PROGRAMME GUIDE

ECCV

EUROPEAN CONFERENCE ON COMPUTER VISION

2024

MAIN CONFERENCE

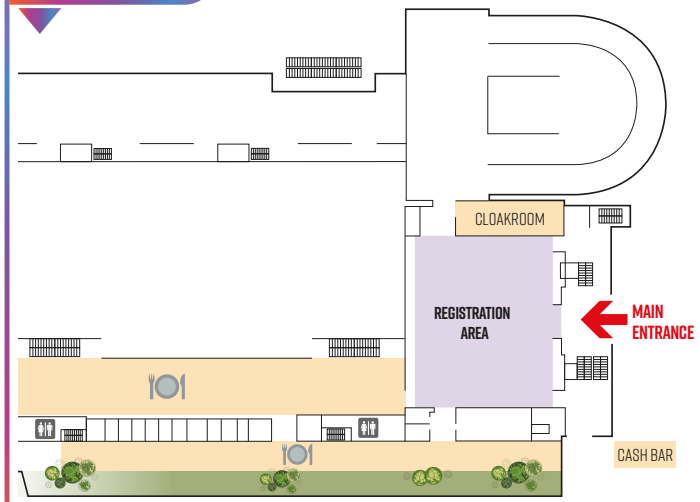
ECCV 2024
EUROPEAN CONFERENCE ON COMPUTER VISION

Level 0 - Exhibition Area

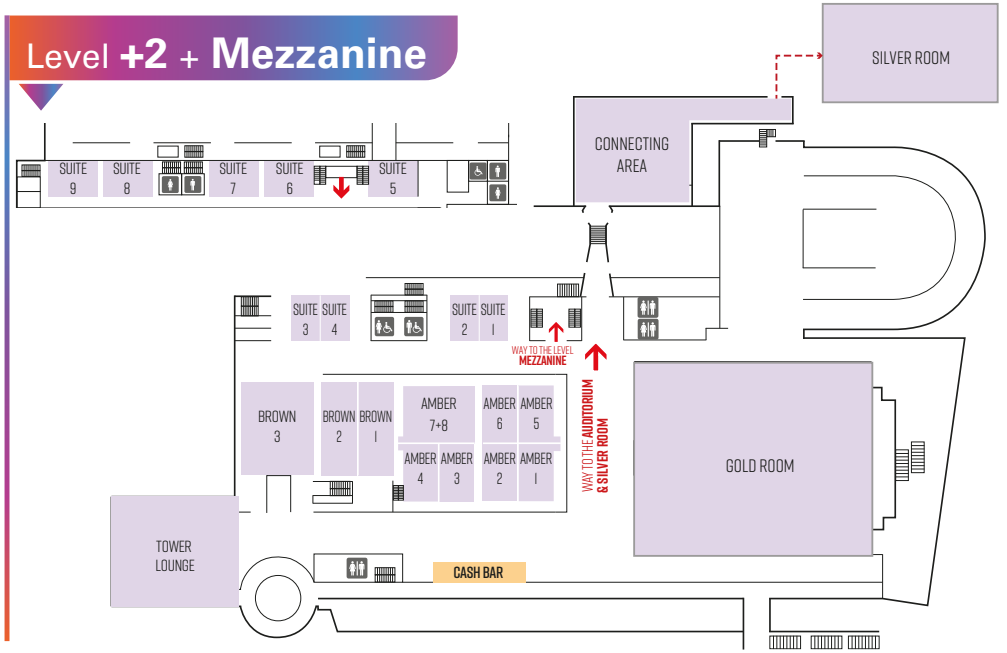
EXHIBITORS

9	3dMD
71	Abaka AI
40	Advex AI
43	Ant Research
34	Apple
33	Baidu
49	Bending Spoons
19+20	ByteDance
73	Cinemersive Labs
51	Covision Media
53+54	Encord
38	EVS Embedded Vision Systems
41	Google Research
31	HPC-AI TECH
52	Huawei
48	INSAIT Institue for Computer Science, AI and Technology
21	IO Industries Inc.
25+26, 27, 28, 29, 30	Istituto Italiano di Tecnologia
84	ItalAI S.r.l.
75	Keylabs
18	LabelBox
44	LatticeFlow AI
39	Lightly
66	Lightning AI
80	Living Optics
83	MDPI AG
62	Meshcapade
42	Meta
17	Move4D
7	OneSource Cloud
8	Parallel Domain
50	Springer Nature
45	SuperAnnotate
72	Tenyks
81	The Institution of Engineering and Technology
46+47	Three Lines of Code (3lc)
74	University of Science and Technology of China
82	Visual Layer
5+6	Voxel51
32	Weights and Biases

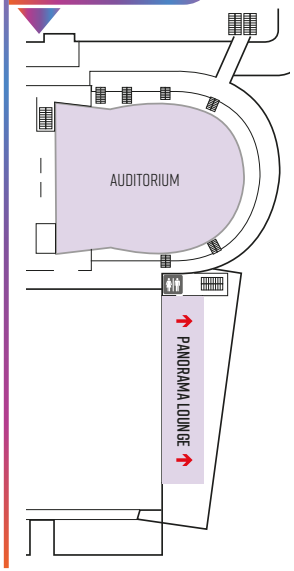
Level +1



Level +2 + Mezzanine



Level +3



ECCV 2024 ORGANIZING COMMITTEE

General Chairs
Andrew Fitzgibbon (Graphcore)
Laura Leal-Taixé (NVIDIA)
Vittorio Murino (University of Verona and University of Genova, Italy)

Program Chairs
Aleš Leonardis (University of Birmingham)
Elisa Ricci (University of Trento)
Gül Varol (Ecole des Ponts ParisTech)
Olga Russakovsky (Princeton University)
Stefan Roth (TU Darmstadt)
Torsten Sattler (Czech Technical University in Prague)

Workshop & Tutorial Chairs
Alessio Del Bue (Istituto Italiano di Tecnologia)
Cristian Canton (Meta AI)
Jordi Pont-Tuset (Google DeepMind)
Tatiana Tommasi (Politecnico di Torino)

Demo Chairs
Hyung Jin Chang (University of Birmingham)
Marco Cristani (University of Verona)

Publication Chairs
Mahmoud Ali (Inria)
Francois Bremond (Inria)
Jovita Lukasik (University of Siegen)
Michael Moeller (University of Siegen)

Poster Chairs
Aljoša Ošep (Carnegie Mellon University)
Zuzana Kukelova (Czech Technical University in Prague)

Diversity Chairs
David Fouhey (New York University)
Rita Cucchiara
(Università di Modena e Reggio Emilia)

Ethics Review Committee
Chloé Bakalar (Meta)
Kate Saenko (Boston University)
Remi Denton (Google Research)
Yisong Yue (Caltech, Asari AI & Latitude AI)

Conference Ombud
Georgia Gkioxari
(California Institute of Technology)
Greg Mori (Borealis AI / SFU)

Social Activities Chair
Giovanni Maria Farinella
(University of Catania, Italy)
Raffaella Lanzarotti
(Università degli Studi di Milano)
Simone Bianco
(University of Milano-Bicocca)

Doctoral Consortium Chairs
Cigdem Beyan (University of Verona, Italy)
Or Litany (NVIDIA / Technion)

Local Chairs
Raffaella Lanzarotti
(Università degli Studi di Milano)
Simone Bianco
(University of Milano-Bicocca)

Industry Liaison Chairs
Cees Snoek (University of Amsterdam)
Shaogang Gong
(Queen Mary University of London)

Publicity and social media chair
Konstantinos Derpanis (York University, Samsung AI Centre Toronto)

Finance Chairs
Gerard Medioni (Amazon)
Nicole Finn (c to c events)

Web Developer
Lee Campbell (Eventhosts)

Tech Chair
Sascha Hornauer (Mines Paris – PSL)

Welcome to ECCV 2024!

It is our great pleasure to welcome you to the 18th European Conference on Computer Vision (ECCV 2024), which will takes place in the dynamic and historic city of Milan, Italy, from September 29th to October 4th, 2024.

As one of the leading global forums for computer vision, machine learning, artificial intelligence, and related fields, ECCV brings together a vibrant community of researchers and practitioners. This year, the program features an exceptional lineup, including keynotes from distinguished speakers, oral and poster sessions, workshops, tutorials, industry demonstrations, and exhibitions. These events offer a fantastic opportunity to engage with cutting-edge research and foster meaningful connections.

ECCV 2024 has attracted an unprecedented number of submissions, reflecting the ever-growing interest and advances in our field. With over 8,585 submissions, we are excited to announce that 2,387 papers have been accepted for publication, thanks to the tireless work of our Program Chairs, Area Chairs (ACs), and an incredible team of expert reviewers. Among these accepted papers, 200 have been selected for oral presentations, showcasing some of the most innovative and impactful research being conducted today.

Though the conference will take place primarily in person, we recognize the importance of inclusivity. For those unable to attend in Milan, we have designed a virtual component that will allow remote participants to access key content, including keynote talks and oral presentations, ensuring that the conference remains accessible to all. Additionally, in our commitment to inclusivity, we are offering travel grants to support attendees from low-income economies or those facing financial difficulties, helping ensure broader participation in ECCV 2024.

Milan, a city that perfectly blends tradition with modernity, is a fitting host for ECCV 2024. We encourage you to take time outside of the conference to explore its many cultural treasures—from the iconic Duomo and world-renowned museums to its vibrant culinary scenery.

The success of ECCV 2024 would not be possible without the extraordinary effort of so many individuals. We extend our deepest appreciation to the organizing committee, the reviewers, the authors, and our sponsors.

Finally, we offer our heartfelt thanks to you – the attendees. Your participation, whether in person or virtually, is what makes ECCV such a special event. We are confident that you will find the conference enriching, engaging, and inspiring.

We eagerly anticipate seeing you in Milan for what promises to be a memorable and impactful ECCV 2024!

General Chairs
Andrew Fitzgibbon (Graphcore)
Laura Leal-Taixé (NVIDIA)
Vittorio Murino
(University of Verona and Genova, Italy)

Program Chairs
Aleš Leonardis (University of Birmingham)
Elisa Ricci (University of Trento)
Gül Varol (Ecole des Ponts ParisTech)
Olga Russakovsky (Princeton University)
Stefan Roth (TU Darmstadt)
Torsten Sattler (Czech Technical University in Prague)

Please visit the ECCV 2024 Virtual Conference Website:

ECCV 2024 AREA CHAIRS

A. Sophia Koepke
Abby Stylianou
Abhinav Shrivastava
Adriana Kovashka
Ahmet Iscen
Aishwarya Agrawal
Ajay Kumar
Ajmal Mian
Akihiro Sugimoto
Aliaksandr Siarohin
Alireza Fathi
Aljosa Osep
Andre Araujo
Andrea Fusiello
Andrea Vedaldi
Andrei Bursuc
Andrés Bruhn
Andrew Owens
Andrew Zisserman
Angela Dai
Angela Yao
Anh Tran
Anna Khoreva
Anna Rohrbach
Anoop Cherian
Anthony Hoogs
Antitza Dantcheva
Antoni Chan
Anurag Mittal
Arsha Nagrani
Aswin Sankaranarayanan
Atsuto Maki
Ayan Chakrabarti
Ayellet Tal
Bastian Leibe
Basura Fernando
Benjamin Busam
Benjamin Kimia
Bernt Schiele
Bharat Bhatnagar
Bin Fan
Bjorn Stenger
Bo Wang
Bohyung Han
Bolei Zhou
Boxin Shi
Brian Price
Bryan Plummer
Bryan Russell
Bumsub Ham
C.V. Jawahar
Cristian Canton
Carl Olsson
Carsten Rother
Chang D. Yoo
Charless Fowlkes
Chen Sun
Chen Change Loy
Chi-Keung Tang
Christian Rupprecht
Christian Wolf
Christoph Feichtenhofer
Chuang Gan
Cigdem Beyan
Concetto Spampinato
Cordelia Schmid
Dan Xu
Daniel Zoran
David Fouhey
David Picard
Davide Modolo
Davide Moltisanti
Deqing Sun
Despoina Paschalidou
Deva Ramanan
Devendra Singh Chaplot
Diane Larlus
Dima Damen
Dimitrios Tzionas
Dimitris Samaras
Djamila Aouada
Du Tran
Eddy Ilg
Edmond Boyer
Efstratios Gavves
Ehsan Adeli
Eli Shechtman
Enrique Dunn
Eric Brachmann
Evan Shelhamer
Evangelos Kalogerakis
Fabio Galasso
Fabio Poiesi
Fahad Shahbaz Khan
Fatih Porikli
Fatma Guney
Federica Bogo
Federico Tombari
Felix Heide
Feng Lu
Francesc Moreno
Francesca Odone
Fredrik Kahl
Friedrich Fraundorfer
Fuxin Li
Gabriel Brostow
Gang Hua
Gemma Roig
Georgios Pavlakos
Gerard Pons-Moll
Gim Hee Lee
Giorgos Tolias
Giovanni Maria Farinella
Golnaz Ghiasi
Greg Mori
Gunhee Kim
Guo-Jun Qi
Haibin Ling
Hajime Nagahara
Hamed Pirsiavash
Hamid Rezatofighi
Hazel Doughty
Hedvig Kjellström
Helge Rhodin
Heng Wang
Hengshuang Zhao
Hilde Kuehne
Hiroshi Kawasaki
Hisham Cholakkal
Holger Caesar
Hongbin Zha
Hongdong Li
Hossein Rahmani
Huan Fu
Huijuan Xu
Iasonas Kokkinos
Iddo Drori
Ilke Demir
In Kyu Park
Iro Armeni
Iro Laina
Ishan Misra
Ismini Lourentzou
Jaesik Park
James Tompkin
Jan van Gemert
Jana Kosecka
Javier Vazquez-Corral
Jean Ponce
Jean-Francois Lalonde
Jean-Marc Odobez
Jia-Bin Huang
Jiajun Wu
Jian Sun
Jianbo Shi
Jianfei Cai
Jiangxin Dong
Jianwen Xie
Jianxin Wu
Jiaolong Yang
Jiaya Jia
Jiaying Liu
Jifeng Dai
Jing Zhang
Jingdong Wang
Jingya Wang
Jingyi Yu
Jinshan Pan
Jiri Matas
Joachim Denzler
Joao Carreira
Jongwoo Lim
Joost van de Weijer
Jordi Pont-Tuset
Jose Alvarez
Jose Dolz
Jun Liu
Jungseock Joo
Junhwa Hur
Junmo Kim
Junsong Yuan
Jürgen Gall
Justus Thies
Kai Han
Kaiyang Zhou
Kaiyu Yang
Kaleem Siddiqi
Karl Åström
Karteek Alahari
Katerina Fragkiadaki
Katherine Bouman
Kenneth Marino
Kevis-Kokitsi Maninis
Kiriakos Kutulakos
Ko Nishino
Konstantinos Derpanis
Kostas Daniilidis
Krishna Kumar Singh
Krystian Mikolajczyk
Kwang Moo Yi
Lamberto Ballan
Laszlo Jeni
Laura Sevilla-Lara
Laurens van der Maaten
Laurent Kneip
Le Lu
Lei Zhang
Lei Zhu
Leonid Sigal
Liang Zheng
Liang-Chieh Chen
Liangliang Nan
Liangyan Gui
Lianli Gao
Linchao Bao
Linchao Zhu
Linjie Yang
Loic Landrieu
Long Chen
Long Quan
Lu Sheng
Lu Yuan
Lubomir Bourdev
Luca Magri
Luisa Verdoliva
Luping Zhou
Mahdi Hosseini
Makarand Tapaswi
Mang Ye
Manmohan Chandraker
manohar paluri
Manolis Savva
Marc Pollefeys
Marcella Cornia
Marcello Pelillo
Marco Cristani
Marcus Rohrbach

