Visual-Language Pretraining. 20814-20824



Zhimin Sun, Shen Chen, Taiping Yao, Bangjie Yin, Ran Yi, Shouhong Ding, Lizhuang Ma:

Contrastive Pseudo Learning for Open-World DeepFake Attribution. 20825-20835



Chaitanya Ahuja, Pratik Joshi, Ryo Ishii, Louis-Philippe Morency:

Continual Learning for Personalized Co-Speech Gesture Generation. 20836-20846



Wencan Cheng, Jong Hwan Ko:

HandR²N²: Iterative 3D Hand Pose Estimation Using a Residual Recurrent Neural Network. 20847-20856



Siddharth Gururani, Arun Mallya, Ting-Chun Wang, Rafael Valle, Ming-Yu Liu:

SPACE: Speech-driven Portrait Animation with Controllable Expression.

20857-20866



Artem Sevastopolsky, Yury Malkov, Nikita Durasov, Luisa Verdoliva, Matthias Nießner:

How to Boost Face Recognition with StyleGAN? 20867-20877

-      Samy Tafasca, Anshul Gupta, Jean-Marc Odobez : **ChildPlay: A New Benchmark for Understanding Children's Gaze Behaviour.** 20878-20899
-     Trevine Oorloff, Yaser Yacoob: **Robust One-Shot Face Video Re-enactment using Hybrid Latent Spaces of StyleGAN2.** 20890-20900
-     Shubhra Aich, Jesús Ruiz-Santaquiteria, Zhenyu Lu, Prachi Garg, K. J. Joseph, Alvaro Fernandez Garcia, Vineeth N. Balasubramanian, Kenrick Kin, Chengde Wan, Necati Cihan Camgöz, Shugao Ma, Fernando De la Torre: **Data-Free Class-Incremental Hand Gesture Recognition.** 20901-20910
-     Yunan Li, Huizhou Chen, Guanwen Feng, Qiguang Miao: **Learning Robust Representations with Information Bottleneck and Memory Network for RGB-D-based Gesture Recognition.** 20911-20921
-     Xiaotian Li, Xiang Zhang, Taoyue Wang, Lijun Yin: **Knowledge-Spreader: Learning Semi-Supervised Facial Action Dynamics by Consistifying Knowledge Granularity.** 20922-20932
-     Yang Wu, Zhiwei Ge, Yuhao Luo, Lin Liu, Sulong Xu: **Face Clustering via Graph Convolutional Networks with Confidence Edges.** 20933-20942
-     Shuai Yang, Liming Jiang, Ziwei Liu, Chen Change Loy: **StyleGANEX: StyleGAN-Based Manipulation Beyond Cropped Aligned Faces.** 20943-20953
-     Nicolas Larue, Ngoc-Son Vu, Vitomir Struc, Peter Peer, Vassilis Christophides: **SeeABLE: Soft Discrepancies and Bounded Contrastive Learning for Exposing Deepfakes.** 20954-20964
-     Zhizhong Huang, Siteng Ma, Junping Zhang, Hongming Shan: **Adaptive Nonlinear Latent Transformation for Conditional Face Editing.** 20965-20974
-     Peiji Yang, Huawei Wei, Yicheng Zhong, Zhisheng Wang: **Semi-supervised Speech-driven 3D Facial Animation via Cross-modal Encoding.** 20975-20984
-     Zhipeng Yu, Jiaheng Liu, Haoyu Qin, Yichao Wu, Kun Hu, Jiayi Tian, Ding Liang: **ICD-Face: Intra-class Compactness Distillation for Face Recognition.** 20985-20995
-     Huaiwen Zhang, Zihang Guo, Yang Yang, Xin Liu, De Hu: **C²ST: Cross-modal Contextualized Sequence Transduction for Continuous Sign Language Recognition.** 20996-21005
-     Ramin Nakhli, Allen W. Zhang, Ali Khajegili Mirabadi, Katherine Rich, Maryam Asadi, C. Blake Gilks, Hossein Farahani, Ali Bashashati: **CO-PILOT: Dynamic Top-Down Point Cloud with Conditional Neighborhood Aggregation for Multi-Gigapixel Histopathology Image Representation.** 21006-21016

-      Yang Liu, Jiayu Huo , Jingjing Peng, Rachel Sparks , Prokar Dasgupta, Alejandro Granados, Sébastien Ourselin:
SKiT: a Fast Key Information Video Transformer for Online Surgical Phase Recognition. 21017-21027
-     Yanfeng Zhou, Jiaxing Huang, Chenlong Wang, Le Song, Ge Yang:
XNet: Wavelet-Based Low and High Frequency Fusion Networks for Fully- and Semi-Supervised Semantic Segmentation of Biomedical Images. 21028-21039
-     Arne Schmidt, Pablo Morales-Álvarez, Rafael Molina:
Probabilistic Modeling of Inter- and Intra-observer Variability in Medical Image Segmentation. 21040-21049
-     Xiaoyu Liu, Wei Huang, Zhiwei Xiong, Shenglong Zhou, Yueyi Zhang, Xuejin Chen, Zheng-Jun Zha, Feng Wu:
Learning Cross-Representation Affinity Consistency for Sparsely Supervised Biomedical Instance Segmentation. 21050-21060
-     Yongheng Sun , Fan Wang, Jun Shu, Haifeng Wang, Li Wang, Deyu Meng, Chunfeng Lian:
Dual Meta-Learning with Longitudinally Generalized Regularization for One-Shot Brain Tissue Segmentation Across the Human Lifespan. 21061-21071
-     Hwiwon Jeong, Heejoon Byun, Dong Un Kang, Jongho Lee:
BlindHarmony: "Blind" Harmonization for MR Images via Flow model. 21072-21082
-     Zhanghexuan Ji, Dazhou Guo, Puyang Wang, Ke Yan, Le Lu, Minfeng Xu, Qifeng Wang, Jia Ge, Mingchen Gao, Xianghua Ye, Dakai Jin:
Continual Segment: Towards a Single, Unified and Non-forgetting Continual Segmentation Model of 143 Whole-body Organs in CT Scans. 21083-21094
-     Jie Liu , Yixiao Zhang, Jieneng Chen, Junfei Xiao, Yongyi Lu, Bennett A. Landman, Yixuan Yuan, Alan L. Yuille, Yucheng Tang, Zongwei Zhou:
CLIP-Driven Universal Model for Organ Segmentation and Tumor Detection. 21095-21107
-     Gefen Dawidowicz, Elad Hirsch, Ayellet Tal:
LIMITR: Leveraging Local Information for Medical Image-Text Representation. 21108-21116
-     Jianan Fan, Dongnan Liu, Hang Chang, Heng Huang, Mei Chen, Weidong Cai:
Taxonomy Adaptive Cross-Domain Adaptation in Medical Imaging via Optimization Trajectory Distillation. 21117-21127
-     Zixuan Chen, Lingxiao Yang, Jian-Huang Lai, Xiaohua Xie:
CuNeRF: Cube-Based Neural Radiance Field for Zero-Shot Medical Image Arbitrary-Scale Super Resolution. 21128-21138
-     Zilong Li, Chenglong Ma, Jie Chen, Junping Zhang, Hongming Shan:
Learning to Distill Global Representation for Sparse-View CT. 21139-21150
-     Qihua Dong, Hao Du, Ying Song, Yan Xu, Jing Liao:
Preserving Tumor Volumes for Unsupervised Medical Image Registration.