Margret Keuper
Maria Vakalopoulou
Martin R. Oswald
Massimiliano Mancini
Mathieu Aubry
Mathieu Salzmann
Matthew Blaschko
Matthew O'Toole
Matthew Trager
Mayank Vatsa
Mei Chen
Miaomiao Liu
Michael Brown
Michael Maire
Michael Niemeyer
Michael Rubinstein
Michael Ryoo
Michael Wray
Michele Merler
Michele Nappi
Mike Zheng Shou
Min Sun
Min H. Kim
Ming-Hsuan Yang
Minh Ha Quang
Minsu Cho
Mohamed Elhoseiny
Mohit Gupta
Nalini Ratha
Nassir Navab
Natalia Neverova
Nathan Jacobs
Naveed Akhtar
Nazli Ikizler-Cinbis
Negar Rostamzadeh
Nicoletta Noceti
Nicu Sebe
Niki Martinel
Niloy Mitra
Ning Yu
Ning Zhang
Nuno Vasconcelos
Octavia Camps
Oisin Mac Aodha
Olga Veksler
Olivia Wiles
Ondrej Chum
Or Litany
Orazio Gallo
Oren Freifeld
Oriane Siméoni
P. J. Narayanan
Pablo Arbelaez
Paolo Rota
Pascal Mettes
Peng Hu
Peter Gehler
Philippos Mordohai
Pietro Moreri
Ping Tan
Piotr Koniusz
Qi Shan
Qi Yu
Qi Zhao
Qianru Sun
Qin Jin
Qixing Huang
Radu Timofte
Ram Nevatia
Raoul de Charette
Renaud Marlet
Rene Vidal
Reza Sabzevari
Richa Singh
Richard Zhang
Robby Tan
Rodrigo Benenson
Rogerio Feris
Rohit Girdhar
Ronen Basri
Roozbeh Mottaghi
Ross Girshick
Ruiping Wang
Ryan Farrell
S. Kevin Zhou
Sagie Benaim
Saining Xie
Salman Khan
Sangdoo Yun
Sara Beery
Sayna Ebrahimi
Sebastiano Vascon
Seon Joo Kim
Serge Belongie
Sergey Tulyakov
Ser-Nam Lim
Shai Bagon
Shalini De Mello
Shang-Hong Lai
Shangzhe Wu
Shaodi You
Sharon Xiaolei Huang
Shengcai Liao
Shenghua Gao
Shenlong Wang
Shiguang Shan
Shizhe Chen
Shubham Tulsiani
Shuo Chen
Si Liu
Sicheng Zhao
Sifei Liu
Silvia Cascianelli
Simon Niklaus
Simone Calderara
Simone Melzi
Simone Schaub-Meyer
Song Bai
Sourav Garg
Srinath Sridhar
Srinivasa Narasimhan
Stella Yu
Stephan Richter
Stéphane Lathuilière
Stephen Gould
Stephen Lin
Subhankar Roy
Subhransu Maji
Sudeep Sarkar
Suha Kwak
Tae-Kyun Kim
Takayuki Okatani
Tal Hassner
Tali Dekel
Tammy Riklin Raviv
Tatiana Tommasi
Tat-Jen Cham
Tat-Jun Chin
Tatsuya Harada
Theo Gevers
Thibaut Durand
Thomas Mensink
Thomas Pock
Tianzhu Zhang
Timo Bolkart
Timothy Hospedales
Ting-Chun Wang
Tolga Birdal
Tomas Pajdla
Tsung-Yi Lin
Varun Jampani
Vasileios Belagiannis
Venkatesh Babu Radhakrishnan
Vicente Ordonez
Vicky Kalogeiton
Vignesh Ramanathan
Vikram V. Ramaswamy
Viktor Larsson
Viktoriia Sharmanska
Vinay Namboodiri
Vincent Lepetit
Vineeth N Balasubramanian
Vladimir Pavlovic
Vladislav Golyanik
Wai-Kin Adams Kong
Wangmeng Zuo
Wei Liu
Weidi Xie
Wei-Shi Zheng
Wenguan Wang
Wieland Brendel
Xavier Giro-i-Nieto
Xi Li
Xi Yin
Xiang Bai
Xiangyu Xu
Xiaodan Liang
Xiaoguang Han
Xiaojun Chang
Xiaolong Wang
Xiaoqian Wang
Xiaoyu Wang
Xilin Chen
Xin Wang
Xinchao Wang
Xuming He
Yadong Mu
Yagiz Aksoy
Yale Song
Yan Yan
Yanchao Yang
Yang Bai
Yang Wang
Yannis Kalantidis
Yanxi Liu
Yaoyao Liu
Yasushi Makihara
Yasutaka Furukawa
Yasuyuki Matsushita
Yebin Liu
Yen-Yu Lin
Yi Fang
Yi Yang
Yiming Wang
Yin Li
Ying Wu
Yizhou Yu
Yogesh Rawat
Yoichi Sato
Yong Jae Lee
Yosi Keller
Yu Li
Yu Wu
Yuchao Dai
Yu-Chiang Frank Wang
Yuki Asano
Yulan Guo
Yung-Yu Chuang
Yunhui Guo
Yuri Boykov
Yutian Lin
Yu-Wing Tai
Yu-Xiong Wang
Zan Gojcic
Zhangyang Wang
Zhaopeng Cui
Zhe Lin
Zhifan Gao
Zhun Zhong
Zicheng Liu
Ziwei Liu
Zorah Laehner
Zuzana Kukelova

TUESDAY, 1ST OCTOBER

07:00 – 18:30
Registration / Badge Pickup

08:00 – 09:00
Welcome Ceremony - Gold Room (live), Auditorium (broadcast), Silver Room (broadcast)

09:00 – 18:00
Exhibition - Level 0

09:00 – 10:30
Oral session 1A: Scene analysis and understanding - Gold Room
Chairs: Serge Belongie; Kenneth Marino

1. Towards Scene Graph Anticipation; Rohith Peddi; Saksham Singh; Saurabh; Parag Singla; Vibhav Gogate*

2. OP-Align: Object-level and Part-level Alignment for Self-supervised Category-level Articulated Object Pose Estimation; *Yuchen Che*; Ryo Furukawa; Asako Kanezakiù*

3. PDiscoFormer: Relaxing Part Discovery Constraints with Vision Transformers; *Ananthu Aniraj*; Cassio F. Dantas; Dino Ienco; Diego Marcos*

4. Bi-directional Contextual Attention for 3D Dense Captioning; Minjung Kim*; Hyung Suk Lim; Soonyoung Lee; Bumsoo Kim*; Gunhee Kim*

5. OmniNOCS: A unified NOCS dataset and model for 3D lifting of 2D objects; Akshay Krishnan*; Abhijit Kundu*; Kevis-Kokitsi Maninis; James Hays; Matthew Brown

6. ABC Easy as 123: A Blind Counter for Exemplar-Free Multi-Class Class-agnostic Counting; Michael A Hobley*; Victor Adrian Prisacariu

7. A Fair Ranking and New Model for Panoptic Scene Graph Generation; Julian Lorenz*; Alexander Pest; Daniel Kienzle; Katja Ludwig; Rainer Lienhart

8. Expanding Scene Graph Boundaries: Fully Open-vocabulary Scene Graph Generation via Visual-Concept Alignment and Retention; *Zuyao Chen; Jinlin Wu; Zhen Lei; Zhaoxiang Zhang; Chang Wen Chen**BEST PAPER CANDIDATE

09:00 – 10:30
Oral session 1B: Autonomous driving - Auditorium
Chairs: Oriane Siméoni; Holger Caesar

1. Making Large Language Models Better Planners with Reasoning-Decision Alignment; Zhijian Huang; Tao Tang; Shaoxiang Chen; *Sihao Lin; Zequn Jie; Lin Ma; Guangrun Wang; Xiaodan Liang**

2. MapTracker: Tracking with Strided Memory Fusion for Consistent Vector HD Mapping; *Jiacheng Chen*; Yuefan Wu; Jiaqi Tan; Hang Ma; Yasutaka Furukawa**

3. M^2Depth: Self-supervised Two-Frame Multi-camera Metric Depth Estimation; *Yingshuang Zou*; Yikang Ding; Xi Qiu; Haoqian Wang*; Haotian Zhang**

4. H-V2X: A Large Scale Highway Dataset for BEV Perception; *Chang Liu*; MingXu zhu; Cong Ma*

5. Adaptive Bounding Box Uncertainties via Two-Step Conformal Prediction; *Alexander Timans*; Christoph-Nikolas Straehle; Kaspar Sakmann; Eric Nalisnick*

6. DriveLM: Driving with Graph Visual Question Answering; *Chonghao Sima*; Katrin Renz; Kashyap Chitta; Li Chen; Zhang Hanxue; Chengen Xie; Jens Beißwenger; Ping Luo; Andreas Geiger; Hongyang Li*

7. RealGen: Retrieval Augmented Generation for Controllable Traffic Scenarios; Wenhao Ding; Yulong Cao; DING ZHAO; Chaowei Xiao; Marco Pavone*

8. Mask2Map: Vectorized HD Map Construction Using Bird's Eye View Segmentation Masks; *Sehwan Choi*; Jun Won Choi; Jungho Kim; Hongjae Shin*

09:00 – 10:30
Oral session 1C: Low-level vision and imaging - Silver Room
Chairs: Javier Vazquez-Corral; Djamila Aouada

1. Integer-Valued Training and Spike-driven Inference Spiking Neural Network for High-performance and Energy-efficient Object Detection; Xinhao Luo; Man Yao; Yuhong Chou; Bo Xu; Guoqi Li* BEST PAPER CANDIDATE

2. Latent Diffusion Prior Enhanced Deep Unfolding for Snapshot Spectral Compressive Imaging; Zongliang Wu*; Ruiying Lu; Ying Fu; Xin Yuan BEST PAPER CANDIDATE

3. SEA-RAFT: Simple, Efficient, Accurate RAFT for Optical Flow; Yihan Wang*; Lahav O Lipson; Jia Deng BEST PAPER CANDIDATE

4. Photon Inhibition for Energy-Efficient Single-Photon Imaging; Lucas J Koerner*; Shantanu Gupta; Atul N Ingle; Mohit Gupta

5. Minimalist Vision with Freeform Pixels; Jeremy Klotz*; Shree Nayar BEST PAPER CANDIDATE

6. Flying with Photons: Rendering Novel Views of Propagating Light; Anagh Malik*; Noah Juravsky; Ryan Po; Gordon Wetzstein; Kiriakos N. Kutulakos; David B. Lindell

7. A Simple Low-bit Quantization Framework for Video Snapshot Compressive Imaging; Miao Cao*; Lishun Wang; Huan Wang; Xin Yuan

8. GazeXplain: Learning to Predict Natural Language Explanations of Visual Scanpaths; Xianyu Chen*; Ming Jiang; Qi Zhao*

09:00 – 12:30
Demo session 1 - Level 0

1. Transforming Retail with Shopic's Vision & AI-Powered Smart Cart; Shlomi Amitai, Eden Shwartz - Shopic

2. EmoVOCA: Speech-Driven Emotional 3D Talking Heads; *Federico Nocentini, Claudio Ferrari, Stefano Berretti University of Firenze*

3. Controllable Face Synthesis with Semantic Latent Diffusion Models; Alex Ergasti, Tomaso Fontanini, Claudio Ferrari, Massimo Bertozzi, Andrea Prati - University of Parma

4. Dynaphos: A VR demo of biologically-plausible simulated phosphene vision for visual cortical prostheses; *Umut Güçlü, Antonio Lozano, Burcu Küçükoğlu, Eleftherios Papadopoulos,Marcel van Gerven, Yağmur Güçlütürk - Radboud University*

5. OPT-IQA: Automated Camera Parameters Tuning Framework with IQA-guided Optimization; *Jan-Henner Roberg, Vladyslav Mosiichuk, Ricardo Silva, Luís Rosado - Fraunhofer Portugal Research*

10:30 – 11:00
HPC-AI Tech Technical Session - Technical Presentation Area (Level 0)
Video Ocean: Democratizing Efficient Video Production for All

10:30 – 11:00
Coffee Break - Exhibition Area (Level 0)

10:30 – 12:30
Poster session 1

1. Out-of-Bounding-Box Triggers: A Stealthy Approach to Cheat Object Detectors; *Tao Lin*; lijia Yu*; Gaojie Jin*; Renjue Li*; Peng Wu*; Lijun Zhang**

2. Fake It till You Make It: Curricular Dynamic Forgery Augmentations towards General Deepfake Detection; *Yuzhen Lin*; Wentang Song; Bin Li*; Yuezun Li; Jiangqun Ni; Han Chen; Qiushi Li*

3. Quantization-Friendly Winograd Transformations for Convolutional Neural Networks; *Vladimir Protsenko*; Vladimir Kryzhanovskiy; Alexander Filippov*

4. AdversariaLeak: External Information Leakage Attack Using Adversarial Samples on Face Recognition Systems; *Roye Katzav*; Amit Giloni; Edita Grolman*; Hiroo Saito; Tomoyuki Shibata; Tsukasa Omino; Misaki Komatsu; Yoshikazu Hanatani; Yuval Elovici; Asaf Shabtai*

5. YOLOv9: Learning What You Want to Learn Using Programmable Gradient Information; *Chien-Yao Wang*; I-Hau Yeh; Hong-Yuan Mark Liao*

6. CLR-GAN: Improving GANs Stability and Quality via Consistent Latent Representation and Reconstruction; *Shengke Sun; Ziqian Luan; Zhanshan Zhao*; Shijie Luo; Shuzhen Han**

7. Generalizable Symbolic Optimizer Learning; *Xiaotian Song; Peng Zeng; Yanan Sun*; Andy Song*

8. Nickel and Diming Your GAN: A Dual-Method Approach to Enhancing GAN Efficiency via Knowledge Distillation; *Sangyeop Yeo; Yoojin Jang; Jaejun Yoo**

9. Dataset Distillation by Automatic Training Trajectories; *Dai Liu**; *Jindong Gu**; *Hu Cao; Carsten Trinitis; Martin Schulz**

10. Refine, Discriminate and Align: Stealing Encoders via Sample-Wise Prototypes and Multi-Relational Extraction; *Shuchi Wu**; *Chuan Ma**; *Kang Wei**; *Xiaogang XU; Ming Ding; Yuwen Qian; Di Xiao; Tao Xiang*

11. PYRA: Parallel Yielding Re-Activation for Training-Inference Efficient Task Adaptation; *Yizhe Xiong; Hui Chen**; *Tianxiang Hao; Zijia Lin; Jungong Han; Yuesong Zhang; Guoxin Wang; Yongjun Bao; Guiguang Ding*

12. Fisher Calibration for Backdoor-Robust Heterogeneous Federated Learning; *Wenke Huang; Mang Ye**; *zekun shi; Bo Du**; *Dacheng Tao*

13. Exploiting Supervised Poison Vulnerability to Strengthen Self-Supervised Defense; *Jeremy Styborski**; *Mingzhi Lyu**; *Yi Huang**; *Adams Kong**

14. SSL-Cleanse: Trojan Detection and Mitigation in Self-Supervised Learning; *Mengxin Zheng**; *Jiaqi Xue; Zihao Wang; Xun Chen; Qian Lou; Lei Jiang; Xiaofeng Wang*

15. Learning Non-Linear Invariants for Unsupervised Out-of-Distribution Detection; *Lars Doorenbos**; *Raphael Sznitman; Pablo Márquez Neila*

16. Distill Gold from Massive Ores: Bi-level Data Pruning towards Efficient Dataset Distillation; *Yue Xu; Yong-Lu Li**; *Kaitong Cui; Ziyu Wang; Cewu Lu; Yu-Wing Tai; Chi-Keung Tang*

17. Optimization-based Uncertainty Attribution Via Learning Informative Perturbations; *Hanjing Wang**; *Bashirul Azam Biswas; Qiang Ji*

18. Representation Enhancement-Stabilization: Reducing Bias-Variance of Domain Generalization; *Wei Huang**; *Yilei Shi; Zhitong Xiong; Xiao Xiang Zhu*

19. Deep Feature Surgery: Towards Accurate and Efficient Multi-Exit Networks; *Cheng Gong; Yao Chen**; *Qiuyang Luo; Ye Lu; Tao Li; Yuzhi Zhang; Yufei Sun**; *Le Zhang*

20. Learn to Preserve and Diversify: Parameter-Efficient Group with Orthogonal Regularization for Domain Generalization; *Jiajun Hu; Jian Zhang; Lei Qi**; *Yinghuan Shi**; *Yang Gao*

21. MutDet: Mutually Optimizing Pre-training for Remote Sensing Object Detection; *Ziyue Huang; Yongchao Feng; Qingjie Liu**; *Yunhong Wang*

22. UDA-Bench: Revisiting Common Assumptions in Unsupervised Domain Adaptation Using a Standardized Framework; *Tarun Kalluri**; *Sreyas Ravichandran; Manmohan Chandraker*

23. Versatile Incremental Learning: Towards Class and Domain-Agnostic Incremental Learning; *Min-Yeong Park; Jae-Ho Lee; Gyeong-Moon Park**

24. POA: Pre-training Once for Models of All Sizes; *Yingying Zhang**; *Xin Guo; Jiangwei Lao; Lei Yu; Lixiang Ru; Jian Wang; Guo Ye; HUIMEI HE; Jingdong Chen; Ming Yang**

25. MTaDCS: Moving Trace and Feature Density-based Confidence Sample Selection under Label Noise; *Qingzheng Huang; Xilin He; Xiaole Xian; Qinliang Lin; Weicheng Xie**; *Siyang Song; Linlin Shen; Zitong Yu*

26. Open-set Domain Adaptation via Joint Error based Multi-class Positive and Unlabeled Learning; *Dexuan Zhang**; *Thomas Westfechtel; Tatsuya Harada*

27. Bidirectional Uncertainty-Based Active Learning for Open-Set Annotation; *Chen-Chen Zong; Ye-Wen Wang; Kun-Peng Ning; Hai-Bo Ye; Sheng-Jun Huang**

28. Rethinking Few-shot Class-incremental Learning: Learning from Yourself; *Yu-Ming Tang; Yi-Xing Peng; Jingke Meng**; *Wei-Shi Zheng*

29. Bridge Past and Future: Overcoming Information Asymmetry in Incremental Object Detection; *Qijie Mo; Yipeng Gao; Shenghao Fu; Junkai Yan; Ancong Wu**; *Wei-Shi Zheng**

30. Confidence Self-Calibration for Multi-Label Class-Incremental Learning; *Kaile Du**; *Yifan Zhou; Fan Lyu; Yuyang Li; Chen Lu; Guangcan Liu**

31. Early Preparation Pays Off: New Classifier Pre-tuning for Class Incremental Semantic Segmentation; *Zhengyuan Xie; Haiquan Lu; Jia-wen Xiao; Enguang Wang; Le Zhang; Xialei Liu**

32. Online Continuous Generalized Category Discovery; *Keon-Hee Park; Hakyung Lee; Kyungwoo Song*; Gyeong-Moon Park**

33. Better Regression Makes Better Test-time Adaptive 3D Object Detection; *Jiakang Yuan; Bo Zhang; Kaixiong Gong; Xiangyu Yue; Botian Shi; Yu Qiao; Tao Chen**

34. Bayesian Detector Combination for Object Detection with Crowdsourced Annotations; *Zhi Qin Tan*; Olga Isupova; Gustavo Carneiro; Xiatian Zhu; Yunpeng Li*