-      21151-21161
Ashesh, Alexander Krull, Moises Di Sante, Francesco Silvio Pasqualini, Florian Jug:
μSplit: image decomposition for fluorescence microscopy. 21162-21172
-     Guangyuan Li, Lei Zhao, Jiakai Sun, Zehua Lan, Zhanjie Zhang, Jiafu Chen, Zhijie Lin, Huaizhong Lin, Wei Xing:
Rethinking Multi-Contrast MRI Super-Resolution: Rectangle-Window Cross-Attention Transformer and Arbitrary-Scale Upsampling. 21173-21183
-     Yingxue Xu, Hao Chen:
Multimodal Optimal Transport-based Co-Attention Transformer with Global Structure Consistency for Survival Prediction. 21184-21194
-     Xiaohan Yuan, Cong Liu, Yangang Wang:
4D Myocardium Reconstruction with Decoupled Motion and Shape Model. 21195-21205
-     Steffen Wolf, Manan Lalit, Katie McDole, Jan Funke:
Unsupervised Learning of Object-Centric Embeddings for Cell Instance Segmentation in Microscopy Images. 21206-21215
-     Javier Rodriguez Puigvert, Victor M. Batlle, J. M. M. Montiel, Ruben Martinez-Cantin, Pascal Fua, Juan D. Tardós, Javier Civera:
LightDepth: Single-View Depth Self-Supervision from Illumination Decline. 21216-21226
-     Yuanhong Chen, Fengbei Liu, Hu Wang, Chong Wang, Yuyuan Liu, Yu Tian, Gustavo Carneiro:
BoMD: Bag of Multi-label Descriptors for Noisy Chest X-ray Classification. 21227-21238
-     Pengcheng Lei, Faming Fang, Guixu Zhang, Tieyong Zeng:
Decomposition-Based Variational Network for Multi-Contrast MRI Super-Resolution and Reconstruction. 21239-21249
-     Hongliang He, Jun Wang, Pengxu Wei, Fan Xu, Xiangyang Ji, Chang Liu, Jie Chen:
TopoSeg: Topology-Aware Nuclear Instance Segmentation. 21250-21259
-     Yansheng Qiu, Delin Chen, Hongdou Yao, Yongchao Xu, Zheng Wang:
Scratch Each Other's Back: Incomplete Multi-modal Brain Tumor Segmentation Via Category Aware Group Self-Support Learning. 21260-21269
-     Jieneng Chen, Yingda Xia, Jiawen Yao, Ke Yan, Jianpeng Zhang, Le Lu, Fakai Wang, Bo Zhou, Mingyan Qiu, Qihang Yu, Mingze Yuan, Wei Fang, Yuxing Tang, Minfeng Xu, Jian Zhou, Yuqian Zhao, Qifeng Wang, Xianghua Ye, Xiaoli Yin, Yu Shi, Xin Chen, Jingren Zhou, Alan L. Yuille, Zaiyi Liu, Ling Zhang:
CancerUniT: Towards a Single Unified Model for Effective Detection, Segmentation, and Diagnosis of Eight Major Cancers Using a Large Collection of CT Scans. 21270-21281
-     Xihe Qiu, Shaojie Shi, Xiaoyu Tan, Chao Qu, Zhijun Fang, Hailing Wang, Yongbin Gao, Peixia Wu, Huawei Li:
Gram-based Attentive Neural Ordinary Differential Equations Network for Video Nystagmography Classification. 21282-21291

-      Yanyan Huang, Weiqin Zhao, Shujun Wang, Yu Fu, Yuming Jiang, Lequan Yu: **ConSlide: Asynchronous Hierarchical Interaction Transformer with Breakup-Reorganize Rehearsal for Continual Whole Slide Image Analysis.** 21292-21303
-     Pujin Cheng, Li Lin, Junyan Lyu, Yijin Huang, Wenhan Luo, Xiaoying Tang: **PRIOR: Prototype Representation Joint Learning from Medical Images and Reports.** 21304-21314
-     Chaoyi Wu, Xiaoman Zhang, Ya Zhang , Yanfeng Wang, Weidi Xie: **MedKLIP: Medical Knowledge Enhanced Language-Image Pre-Training for X-ray Diagnosis.** 21315-21326
-     Junjia Huang, Haofeng Li, Xiang Wan, Guanbin Li: **Affine-Consistent Transformer for Multi-Class Cell Nuclei Detection.** 21327-21336
-     Martin J. Menten, Johannes C. Paetzold, Veronika A. Zimmer, Suprosanna Shit, Ivan Ezhov, Robbie Holland, Monika Probst, Julia A. Schnabel, Daniel Rueckert: **A skeletonization algorithm for gradient-based optimization.** 21337-21346
-     Weiyi Wu, Chongyang Gao, Joseph DiPalma, Soroush Vosoughi, Saeed Hassanpour: **Improving Representation Learning for Histopathologic Images with Cluster Constraints.** 21347-21357
-     Aishik Konwer, Xiaoling Hu, Joseph Bae, Xuan Xu, Chao Chen, Prateek Prasanna: **Enhancing Modality-Agnostic Representations via Meta-learning for Brain Tumor Segmentation.** 21358-21368
-     Juzheng Miao, Cheng Chen, Furui Liu, Hao Wei, Pheng-Ann Heng: **CauSSL: Causality-inspired Semi-supervised Learning for Medical Image Segmentation.** 21369-21380
-     Victor Ion Butoi, Jose Javier Gonzalez Ortiz, Tianyu Ma, Mert R. Sabuncu, John V. Guttag, Adrian V. Dalca: **UniverSeg: Universal Medical Image Segmentation.** 21381-21394
-     Qiushi Yang, Wuyang Li, Baopu Li, Yixuan Yuan: **MRM: Masked Relation Modeling for Medical Image Pre-Training with Genetics.** 21395-21405
-     Linhao Qu, Zhiwei Yang, Minghong Duan, Yingfan Ma, Shuo Wang, Manning Wang, Zhijian Song: **Boosting Whole Slide Image Classification from the Perspectives of Distribution, Correlation and Magnification.** 21406-21416
-     Yuwen Pan, Naisong Luo, Rui Sun, Meng Meng, Tianzhu Zhang, Zhiwei Xiong, Yongdong Zhang: **Adaptive Template Transformer for Mitochondria Segmentation in Electron Microscopy Images.** 21417-21427
-     Fengtao Zhou, Hao Chen: **Cross-Modal Translation and Alignment for Survival Analysis.** 21428-21437



- Zhuchen Shao, Yifeng Wang, Yang Chen, Hao Bian, Shaohui Liu, Haoqian Wang, Yongbing Zhang:
LNPL-MIL: Learning from Noisy Pseudo Labels for Promoting Multiple Instance Learning in Whole Slide Image. 21438
- Yating Xu, Conghui Hu, Na Zhao, Gim Hee Lee:
Generalized Few-Shot Point Cloud Segmentation Via Geometric Words. 21449-21458
- Yujiao Shi, Fei Wu, Akhil Perincherry, Ankit Vora, Hongdong Li:
Boosting 3-DoF Ground-to-Satellite Camera Localization Accuracy via Geometry-Guided Cross-View Transformer. 21459-21469
- Minjung Kim, Junseo Koo, Gunhee Kim:
EP2P-Loc: End-to-End 3D Point to 2D Pixel Localization for Large-Scale Visual Localization. 21470-21480
- Shuhong Zheng, Zhipeng Bao, Martial Hebert, Yu-Xiong Wang:
Multi-task View Synthesis with Neural Radiance Fields. 21481-21492
- Yangyang Xu, Yibo Yang, Lefei Zhang:
Multi-Task Learning with Knowledge Distillation for Dense Prediction. 21493-21502
- Qifan Yu, Juncheng Li, Yu Wu, Siliang Tang, Wei Ji, Yuetong Zhuang:
Visually-Prompted Language Model for Fine-Grained Scene Graph Generation in an Open World. 21503-21514
- Ruihao Xia, Chaoqiang Zhao, Meng Zheng, Ziyan Wu, Qiyu Sun, Yang Tang:
CMDA: Cross-Modality Domain Adaptation for Nighttime Semantic Segmentation. 21515-21524
- Yanan Wang, Michihiro Yasunaga, Hongyu Ren, Shinya Wada, Jure Leskovec:
VQA-GNN: Reasoning with Multimodal Knowledge via Graph Neural Networks for Visual Question Answering. 21525-21535
- Zhixiang Wei, Lin Chen, Tao Tu, Pengyang Ling, Huaian Chen, Yi Jin:
Disentangle then Parse: Night-time Semantic Segmentation with Illumination Disentanglement. 21536-21546
- Yunfei Guo, Fei Yin, Xiao-Hui Li, Xudong Yan, Tao Xue, Shuqi Mei, Cheng-Lin Liu:
Visual Traffic Knowledge Graph Generation from Scene Images. 21547-21556
- Danyang Tu, Wei Sun, Guangtao Zhai, Wei Shen:
Agglomerative Transformer for Human-Object Interaction Detection. 21557-21567
- Guangyao Zhou, Nishad Gothsakar, Lirui Wang, Joshua B. Tenenbaum, Dan Gutfreund, Miguel Lázaro-Gredilla, Dileep George, Vikash K. Mansinghka:
3D Neural Embedding Likelihood: Probabilistic Inverse Graphics for Robust 6D Pose Estimation. 21568-21579
- Zijian Zhou , Miaojing Shi, Holger Caesar:
HiLo: Exploiting High Low Frequency Relations for Unbiased Panoptic Scene Graph Generation. 21580-21591

-      Hangjie Yuan, Shiwei Zhang, Xiang Wang, Samuel Albanie, Yining Pan, Tao Feng, Jianwen Jiang, Dong Ni, Yingya Zhang, Deli Zhao:
RLIPv2: Fast Scaling of Relational Language-Image Pre-training. 21592-21604
-     Youquan Liu, Runnan Chen, Xin Li, Lingdong Kong, Yuchen Yang, Zhaoyang Xia, Yeqi Bai, Xinge Zhu, Yuexin Ma, Yikang Li, Yu Qiao, Yuenan Hou:
UniSeg: A Unified Multi-Modal LiDAR Segmentation Network and the OpenPCSeg Codebase. 21605-21616
-     Yuhang Lu, Qi Jiang, Runnan Chen, Yuenan Hou, Xinge Zhu, Yuexin Ma:
See More and Know More: Zero-shot Point Cloud Segmentation via Multi-modal Visual Data. 21617-21627
-     Lin Li, Guikun Chen, Jun Xiao, Yi Yang, Chunping Wang, Long Chen:
Compositional Feature Augmentation for Unbiased Scene Graph Generation. 21628-21638
-     Prashant W. Patil, Sunil Gupta, Santu Rana, Svetha Venkatesh, Subrahmanyam Murala:
Multi-weather Image Restoration via Domain Translation. 21639-21648
-     Aviad Aberdam, David Bensaïd, Alona Golts, Roy Ganz, Oren Nuriel, Royee Tichauer, Shai Mazor, Ron Litman:
CLIPTER: Looking at the Bigger Picture in Scene Text Recognition. 21649-21660
-     Roy Ganz, Oren Nuriel, Aviad Aberdam, Yair Kittenplon, Shai Mazor, Ron Litman:
Towards Models that Can See and Read. 21661-21671
-     Yi Wei, Linqing Zhao, Wenzhao Zheng, Zheng Zhu, Jie Zhou, Jiwen Lu:
SurroundOcc: Multi-Camera 3D Occupancy Prediction for Autonomous Driving. 21672-21683
-     Yuanfeng Ji, Zhe Chen, Enze Xie, Lanqing Hong, Xihui Liu, Zhaoqiang Liu, Tong Lu, Zhenguo Li, Ping Luo:
DDP: Diffusion Model for Dense Visual Prediction. 21684-21695
-     Shengyi Qian, David F. Fouhey:
Understanding 3D Object Interaction from a Single Image. 21696-21706
-     Qianyi Wu, Kaisiyuan Wang, Kejie Li, Jianmin Zheng, Jianfei Cai:
ObjectSDF++: Improved Object-Compositional Neural Implicit Surfaces. 21707-21717
-     Yuanyi Zhong, Anand Bhattad, Yu-Xiong Wang, David A. Forsyth:
Improving Equivariance in State-of-the-Art Supervised Depth and Normal Predictors. 21718-21728
-     Yifang Yin, Wenmiao Hu, Zhenguang Liu, Guanfeng Wang, Shili Xiang, Roger Zimmermann:
CrossMatch: Source-Free Domain Adaptive Semantic Segmentation via Cross-Modal Consistency Training. 21729-21739
-     Yiqing Liang, Eliot Laidlaw, Alexander Meyerowitz, Srinath Sridhar, James Tompkin:



Semantic Attention Flow Fields for Monocular Dynamic Scene Decomposition. 21740-21749