35. Hierarchical Gaussian Mixture Normalizing Flow Modeling for Unified Anomaly Detection; *Xincheng Yao*; Ruoqi Li; Zefeng Qian; lu wang; Chongyang Zhang**

36. Towards Open-World Object-based Anomaly Detection via Self-Supervised Outlier Synthesis; *Brian Kostadinov Shalon Isaac-Medina*; Yona Falinie Abdul Gaus*; Neelanjan Bhowmik; Toby P Breckon*

37. Robust Zero-Shot Crowd Counting and Localization with Adaptive Resolution SAM; *Jia Wan*; Qiangqiang Wu; Wei Lin; Antoni Chan*

38. AdaCLIP: Adapting CLIP with Hybrid Learnable Prompts for Zero-Shot Anomaly Detection; *Yunkang Cao*; Jiangning Zhang; Luca Frittoli; Yuqi Cheng; Weiming Shen*; Giacomo Boracchi*

39. Bucketed Ranking-based Losses for Efficient Training of Object Detectors; *Feyza Yavuz*; Baris Can Cam; Adnan Harun Dogan; Kemal Oksuz; Emre Akbas; Sinan Kalkan*

40. HERGen: Elevating Radiology Report Generation with Longitudinal Data; *Fuying Wang; Shenghui Du; Lequan Yu**

41. Rethinking Unsupervised Outlier Detection via Multiple Thresholding; *Zhonghang Liu*; Panzhong Lu; Guoyang Xie; Zhichao Lu; Wen-Yan Lin*

42. MedRAT: Unpaired Medical Report Generation via Auxiliary Tasks; *Elad Hirsch*; Gefen Dawidowicz; Ayellet Tal*

43. Finding Meaning in Points: Weakly Supervised Semantic Segmentation for Event Cameras; *Hoonhee Cho; Sung-Hoon Yoon; Hyeokjun Kweon; Kuk-Jin Yoon**

44. NOVUM: Neural Object Volumes for Robust Object Classification; *Artur Jesslen*; Guofeng Zhang; Angtian Wang; Wufei Ma; Alan Yuille; Adam Kortylewski*

45. Unsupervised Dense Prediction using Differentiable Normalized Cuts; *Yanbin Liu*; Stephen Gould*

46. Bridging the Pathology Domain Gap: Efficiently Adapting CLIP for Pathology Image Analysis with Limited Labeled Data; *Zhengfeng Lai*; Joohi Chauhan; Brittany N. Dugger; Chen-Nee Chuah*

47. Multistain Pretraining for Slide Representation Learning in Pathology; *Guillaume Jaume*; Anurag J Vaidya*; Andrew Zhang; Andrew Song; Richard J Chen; Sharifa Sahai; Dandan Mo; Emilio Madrigal; Long P Le; Faisal Mahmood**

48. Agglomerative Token Clustering; *Joakim Bruslund Haurum*; Sergio Escalera; Graham W. Taylor*; Thomas B. Moeslund*

49. A Rotation-invariant Texture ViT for Fine-Grained Recognition of Esophageal Cancer Endoscopic Ultrasound Images; *Tianyi Liu; Shuaishuai S Zhuang; Jiacheng Nie; Geng Chen ; Yusheng Guo; Guangquan Zhou*; Jean-Louis Coatrieux; Yang Chen**

50. Semi-supervised Segmentation of Histopathology Images with Noise-Aware Topological Consistency; *Meilong Xu*; Xiaoling Hu; Saumya Gupta; Shahira Abousamra; Chao Chen*

51. The Devil is in the Statistics: Mitigating and Exploiting Statistics Difference for Generalizable Semi-supervised Medical Image Segmentation; *Muyang Qiu; Jian Zhang; Lei Qi; Qian Yu; Yinghuan Shi*; Yang Gao*

52. Self-supervised co-salient object detection via feature correspondences at multiple scales; *Souradeep Chakraborty*; Dimitris Samaras*

53. Recursive Visual Programming; *Jiaxin Ge*; Sanjay Subramanian; Baifeng Shi; Roei Herzig; Trevor Darrell*

54. FREST: Feature RESToration for Semantic Segmentation under Multiple Adverse Conditions; *Sohyun Lee; Namyup Kim; Sungyeon Kim; Suha Kwak**

55. Open-Vocabulary SAM: Segment and Recognize Twenty-thousand Classes Interactively; *Haobo Yuan; Xiangtai Li*; Chong Zhou; Yining Li; Kai Chen; Chen Change Loy*

56. Evaluating the Adversarial Robustness of Semantic Segmentation: Trying Harder Pays Off; *Levente Halmosi; Bálint Mohos; Márk Jelasity**

57. Progressive Proxy Anchor Propagation for Unsupervised Semantic Segmentation; *Hyun Seok Seong; WonJun Moon; SuBeen Lee; Jae-Pil Heo**

58. IRSAM: Advancing Segment Anything Model for Infrared Small Target Detection; *Mingjin Zhang; Yuchun Wang*; Jie Guo*; Yunsong Li; Xinbo Gao; Jing Zhang*

59. Open-Vocabulary RGB-Thermal Semantic Segmentation; *GuoQiang Zhao; JunJie Huang; Xiaoyun Yan*; Zhaojing Wang; Junwei Tang; Yangjun Ou; Xinrong Hu; Tao Peng*

60. SPIN: Hierarchical Segmentation with Subpart Granularity in Natural Images; *Josh David Myers-Dean*; Jarek T Reynolds; Brian Price; Yifei Fan; Danna Gurari*

61. Pseudo-Embedding for Generalized Few-Shot Point Cloud Segmentation; *Chih-Jung Tsai; Hwann-Tzong Chen*; Tyng-Luh Liu*

62. Region-aware Distribution Contrast: A Novel Approach to Multi-Task Partially Supervised Learning; *Meixuan Li; Tianyu Li; Guoqing Wang*; Peng Wang; Yang Yang; Jie Zou*

63. PartSTAD: 2D-to-3D Part Segmentation Task Adaptation; *Hyunjin Kim; Minhyuk Sung**

64. SHINE: Saliency-aware HIerarchical NEgative Ranking for Compositional Temporal Grounding; *Zixu Cheng*; Yujiang Pu*; Shaogang Gong; Parisa Kordjamshidi; Yu Kong*

65. CPM: Class-conditional Prompting Machine for Audio-visual Segmentation; *Yuanhong Chen*; Chong Wang; Yuyuan Liu; Hu Wang; Gustavo Carneiro*

66. Large-Scale Multi-Hypotheses Cell Tracking Using Ultrametric Contours Maps; *Jordão Bragantini*; Merlin Lange; Loïc A Royer*

67. OpenIns3D: Snap and Lookup for 3D Open-vocabulary Instance Segmentation; *Zhening Huang; Xiaoyang Wu; Xi Chen; Hengshuang Zhao*; Lei Zhu; Joan Lasenby**

68. Towards Dual Transparent Liquid Level Estimation in Biomedical Lab: Dataset, Methods and Practice; *Xiayu Wang; Ke Ma; Ruiyun Zhong; Xinggang Wang; Yi Fang; Yang Xiao; Tian Xia**

69. Efficient Unsupervised Visual Representation Learning with Explicit Cluster Balancing; *Ioannis Maniadis Metaxas*; Georgios Tzimiropoulos; Ioannis Patras*

70. 3D Weakly Supervised Semantic Segmentation with 2D Vision-Language Guidance; *Xiaoxu Xu; Yitian Yuan; Jinlong Li; Qiudan Zhang; Zequn Jie; Lin Ma; Hao Tang; Nicu Sebe; Xu Wang**

71. Robustness Preserving Fine-tuning using Neuron Importance; *Guangrui Li; Rahul Duggal*; Aaditya Singh; Kaustav Kundu; Bing Shuai; Jonathan Wu*

72. Grounding DINO: Marrying DINO with Grounded Pre-Training for Open-Set Object Detection; *Shilong Liu*; Zhaoyang Zeng; Tianhe Ren; Feng Li; Hao Zhang; Jie Yang; Qing Jiang; Chunyuan Li; Jianwei Yang; Hang Su; Jun Zhu; Lei Zhang**

73. Unlocking Textual and Visual Wisdom: Open-Vocabulary 3D Object Detection Enhanced by Comprehensive Guidance from Text and Image; *Pengkun Jiao*; Na Zhao*; Jingjing Chen; Yu-Gang Jiang*

74. Contrastive ground-level image and remote sensing pre-training improves representation learning for natural world imagery; *Andy V Huynh*; Lauren Gillespie; Jael Lopez-Saucedo; Claire Tang; Rohan Sikand; Moisés Expósito-Alonso*

75. Expanding Scene Graph Boundaries: Fully Open-vocabulary Scene Graph Generation via Visual-Concept Alignment and Retention; *Zuyao Chen; Jinlin Wu; Zhen Lei; Zhaoxiang Zhang; Chang Wen Chen** BEST PAPER CANDIDATE

76. Multimodal Label Relevance Ranking via Reinforcement Learning; *Taian Guo; Taolin Zhang; Haoqian Wu; Hanjun Li; Ruizhi Qiao*; Xing Sun*

77. Open-Set Recognition in the Age of Vision-Language Models; *Dimity Miller*; Niko Suenderhauf; Alex Kenna; Keita Mason*

78. A Simple Background Augmentation Method for Object Detection with Diffusion Model; *Yuhang Li; Xin Dong; Chen Chen; Weiming Zhuang; Lingjuan Lyu**

79. Embedding-Free Transformer with Inference Spatial Reduction for Efficient Semantic Segmentation; *Hyunwoo Yu; Yubin Cho; Beoungwoo Kang; Seunghun Moon; Kyeongbo Kong; Suk-Ju Kang**

80. Textual Knowledge Matters: Cross-Modality Co-Teaching for Generalized Visual Class Discovery; *Haiyang Zheng; Nan Pu; Wenjing Li*; Nicu Sebe; Zhun Zhong**

81. Multi-Label Cluster Discrimination for Visual Representation Learning; *Xiang An; Kaicheng Yang; Xiangzi Dai; Ziyong Feng; Jiankang Deng**

82. Online Zero-Shot Classification with CLIP; *Qi Qian*; Juhua Hu*

83. MultiDelete for Multimodal Machine Unlearning; *Jiali Cheng*; Hadi Amiri*

84. WRIM-Net: Wide-Ranging Information Mining Network for Visible-Infrared Person Re-Identification; *Yonggan Wu; Ling-Chao Meng*; Yuan Zichao; Sixian Chan; Hong-Qiang Wang**

85. Open-Vocabulary 3D Semantic Segmentation with Text-to-Image Diffusion Models; *Xiaoyu Zhu*; Hao Zhou; Pengfei Xing; Long Zhao; Hao Xu; Junwei Liang; Alexander G. Hauptmann; Ting Liu; Andrew Gallagher*

86. When Pedestrian Detection Meets Multi-Modal Learning: Generalist Model and Benchmark Dataset; *Yi Zhang; Wang Zeng; Sheng Jin; Chen Qian*; Ping Luo; Wentao Liu*

87. DataDream: Few-shot Guided Dataset Generation; *Jae Myung Kim*; Jessica Bader; Stephan Alaniz; Cordelia Schmid; Zeynep Akata*

88. Semantic Residual Prompts for Continual Learning; *Martin Menabue*; Emanuele Frascaroli; Matteo Boschini; Enver Sangineto; Lorenzo Bonicelli; Angelo Porrello*; SIMONE CALDERARA*

89. ViC-MAE: Self-Supervised Representation Learning from Images and Video with Contrastive Masked Autoencoders; *Jefferson Hernandez*; Ruben Villegas; Vicente Ordonez*

90. A Unified Image Compression Method for Human Perception and Multiple Vision Tasks; *Sha Guo; Lin Sui; Chen-Lin Zhang; Zhuo Chen; Wenhan Yang; Lingyu Duan**

91. Encapsulating Knowledge in One Prompt; *Qi Li*; Runpeng Yu*; Xinchao Wang**

92. Stripe Observation Guided Inference Cost-free Attention Mechanism; *Zhongzhan Huang*; Shanshan Zhong; Wushao Wen; Jinghui Qin; Liang Lin**

93. Agent Attention: On the Integration of Softmax and Linear Attention; *Dongchen Han; Tianzhu Ye; Yizeng Han; Zhuofan Xia; Siyuan Pan; Pengfei Wan; Shiji Song; Gao Huang**

94. Good Teachers Explain: Explanation-Enhanced Knowledge Distillation; *Amin Parchami-Araghi*; Moritz Böhle; Sukrut Rao; Bernt Schiele*

95. Graph Neural Network Causal Explanation via Neural Causal Models; *Arman Behnam*; Binghui Wang*

96. Understanding Multi-compositional learning in Vision and Language models via Category Theory; *Sotirios Panagiotis Chytas*; Hyunwoo J Kim; Vikas Singh*

97. Weak-to-Strong Compositional Learning from Generative Models for Language-based Object Detection; *Kwanyong Park; Kuniaki Saito; Donghyun Kim**

98. This Probably Looks Exactly Like That: An Invertible Prototypical Network; *Zachariah Carmichael*; Timothy P Redgrave; Daniel Gonzalez Cedre; Walter Scheirer*

99. DEPICT: Diffusion-Enabled Permutation Importance for Image Classification Tasks; *Sarah Jabbour*; Gregory Kondas; Ella Kazerooni; Michael Sjoding; David Fouhey; Jenna Wiens*

100. ViG-Bias: Visually Grounded Bias Discovery and Mitigation; *Badr-Eddine Marani*; Mohamed Hanini; Nihitha Malayarukil; Stergios Christodoulidis; Maria Vakalopoulou; Enzo Ferrante*

101. Do text-free diffusion models learn discriminative visual representations?; *Soumik Mukhopadhyay*; Matthew A Gwilliam*; Yosuke Yamaguchi; Vatsal Agarwal; Namitha Padmanabhan; Archana Swaminathan; Tianyi Zhou; Jun Ohya; Abhinav Shrivastava*

102. Diff-Tracker: Text-to-Image Diffusion Models are Unsupervised Trackers; *Zhengbo Zhang*; Li Xu; Duo Peng; Hossein Rahmani; Jun Liu**

103. Unveiling and Mitigating Memorization in Text-to-image Diffusion Models through Cross Attention; *Jie Ren*; Yaxin Li; Shenglai Zeng; Han Xu; Lingjuan Lyu; Yue Xing; Jiliang Tang*

104. Receler: Reliable Concept Erasing of Text-to-Image Diffusion Models via Lightweight Erasers; *Chi-Pin Huang*; Kai-Po Chang; Chung-Ting Tsai; Yung-Hsuan Lai; Fu-En Yang; Yu-Chiang Frank Wang*

105. HybridBooth: Hybrid Prompt Inversion for Efficient Subject-Driven Generation; *Shanyan Guan; Yanhao Ge; Ying Tai*; Jian Yang; Wei Li; Mingyu You**

106. PixArt-Sigma: Weak-to-Strong Training of Diffusion Transformer for 4K Text-to-Image Generation; *Junsong Chen; Chongjian GE; Enze Xie*; Yue Wu; Lewei Yao; Xiaozhe Ren; Zhongdao Wang; Ping Luo; Huchuan Lu; Zhenguo Li*

107. Diffusion Models as Data Mining Tools; *Ioannis Siglidis*; Aleksander Holynski; Alexei A. Efros; Mathieu Aubry; Shiry Ginosar*

108. Improving Geo-diversity of Generated Images with Contextualized Vendi Score Guidance; *Reyhane Askari Hemmat*; Melissa Hall*; Alicia Yi Sun; Candace Ross; Michal Drozdzal; Adriana Romero-Soriano*

109. NeuroPictor: Refining fMRI-to-Image Reconstruction via Multi-individual Pretraining and Multi-level Modulation; *Jingyang Huo; Yikai Wang; Yanwei Fu*; Xuelin Qian; Chong Li; Yun Wang; Jianfeng Feng*

110. Infinite-ID: Identity-preserved Personalization via ID-semantics Decoupling Paradigm; *Yi Wu; Ziqiang Li; Heliang Zheng; Chaoyue Wang*; Bin Li**

111. Investigating Style Similarity in Diffusion Models; *Gowthami Somepalli*; Anubhav Gupta; Kamal Gupta; Shramay Palta; Micah Goldblum; Jonas A. Geiping; Abhinav Shrivastava; Tom Goldstein*

112. Diffusion Soup: Model Merging for Text-to-Image Diffusion Models; *Benjamin J Biggs*; Arjun Seshadri; Yang Zou; Achin Jain; Aditya Golatkar; Yusheng Xie; Alessandro Achille; Ashwin Swaminathan; Stefano Soatto*

113. GazeXplain: Learning to Predict Natural Language Explanations of Visual Scanpaths; *Xianyu Chen*; Ming Jiang; Qi Zhao**

114. GarmentAligner: Text-to-Garment Generation via Retrieval-augmented Multi-level Corrections; *Shiyue Zhang; Zheng Chong; Xujie Zhang; Hanhui Li; Yuhao Cheng; yiqiang yan; Xiaodan Liang**

115. WeCromCL: Weakly Supervised Cross-Modality Contrastive Learning for Transcription-only Supervised Text Spotting; *Jingjing Wu; Zhengyao Fang; Pengyuan Lyu; Chengquan Zhang; Fanglin Chen; Guangming Lu; Wenjie Pei**

116. WAS: Dataset and Methods for Artistic Text Segmentation; *Xudong Xie; Yuzhe Li; Yang Liu; Zhifei Zhang; Zhaowen Wang; Wei Xiong; Xiang Bai**