- Ziqiong Lu, Linxi Huan, Qiyuan Ma, Xianwei Zheng:
Holistic Geometric Feature Learning for Structured Reconstruction. 21750-21760
- Zhuo Zheng, Shiqi Tian, Ailong Ma, Liangpei Zhang, Yanfei Zhong:
Scalable Multi-Temporal Remote Sensing Change Data Generation via Simulating Stochastic Change Process. 21761-21770
- Hanrong Ye, Dan Xu:
TaskExpert: Dynamically Assembling Multi-Task Representations with Memorial Mixture-of-Experts. 21771-21780
- Shuai He, Anlong Ming, Yaqi Li, Jinyuan Sun, Shuntian Zheng, Huadong Ma:
Thinking Image Color Aesthetics Assessment: Models, Datasets and Benchmarks. 21781-21790
- Tao Han, Lei Bai, Lingbo Liu, Wanli Ouyang:
STEERER: Resolving Scale Variations for Counting and Localization via Selective Inheritance Learning. 21791-21802
- Francesco Tonini, Nicola Dall'Asen, Cigdem Beyan, Elisa Ricci:
Object-aware Gaze Target Detection. 21803-21812
- Jungbeom Lee, Sungjin Lee, Jinseok Nam, Seunghak Yu, Jaeyoung Do, Tara Taghavi:
Weakly Supervised Referring Image Segmentation with Intra-Chunk and Inter-Chunk Consistency. 21813-21824
- Gopika Sudhakaran, Devendra Singh Dhami, Kristian Kersting, Stefan Roth:
Vision Relation Transformer for Unbiased Scene Graph Generation. 21825-21836
- Haoang Li, Jinhua Dong, Binghui Wen, Ming Gao, Tianyu Huang, Yun-Hui Liu, Daniel Cremers:
DDIT: Semantic Scene Completion via Deformable Deep Implicit Templates. 21837-21847
- Huan-ang Gao, Beiwen Tian, Pengfei Li, Hao Zhao, Guyue Zhou:
DQS3D: Densely-matched Quantization-aware Semi-supervised 3D Detection. 21848-21858
- Mingyue Dong, Linxi Huan, Hanjiang Xiong, Shuhan Shen, Xianwei Zheng:
Shape Anchor Guided Holistic Indoor Scene Understanding. 21859-21869
- Sayan Deb Sarkar, Ondrej Miksik, Marc Pollefeys, Daniel Barath, Iro Armeni:
SGAligner: 3D Scene Alignment with Scene Graphs. 21870-21880
- Jianzong Wu, Xiangtai Li, Henghui Ding, Xia Li, Guangliang Cheng, Yunhai Tong, Chen Change Loy:
Betrayed by Captions: Joint Caption Grounding and Generation for Open Vocabulary Instance Segmentation. 21881-21891
- Jiang-Tian Zhai, Qi Zhang, Tong Wu, Xing-Yu Chen, Jiang-Jiang Liu, Ming-Ming Cheng:



- SLAN: Self-Locator Aided Network for Vision-Language Understanding.**
21892-21901
Sifan Long, Zhen Zhao, Junkun Yuan, Zichang Tan, Jiangjiang Liu, Luping Zhou, Shengsheng Wang, Jingdong Wang:
- Task-Oriented Multi-Modal Mutual Learning for Vision-Language Models.**
21902-21912
Kan Wu, Houwen Peng, Zhenghong Zhou, Bin Xiao, Mengchen Liu, Lu Yuan, Hong Xuan, Michael Valenzuela, Xi Stephen Chen, Xinggang Wang, Hongyang Chao, Han Hu:
- TinyCLIP: CLIP Distillation via Affinity Mimicking and Weight Inheritance.**
21913-21923
Nina Shvetsova, Anna Kukleva, Bernt Schiele, Hilde Kuehne:
In-Style: Bridging Text and Uncurated Videos with Style Transfer for Text-Video Retrieval. 21924-21935
- Sirnam Swetha, Mamshad Nayeem Rizve, Nina Shvetsova, Hilde Kuehne, Mubarak Shah:
Preserving Modality Structure Improves Multi-Modal Learning. 21936-21946
- Eulrang Cho, Jooyeon Kim, Hyunwoo J. Kim:
Distribution-Aware Prompt Tuning for Vision-Language Models. 21947-21956
- Yiran Qin, Chaoqun Wang, Zijian Kang, Ningning Ma, Zhen Li, Ruimao Zhang:
SupFusion: Supervised LiDAR-Camera Fusion for 3D Object Detection.
21957-21967
- Yuanzhi Wang, Zhen Cui, Yong Li:
Distribution-Consistent Modal Recovering for Incomplete Multimodal Learning. 21968-21977
- Yin Wang, Zhiying Leng, Frederick W. B. Li , Shun-Cheng Wu, Xiaohui Liang :
Fg-T2M: Fine-Grained Text-Driven Human Motion Generation via Diffusion Model. 21978-21987
- Zhiyu Zhu, Junhui Hou , Dapeng Oliver Wu :
Cross-modal Orthogonal High-rank Augmentation for RGB-Event Transformer-trackers. 21988-21998
- Mustafa Shukor, Corentin Dancette, Matthieu Cord:
eP-ALM: Efficient Perceptual Augmentation of Language Models. 21999-22012
- Fengyu Yang, Jiacheng Zhang, Andrew Owens:
Generating Visual Scenes from Touch. 22013-22023
- Xi Wei, Zhangxiang Shi, Tianzhu Zhang, Xiaoyuan Yu, Lei Xiao:
Multimodal High-order Relation Transformer for Scene Boundary Detection. 22024-22033
- Mia Chiquier, Carl Vondrick:
Muscles in Action. 22034-22044

-      Fei Ye, Adrian G. Bors:
Self-Evolved Dynamic Expansion Model for Task-Free Continual Learning. 22045-22055
-     Gengyuan Zhang, Jisen Ren, Jindong Gu, Volker Tresp:
Multi-event Video-Text Retrieval. 22056-22066
-     Fang Liu, Yuhao Liu, Yuqiu Kong, Ke Xu, Lihe Zhang, Baocai Yin, Gerhard P. Hancke, Rynson W. H. Lau:
Referring Image Segmentation Using Text Supervision. 22067-22077
-     Xiaobao Guo, Nithish Muthuchamy Selvaraj, Zitong Yu, Adams Wai-Kin Kong, Bingquan Shen, Alex C. Kot:
Audio-Visual Deception Detection: DOLOS Dataset and Parameter-Efficient Crossmodal Learning. 22078-22088
-     Shuai Tan, Bin Ji, Ye Pan:
EMMN: Emotional Motion Memory Network for Audio-driven Emotional Talking Face Generation. 22089-22099
-     Tianyu Huang, Bowen Dong, Yunhan Yang, Xiaoshui Huang, Rynson W. H. Lau, Wanli Ouyang, Wangmeng Zuo:
CLIP2Point: Transfer CLIP to Point Cloud Classification with Image-Depth Pre-Training. 22100-22110
-     Xiuzhe Wu, Pengfei Hu, Yang Wu, Xiaoyang Lyu, Yan-Pei Cao, Ying Shan, Wenming Yang, Zhongqian Sun, Xiaojuan Qi:
Speech2Lip: High-fidelity Speech to Lip Generation by Learning from a Short Video. 22111-22120
-     Xinchi Deng, Han Shi, Runhui Huang, Changlin Li, Hang Xu, Jianhua Han, James T. Kwok, Shen Zhao, Wei Zhang, Xiaodan Liang:
GrowCLIP: Data-aware Automatic Model Growing for Large-scale Contrastive Language-Image Pre-training. 22121-22132
-     Ziliang Chen, Xin Huang, Quanlong Guan, Liang Lin, Weiqi Luo:
A Retrospect to Multi-prompt Learning across Vision and Language. 22133-22144
-     Zhi-Qi Cheng, Qi Dai, Alexander G. Hauptmann:
ChartReader: A Unified Framework for Chart Derendering and Comprehension without Heuristic Rules. 22145-22156
-     Hong Li, Xingyu Li, Pengbo Hu, Yinuo Lei, Chunxiao Li, Yi Zhou:
Boosting Multi-modal Model Performance with Adaptive Gradient Modulation. 22157-22167
-     Maya Varma, Jean-Benoit Delbrouck, Sarah M. Hooper, Akshay Chaudhari , Curtis P. Langlotz:
ViLLA: Fine-Grained Vision-Language Representation Learning from Real-World Data. 22168-22178
-     Xiang Li, Jinglu Wang, Xiaohao Xu, Xiao Li, Bhiksha Raj, Yan Lu:
Robust Referring Video Object Segmentation with Cyclic Structural Consensus. 22179-22188

-      Rui Chen, Yongwei Chen, Ningxin Jiao, Kui Jia:
Fantasia3D: Disentangling Geometry and Appearance for High-quality Text-to-3D Content Creation. 22189-22199
-     Hongguang Zhu, Yunchao Wei, Xiaodan Liang, Chunjie Zhang, Yao Zhao:
CTP: Towards Vision-Language Continual Pretraining via Compatible Momentum Contrast and Topology Preservation. 22200-22210
-     Haibiao Xuan, Xiongzheng Li, Jinsong Zhang, Hongwen Zhang, Yebin Liu, Kun Li:
Narrator: Towards Natural Control of Human-Scene Interaction Generation via Relationship Reasoning. 22211-22221
-     Yu-Tong Cao, Ye Shi, Baosheng Yu, Jingya Wang, Dacheng Tao:
Knowledge-Aware Federated Active Learning with Non-IID Data. 22222-22232
-     Qin Liu, Zhenlin Xu, Gedas Bertasius, Marc Niethammer:
SimpleClick: Interactive Image Segmentation with Simple Vision Transformers. 22233-22243
-     You Huang, Hao Yang, Ke Sun, Shengchuan Zhang, Liujuan Cao, Guannan Jiang, Rongrong Ji:
InterFormer Real-time Interactive Image Segmentation. 22244-22254
-     Yifeng Huang, Viresh Ranjan, Minh Hoai:
Interactive Class-Agnostic Object Counting. 22255-22265
-     Otilia Stretcu, Edward Vendrow, Kenji Hata, Krishnamurthy Viswanathan, Vittorio Ferrari, Sasan Tavakkol, Wenlei Zhou, Aditya Avinash, Enming Luo, Neil Gordon Alldrin, MohammadHossein Bateni, Gabriel Berger, Andrew Bunner, Chun-Ta Lu, Javier A Rey, Giulia DeSalvo, Ranjay Krishna, Ariel Fuxman:
Agile Modeling: From Concept to Classifier in Minutes. 22266-22277
-     Seong Min Kye, Kwanghee Choi, Hyeongmin Byun, Buru Chang:
TiDAL: Learning Training Dynamics for Active Learning. 22278-22288
-     Jizhe Zhou, Xiaochen Ma, Xia Du, Ahmed Y. Al Hammadi, Wentao Feng:
Pre-training-free Image Manipulation Localization through Non-Mutually Exclusive Contrastive Learning. 22289-22299
-     Alexander Black, Simon Jenni, Tu Bui, Md. Mehrab Tanjim, Stefano Petrangeli, Ritwik Sinha, Viswanathan Swaminathan, John P. Collomosse:
VADER: Video Alignment Differencing and Retrieval. 22300-22310
-     Xin Deng, Chao Gao, Mai Xu:
PIRNet: Privacy-Preserving Image Restoration Network via Wavelet Lifting. 22311-22320
-     Binh Minh Le, Simon S. Woo:
Quality-Agnostic Deepfake Detection with Intra-model Collaborative Learning. 22321-22332
-     Yuanhao Zhai, Tianyu Luan, David S. Doermann, Junsong Yuan:
Towards Generic Image Manipulation Detection with Weakly-Supervised Self-Consistency Learning. 22333-22343

-      Ziyuan Luo, Qing Guo, Ka Chun Cheung, Simon See, Renjie Wan :
CopyRNeRF: Protecting the CopyRight of Neural Radiance Fields. 22344-22354
-     Zhiyuan Yan, Yong Zhang, Yanbo Fan, Baoyuan Wu:
UCF: Uncovering Common Features for Generalizable Deepfake Detection. 22355-22366
-     Zhihao Sun, Haoran Jiang, Danding Wang, Xirong Li, Juan Cao:
SAFL-Net: Semantic-Agnostic Feature Learning Network with Auxiliary Plugins for Image Manipulation Detection. 22367-22376
-     Xiaoxiao Hu, Qichao Ying, Zhenxing Qian, Sheng Li, Xinpeng Zhang:
DRAW: Defending Camera-shot RAW against Image Manipulation. 22377-22387
-     Zhendong Wang, Jianmin Bao, Wengang Zhou, Weilun Wang, Hezhen Hu, Hong Chen, Houqiang Li:
DIRE for Diffusion-Generated Image Detection. 22388-22398
-     Kaixiang Ji, Feng Chen, Xin Guo, Yadong Xu, Jian Wang, Jingdong Chen:
Uncertainty-guided Learning for Improving Image Manipulation Detection. 22399-22408
-     Pierre Fernandez, Guillaume Couairon, Hervé Jégou, Matthijs Douze, Teddy Furon:
The Stable Signature: Rooting Watermarks in Latent Diffusion Models. 22409-22420
-     Haoqi Wang, Zhizhong Li, Wayne Zhang:
Get the Best of Both Worlds: Improving Accuracy and Transferability by Grassmann Class Representation. 22421-22430
-     Minghan Zhu, Shizhong Han, Maani Ghaffari, Hong Cai, Fatih Porikli, Shubhankar Borse:
4D Panoptic Segmentation as Invariant and Equivariant Field Prediction. 22431-22441
-     Pierre Gleize, Weiyao Wang, Matt Feiszli:
SiLK: Simple Learned Keypoints. 22442-22451
-     Mohammad Zohaib, Alessio Del Bue:
SC3K: Self-supervised and Coherent 3D Keypoints Estimation from Rotated, Noisy, and Decimated Point Cloud Data. 22452-22462
-     Zhixiang Min, Juan Carlos Dibene, Enrique Dunn:
Geometric Viewpoint Learning with Hyper-Rays and Harmonics Encoding. 22463-22473
-     Congyi Zhang, Guying Lin, Lei Yang, Xin Li, Taku Komura, Scott Schaefer, John Keyser, Wenping Wang:
Surface Extraction from Neural Unsigned Distance Fields. 22474-22483
-     Avishkar Saha, Oscar Mendez, Chris Russell, Richard Bowden:
Learning Adaptive Neighborhoods for Graph Neural Networks. 22484-22493