117. Elegantly Written: Disentangling Writer and Character Styles for Enhancing Online Chinese Handwriting; *Yu Liu; Fatimah binti Khalid; Lei Wang; Youxi Zhang; Cunrui Wang**

118. One-Shot Diffusion Mimicker for Handwritten Text Generation; *Gang Dai; Yifan Zhang; Quhui Ke; Qiangya Guo; Shuangping Huang**

119. Bi-directional Contextual Attention for 3D Dense Captioning; *Minjung Kim*; Hyung Suk Lim; Soonyoung Lee; Bumsoo Kim*; Gunhee Kim**

120. BRIDGE: Bridging Gaps in Image Captioning Evaluation with Stronger Visual Cues; *Sara Sarto*; Marcella Cornia; Lorenzo Baraldi; Rita Cucchiara*

121. Controllable Contextualized Image Captioning: Directing the Visual Narrative through User-Defined Highlights; *Shunqi Mao*; Chaoyi Zhang; Hang Su; Hwanjun Song; Igor Shalyminov; Weidong Cai*

122. AddressCLIP: Empowering Vision-Language Models for City-wide Image Address Localization; *Shixiong Xu; Chenghao Zhang; Lubin Fan*; Gaofeng Meng*; SHIMING XIANG; Jieping Ye*

123. Visual Grounding for Object-Level Generalization in Reinforcement Learning; *Haobin Jiang; Zongqing Lu**

124. CLAP: Isolating Content from Style through Contrastive Learning with Augmented Prompts; *Yichao Cai*; Yuhang Liu; Zhen Zhang; Javen Qinfeng Shi*

125. Tracking Meets LoRA: Faster Training, Larger Model, Stronger Performance; *Liting Lin; Heng Fan; Zhipeng Zhang; Yaowei Wang*; Yong Xu; Haibin Ling**

126. Synergy of Sight and Semantics: Visual Intention Understanding with CLIP; *Qu Yang; Mang Ye*; Dacheng Tao*

127. Depicting Beyond Scores: Advancing Image Quality Assessment through Multi-modal Language Models; *Zhiyuan You*; Zheyuan Li; Jinjin Gu*; Zhenfei Yin; Tianfan Xue*; Chao Dong**

128. PartCraft: Crafting Creative Objects by Parts; *Kam Woh Ng*; Xiatian Zhu; Yi-Zhe Song; Tao Xiang*

130. Elevating All Zero-Shot Sketch-Based Image Retrieval Through Multimodal Prompt Learning; *Mainak Singha*; Ankit Jha; Divyam Gupta; Pranav Singla; Biplab Banerjee*

131. Mismatch Quest: Visual and Textual Feedback for Image-Text Misalignment; *Brian Gordon*; Yonatan Bitton*; Yonatan Shafir; Roopal Garg; Xi Chen; Dani Lischinski; Daniel Cohen-Or; Idan Szpektor*

132. Adversarial Prompt Tuning for Vision-Language Models; *Jiaming Zhang; Xingjun Ma*; Xin Wang; Lingyu Qiu; Jiaqi Wang; Yu-Gang Jiang; Jitao Sang**

133. FlexAttention for Efficient High-Resolution Vision-Language Models; *Junyan Li*; Delin Chen; Tianle Cai; Peihao Chen; Yining Hong; Zhenfang Chen; Yikang Shen; Chuang Gan*

134. HaloQuest: A Visual Hallucination Dataset for Advancing Multimodal Reasoning; *Zhecan Wang; Garrett Bingham*; Adams Wei Yu; Quoc V. Le; Thang Luong; Golnaz Ghiasi*

135. Multiscale Graph Texture Network; *Ravishankar Evani*; Deepu Rajan; Shangbo Mao*

136. MM1: Methods, Analysis & Insights from Multimodal LLM Pre-training; *Brandon McKinzie; Zhe Gan; Jean-Philippe Fauconnier; Samuel Dodge; Bowen Zhang; Philipp Dufter; Dhruti Shah; Futang Peng; Anton Belyi; Max A Schwarzer; Hongyu Hè; Xianzhi Du; Haotian Zhang; Karanjeet Singh; Doug Kang; Tom Gunter; Xiang Kong; Aonan Zhang; Jianyu Wang; Chong Wang; Nan Du; Tao Lei; Sam Wiseman; Mark Lee; Zirui Wang; Ruoming Pang; Peter Grasch; Alexander Toshev*; Yinfei Yang*

137. VisionLLaMA: A Unified LLaMA Backbone for Vision Tasks; *Xiangxiang Chu*; Jianlin Su; Bo Zhang*; Chunhua Shen*

138. Any2Point: Empowering Any-modality Transformers for Efficient 3D Understanding; *Yiwen Tang; Ray Zhang; Jiaming Liu; Zoey Guo; Bin Zhao*; Zhigang Wang; Dong Wang*; Peng Gao; Hongsheng Li; Xuelong Li*

139. BaSIC: BayesNet Structure Learning for Computational Scalable Neural Image Compression; *Yufeng Zhang; Hang Yu; Shizhan Liu; Wenrui Dai; Weiyao Lin**

140. REVISION: Rendering Tools Enable Spatial Fidelity in Vision-Language Models; *Agneet Chatterjee*; Yiran Luo; Tejas Gokhale; Yezhou Yang; Chitta R Baral*

141. CAT: Enhancing Multimodal Large Language Model to Answer Questions in Dynamic Audio-Visual Scenarios; *Qilang Ye; Zitong Yu*; Rui Shao; Xinyu Xie; Philip Torr; Xiaochun Cao*

142. Meta-optimized Angular Margin Contrastive Framework for Video-Language Representation Learning; *Thong Thanh Nguyen*; Yi Bin; Xiaobao Wu; Xinshuai Dong; Zhiyuan Hu; Khoi M Le; Cong-Duy Nguyen; See Kiong Ng; Anh Tuan Luu*

143. Multi-Sentence Grounding for Long-term Instructional Video; *Zeqian Li; Qirui Chen; Tengda Han; Ya Zhang; Yan-Feng Wang; Weidi Xie**

144. FTBC: Forward Temporal Bias Correction for Optimizing ANN-SNN Conversion; *Xiaofeng Wu*; Velibor Bojkovic; Bin Gu*; Kun Suo; Kai Zou*

145. ABC Easy as 123: A Blind Counter for Exemplar-Free Multi-Class Class-agnostic Counting; *Michael A Hobley*; Victor Adrian Prisacariu*

146. PDiscoFormer: Relaxing Part Discovery Constraints with Vision Transformers; *Ananthu Aniraj*; Cassio F. Dantas; Dino Ienco; Diego Marcos*

147. Integer-Valued Training and Spike-driven Inference Spiking Neural Network for High-performance and Energy-efficient Object Detection; *Xinhao Luo; Man Yao; Yuhong Chou; Bo Xu; Guoqi Li** BEST PAPER CANDIDATE

148. A Simple Low-bit Quantization Framework for Video Snapshot Compressive Imaging; *Miao Cao*; Lishun Wang; Huan Wang; Xin Yuan*

149. Latent Diffusion Prior Enhanced Deep Unfolding for Snapshot Spectral Compressive Imaging; *Zongliang Wu*; Ruiying Lu; Ying Fu; Xin Yuan* BEST PAPER CANDIDATE

150. Photon Inhibition for Energy-Efficient Single-Photon Imaging; *Lucas J Koerner*; Shantanu Gupta; Atul N Ingle; Mohit Gupta*

151. Minimalist Vision with Freeform Pixels; *Jeremy Klotz*; Shree Nayar* BEST PAPER CANDIDATE

152. SEA-RAFT: Simple, Efficient, Accurate RAFT for Optical Flow; *Yihan Wang*; Lahav O Lipson; Jia Deng* BEST PAPER CANDIDATE

153. M^2Depth: Self-supervised Two-Frame Multi-camera Metric Depth Estimation; *Yingshuang Zou*; Yikang Ding; Xi Qiu; Haoqian Wang*; Haotian Zhang**

154. Adaptive Bounding Box Uncertainties via Two-Step Conformal Prediction; *Alexander Timans*; Christoph-Nikolas Straehle; Kaspar Sakmann; Eric Nalisnick*

155. MapTracker: Tracking with Strided Memory Fusion for Consistent Vector HD Mapping; *Jiacheng Chen*; Yuefan Wu; Jiaqi Tan; Hang Ma; Yasutaka Furukawa**

156. Mask2Map: Vectorized HD Map Construction Using Bird's Eye View Segmentation Masks; *Sehwan Choi*; Jun Won Choi; Jungho Kim; Hongjae Shin*

157. H-V2X: A Large Scale Highway Dataset for BEV Perception; *Chang Liu*; MingXu zhu; Cong Ma*

158. Towards Scene Graph Anticipation; *Rohith Peddi*; Saksham Singh; Saurabh .; Parag Singla; Vibhav Gogate*

159. RealGen: Retrieval Augmented Generation for Controllable Traffic Scenarios; *Wenhao Ding*; Yulong Cao; DING ZHAO; Chaowei Xiao; Marco Pavone*

160. DriveLM: Driving with Graph Visual Question Answering; *Chonghao Sima*; Katrin Renz; Kashyap Chitta; Li Chen; Zhang Hanxue; Chengen Xie; Jens Beißwenger; Ping Luo; Andreas Geiger; Hongyang Li*

161. Making Large Language Models Better Planners with Reasoning-Decision Alignment; *Zhijian Huang; Tao Tang; Shaoxiang Chen; Sihao Lin; Zequn Jie; Lin Ma; Guangrun Wang; Xiaodan Liang**

162. Synchronization is All You Need: Exocentric-to-Egocentric Transfer for Temporal Action Segmentation with Unlabeled Synchronized Video Pairs; *Camillo Quattrocchi*; Antonino Furnari; Daniele Di Mauro; Mario Valerio Giuffrida; Giovanni Maria Farinella*

163. Walker: Self-supervised Multiple Object Tracking by Walking on Temporal Object Appearance Graphs; *Mattia Segù*; Luigi Piccinelli; Siyuan Li; Luc Van Gool; Fisher Yu; Bernt Schiele*

164. Lost and Found: Overcoming Detector Failures in Online Multi-Object Tracking; *Lorenzo Vaquero*; Yihong Xu; Xavier Alameda-Pineda; Victor M. Brea; Manuel Mucientes*

165. Reliable Spatial-Temporal Voxels For Multi-Modal Test-Time Adaptation; *Haozhi Cao; Yuecong Xu; Jianfei Yang*; Pengyu Yin; Xingyu Ji; Shenghai Yuan; Lihua Xie*

166. Towards More Practical Group Activity Detection: A New Benchmark and Model; *Dongkeun Kim; Youngkil Song; Minsu Cho; Suha Kwak**

167. Listen to Look into the Future: Audio-Visual Egocentric Gaze Anticipation; *Bolin Lai*; Fiona Ryan; Wenqi Jia; Miao Liu; James M Rehg*

168. Learning by Aligning 2D Skeleton Sequences and Multi-Modality Fusion; *Quoc-Huy Tran*; Muhammad Ahmed; Murad Popattia; Muhammad Hassan Ahmed; Andrey Konin; Zeeshan Zia*

169. Event-based Head Pose Estimation: Benchmark and Method; *Jiahui Yuan*; Hebei Li; Yansong Peng; Jin Wang; Yuheng Jiang; Yueyi Zhang*; Xiaoyan Sun*

170. CrossGLG: LLM Guides One-shot Skeleton-based 3D Action Recognition in a Cross-level Manner; *Tingbing Yan; Wenzheng Zeng*; Yang Xiao*; Xingyu Tong; Bo Tan; Zhiwen Fang; Zhiguo Cao; Joey Tianyi Zhou*

171. Masked Video and Body-worn IMU Autoencoder for Egocentric Action Recognition; *Mingfang Zhang; Yifei Huang*; Ruicong Liu; Yoichi Sato*

172. E3V-K5: An Authentic Benchmark for Redefining Video-Based Energy Expenditure Estimation; *Shengxuming Zhang; Lei Jin; Yifan Wang; Xinyu Wang; Xu Wen; Zunlei Feng*; Mingli Song*

173. DailyDVS-200: A Comprehensive Benchmark Dataset for Event-Based Action Recognition; *Qi Wang; Zhou Xu; Yuming Lin; Jingtao Ye; Hongsheng Li; Guangming Zhu; Syed Afaq Ali Shah; Mohammed Bennamoun; Liang Zhang**

174. DetailSemNet: Elevating Signature Verification through Detail-Semantic Integration; *Meng-Cheng Shih*; Tsai-Ling Huang; Yu-Heng Shih; Hong-Han Shuai; Hsuan-Tung Liu; Yi-Ren Yeh; Ching-Chun Huang**

175. X-Pose: Detecting Any Keypoints; *Jie Yang; Ailing Zeng*; Ruimao Zhang*; Lei Zhang*

176. EgoLifter: Open-world 3D Segmentation for Egocentric Perception; *Qiao Gu*; Zhaoyang Lv*; Duncan Frost; Simon Green; Julian Straub; Chris Sweeney**

177. GTPT: Group-based Token Pruning Transformer for Efficient Human Pose Estimation; *Haonan Wang; Jie Liu*; Jie Tang; Gangshan Wu; Bo Xu; Yanbing Chou; Yong Wang*

178. Diffusion Reward: Learning Rewards via Conditional Video Diffusion; *Tao Huang*; Guangqi Jiang; Yanjie Ze; Huazhe Xu**

179. m&m's: A Benchmark to Evaluate Tool-Use for multi-step multi-modal Tasks; *Zixian Ma*; Weikai Huang; Jieyu Zhang; Tanmay Gupta; Ranjay Krishna*

180. OP-Align: Object-level and Part-level Alignment for Self-supervised Category-level Articulated Object Pose Estimation; *Yuchen Che*; Ryo Furukawa; Asako Kanezaki*

181. Omni6DPose: A Benchmark and Model for Universal 6D Object Pose Estimation and Tracking; *Jiyao Zhang; Weiyao Huang; Bo Peng; Mingdong Wu; Fei Hu; Zijian Chen; Bo Zhao; Hao Dong**

182. Pseudo-keypoint RKHS Learning for Self-supervised 6DoF Pose Estimation; *Yangzheng Wu*; Michael Alan Greenspan*

183. OmniNOCS: A unified NOCS dataset and model for 3D lifting of 2D objects; *Akshay Krishnan*; Abhijit Kundu*; Kevis-Kokitsi Maninis; James Hays; Matthew Brown*

184. FreeZe: Training-free zero-shot 6D pose estimation with geometric and vision foundation models; *Andrea Caraffa*; Davide Boscaini; Amir Hamza; Fabio Poiesi*

185. Shape-guided Configuration-aware Learning for Endoscopic-image-based Pose Estimation of Flexible Robotic Instruments; *Yiyao Ma*; Kai Chen*; Hon-Sing Tong; Ruofeng Wei; Yui-Lun Ng; Ka-Wai Kwok*; Qi Dou**

186. Large Motion Model for Unified Multi-Modal Motion Generation; *Mingyuan Zhang*; Daisheng Jin; Chenyang Gu; Fangzhou Hong; Zhongang Cai; Jingfang Huang; Chongzhi Zhang; Xinying Guo; Lei Yang; Ying He; Ziwei Liu**

187. PoseAugment: Generative Human Pose Data Augmentation with Physical Plausibility for IMU-based Motion Capture; *Zhuojun Li*; Chun Yu*; Chen Liang; Yuanchun Shi*

188. HUMOS: Human Motion Model Conditioned on Body Shape; *Shashank Tripathi*; Omid Taheri; Christoph Lassner*; Michael J. Black*; Daniel Holden*; Carsten Stoll**

189. SignAvatars: A Large-scale 3D Sign Language Holistic Motion Dataset and Benchmark; *Zhengdi Yu; Shaoli Huang*; yongkang cheng; Tolga Birdal*

190. Text Motion Translator: A Bi-Directional Model for Enhanced 3D Human Motion Generation from Open-Vocabulary Descriptions; *Yijun Qian*; Jack Urbanek; Alexander Hauptmann; Jungdam Won*

191. A Fair Ranking and New Model for Panoptic Scene Graph Generation; *Julian Lorenz*; Alexander Pest; Daniel Kienzle; Katja Ludwig; Rainer Lienhart*

192. Realistic Human Motion Generation with Cross-Diffusion Models; *Zeping Ren; Shaoli Huang*; Xiu Li**

193. TRAM: Global Trajectory and Motion of 3D Humans from in-the-wild Videos; *Yufu Wang*; Ziyun Wang; Lingjie Liu; Kostas Daniilidis*

194. Parameterized Quasi-Physical Simulators for Dexterous Manipulations Transfer; *Xueyi Liu*; Kangbo Lyu; jieqiong zhang; Tao Du; Li Yi**

195. How Video Meetings Change Your Expression; *Sumit Sarin*; Utkarsh Mall; Purva Tendulkar; Carl Vondrick*

196. Divide and Fuse: Body Part Mesh Recovery from Partially Visible Human Images; *Tianyu Luan; Zhongpai Gao; Luyuan Xie; Abhishek Sharma; Hao Ding; Benjamin Planche; Meng Zheng; Ange Lou; Terrence Chen; Junsong Yuan; Ziyan Wu**

197. Defect Spectrum: A Granular Look of Large-scale Defect Datasets with Rich Semantics; *Shuai Yang; ZhiFei Chen; Pengguang Chen; Xi Fang; Yixun Liang; Shu Liu*; Yingcong Chen**

198. UniTalker: Scaling up Audio-Driven 3D Facial Animation through A Unified Model; *Xiangyu Fan*; Jiaqi Li; Zhiqian Lin; Weiye Xiao; Lei Yang**

199. Generating Human Interaction Motions in Scenes with Text Control; *Hongwei Yi*; Justus Thies; Michael J. Black; Xue Bin Peng; Davis Rempe**

200. DreamStruct: Understanding Slides and User Interfaces via Synthetic Data Generation; *Yi-Hao Peng*; Faria Huq; Yue Jiang; Jason Wu; Xin Yue Li; Jeffrey Bigham; Amy Pavel*

201. HeadStudio: Text to Animatable Head Avatars with 3D Gaussian Splatting; *Zhenglin Zhou*; Fan Ma; Hehe Fan; Zongxin Yang; Yi Yang*

202. SceneScript: Reconstructing Scenes With An Autoregressive Structured Language Model; *Armen Avetisyan*; Christopher Xie; Henry Howard-Jenkins; Tsun-Yi Yang; Samir Aroudj; Suvam Patra; Fuyang Zhang; Luke Holland; Duncan Frost; Campbell Orme; Jakob Engel; Edward Miller; Richard Newcombe; Vasileios Balntas*

203. DiffuMatting: Synthesizing Arbitrary Objects with Matting-level Annotation; *Xiaobin Hu; Xu Peng; Donghao Luo*; Xiaozhong Ji; Jinlong Peng; ZhengKai Jiang; Jiangning Zhang; Taisong Jin*; Chengjie Wang; Rongrong Ji*

204. AnimatableDreamer: Text-Guided Non-rigid 3D Model Generation and Reconstruction with Canonical Score Distillation; *Xinzhou Wang; Yikai Wang*; Junliang Ye; Fuchun Sun*; Zhengyi Wang; Ling Wang; Pengkun Liu; Kai Sun; Xintong Wang; Xie wende; Fangfu Liu; Bin He*

205. AnyHome: Open-Vocabulary Large-Scale Indoor Scene Generation with First-Person View Exploration; *Rao Fu*; Zehao Wen; Zichen Liu ; Srinath Sridhar*

206. Topo4D: Topology-Preserving Gaussian Splatting for High-Fidelity 4D Head Capture; *Xuanchen Li; Yuhao Cheng; Xingyu Ren; Haozhe Jia; Di Xu; Wenhan Zhu; Yichao Yan**

207. MeshSegmenter: Zero-Shot Mesh Segmentation via Texture Synthesis; *Ziming Zhong*; Yanyu Xu; Jing Li; Jiale Xu; Zhengxin Li; Chaohui Yu; Shenghua Gao*

208. TPA3D: Triplane Attention for Fast Text-to-3D Generation; *Bin-Shih Wu*; Hong-En Chen*; Sheng-Yu Huang; Yu-Chiang Frank Wang*

209. Text2LiDAR: Text-guided LiDAR Point Clouds Generation via Equirectangular Transformer; *Yang Wu*; Kaihua Zhang; Jianjun Qian; Jin Xie*; Jian Yang*

210. MagicMirror: Fast and High-Quality Avatar Generation with Constrained Search Space; *Armand Comas; Di Qiu*; Menglei Chai; Marcel C. Bühler; Amit Raj; Ruiqi Gao; Qiangeng Xu; Mark J Matthews; Paulo Gotardo; Sergio Orts-Escolano; Thabo Beeler*

211. SENC: Handling Self-collision in Neural Cloth Simulation; *Zhouyingcheng Liao*; Sinan Wang; Taku Komura*

212. WordRobe: Text-Guided Generation of Textured 3D Garments; *Astitva Srivastava*; Pranav Manu; Amit Raj; Varun Jampani; Avinash Sharma*

213. Face Adapter for Pre-Trained Diffusion Models with Fine-Grained ID and Attribute Control; *Yue Han*; Junwei Zhu; Keke He; Xu Chen; Yanhao Ge; Wei Li; Xiangtai Li; Jiangning Zhang; Chengjie Wang; Yong Liu*

214. HumanRefiner: Benchmarking Abnormal Human Generation and Refining with Coarse-to-fine Pose-Reversible Guidance; *Guian Fang*; Wenbiao Yan; Yuanfan Guo; Jianhua Han; Zutao Jiang; Hang Xu; Shengcai Liao; Xiaodan Liang*

215. DreamDrone: Text-to-Image Diffusion Models are Zero-shot Perpetual View Generators; *Hanyang Kong*; Dongze Lian; Michael Bi Mi; Xinchao Wang**

216. MoVideo: Motion-Aware Video Generation with Diffusion Models; *Jingyun Liang*; Yuchen Fan; Kai Zhang*; Radu Timofte; Luc Van Gool; Rakesh Ranjan*

217. PoseCrafter: One-Shot Personalized Video Synthesis Following Flexible Pose Control; *Yong Zhong; Min Zhao; Zebin You; Xiaofeng Yu; Changwang Zhang; Chongxuan Li**

218. BlenderAlchemy: Editing 3D Graphics with Vision-Language Models; *Ian Huang*; Guandao Yang; Leonidas Guibas*

219. DreamMotion: Space-Time Self-Similar Score Distillation for Zero-Shot Video Editing; *Hyeonho Jeong; Jinho Chang; Geon Yeong Park; Jong Chul Ye**

220. Masked Generative Video-to-Audio Transformers with Enhanced Synchronicity; *Santiago Pascual; Chunghsin YEH*; Ioannis Tsiamas; Joan Serrà*

221. DreamMesh: Jointly Manipulating and Texturing Triangle Meshes for Text-to-3D Generation; *Haibo Yang; Yang Chen; Yingwei Pan*; Ting Yao; Zhineng Chen; Zuxuan Wu; Yu-Gang Jiang; Tao Mei*

222. Diffusion Models are Geometry Critics: Single Image 3D Editing Using Pre-Trained Diffusion Priors; *Ruicheng Wang*; Jianfeng Xiang; Jiaolong Yang; Xin Tong*

223. Videoshop: Localized Semantic Video Editing with Noise-Extrapolated Diffusion Inversion; *Xiang Fan*; Anand Bhattad; Ranjay Krishna*

224. DECOLLAGE: 3D Detailization by Controllable, Localized, and Learned Geometry Enhancement; *Qimin Chen*; Zhiqin Chen; Vladimir G. Kim; Noam Aigerman; Hao Zhang; Siddhartha Chaudhuri*

225. Nuvo: Neural UV Mapping for Unruly 3D Representations; *Pratul Srinivasan*; Stephan J Garbin; Dor Verbin; Jonathan T Barron; Ben Mildenhall*

226. FreeInit: Bridging Initialization Gap in Video Diffusion Models; *Tianxing Wu*; Chenyang Si; Yuming Jiang; Ziqi Huang; Ziwei Liu*

227. ReNoise: Real Image Inversion Through Iterative Noising; *Daniel Garibi*; Or Patashnik; Andrey Voynov; Hadar Averbuch-Elor; Danny Cohen-Or*

228. A Diffusion Model for Simulation Ready Coronary Anatomy with Morpho-skeletal Control; *Karim Kadry*; Shreya Gupta; Jonas Sogbadji; Michiel Schaap; Kersten Petersen; Takuya Mizukami; Carlos Collet; Farhad R. Nezami; Elazer R Edelman*

229. FreeCompose: Generic Zero-Shot Image Composition with Diffusion Prior; *Zhekai Chen; Wen Wang; Zhen Yang; Zeqing Yuan; Hao Chen*; Chunhua Shen**

230. Timestep-Aware Correction for Quantized Diffusion Models; *Yuzhe Yao; Feng Tian; Jun Chen*; Haonan Lin; Guang Dai; Yong Liu; Jingdong Wang*

231. Noise Calibration: Plug-and-play Content-Preserving Video Enhancement using Pre-trained Video Diffusion Models; *Qinyu Yang; Haoxin Chen; Yong Zhang*; Menghan Xia; Xiaodong Cun; Zhixun Su*; Ying Shan*

232. ZeST: Zero-Shot Material Transfer from a Single Image; *Ta-Ying Cheng; Prafull Sharma; Andrew Markham; Niki Trigoni; Varun Jampani**

233. View-Consistent 3D Editing with Gaussian Splatting; *Yuxuan Wang*; Xuanyu Yi; Zike Wu; Na Zhao; Long Chen; Hanwang Zhang*

234. Inf-DiT: Upsampling any-resolution image with memory-efficient diffusion transformer.; *Zhuoyi Yang*; Heyang Jiang; Wenyi Hong; Jiayan Teng; Wendi Zheng; Yuxiao Dong; Ming Ding; Jie Tang*

235. PolyOculus: Simultaneous Multi-view Image-based Novel View Synthesis; *Jason J. Yu*; Tristan Aumentado-Armstrong; Fereshteh Forghani; Konstantinos G. Derpanis; Marcus A. Brubaker*

236. Taming Latent Diffusion Model for Neural Radiance Field Inpainting; *Chieh Hubert Lin*; Changil Kim; Jia-Bin Huang; Qinbo Li; Chih-Yao Ma; Johannes Kopf; Ming-Hsuan Yang; Hung-Yu Tseng*

237. SNeRV: Spectra-preserving Neural Representation for Video; *Jina Kim*; Jihoo Lee*; Jewon Kang**

238. Learning Equilibrium Transformation for Gamut Expansion and Color Restoration; *Jun Xiao*; Changjian Shui; Zhi-Song Liu; Qian Ye; Kin-Man Lam*

239. HiFi-123: Towards High-fidelity One Image to 3D Content Generation; *Wangbo Yu*; Li Yuan; Yan-Pei Cao; Xiangjun Gao; Xiaoyu Li; Wenbo Hu; Long Quan; Ying Shan; Yonghong Tian*

240. Combining Generative and Geometry Priors for Wide-Angle Portrait Correction; *Lan Yao; Chaofeng Chen; Xiaoming Li*; Zifei Yan; Wangmeng Zuo*

241. Energy-Clibrated VAE with Test Time Free Lunch; *Yihong Luo; Siya Qiu; Xingjian Tao; Yujun Cai; Jing Tang**

242. CMTA: Cross-Modal Temporal Alignment for Event-guided Video Deblurring; *Taewoo Kim; Hoonhee Cho; Kuk-Jin Yoon**

243. SMFANet: A Lightweight Self-Modulation Feature Aggregation Network for Efficient Image Super-Resolution; *mingjun zheng; Long Sun; Jiangxin Dong; Jinshan Pan**

244. MegaScenes: Scene-Level View Synthesis at Scale; *Joseph Tung; Gene Chou*; Ruojin Cai; Guandao Yang; Kai Zhang; Gordon Wetzstein; Bharath Hariharan; Noah Snavely*

245. Continual Learning for Remote Physiological Measurement: Minimize Forgetting and Simplify Inference; *Qian Liang; Yan Chen; Yang Hu**

246. Flying with Photons: Rendering Novel Views of Propagating Light; *Anagh Malik*; Noah Juravsky; Ryan Po; Gordon Wetzstein; Kiriakos N. Kutulakos; David B. Lindell*

247. Personalized Video Relighting With an At-Home Light Stage; *Jun Myeong Choi*; Max Christman; Roni Sengupta*

248. Deblur e-NeRF: NeRF from Motion-Blurred Events under High-speed or Low-light Conditions; *Weng Fei Low*; Gim Hee Lee*

249. BlazeBVD: Make Scale-Time Equalization Great Again for Blind Video Deflickering; *Xinmin Qiu; Congying Han; Zicheng Zhang; Bonan Li*; Tiande Guo; Pingyu Wang; Xuecheng Nie*

250. FlashSplat: 2D to 3D Gaussian Splatting Segmentation Solved Optimally; *Qiuhong Shen; Xingyi Yang; Xinchao Wang**

251. Solving the inverse problem of microscopy deconvolution with a residual Beylkin-Coifman-Rokhlin neural network; *Rui Li; Mikhail Kudryashev; Artur Yakimovich**

252. Towards Robust Full Low-bit Quantization of Super Resolution Networks; *Denis S. Makhov*; Irina Zhelavskaya; Ruslan Ostapets; Dehua Song; Kirill Solodskikh*

253. Blind image deblurring with noise-robust kernel estimation; *Chanseok Lee*; Jeongsol Kim; Seungmin Lee; Jaehwang Jung; Yunje Cho; Taejoong Kim; Taeyong Jo; Myungjun Lee; Mooseok Jang**

254. Fast Context-Based Low-Light Image Enhancement via Neural Implicit Representations; *Tomáš Chobola*; Yu Liu; Hanyi Zhang; Julia A Schnabel; Tingying Peng**

255. Retrieval Robust to Object Motion Blur; *Rong Zou; Marc Pollefeys; Denys Rozumnyi**

256. Asymmetric Mask Scheme for Self-Supervised Real Image Denoising; *Xiangyu Liao*; Tianheng Zheng; Jiayu Zhong; Pingping Zhang; Chao Ren**

257. Vamos: Versatile Action Models for Video Understanding; *Shijie Wang*; Qi Zhao; Minh Quan Do; Nakul Agarwal; Kwonjoon Lee; Chen Sun*

258. A New Dataset and Framework for Real-World Blurred Images Super-Resolution; *Rui Qin; Ming Sun; Chao Zhou; Bin Wang**

259. Prompt-Based Test-Time Real Image Dehazing: A Novel Pipeline; *Zixuan Chen; Zewei He*; Ziqian Lu; Xuecheng Sun; Zheming Lu*

260. Compress3D: a Compressed Latent Space for 3D Generation from a Single Image; *Bowen Zhang*; Tianyu Yang*; Yu Li; Lei Zhang; Xi Zhao**

261. HoloADMM: High-Quality Holographic Complex Field Recovery; *Mazen Mel*; Paul Springer; Pietro Zanuttigh; Haitao Zhou; Alexander Gatto*

262. ReMatching: Low-Resolution Representations for Scalable Shape Correspondence; *Filippo Maggioli*; Daniele Baieri; Emanuele Rodola; Simone Melzi*

263. Learning Dual-Level Deformable Implicit Representation for Real-World Scale Arbitrary Super-Resolution; *Zhiheng Li; Muheng Li; Jixuan Fan; Lei Chen*; Yansong Tang; Jiwen Lu; Jie Zhou*

264. GaussianImage: 1000 FPS Image Representation and Compression by 2D Gaussian Splatting; *Xinjie Zhang; Xingtong Ge; Tongda Xu; Dailan He; Yan Wang; Hongwei Qin; Guo Lu; Jing Geng*; Jun Zhang**

265. 3DFG-PIFu: 3D Feature Grids for Human Digitization from Sparse Views; *Kennard Yanting Chan*; Fayao Liu; Guosheng Lin; Chuan Sheng Foo; Weisi Lin*

266. DiffuX2CT: Diffusion Learning to Reconstruct CT Images from Biplanar X-Rays; *Xuhui Liu; Zhi Qiao; Runkun Liu; Hong Li; Xiantong Zhen*; Zhen Qian; Juan Zhang*; Baochang Zhang*

267. Adaptive Compressed Sensing with Diffusion-Based Posterior Sampling; *Noam Elata*; Tomer Michaeli; Michael Elad*

268. CityGaussian: Real-time High-quality Large-Scale Scene Rendering with Gaussians; *Yang Liu; Chuanchen Luo; Lue Fan; Naiyan Wang; Junran Peng*; Zhaoxiang Zhang**

269. GRIDS: Grouped Multiple-Degradation Restoration with Image Degradation Similarity; *Shuo Cao; Yihao Liu; Wenlong Zhang; Yu Qiao; Chao Dong**

270. Goldfish: Vision-Language Understanding of Arbitrarily Long Videos; *Kirolos Ataallah*; Xiaoqian shen; Eslam mohamed abdelrahman*; Essam Sleiman; Mingchen Zhuge; Jian Ding; Deyao Zhu; Jürgen Schmidhuber; Mohamed Elhoseiny*

271. Learning Unsigned Distance Functions from Multi-view Images with Volume Rendering Priors; *Wenyuan Zhang; Kanle Shi; Yu-Shen Liu*; Zhizhong Han*

272. MambaIR: A Simple Baseline for Image Restoration with State-Space Model; *Hang Guo*; Jinmin Li; Tao Dai*; Zhihao Ouyang; Xudong Ren; Shu-Tao Xia*

273. GMT: Enhancing Generalizable Neural Rendering via Geometry-Driven Multi-Reference Texture Transfer; *Youngho Yoon; Hyun-Kurl Jang; Kuk-Jin Yoon**

274. Differentiable Convex Polyhedra Optimization from Multi-view Images; *Daxuan Ren*; Haiyi Mei; Hezi Shi; Jianmin Zheng; Jianfei Cai; Lei Yang*

275. MaRINeR: Enhancing Novel Views by Matching Rendered Images with Nearby References; *Lukas Bösiger*; Mihai Dusmanu; Marc Pollefeys; Zuria Bauer*

276. Long-range Turbulence Mitigation: A Large-scale Dataset and A Coarse-to-fine Framework; *Shengqi Xu; Run Sun; Yi Chang*; Shuning Cao; Xueyao Xiao; Luxin Yan*

277. Panel-Specific Degradation Representation for Raw Under-Display Camera Image Restoration; *Youngjin Oh*; Keuntek Lee; Jooyoung Lee; Dae-Hyun Lee; Nam Ik Cho*

278. MVSGaussian: Fast Generalizable Gaussian Splatting Reconstruction from Multi-View Stereo; *Tianqi Liu; Guangcong Wang; Shoukang Hu; Liao Shen; Xinyi Ye; Yuhang Zang; Zhiguo Cao*; Wei Li; Ziwei Liu*