-      Qingyang Wang, Michael A. Powell, Ali Geisa, Eric Bridgeford, Carey E. Priebe, Joshua T. Vogelstein:
Why do networks have inhibitory/negative connections? 22494-22502
-     Mingdeng Cao, Xintao Wang, Zhongang Qi, Ying Shan, Xiaohu Qie, Yinqiang Zheng:
MasaCtrl: Tuning-Free Mutual Self-Attention Control for Consistent Image Synthesis and Editing. 22503-22513
-     Shuyi Jiang, Daochang Liu, Dingquan Li, Chang Xu:
Personalized Image Generation for Color Vision Deficiency Population. 22514-22523
-     Yingyan Xu, Gaspard Zoss, Prashanth Chandran, Markus H. Gross , Derek Bradley, Paulo F. U. Gotardo:
ReNeRF: Relightable Neural Radiance Fields with Nearfield Lighting. 22524-22534
-     Jing Zhao , Heliang Zheng, Chaoyue Wang, Long Lan, Wenjing Yang:
MagicFusion: Boosting Text-to-Image Generation Performance by Fusing Diffusion Models. 22535-22545
-     Gwanghyun Kim, Ji Ha Jang, Se Young Chun:
PODIA-3D: Domain Adaptation of 3D Generative Model Across Large Domain Gap Using Pose-Preserved Text-to-Image Diffusion. 22546-22555
-     Peipei Li, Rui Wang, Huaibo Huang, Ran He, Zhaofeng He:
Pluralistic Aging Diffusion Autoencoder. 22556-22566
-     Zezeng Li, Shenghao Li, Zhanpeng Wang, Na Lei, Zhongxuan Luo, Xianfeng David Gu:
DPM-OT: A New Diffusion Probabilistic Model Based on Optimal Transport. 22567-22576
-     Yuan Gan, Zongxin Yang, Xihang Yue, Lingyun Sun, Yi Yang:
Efficient Emotional Adaptation for Audio-Driven Talking-Head Generation. 22577-22588
-     Puntawat Ponglertnapakorn, Nontawat Tritrong, Supasorn Suwajanakorn:
DiFaReli: Diffusion Face Relighting. 22589-22600
-     Yuting Xu, Jian Liang, Gengyun Jia, Ziming Yang, Yanhao Zhang, Ran He:
TALL: Thumbnail Layout for Deepfake Video Detection. 22601-22611
-     Binbin Yang, Yi Luo, Ziliang Chen, Guangrun Wang, Xiaodan Liang, Liang Lin:
LAW-Diffusion: Complex Scene Generation by Diffusion with Layouts. 22612-22622
-     Johanna Karras, Aleksander Holynski, Ting-Chun Wang, Ira Kemelmacher-Shlizerman:
DreamPose: Fashion Image-to-Video Synthesis via Stable Diffusion. 22623-22633
-     Nupur Kumari, Bingliang Zhang, Sheng-Yu Wang, Eli Shechtman, Richard Zhang, Jun-Yan Zhu:
Ablating Concepts in Text-to-Image Diffusion Models. 22634-22645

-      Yu Chen, Gim Hee Lee:
DReg-NeRF: Deep Registration for Neural Radiance Fields. 22646-22656
-     Lingxiao Li, Yi Zhang, Shuhui Wang:
The Euclidean Space is Evil: Hyperbolic Attribute Editing for Few-shot Image Generation. 22657-22667
-     Idan Schwartz, Vésteinn Snæbjarnarson, Hila Chefer, Serge J. Belongie , Lior Wolf, Sagie Benaim:
Discriminative Class Tokens for Text-to-Image Diffusion Models. 22668-22678
-     Bin Cheng, Zuhao Liu, Yunbo Peng, Yue Lin:
General Image-to-Image Translation with One-Shot Image Guidance. 22679-22689
-     Yuming Jiang, Shuai Yang, Tong Liang Koh, Wayne Wu, Chen Change Loy, Ziwei Liu:
Text2Performer: Text-Driven Human Video Generation. 22690-22700
-     Kibeom Hong, Seogkyu Jeon, Junsoo Lee, Namhyuk Ahn, Kunhee Kim, Pilhyeon Lee, Daesik Kim, Youngjung Uh, Hyeran Byun:
AesPA-Net: Aesthetic Pattern-Aware Style Transfer Networks. 22701-22710
-     Xiao Han, Xiatian Zhu, Jiankang Deng, Yi-Zhe Song, Tao Xiang:
Controllable Person Image Synthesis with Pose-Constrained Latent Diffusion. 22711-22720
-     Saman Motamed, Jianjin Xu, Chen Henry Wu, Christian Häne, Jean-Charles Bazin, Fernando De la Torre:
PATMAT: Person Aware Tuning of Mask-Aware Transformer for Face inpainting. 22721-22730
-     Zhi Li, Pengfei Wei, Xiang Yin, Zejun Ma, Alex C. Kot:
Virtual Try-On with Pose-Garment Keypoints Guided Inpainting. 22731-22740
-     Chuanxia Zheng, Andrea Vedaldi:
Online Clustered Codebook. 22741-22750
-     Chieh Hubert Lin, Hsin-Ying Lee, Willi Menapace, Menglei Chai, Aliaksandr Siarohin, Ming-Hsuan Yang, Sergey Tulyakov:
InfiniCity: Infinite-Scale City Synthesis. 22751-22761
-     Junshu Tang, Tengfei Wang, Bo Zhang, Ting Zhang, Ran Yi, Lizhuang Ma, Dong Chen:
Make-It-3D: High-Fidelity 3D Creation from A Single Image with Diffusion Prior. 22762-22772
-     Xiaoyu Zhou, Zhiwei Lin, Xiaojun Shan, Yongtao Wang, Deqing Sun, Ming-Hsuan Yang:
SAMPLING: Scene-adaptive Hierarchical Multiplane Images Representation for Novel View Synthesis from a Single Image. 22773-22783
-     Taekyung Ki, Dongchan Min:
StyleLipSync: Style-based Personalized Lip-sync Video Generation. 22784-

-      22793
Yuhan Wang , Liming Jiang, Chen Change Loy:
StyleInV: A Temporal Style Modulated Inversion Network for Unconditional Video Generation. 22794-22804
-     Kyungmin Jo, Wonjoon Jin, Jaegul Choo, Hyunjoon Lee, Sunghyun Cho:
3D-Aware Generative Model for Improved Side-View Image Synthesis.
22805-22815
-     Serin Yang, Hyunmin Hwang, Jong Chul Ye:
Zero-Shot Contrastive Loss for Text-Guided Diffusion Image Style Transfer.
22816-22825
-     Seunghyeon Seo, Yeonjin Chang, Nojun Kwak:
FlipNeRF: Flipped Reflection Rays for Few-shot Novel View Synthesis.
22826-22836
-     Jean Prost, Antoine Houdard, Andrés Almansa, Nicolas Papadakis:
Inverse problem regularization with hierarchical variational autoencoders.
22837-22848
-     Hyunsu Kim, Gayoung Lee, Yunjey Choi, Jin-Hwa Kim, Jun-Yan Zhu:
3D-aware Blending with Generative NeRFs. 22849-22861
-     Youjia Zhang, Teng Xu, Junqing Yu, Yuteng Ye, Yanqing Jing, Junle Wang, Jingyi Yu, Wei Yang:
NeMF: Inverse Volume Rendering with Neural Microflake Field. 22862-22872
-     Songwei Ge, Seungjun Nah, Guilin Liu, Tyler Poon, Andrew Tao, Bryan Catanzaro, David Jacobs, Jia-Bin Huang, Ming-Yu Liu, Yogesh Balaji:
Preserve Your Own Correlation: A Noise Prior for Video Diffusion Models.
22873-22884
-     Junting Dong, Qi Fang, Tianshuo Yang, Qing Shuai, Chengyu Qiao, Sida Peng:
iVS-Net: Learning Human View Synthesis from Internet Videos. 22885-22894
-     Qiushan Guo, Chuofan Ma, Yi Jiang, Zehuan Yuan, Yizhou Yu, Ping Luo:
EGC: Image Generation and Classification via a Diffusion Energy-Based Model. 22895-22905
-     Wenpeng Xiao, Wentao Liu, Yitong Wang, Bernard Ghanem , Bing Li:
Automatic Animation of Hair Blowing in Still Portrait Photos. 22906-22918
-     Animesh Karnewar, Niloy J. Mitra, Andrea Vedaldi, David Novotný:
HoloFusion: Towards Photo-realistic 3D Generative Modeling. 22919-22928
-     Bo Zhang, Jiacheng Sui, Li Niu:
Foreground Object Search by Distilling Composite Image Feature. 22929-22938
-     Honglin He, Zhuoqian Yang, Shikai Li, Bo Dai, Wayne Wu:
OrthoPlanes: A Novel Representation for Better 3D-Awareness of GANs.
22939-22950