279. Ray-Distance Volume Rendering for Neural Scene Reconstruction; *Ruihong Yin*; Yunlu Chen; Sezer Karaoglu; Theo Gevers*

280. Surˆ2f: A Hybrid Representation for High-Quality and Efficient Surface Reconstruction from Multi-view Images; *Zhangjin Huang*; Zhihao Liang; Kui Jia**

281. Efficient Depth-Guided Urban View Synthesis; *sheng miao*; Jiaxin Huang; Dongfeng Bai; Weichao Qiu; Liu Bingbing; Andreas Geiger; Yiyi Liao*

282. Rethinking Directional Parameterization in Neural Implicit Surface Reconstruction; *Zijie Jiang*; Tianhan Xu*; Hiroharu Kato*

283. SAH-SCI: Self-Supervised Adapter for Efficient Hyperspectral Snapshot Compressive Imaging; *Haijin Zeng; Yuxi Liu; Yongyong Chen*; Youfa Liu; Chong Peng; Jingyong Su*

284. UNIKD: UNcertainty-Filtered Incremental Knowledge Distillation for Neural Implicit Representation; *Mengqi Guo*; Chen Li; Hanlin Chen; Gim Hee Lee*

285. RAW-Adapter: Adapting Pretrained Visual Model to Camera RAW Images; *Ziteng Cui*; Tatsuya Harada*

286. Forest2Seq: Revitalizing Order Prior for Sequential Indoor Scene Synthesis; *Qi Sun*; Hang Zhou; Wengang Zhou; Li Li; Houqiang Li*

287. Text-Conditioned Resampler For Long Form Video Understanding; *Bruno Korbar*; Yongqin Xian; Alessio Tonioni; Andrew Zisserman; Federico Tombari*

288. Regularizing Dynamic Radiance Fields with Kinematic Fields; *Woobin Im; Geonho Cha; Sebin Lee; Jumin Lee; Juhyeong Seon; Dongyoon Wee; Sungeui Yoon**

289. Computing the Lipschitz constant needed for fast scene recovery from CASSI measurements; *Niels Chr Overgaard*; Anders Holst*

290. I2-SLAM: Inverting Imaging Process for Robust Photorealistic Dense SLAM; *Gwangtak Bae; Changwoon Choi; Hyeongjun Heo; Sang Min Kim; Young Min Kim**

291. Temporally Consistent Stereo Matching; *Jiaxi Zeng*; Chengtang Yao; Yuwei Wu*; Yunde Jia*

292. Diffusion Models for Monocular Depth Estimation: Overcoming Challenging Conditions; *Fabio Tosi; Pierluigi Zama Ramirez; Matteo Poggi**

293. Fundamental Matrix Estimation Using Relative Depths; *Yaqing Ding*; Václav Vávra; Snehal Bhayani; Qianliang Wu; Jian Yang; Zuzana Kukelova*

294. Mitigating Perspective Distortion-induced Shape Ambiguity in Image Crops; *Aditya Prakash*; Arjun Gupta; Saurabh Gupta*

295. GENIXER: Empowering Multimodal Large Language Models as a Powerful Data Generator; *Henry Hengyuan Zhao*; Pan Zhou*; Mike Zheng Shou**

296. Learning to Make Keypoints Sub-Pixel Accurate; *Shinjeong Kim*; Marc Pollefeys; Daniel Barath*

297. Track Everything Everywhere Fast and Robustly; *Yunzhou Song; Jiahui Lei*; Ziyun Wang; Lingjie Liu; Kostas Daniilidis*

298. VideoMamba: State Space Model for Efficient Video Understanding; *Kunchang Li*; Xinhao Li; Yi Wang*; Yinan He; Yali Wang*; Limin Wang*; Yu Qiao**

299. iMatching: Imperative Correspondence Learning; *Zitong Zhan*; Dasong Gao; Yun-Jou Lin; Youjie Xia; Chen Wang**

300. Vision-Language Action Knowledge Learning for Semantic-Aware Action Quality Assessment; *Huangbiao Xu; Xiao Ke*; Yuezhou Li; Rui Xu; Huanqi Wu; Xiaofeng Lin; Wenzhong Guo*

301. D-SCo: Dual-Stream Conditional Diffusion for Monocular Hand-Held Object Reconstruction; *Bowen Fu*; Gu Wang*; Chenyangguang Zhang; Yan Di; Ziqin Huang; Zhiying Leng; Fabian Manhardt; Xiangyang Ji*; Federico Tombari**

302. CountFormer: Multi-View Crowd Counting Transformer; *Hong Mo*; Xiong Zhang*; Jianchao Tan; Cheng Yang; Qiong Gu; Bo Hang; Wenqi Ren*

303. Easing 3D Pattern Reasoning with Side-view Features for Semantic Scene Completion; *Linxi Huan; Mingyue Dong; Linwei Yue; Shuhan Shen; Xianwei Zheng**

304. Hiding Imperceptible Noise in Curvature-Aware Patches for 3D Point Cloud Attack; *Mingyu Yang*; Daizong Liu; Keke Tang; Pan Zhou; Lixing Chen; Junyang Chen*

305. DatasetNeRF: Efficient 3D-aware Data Factory with Generative Radiance Fields; *Yu Chi*; Fangneng Zhan; Sibo Wu; Christian Theobalt; Adam Kortylewski*

306. GlobalPointer: Large-Scale Plane Adjustment with Bi-Convex Relaxation; *Bangyan Liao; Zhenjun Zhao; Lu Chen; Haoang Li; Daniel Cremers; Peidong Liu**

307. Explicitly Guided Information Interaction Network for Cross-modal Point Cloud Completion; *Xu Hang; Chen Long; Wenxiao Zhang*; Yuan Liu; Zhen Cao; Zhen Dong; Bisheng Yang*

308. WindPoly: Polygonal Mesh Reconstruction via Winding Numbers; *Xin He; Chenlei Lv; Pengdi Huang; Hui Huang**

309. Diffusion Bridges for 3D Point Cloud Denoising; *Mathias Vogel Hüni; Keisuke Tateno; Marc Pollefeys; Federico Tombari; Marie-Julie Rakotosaona; Francis Engelmann**

310. PolyRoom: Room-aware Transformer for Floorplan Reconstruction; *Yuzhou Liu; Lingjie Zhu; Xiaodong Ma; Hanqiao Ye; Xiang Gao; Xianwei Zheng; Shuhan Shen**

311. Towards a Density Preserving Objective Function for Learning on Point Sets; *Haritha Jayasinghe*; Ioannis Brilakis*

312. Syn-to-Real Domain Adaptation for Point Cloud Completion via Part-based Approach; *Yunseo Yang; Jihun Kim; Kuk-Jin Yoon**

313. Monocular Occupancy Prediction for Scalable Indoor Scenes; *Hongxiao Yu; Yuqi Wang; Yuntao Chen; Zhaoxiang Zhang**

314. RoDUS: Robust Decomposition of Static and Dynamic Elements in Urban Scenes; *Thang-Anh-Quan Nguyen*; Luis G Roldao Jimenez*; Nathan Piasco*; Moussab Bennehar*; Dzmitry Tsishkou**

315. Remove Projective LiDAR Depthmap Artifacts via Exploiting Epipolar Geometry; *Shengjie Zhu*; Girish Chandar Ganesan; Abhinav Kumar; Xiaoming Liu*

316. GroCo: Ground Constraint for Metric Self-Supervised Monocular Depth; *Aurélien Cecille*; Stefan Duffner; Franck Davoine; Thibault Neveu; Rémi Agier*

317. MapDistill: Boosting Efficient Camera-based HD Map Construction via Camera-LiDAR Fusion Model Distillation; *Xiaoshuai Hao*; Ruikai Li; Hui Zhang; Rong Yin; Dingzhe Li; Sangil Jung; Seung-In Park; ByungIn Yoo; Haimei Zhao; Jing Zhang*

318. TCC-Det: Temporarily consistent cues for weakly-supervised 3D detection; *Jan Skvrna*; Lukáš Neumann*

319. T-MAE: Temporal Masked Autoencoders for Point Cloud Representation Learning; *Weijie Wei*; Fatemeh Karimi Nejadasl; Theo Gevers; Martin R. Oswald**

320. Spatial-Temporal Multi-level Association for Video Object Segmentation; *Deshui Miao; Xin Li; Zhenyu He*; Huchuan Lu; Ming-Hsuan Yang*

321. 4D Contrastive Superflows are Dense 3D Representation Learners; *Xiang Xu*; Lingdong Kong; Hui Shuai; Wenwei Zhang; Liang Pan; Kai Chen; Ziwei Liu; Qingshan Liu**

322. nuCraft: Crafting High Resolution 3D Semantic Occupancy for Unified 3D Scene Understanding; *Benjin Zhu*; zhe wang; Hongsheng Li**

323. Trackastra: Transformer-based cell tracking for live-cell microscopy; *Benjamin Gallusser; Martin Weigert**

324. SeFlow: A Self-Supervised Scene Flow Method in Autonomous Driving; *Qingwen Zhang*; Yi Yang; Peizheng Li; Olov Andersson; Patric Jensfelt*

325. CARFF: Conditional Auto-encoded Radiance Field for 3D Scene Forecasting; *Jiezhi Yang*; Khushi P Desai*; Charles Packer*; Harshil bhatia; Nicholas Rhinehart; Rowan McAllister; Joseph E Gonzalez**

326. TrafficNight : An Aerial Multimodal Benchmark For Nighttime Vehicle Surveillance; *Guoxing Zhang; Yiming Liu; xiaoyu yang; Chao Huang*; HUANG Hailong*

327. CARB-Net: Camera-Assisted Radar-Based Network for Vulnerable Road User Detection; *Wei-Yu Lee*; Martin Dimitrievski; David Van Hamme; Jan Aelterman; Ljubomir Jovanov; Wilfried Philips*

328. Cross-view image geo-localization with Panorama-BEV Co-Retrieval Network; *Junyan Ye; Zhutao Lv; Weijia Li*; Jinhua Yu; Haote Yang; Huaping Zhong; Conghui He**

329. Online Temporal Action Localization with Memory-Augmented Transformer; *Youngkil Song; Dongkeun Kim; Minsu Cho; Suha Kwak**

330. Gated Temporal Diffusion for Stochastic Long-term Dense Anticipation; *Olga Zatsarynna*; Emad Bahrami*; Yazan Abu Farha; Gianpiero Francesca; Jürgen Gall**

331. Neural Volumetric World Models for Autonomous Driving; *Zanming Huang*; Jimuyang Zhang*; Eshed Ohn-Bar**

332. RepVF: A Unified Vector Fields Representation for Multi-task 3D Perception; *Jianbing Shen; Chunliang Li; Wencheng Han; Junbo Yin; Sanyuan Zhao**

334. CityGuessr: City-Level Video Geo-Localization on a Global Scale; *Parth Parag Kulkarni*; Gaurav Kumar Nayak; Mubarak Shah*

335. Risk-Aware Self-Consistent Imitation Learning for Trajectory Planning in Autonomous Driving; *Yixuan Fan*; Ya-Li Li; Shengjin Wang**

336. Stepwise Multi-grained Boundary Detector for Point-supervised Temporal Action Localization; *Mengnan Liu; Le Wang*; Sanping Zhou; Kun Xia; Qi Wu; Qilin Zhang; Gang Hua*

337. Safe-Sim: Safety-Critical Closed-Loop Traffic Simulation with Diffusion-Controllable Adversaries; *Wei-Jer Chang*; Francesco Pittaluga; Masayoshi Tomizuka; Wei Zhan; Manmohan Chandraker*

338. Occluded Gait Recognition with Mixture of Experts: An Action Detection Perspective; *Panjian Huang; Yunjie Peng; Saihui Hou*; Chunshui Cao; Xu Liu; Zhiqiang He; Yongzhen Huang**

339. SemTrack: A Large-scale Dataset for Semantic Tracking in the Wild; *Pengfei Wang; Xiaofei Hui; Jing Wu; Zile Yang; Kian Eng Ong; Xinge Zhao; Beijia Lu; Dezhao Huang; Evan Ling; Weiling Chen; Keng Teck Ma; Minhoe Hur; Jun Liu**

340. Progressive Pretext Task Learning for Human Trajectory Prediction; *Xiaotong Lin; Tianming Liang; Jianhuang Lai; Jian-Fang Hu**

341. Dolphins: Multimodal Language Model for Driving; *Yingzi Ma; Yulong Cao; Jiachen Sun; Marco Pavone; Chaowei Xiao**

342. PRET: Planning with Directed Fidelity Trajectory for Vision and Language Navigation; *Renjie Lu; Jingke Meng*; WEI-SHI ZHENG*

343. LingoQA: Video Question Answering for Autonomous Driving; *Ana-Maria Marcu*; Long Chen; Jan Hünermann; Alice Karnsund; Benoit Hanotte; Prajwal Chidananda; Saurabh Nair; Vijay Badrinarayanan; Alex Kendall; Jamie Shotton; Elahe Arani; Oleg Sinavski*

344. LLM as Copilot for Coarse-grained Vision-and-Language Navigation; *Yanyuan Qiao*; Qianyi Liu; Jiajun Liu; Jing Liu; Qi Wu*

12:00 – 13:30
Speed Mentoring – Space 4

12:00 – 14:00
Doctoral Consortium – Brown 3

Antonio Alliegro - Politecnico di Torino
Camillo Quattrocchi - University of Catania
Changhoon Kim - Arizona State University
Deepti B Hegde - Johns Hopkins University
Denys Rozumnyi - ETH Zurich
Evin Pınar Örnek - TU Munich
Ginger D Delmas - Naver Labs Europe
Guénolé Fiche - CentraleSupélec
Hyeokjun Kweon - KAIST
Irit Chelly - Ben Gurion University of the Negev
Ishan Rajendrakumar Dave - University of Central Florida
Ivona Najdenkoska - University of Amsterdam
Jaehui Hwang - Yonsei University
Janis Keuper - Offenburg University
Jia-Wei Liu - National University of Singapore
Julia Grabinski - University of Mannheim
Kai Chen - The Chinese University of Hong Kong
Manasi Muglikar - University of Zurich
Mathis Petrovich - Ecole des Ponts
Mina Ghadimi Atigh - University of Amsterdam
Parsa Mirdehghan - University of Toronto
Rosario Leonardi - University of Catania
Rui Gong - ETH Zurich
Sayanton V. Dibbo - Dartmouth College
Seunggeun Chi - Purdue University
Shahaf E Finder - Ben-Gurion University
Sung-Hoon Yoon - KAIST
Surbhi Mittal - Indian Institute of Technology, Jodhpur
Tuan-Anh Vu - The Hong Kong University of Science and Technology
Weihao Xia - University College London
Yeji Song - Seoul National University
Zhecan Wang - Columbia University
Zhiyang Dou - University of Pennsylvania
Zipeng Xu - University of Trento

12:30 – 13:30
Lunch – Exhibition Area (Level 0) & Balcony Level 1

13:30 – 15:30
Oral session 2A: Generative models I - Gold Room
Chairs: Hedvig Kjellström; Jianfei Cai

1. EDTalk: Efficient Disentanglement for Emotional Talking Head Synthesis; *Shuai Tan*; Bin Ji; Mengxiao Bi; ye pan**

2. TexDreamer: Towards Zero-Shot High-Fidelity 3D Human Texture Generation; *Yufei Liu; Junwei Zhu; Junshu Tang; Shijie Zhang; Jiangning Zhang; Weijian Cao; Chengjie Wang; Yunsheng Wu; Dongjin Huang**

3. LGM: Large Multi-View Gaussian Model for High-Resolution 3D Content Creation; *Jiaxiang Tang*; Zhaoxi Chen; Xiaokang Chen; Tengfei Wang; Gang Zeng; Ziwei Liu*

4. FlashTex: Fast Relightable Mesh Texturing with LightControlNet; *Kangle Deng*; Timothy Omernick; Alexander B Weiss; Deva Ramanan; Jun-Yan Zhu; Tinghui Zhou; Maneesh Agrawala*

5. TextDiffuser-2: Unleashing the Power of Language Models for Text Rendering; *Jingye Chen*; Yupan Huang; Tengchao Lv; Lei Cui; Qifeng Chen; Furu Wei*

6. LLMGA: Multimodal Large Language Model based Generation Assistant; *bin xia*; Shiyin Wang; Yingfan Tao; Yitong Wang; Jiaya Jia*

7. Accelerating Image Generation with Sub-path Linear Approximation Model; *Chen Xu; Tianhui Song; Weixin Feng; Xubin Li; Tiezheng Ge; Bo Zheng; Limin Wang**

8. SphereHead: Stable 3D Full-head Synthesis with Spherical Tri-plane Representation; *Heyuan Li*; Ce Chen; Tianhao Shi; Yuda Qiu; Sizhe An; Guanying CHEN; Xiaoguang Han**

9. Bridging the Gap: Studio-like Avatar Creation from a Monocular Phone Capture; *ShahRukh Athar*; Shunsuke Saito; Stanislav Pidhorskyi; Zhengyu Yang; Chen Cao*

10. Zero-Shot Detection of AI-Generated Images; *Davide Cozzolino; Glovanni Poggi; Matthias Niessner; Luisa Verdoliva**

11. Action2Sound: Ambient-Aware Generation of Action Sounds from Egocentric Videos; *Changan Chen*; Puyuan Peng; Ami Baid; Zihui Xue; Wei-Ning Hsu; David Harwath; Kristen Grauman*

13:30 – 15:30
Oral session 2B: Recognition - Auditorium
Chairs: Jordi Pont-Tuset; Sara Beery
1. Efficient Bias Mitigation Without Privileged Information; *Mateo Espinosa Zarlenga*; Swami Sankaranarayanan; Jerone T. A. Andrews; Zohreh Shams; Mateja Jamnik; Alice Xiang* BEST PAPER CANDIDATE

2. Fast Diffusion-Based Counterfactuals for Shortcut Removal and Generation; *Nina Weng*; Paraskevas Pegios; Eike Petersen; Aasa Feragen; Siavash Arjomand Bigdeli*

3. MobileNetV4: Universal Models for the Mobile Ecosystem; *Danfeng Qin*; Chas H Leichner; Manolis Delakis; Marco Fornoni; Shixin Luo; Fan Yang; Weijun Wang; Colby Banbury; Chengxi Ye; Berkin Akin; Vaibhav Aggarwal; Tenghui Zhu; Daniele Moro; Andrew Howard*

4. Momentum Auxiliary Network for Supervised Local Learning; *Junhao Su; Changpeng Cai; Feiyu Zhu; Chenghao He; Xiaojie Xu; Dongzhi Guan*; Chenyang Si**

5. From Fake to Real: Pretraining on Balanced Synthetic Images to Prevent Spurious Correlations in Image Recognition; *Maan Qraitem*; Kate Saenko; Bryan A. Plummer*

6. Dataset Enhancement with Instance-Level Augmentations; *Orest Kupyn*; Christian Rupprecht*

7. Adaptive Parametric Activation; *Konstantinos P Alexandridis*; Jiankang Deng; Anh Nguyen; Shan Luo*

8. Relation DETR: Exploring Explicit Position Relation Prior for Object Detection; *Xiuquan Hou; Meiqin Liu*; Senlin Zhang; Ping Wei; Badong Chen; Xuguang Lan*

9. Projecting Points to Axes: Oriented Object Detection via Point-Axis Representation; *Zeyang Zhao; Qilong Xue; Yifan Bai; Yuhang He; Xing Wei*; Yihong Gong*

10. CLIFF: Continual Latent Diffusion for Open-Vocabulary Object Detection; *Wuyang Li; Xinyu Liu; Jiayi Ma; Yixuan Yuan**

11. On Calibration of Object Detectors: Pitfalls, Evaluation and Baselines; *Selim Kuzucu*; Kemal Oksuz*; Jonathan Sadeghi; Puneet Dokania*

13:30 – 15:30 Oral session 2C: Multi-view and visual odometry - Silver Room

Chairs: Dan Xu; Laurent Kneip

1. Physics-Free Spectrally Multiplexed Photometric Stereo under Unknown Spectral Composition; *Satoshi Ikehata*; Yuta Asano*

2. COMO: Compact Mapping and Odometry; *Eric Dexheimer*; Andrew Davison*

3. Smoothness, Synthesis, and Sampling: Re-thinking Unsupervised Multi-View Stereo with DIV Loss; *Alex Rich*; Noah Stier; Pradeep Sen; Tobias Hollerer*

4. ADen: Adaptive Density Representations for Sparse-view Camera Pose Estimation; *Hao Tang; Weiyao Wang; Pierre Gleize; Matt Feiszli**

5. SPVLoc: Semantic Panoramic Viewport Matching for 6D Camera Localization in Unseen Environments; *Niklas Gard*; Anna Hilsmann; Peter Eisert*

6. Six-Point Method for Multi-Camera Systems with Reduced Solution Space; *Banglei Guan; Ji Zhao*; Laurent Kneip*

7. Scene Coordinate Reconstruction: Posing of Image Collections via Incremental Learning of a Relocalizer; *Eric Brachmann*; Jamie Wynn; Shuai Chen; Tommaso Cavallari; Aron Monszpart; Daniyar Turmukhambetov; Victor Adrian Prisacariu*

8. Grounding Image Matching in 3D with MASt3R; *Vincent Leroy*; Yohann Cabon; Jerome Revaud*

9. ConDense: Consistent 2D-3D Pre-training for Dense and Sparse Features from Multi-View Images; *Xiaoshuai Zhang*; Zhicheng Wang; Howard Zhou; Soham Ghosh; Danushen L Gnanapragasam; Varun Jampani; Hao Su; Leonidas Guibas*

10. Correspondences of the Third Kind: Camera Pose Estimation from Object Reflection; *Kohei Yamashita*; Vincent Lepetit; Ko Nishino*

11. Camera Calibration using a Collimator System; *Shunkun Liang; Banglei Guan*; Zhenbao Yu; Pengju Sun; Yang Shang*

14:30 – 18:00 Demo session 2 - Level 0
1. OpenCity: Open-Vocabulary Attribution of 3D Buildings in City-Scale Photogrammetric Meshes; *Hakeem Frank, Caleb Buffa, Justin Chae, Justin Snider, Patrick Tutzauer, Dmitry Kudinov - Environmental Systems Research Institute*

2. Visual Place Recognition using 3D City Models; *Lorenz Junglas, Gabriele Berton, Thomas Pollok, Carlo Masone, Barbara Caputo - Karlsruhe Institute of Technology*

3. Leveraging Computer Vision on the Ski Slopes; *Matteo Dunnhofer, Christian Micheloni - University of Udine*4. A Tool for Collecting Spatio-temporally Sparse Point Annotations for Video Object Segmentation; *Idil Esen Zulfikar, Sabarinath Mahadevan, Paul Voigtlaender - RWTH Aachen University*

4. A Tool for Collecting Spatio-temporally Sparse Point Annotations for Video Object Segmentation; *Idil Esen Zulfikar, Sabarinath Mahadevan, Paul Voigtlaender - RWTH Aachen University*

5. AR Deployment and Scene Modelling on Your Phone; Kourosh Riahidehkordi, Mahtab Dahaghin, Myrna Castillo, Matteo Toso, Alessio Del Bue - Istituto Italiano di Tecnologia

15:30 – 16:30
Keynote Lecture - Gold Room (live), Auditorium (broadcast), Silver Room (broadcast)
Synthesia: From computer vision research to real-world AI avatars; *Lourdes Agapito; Vittorio Ferrari*

16:30 – 17:00
Lightning AI Technical Session - Technical Presentation Area (Level 0)
Building Scalable AI for Real-World Business

16:30 – 17:00
Coffee Break - Exhibition Area (Level 0)

16:30 – 18:30
Poster session 2
1. LaWa: Using Latent Space for In-Generation Image Watermarking; *Ahmad Rezaei*; Mohammad Akbari*; Saeed Ranjbar Alvar; Arezou Fatemi; Yong Zhang**

2. Delving into Adversarial Robustness on Document Tampering Localization; *Huiru Shao; Zhuang Qian; Kaizhu Huang; Wei Wang; Xiaowei Huang; Qiufeng Wang**

3. Contrasting Deepfakes Diffusion via Contrastive Learning and Global-Local Similarities; *Lorenzo Baraldi*; Federico Cocchi; Marcella Cornia; Lorenzo Baraldi; Alessandro Nicolosi; Rita Cucchiara*

4. Rethinking Data Bias: Dataset Copyright Protection via Embedding Class-wise Hidden Bias; *Jinhyeok Jang*; ByungOk Han; Jaehong Kim; Chan-Hyun Youn*

5. Generalizable Facial Expression Recognition; *Yuhang Zhang; Xiuqi Zheng; Chenyi Liang; Jiani Hu*; Weihong Deng*

6. Catastrophic Overfitting: A Potential Blessing in Disguise; *MN Zhao; Lihe Zhang*; Yuqiu Kong; Baocai Yin*

7. Prediction Exposes Your Face: Black-box Model Inversion via Prediction Alignment; *Yufan Liu*; Wanqian Zhang; Dayan Wu; Zheng Lin; jingzi Gu; Weiping Wang*

8. CatchBackdoor: Backdoor Detection via Critical Trojan Neural Path Fuzzing; *Haibo Jin; Ruoxi Chen; Jinyin Chen; Haibin Zheng; Yang Zhang; Haohan Wang**

9. Cocktail Universal Adversarial Attack on Deep Neural Networks; *Shaoxin Li*; Xiaofeng Liao; Xin Che; Xintong Li; Yong Zhang; Lingyang Chu**

10. Unveiling Privacy Risks in Stochastic Neural Networks Training: Effective Image Reconstruction from Gradients; *Yiming Chen*; Xiangyu Yang; Nikos Deligiannis*

11. SuperFedNAS: Cost-Efficient Federated Neural Architecture Search for On-Device Inference; *Alind Khare*; Animesh Agrawal; Aditya Annavajjala; Payman Behnam; Myungjin Lee; Hugo M Latapie; Alexey Tumanov*

12. Preventing Catastrophic Overfitting in Fast Adversarial Training: A Bi-level Optimization Perspective; *Zhaoxin Wang*; Handing Wang*; Cong Tian; Yaochu Jin*

13. SkyMask: Attack-agnostic Robust Federated Learning with Fine-grained Learnable Masks; *Peishen Yan; Hao Wang; Tao Song*; Yang Hua; Ruhui Ma; Ningxin Hu; Mohammad Reza Haghighat; Haibing Guan*

14. From Fake to Real: Pretraining on Balanced Synthetic Images to Prevent Spurious Correlations in Image Recognition; *Maan Qraitem*; Kate Saenko; Bryan A. Plummer*

15. Continuous Memory Representation for Anomaly Detection; *Joo Chan Lee*; Taejune Kim; Eunbyung Park*; Simon S Woo*; Jong Hwan Ko**

16. Learning Anomalies with Normality Prior for Unsupervised Video Anomaly Detection; *Haoyue Shi; Le Wang*; Sanping Zhou; Gang Hua; Wei Tang*

17. Uncertainty Calibration with Energy Based Instance-wise Scaling in the Wild Dataset; *Mijoo Kim; Junseok Kwon**

18. Few-Shot Anomaly-Driven Generation for Anomaly Classification and Segmentation; *Guan Gui; Bin-Bin Gao*; Jun Liu; Chengjie Wang; Yunsheng Wu*

19. FlowCon: Out-of-Distribution Detection using Flow-based Contrastive Learning; *Saandeep Aathreya*; Shaun Canavan**

20. EntAugment: Entropy-Driven Adaptive Data Augmentation Framework for Image Classification; *Suorong Yang*; Furao Shen*; Jian Zhao*

21. PixOOD: Pixel-Level Out-of-Distribution Detection; *Tomas Vojir*; Jan Sochman; Jiri Matas*

22. MobileNetV4: Universal Models for the Mobile Ecosystem; *Danfeng Qin*; Chas H Leichner; Manolis Delakis; Marco Fornoni; Shixin Luo; Fan Yang; Weijun Wang; Colby Banbury; Chengxi Ye; Berkin Akin; Vaibhav Aggarwal; Tenghui Zhu; Daniele Moro; Andrew Howard*

23. SiT: Exploring Flow and Diffusion-based Generative Models with Scalable Interpolant Transformers; *Nanye Ma*; Mark Goldstein; Michael Albergo; Nicholas M Boffi; Eric Vanden-Eijnden*; Saining Xie**

24. Adaptive Parametric Activation; *Konstantinos P Alexandridis*; Jiankang Deng; Anh Nguyen; Shan Luo*

25. Deep Nets with Subsampling Layers Unwittingly Discard Useful Activations at Test-Time; *Chiao-An Yang*; Ziwei Liu; Raymond Yeh*

26. Momentum Auxiliary Network for Supervised Local Learning; *Junhao Su; Changpeng Cai; Feiyu Zhu; Chenghao He; Xiaojie Xu; Dongzhi Guan*; Chenyang Si**

27. Efficient Bias Mitigation Without Privileged Information; *Mateo Espinosa Zarlenga*; Swami Sankaranarayanan; Jerone T. A. Andrews; Zohreh Shams; Mateja Jamnik; Alice Xiang* BEST PAPER CANDIDATE

28. PredBench: Benchmarking Spatio-Temporal Prediction across Diverse Disciplines; *ZiDong Wang*; Zeyu Lu*; Di Huang*; Tong He; Xihui Liu; Wanli Ouyang; Lei Bai**

29. Disentangling Masked Autoencoders for Unsupervised Domain Generalization; *An Zhang*; Han Wang; Xiang Wang; Tat-Seng Chua*

30. Adversarially Robust Distillation by Reducing the Student-Teacher Variance Gap; *Junhao Dong; Piotr Koniusz*; Junxi Chen; Yew-Soon Ong**

31. MMEarth: Exploring Multi-Modal Pretext Tasks For Geospatial Representation Learning; *Vishal Nedungadi; Ankit Kariryaa; Stefan Oehmcke; Serge Belongie; Christian Igel; Nico Lang**

32. Assessing Sample Quality via the Latent Space of Generative Models; *Jingyi Xu*; Hieu Le; Dimitris Samaras*

33. FedRA: A Random Allocation Strategy for Federated Tuning to Unleash the Power of Heterogeneous Clients; *Shangchao Su; Bin Li*; Xiangyang Xue*

34. Markov Knowledge Distillation: Make Nasty Teachers trained by Self-undermining Knowledge Distillation Fully Distillable; *En-hui Yang; Linfeng Ye**

35. Iterative Ensemble Training with Anti-Gradient Control for Mitigating Memorization in Diffusion Models; *Xiao Liu; Xiaoliu Guan; Yu Wu*; Jiaxu Miao**

36. Mixture of Efficient Diffusion Experts Through Automatic Interval and Sub-Network Selection; *Alireza Ganjdanesh*; Yan Kang; Yuchen Liu; Richard Zhang; Zhe Lin; Heng Huang*

37. Improving 3D Semi-supervised Learning by Effectively Utilizing All Unlabelled Data; *Sneha Paul*; Zachary Patterson; Nizar Bouguila*

38. Information Bottleneck Based Data Correction in Continual Learning; *Shuai Chen; mingyi zhang; Junge Zhang*; Kaiqi Huang**

39. Scaling Backwards: Minimal Synthetic Pre-training?; *Ryo Nakamura*; Ryu Tadokoro*; Ryosuke Yamada*; Yuki M Asano*; Iro Laina*; Christian Rupprecht*; Nakamasa Inoue*; Rio Yokota*; Hirokatsu Kataoka**

40. Distributionally Robust Loss for Long-Tailed Multi-Label Image Classification; *Dekun Lin*; Zhe Cui; Rui Chen; Tailai Peng; xinran xie; Xiaolin Qin*

41. Bad Students Make Great Teachers: Active Learning Accelerates Large-Scale Visual Understanding; *Talfan Evans*; Shreya Pathak; Hamza Merzic; Jonathan Richard Schwarz; Ryutaro Tanno; Olivier Henaff**

42. GKGNet: Group K-Nearest Neighbor based Graph Convolutional Network for Multi-Label Image Recognition; *Ruijie Yao; Sheng Jin; Lumin Xu; Wang Zeng; Wentao Liu; Chen Qian*; Ping Luo; Ji Wu**

43. Generalized Coverage for More Robust Low-Budget Active Learning; *Wonho Bae; Junhyug Noh; Danica J. Sutherland**

44. Modality Translation for Object Detection Adaptation without forgetting prior knowledge; *Heitor Rapela Medeiros*; Masih Aminbeidokhti; Fidel A Guerrero Pena; David Latortue; Eric Granger; Marco Pedersoli*

45. Category Adaptation Meets Projected Distillation in Generalized Continual Category Discovery; *Grzegorz Rypeść*; Daniel Marczak; Sebastian Cygert; Tomasz Trzcinski; Bartlomiej Twardowski*

46. CroMo-Mixup: Augmenting Cross-Model Representations for Continual Self-Supervised Learning; *Erum Mushtaq*; Duygu Nur Yaldiz; Yavuz Faruk Bakman; Jie Ding; Chenyang Tao; Dimitrios Dimitriadis; Salman Avestimehr*

47. Dyn-Adapter: Towards Disentangled Representation for Efficient Visual Recognition; *Yurong Zhang*; Honghao Chen; Zhang Xinyu; Xiangxiang Chu; Li Song*

48. Class-Incremental Learning with CLIP: Adaptive Representation Adjustment and Parameter Fusion; *Linlan Huang; Xusheng Cao; Haori Lu; Xialei Liu**

49. Beyond Prompt Learning: Continual Adapter for Efficient Rehearsal-Free Continual Learning; *Xinyuan Gao; Songlin Dong; Yuhang He*; Qiang Wang; Yihong Gong*

50. Training-Free Model Merging for Multi-target Domain Adaptation; *Wenyi Li; Huan-ang Gao; Mingju Gao; Beiwen Tian; Rong Zhi; Hao Zhao**

51. Robust Nearest Neighbors for Source-Free Domain Adaptation under Class Distribution Shift; *Antonio Tejero-de-Pablos*; Riku Togashi; Mayu Otani; Shin'ichi Satoh*

52. Learning the Unlearned: Mitigating Feature Suppression in Contrastive Learning; *Jihai Zhang; Xiang Lan; Xiaoye Qu; Yu Cheng; Mengling Feng*; Bryan Hooi**

53. A Multimodal Benchmark Dataset and Model for Crop Disease Diagnosis; *Xiang Liu; Zhaoxiang Liu*; Huan Hu; Zezhou Chen; Kohou Wang; Kai Wang; Shiguo Lian**

54. Improving Intervention Efficacy via Concept Realignment in Concept Bottleneck Models; *Nishad Singhi*; Jae Myung Kim; Karsten Roth; Zeynep Akata*

55. Enhancing Source-Free Domain Adaptive Object Detection with Low-confidence Pseudo Label Distillation; *Ilhoon Yoon; Hyeongjun Kwon; Jin Kim; Junyoung Park; Hyunsung Jang; Kwanghoon Sohn**