-      Zhuoqian Yang, Shikai Li, Wayne Wu, Bo Dai:
3DHumanGAN: 3D-Aware Human Image Generation with 3D Pose Mapping. 22951-22962
-     Yunfei Liu, Lijian Lin, Fei Yu, Changyin Zhou, Yu Li:
MODA: Mapping-Once Audio-driven Portrait Animation with Dual Atentions. 22963-22972
-     Zhuofan Zhang, Zhen Liu, Ping Tan, Bing Zeng, Shuaicheng Liu:
Minimum Latency Deep Online Video Stabilization. 22973-22982
-     Wenhao Chai, Xun Guo, Gaoang Wang, Yan Lu:
StableVideo: Text-driven Consistency-aware Diffusion Video Editing. 22983-22993
-     Or Patashnik, Daniel Garabi, Idan Azuri, Hadar Averbuch-Elor, Daniel Cohen-Or:
Localizing Object-level Shape Variations with Text-to-Image Diffusion Models. 22994-23004
-     Fa-Ting Hong, Dan Xu:
Implicit Identity Representation Conditioned Memory Compensation Network for Talking Head Video Generation. 23005-23015
-     Mingjin Zhang, Chi Zhang, Qiming Zhang, Jie Guo, Xinbo Gao, Jing Zhang:
ESSAformer: Efficient Transformer for Hyperspectral Image Super-resolution. 23016-23027
-     Can Qin, Ning Yu, Chen Xing, Shu Zhang, Zeyuan Chen, Stefano Ermon, Yun Fu, Caiming Xiong, Ran Xu:
GlueGen: Plug and Play Multi-modal Encoders for X-to-image Generation. 23028-23039
-     Quewei Li, Feichao Li, Jie Guo, Yanwen Guo:
UHDNeRF: Ultra-High-Definition Neural Radiance Fields. 23040-23051
-     Mingrui Zhu, Xiao He, Nannan Wang, Xiaoyu Wang, Xinbo Gao:
All-to-key Attention for Arbitrary Style Transfer. 23052-23062
-     Ahmet Burak Yildirim, Hamza Pehlivan, Bahri Batuhan Bilecen, Aysegul Dundar:
Diverse Inpainting and Editing with GAN Inversion. 23063-23073
-     Yi-Hsin Chen, Si-Cun Chen, Yi-Hsin Chen, Yen-Yu Lin, Wen-Hsiao Peng:
MoTIF: Learning Motion Trajectories with Local Implicit Neural Functions for Continuous Space-Time Video Super-Resolution. 23074-23084
-     Umar Iqbal, Akin Caliskan, Koki Nagano, Sameh Khamis, Pavlo Molchanov, Jan Kautz:
RANA: Relightable Articulated Neural Avatars. 23085-23096
-     Xujie Zhang, Binbin Yang, Michael C. Kampffmeyer, Wenqing Zhang, Shiyue Zhang, Guansong Lu, Liang Lin, Hang Xu, Xiaodan Liang:
DiffCloth: Diffusion Based Garment Synthesis and Manipulation via Structural Cross-modal Semantic Alignment. 23097-23106
-     Shanghua Gao, Pan Zhou, Ming-Ming Cheng, Shuicheng Yan:
Masked Diffusion Transformer is a Strong Image Synthesizer. 23107-23116

-      Jiwen Yu, Yinhuai Wang, Chen Zhao, Bernard Ghanem , Jian Zhang:
FreeDoM: Training-Free Energy-Guided Conditional Diffusion Model. 23117-23127
-     Zhipeng Cai, Matthias Müller:
CLNeRF: Continual Learning Meets NeRF. 23128-23137
-     Tianyi Chu, Jiafu Chen, Jiakai Sun, Shuobin Lian, Zhizhong Wang, Zhiwen Zuo, Lei Zhao, Wei Xing, Dongming Lu:
Rethinking Fast Fourier Convolution in Image Inpainting. 23138-23148
-     Duygu Ceylan, Chun-Hao Paul Huang, Niloy J. Mitra:
Pix2Video: Video Editing using Image Diffusion. 23149-23160
-     Yu Qiao, Bo Dong, Ao Jin, Yu Fu, Seung-Hwan Baek, Felix Heide, Pieter Peers, Xiaopeng Wei, Xin Yang:
Multi-view Spectral Polarization Propagation for Video Glass Segmentation. 23161-23171
-     Guillaume Le Moing, Jean Ponce, Cordelia Schmid:
WALDO: Future Video Synthesis using Object Layer Decomposition and Parametric Flow Prediction. 23172-23184
-     Eric Ming Chen, Sidhanth Holalkere, Ruyu Yan, Kai Zhang, Abe Davis:
Ray Conditioning: Trading Photo-consistency for Photo-realism in Multi-view Image Generation. 23185-23194
-     Jaewoong Lee, Sangwon Jang, Jaehyeong Jo, Jaehong Yoon, Yunji Kim, Jin-Hwa Kim, Jung-Woo Ha, Sung Ju Hwang:
Text-Conditioned Sampling Framework for Text-to-Image Generation with Masked Generative Models. 23195-23205
-     Aram Davtyan , Sepehr Sameni, Paolo Favaro:
Efficient Video Prediction via Sparsely Conditioned Flow Matching. 23206-23217
-     Pinaki Nath Chowdhury, Ayan Kumar Bhunia, Aneeshan Sain, Subhadeep Koley, Tao Xiang, Yi-Zhe Song:
Democratising 2D Sketch to 3D Shape Retrieval Through Pivoting. 23218-23229
-     Chun-Mei Feng, Kai Yu, Nian Liu, Xinxing Xu, Salman Khan, Wangmeng Zuo:
Towards Instance-adaptive Inference for Federated Learning. 23230-23239
-     Yi-Hsin Chen, Ying-Chieh Weng, Chia-Hao Kao, Cheng Chien, Wei-Chen Chiu, Wen-Hsiao Peng:
TransTIC: Transferring Transformer-based Image Compression from Human Perception to Machine Perception. 23240-23250
-     Zhi-Kai Huang, Wei-Ting Chen, Yuan-Chun Chiang, Sy-Yen Kuo, Ming-Hsuan Yang:
Counting Crowds in Bad Weather. 23251-23262
-     Chenfeng Xu, Bichen Wu, Ji Hou, Sam S. Tsai, Ruilong Li, Jialiang Wang, Wei Zhan, Zijian He, Peter Vajda, Kurt Keutzer, Masayoshi Tomizuka:
NeRF-Det: Learning Geometry-Aware Volumetric Representation for Multi-View 3D Object Detection. 23263-23273

-      Najmeh Sadoughi, Xinyu Li, Avijit Vajpayee, David Fan, Bing Shuai, Hector J. Santos-Villalobos, Vimal Bhat, Rohith MV:
MEGA: Multimodal Alignment Aggregation and Distillation For Cinematic Video Segmentation. 23274-23283
-     Nanxuan Zhao, Shengqi Dang, Hexun Lin, Yang Shi, Nan Cao:
Bring Clipart to Life. 23284-23293
-     Sunwook Hwang, Youngseok Kim, Seongwon Kim, Saewoong Bahk, Hyung-Sin Kim:
UpCycling: Semi-supervised 3D Object Detection without Sharing Raw-level Unlabeled Scenes. 23294-23304
-     Yijie Lin, Mouxing Yang, Jun Yu, Peng Hu, Changqing Zhang, Xi Peng:
Graph Matching with Bi-level Noisy Correspondence. 23305-23314
-     Woosang Shin, Jonghyeon Lee, Taehan Lee, Sangmoon Lee, Jong Pil Yun:
Anomaly Detection using Score-based Perturbation Resilience. 23315-23325
-     Kun Yang, Dingkang Yang, Jingyu Zhang, Mingcheng Li, Yang Liu, Jing Liu, Hanqi Wang, Peng Sun, Liang Song:
Spatio-Temporal Domain Awareness for Multi-Agent Collaborative Perception. 23326-23335
-     Alberto Baldrati, Davide Morelli, Giuseppe Cartella, Marcella Cornia, Marco Bertini, Rita Cucchiara:
Multimodal Garment Designer: Human-Centric Latent Diffusion Models for Fashion Image Editing. 23336-23345
-     Zhihong Chen, Shizhe Diao, Benyou Wang, Guanbin Li, Xiang Wan:
Towards Unifying Medical Vision-and-Language Pre-training via Soft Prompts. 23346-23356
-     Weiming Zhuang, Yonggang Wen, Lingjuan Lyu, Shuai Zhang:
MAS: Towards Resource-Efficient Federated Multiple-Task Learning. 23357-23367
-     Jinglun Li, Xinyu Zhou, Pinxue Guo, Yixuan Sun, Yiwen Huang, Weifeng Ge, Wenqiang Zhang:
Hierarchical Visual Categories Modeling: A Joint Representation Learning and Density Estimation Framework for Out-of-Distribution Detection. 23368-23378
-     Siao Liu, Zhaoyu Chen, Yang Liu, Yuzheng Wang, Dingkang Yang, Zhile Zhao, Ziqing Zhou, Xie Yi, Wei Li, Wenqiang Zhang, Zhongxue Gan:
Improving Generalization in Visual Reinforcement Learning via Conflict-aware Gradient Agreement Augmentation. 23379-23389
-     Linfeng Zhang, Kaisheng Ma:
Tiny Updater: Towards Efficient Neural Network-Driven Software Updating. 23390-23402
-     Zhicheng Zhang, Shengzhe Liu, Jufeng Yang:
Multiple Planar Object Tracking. 23403-23413

-      Geng Lin , Chen Gao, Jia-Bin Huang, Changil Kim, Yipeng Wang, Matthias Zwicker, Ayush Saraf:
OmnimatteRF: Robust Omnimatte with 3D Background Modeling. 23414-23423
-     Changsong Wen, Xin Zhang, Xingxu Yao, Jufeng Yang:
Ordinal Label Distribution Learning. 23424-23434
-     Yichao Cao, Qingfei Tang, Feng Yang, Xiu Su, Shan You, Xiaobo Lu, Chang Xu:
Re-mine, Learn and Reason: Exploring the Cross-modal Semantic Correlations for Language-guided HOI detection. 23435-23446
-     Zhixuan Li, Weining Ye, Juan Terven, Zachary Bennett, Ying Zheng, Tingting Jiang, Tiejun Huang:
MUVA: A New Large-Scale Benchmark for Multi-view Amodal Instance Segmentation in the Shopping Scenario. 23447-23456
-     Ye Chen, Bingbing Ni, Xuanhong Chen, Zhangli Hu:
Editable Image Geometric Abstraction via Neural Primitive Assembly. 23457-23466
-     Manuel S. Drehwald, Sagi Eppel, Jolina Li, Han Hao, Alán Aspuru-Guzik:
One-shot recognition of any material anywhere using contrastive learning with physics-based rendering. 23467-23476
-     Weiyue Zhao, Xin Li, Zhan Peng, Xianrui Luo, Xinyi Ye, Hao Lu, Zhiguo Cao:
Fast Full-frame Video Stabilization with Iterative Optimization. 23477-23487
-     Bohai Gu, Heng Fan, Libo Zhang:
Two Birds, One Stone: A Unified Framework for Joint Learning of Image and Video Style Transfers. 23488-23497
-     Bing Cao, Yiming Sun, Pengfei Zhu, Qinghua Hu:
Multi-modal Gated Mixture of Local-to-Global Experts for Dynamic Image Fusion. 23498-23507
-     Samuel Wilson, Tobias Fischer , Feras Dayoub , Dimity Miller , Niko Sünderhauf :
SAFE: Sensitivity-Aware Features for Out-of-Distribution Object Detection. 23508-23519
-     Can Zhang, Gim Hee Lee:
GeT: Generative Target Structure Debiasing for Domain Adaptation. 23520-23531
-     Tianyi Wei, Dongdong Chen, Wenbo Zhou, Jing Liao, Weiming Zhang, Gang Hua, Nenghai Yu:
HairCLIPv2: Unifying Hair Editing via Proxy Feature Blending. 23532-23542
-     Qichen Fu, Xingyu Liu, Ran Xu, Juan Carlos Niebles, Kris M. Kitani:
Deformer: Dynamic Fusion Transformer for Robust Hand Pose Estimation. 23543-23554
-     Fangyun Wei, Yutong Chen:
Improving Continuous Sign Language Recognition with Cross-Lingual Signs.



23555-23564



Jiawei Lin, Jiaqi Guo, Shizhao Sun, Weijiang Xu, Ting Liu, Jian-Guang Lou, Dongmei Zhang:

A Parse-Then-Place Approach for Generating Graphic Layouts from Textual Descriptions. 23565-23574



Tzofi Klinghoffer, Kushagra Tiwary, Nikhil Behari, Bhavya Agrawalla, Ramesh Raskar:

DISeR: Designing Imaging Systems with Reinforcement Learning. 23575-23585



Wei Liao:

Segmentation of Tubular Structures Using Iterative Training with Tailored Samples. 23586-23595



Urbano Miguel Nunes, Laurent Udo Perrinet, Sio-Hoi Ieng:

Time-to-Contact Map by Joint Estimation of Up-to-Scale Inverse Depth and Global Motion using a Single Event Camera. 23596-23606



last updated on 2024-05-27 20:45 CEST by the dblp team



all metadata released as open data under CC0 1.0 license

see also: [Terms of Use](#) | [Privacy Policy](#) | [Imprint](#)

dblp was originally created in 1993 at:



since 2018, dblp has been operated and maintained by:



the dblp computer science bibliography is funded and supported by:



Nationale
Forschungsdaten
Infrastruktur