56. CoDA: Instructive Chain-of-Domain Adaptation with Severity-Aware Visual Prompt Tuning; *ZiYang Gong; FuHao Li; Yupeng Deng; Deblina Bhattacharjee; Xianzheng Ma*; Xiangwei Zhu*; Zhenming Ji**

57. Improving Zero-shot Generalization of Learned Prompts via Unsupervised Knowledge Distillation; *Marco Mistretta*; Alberto Baldrati; Marco Bertini; Andrew D. Bagdanov*

58. SMILe: Leveraging Submodular Mutual Information For Robust Few-Shot Object Detection; *Anay Majee*; Ryan X Sharp; Rishabh Iyer**

59. Attention-Challenging Multiple Instance Learning for Whole Slide Image Classification; *Yunlong Zhang*; Honglin Li; YUXUAN SUN; Chenglu Zhu; Sunyi Zheng; Lin Yang**

60. Semantic-guided Robustness Tuning for Few-Shot Transfer Across Extreme Domain Shift; *kangyu xiao*; Zilei Wang; junjie li*

61. GTP-4o: Modality-prompted Heterogeneous Graph Learning for Omni-modal Biomedical Representation; *Chenxin Li*; Xinyu Liu; Cheng Wang; Yifan Liu; Weihao Yu; Jing Shao; Yixuan Yuan*

62. Identity-Consistent Diffusion Network for Grading Knee Osteoarthritis Progression in Radiographic Imaging; *Wenhua Wu; Kun Hu*; Wenxi Yue; Wei Li; Milena Simic; Changyang Li; Wei Xiang; Zhiyong Wang*

63. Make a Strong Teacher with Label Assistance: A Novel Knowledge Distillation Approach for Semantic Segmentation; *Shoumeng Qiu; Jie Chen; Xinrun Li; Ru Wan; Xiangyang Xue; Jian Pu**

64. Mitigating Background Shift in Class-Incremental Semantic Segmentation; *Gilhan Park; WonJun Moon; SuBeen Lee; Tae-Young Kim; Jae-Pil Heo**

65. Zero-shot Object Counting with Good Exemplars; *Huilin Zhu; Jingling Yuan; Zhengwei Yang; Yu Guo; Xian Zhong*; Zheng Wang; Shengfeng He**

66. AugDETR: Improving Multi-scale Learning for Detection Transformer; *Jinpeng Dong; Yutong Lin; Chen Li; Sanping Zhou; Nanning Zheng**

67. DeTra: A Unified Model for Object Detection and Trajectory Forecasting; *Sergio Casas*; Ben T Agro; Jiageng Mao; Thomas Gilles; ALEXANDER Y CUI; Enxu Li; Raquel Urtasun*

68. RoScenes: A Large-scale Multi-view 3D Dataset for Roadside Perception; *Xiaosu Zhu; Hualian Sheng; Sijia Cai; Bing Deng; Shaopeng Yang; Qiao Liang; Ken Chen; Lianli Gao; Jingkuan Song*; Jieping Ye**

69. MARs: Multi-view Attention Regularizations for Patch-based Feature Recognition of Space Terrain; *Timothy Chase Jr*; Karthik Dantu*

70. ScribblePrompt: Fast and Flexible Interactive Segmentation for Any Biomedical Image; *Hallee E. Wong*; Marianne Rakic; John Guttag; Adrian V. Dalca*

71. NePhi: Neural Deformation Fields for Approximately Diffeomorphic Medical Image Registration; *Lin Tian*; Thomas H Greer; Raul San Jose Estepar; Roni Sengupta; Marc Niethammer*

72. PMT: Progressive Mean Teacher via Exploring Temporal Consistency for Semi-Supervised Medical Image Segmentation; *Ning Gao; Sanping Zhou*; Le Wang; Nanning Zheng*

73. Train Till You Drop: Towards Stable and Robust Source-free Unsupervised 3D Domain Adaptation; *Björn Michele*; Alexandre Boulch; Tuan-Hung VU; Gilles Puy; Renaud Marlet; Nicolas Courty*

74. TAG: Text Prompt Augmentation for Zero-Shot Out-of-Distribution Detection; *Xixi Liu*; Christopher Zach*

75. Explain via Any Concept: Concept Bottleneck Model with Open Vocabulary Concepts; *Andong Tan; Fengtao Zhou; Hao Chen**

76. Region-centric Image-Language Pretraining for Open-Vocabulary Detection; *Dahun Kim*; Anelia Angelova; Weicheng Kuo*

77. CLIFF: Continual Latent Diffusion for Open-Vocabulary Object Detection; *Wuyang Li; Xinyu Liu; Jiayi Ma; Yixuan Yuan**

78. Open-Vocabulary Camouflaged Object Segmentation; *Youwei Pang; Xiaoqi Zhao; JiaMing Zuo; Lihe Zhang*; Huchuan Lu*

79. LiFT: A Surprisingly Simple Lightweight Feature Transform for Dense ViT Descriptors; *Saksham Suri*; Matthew Walmer; Kamal Gupta; Abhinav Shrivastava*

80. Integration of Global and Local Representations for Fine-grained Cross-modal Alignment; *Seungwan Jin; Hoyoung Choi; Taehyung Noh; Kyungsik Han**

81. Three Things We Need to Know About Transferring Stable Diffusion to Visual Dense Prediciton Tasks; *Manyuan Zhang*; Guanglu Song; Xiaoyu Shi; Yu Liu; Hongsheng Li*

82. OTSeg: Multi-prompt Sinkhorn Attention for Zero-Shot Semantic Segmentation; *Kwanyoung Kim; Yujin Oh; Jong Chul Ye**

83. Grid-Attention: Enhancing Computational Efficiency of Large Vision Models without Fine-Tuning; *Pengyu Li*; biao wang; Tianchu Guo; Xian-Sheng Hua*

84. Superpixel-informed Implicit Neural Representation for Multi-Dimensional Data; *Jia-Yi Li; Xi-Le Zhao*; Jian-Li Wang; Chao Wang; Min Wang*

85. DEAL: Disentangle and Localize Concept-level Explanations for VLMs; *Tang Li*; Mengmeng Ma; Xi Peng*

86. Comprehensive Attribution: Inherently Explainable Vision Model with Feature Detector; *Xianren Zhang; Dongwon Lee; Suhang Wang**

87. Pseudo-RIS: Distinctive Pseudo-supervision Generation for Referring Image Segmentation; *Seonghoon Yu; Paul Hongsuck Seo*; Jeany Son**

88. Textual Grounding for Open-vocabulary Visual Information Extraction in Layout-diversified Documents; *Mengjun Cheng; Chengquan Zhang; Chang Liu*; Yuke Li; Bohan Li; Kun Yao; Xiawu Zheng; Rongrong Ji; Jie Chen*

89. Fairness-aware Vision Transformer via Debiased Self-Attention; *Yao Qiang; Chengyin Li; Prashant Khanduri; Dongxiao Zhu**

90. Exploring Phrase-Level Grounding with Text-to-Image Diffusion Model; *Danni Yang; Ruohan Dong; Jiayi Ji; Yiwei Ma; Haowei Wang; Xiaoshuai Sun*; Rongrong Ji*

91. Rethinking and Improving Visual Prompt Selection for In-Context Learning Segmentation Framework; *Wei Suo; Lanqing Lai; Mengyang Sun; Hanwang Zhang; Peng Wang*; Yanning Zhang*

92. GRiT: A Generative Region-to-text Transformer for Object Understanding; *Jialian Wu*; Jianfeng Wang; Zhengyuan Yang; Zhe Gan; Zicheng Liu; Junsong Yuan; Lijuan Wang*

93. PSALM: Pixelwise Segmentation with Large Multi-modal Model; *Zheng Zhang; yeyao ma; Enming Zhang; Xiang Bai**

94. SPHINX: A Mixer of Weights, Visual Embeddings and Image Scales for Multi-modal Large Language Models; *Ziyi Lin; Dongyang Liu; Renrui Zhang; Peng Gao*; Longtian Qiu; Han Xiao; Han Qiu; Wenqi Shao; Keqin Chen; Jiaming Han; Siyuan Huang; Yichi Zhang; Xuming He; Yu Qiao*; Hongsheng Li**

95. View Selection for 3D Captioning via Diffusion Ranking; *Tiange Luo*; Justin Johnson; Honglak Lee*

96. ShareGPT4V: Improving Large Multi-Modal Models with Better Captions; *Lin Chen*; Jinsong Li; Xiaoyi Dong; Pan Zhang; Conghui He; Jiaqi Wang; Feng Zhao*; Dahua Lin**

97. MyVLM: Personalizing VLMs for User-Specific Queries; *Yuval Alaluf*; Elad Richardson; Sergey Tulyakov; Kfir Aberman; Danny Cohen-Or*

98. SegVG: Transferring Object Bounding Box to Segmentation for Visual Grounding; *Weitai Kang*; Gaowen Liu; Mubarak Shah; Yan Yan*

99. Safe-CLIP: Removing NSFW Concepts from Vision-and-Language Models; *Samuele Poppi*; Tobia Poppi*; Federico Cocchi; Marcella Cornia; Lorenzo Baraldi; Rita Cucchiara*

100. LookupViT: Compressing visual information to a limited number of tokens; *Rajat Koner; Gagan Jain; Sujoy Paul*; Volker Tresp; Prateek Jain*

101. IVTP: Instruction-guided Visual Token Pruning for Large Vision-Language Models; *Kai Huang*; Hao Zou; Ye Xi; Bochen Wang; Zhen Xie; Liang Yu*

102. Instruction Tuning-free Visual Token Complement for Multimodal LLMs; *Dongsheng Wang*; Jiequan Cui; Miaoge Li; Wang Lin; Bo Chen; Hanwang Zhang*

103. Contrastive Region Guidance: Improving Grounding in Vision-Language Models without Training; *David Wan*; Jaemin Cho; Elias Stengel-Eskin; Mohit Bansal*

104. Visual Alignment Pre-training for Sign Language Translation; *Peiqi Jiao; Yuecong Min; Xilin Chen**

105. AdaLog: Post-Training Quantization for Vision Transformers with Adaptive Logarithm Quantizer; *Zhuguanyu Wu; Jiaxin Chen*; Hanwen Zhong; Di Huang; Yunhong Wang*

106. FineMatch: Aspect-based Fine-grained Image and Text Mismatch Detection and Correction; *Hang Hua*; Jing Shi; Kushal Kafle; Simon Jenni; Daoan Zhang; John Collomosse; Scott Cohen; Jiebo Luo*

107. From Pixels to Objects: A Hierarchical Approach for Part and Object Segmentation Using Local and Global Aggregation; *Yunfei Xie*; Cihang Xie; Alan Yuille; Jieru Mei*

108. LG-Gaze: Learning Geometry-aware Continuous Prompts for Language-Guided Gaze Estimation; *Pengwei Yin*; Jingjing Wang; Guanzhong Zeng; Di Xie; Jiang Zhu*

109. LASS3D: Language-Assisted Semi-Supervised 3D Semantic Segmentation with Progressive Unreliable Data Exploitation; *Jianan Li*; Qiulei Dong**

110. Four Ways to Improve Verbo-visual Fusion for Dense 3D Visual Grounding; *Ozan Unal*; Christos Sakaridis; Suman Saha; Luc Van Gool*

111. On the Viability of Monocular Depth Pre-training for Semantic Segmentation; *Dong Lao*; Fengyu Yang; Daniel Wang; Hyoungseob Park; Samuel Lu; Alex Wong; Stefano Soatto*

112. SceneVerse: Scaling 3D Vision-Language Learning for Grounded Scene Understanding; *Baoxiong Jia*; Yixin Chen; Huangyue Yu; Yan Wang; Xuesong Niu; Tengyu Liu; Qing Li; Siyuan Huang*

113. NavGPT-2: Unleashing Navigational Reasoning Capability for Large Vision-Language Models; *Gengze Zhou*; Yicong Hong; Zun Wang; Xin Eric Wang; Qi Wu*

114. OAT: Object-Level Attention Transformer for Gaze Scanpath Prediction; *Yini Fang*; Jingling Yu; Haozheng Zhang; Ralf van der Lans; Bertram E Shi*

115. Point-supervised Panoptic Segmentation via Estimating Pseudo Labels from Learnable Distance; *Jing Li; Junsong Fan*; Zhaoxiang Zhang**

116. OneVOS: Unifying Video Object Segmentation with All-in-One Transformer Framework; *Wanyun Li; Pinxue Guo; Xinyu Zhou; Lingyi Hong; Yangji He; Xiangyu Zheng; Wei Zhang*; Wenqiang Zhang**

117. INTRA: Interaction Relationship-aware Weakly Supervised Affordance Grounding; *Ji Ha Jang; Hoigi Seo; Se Young Chun**

118. DAMSDet: Dynamic Adaptive Multispectral Detection Transformer with Competitive Query Selection and Adaptive Feature Fusion; *Junjie Guo*; Chenqiang Gao*; Fangcen Liu; Deyu Meng; Xinbo Gao*

119. Diffusion Model for Robust Multi-Sensor Fusion in 3D Object Detection and BEV Segmentation; *Duy Tho Le*; Hengcan Shi*; Jianfei Cai; Hamid Rezatofighi*

120. EAS-SNN: End-to-End Adaptive Sampling and Representation for Event-based Detection with Recurrent Spiking Neural Networks; *Ziming Wang; Ziling Wang; Huaning Li; Lang Qin; Runhao Jiang; De Ma*; Huajin Tang**

121. Quality Assured: Rethinking Annotation Strategies in Imaging AI; *Tim Rädsch*; Annika Reinke; Vivienn Weru; Minu D. Tizabi; Nicholas Heller; Fabian Isensee; Annette Kopp-Schneider; Lena Maier-Hein**

122. Are Synthetic Data Useful for Egocentric Hand-Object Interaction Detection?; *Rosario Leonardi*; Antonino Furnari; Francesco Ragusa; Giovanni Maria Farinella*

123. Find n' Propagate: Open-Vocabulary 3D Object Detection in Urban Environments; *Djamahl Etchegaray*; Zi Helen Huang; Tatsuya Harada; Yadan Luo*

124. Centering the Value of Every Modality: Towards Efficient and Resilient Modality-agnostic Semantic Segmentation; *Xu Zheng*; Yuanhuiyi Lyu; jiazhou zhou; Lin Wang**

125. 3x2: 3D Object Part Segmentation by 2D Semantic Correspondences; *Anh Thai*; Weiyao Wang; Hao Tang; Stefan Stojanov; James M Rehg; Matt Feiszli*

126. MonoTTA: Fully Test-Time Adaptation for Monocular 3D Object Detection; *Hongbin Lin; Yifan Zhang; Shuaicheng Niu; Shuguang Cui; Zhen Li**

127. Multi-modal Crowd Counting via a Broker Modality; *Haoliang Meng; Xiaopeng Hong*; Chenhao Wang; Miao Shang; Wangmeng Zuo*

129. SparseRadNet: Sparse Perception Neural Network on Subsampled Radar Data; *Jialong Wu*; Mirko Meuter; Markus Schoeler; Matthias Rottmann*

130. Multi-HMR: Multi-Person Whole-Body Human Mesh Recovery in a Single Shot; *Fabien Baradel*; Thomas LUCAS; Matthieu Armando; Salma Galaaoui; Romain Brégier; Philippe Weinzaepfel; Gregory Rogez*

131. Urban Waterlogging Detection: A Challenging Benchmark and Large-Small Model Co-Adapter; *Suqi Song; Chenxu Zhang; Peng Zhang; Pengkun Li; Fenglong Song; Lei Zhang**

132. FSD-BEV: Foreground Self-Distillation for Multi-view 3D Object Detection; *Zheng Jiang; Jinqing Zhang; Yanan Zhang; Qingjie Liu*; Zhenghui HU*; Baohui Wang; Yunhong Wang*

133. LayeredFlow: A Real-World Benchmark for Non-Lambertian Multi-Layer Optical Flow; *Hongyu Wen*; Erich Liang; Jia Deng*

134. OPEN: Object-wise Position Embedding for Multi-view 3D Object Detection; *Jinghua Hou; Tong Wang; Xiaoqing Ye; Zhe Liu; Shi Gong; Xiao Tan; Errui Ding; Jingdong Wang; Xiang Bai**

135. Unsupervised Exposure Correction; *Ruodai Cui*; Li Niu; Guosheng Hu*

136. Improving Domain Generalization in Self-Supervised Monocular Depth Estimation via Stabilized Adversarial Training; *Yuanqi Yao*; Gang Wu; Kui Jiang; Siao Liu; Jian Kuai; Xianming Liu; Junjun Jiang**

137. RING-NeRF : Rethinking Inductive Biases for Versatile and Efficient Neural Fields; *Doriand Petit*; Steve Bourgeois; Dumitru Pavel; Vincent Gay-Bellile; Florian Chabot; Loïc Barthe*

138. Spectral Subsurface Scattering for Material Classification; *Haejoon Lee*; Aswin Sankaranarayanan*

139. Learning to Adapt SAM for Segmenting Cross-domain Point Clouds; *Xidong Peng; Runnan Chen; Feng Qiao; Lingdong Kong; Youquan Liu; Yujing Sun; Tai Wang; Xinge Zhu*; Yuexin Ma**

140. Sparse Refinement for Efficient High-Resolution Semantic Segmentation; *Zhijian Liu; Zhuoyang Zhang; Samir Khaki; Shang Yang; Haotian Tang; Chenfeng Xu; Kurt Keutzer; Song Han**

141. Towards Natural Language-Guided Drones: GeoText-1652 Benchmark with Spatial Relation Matching; *Meng Chu; Zhedong Zheng*; Wei Ji; Tingyu Wang; Tat-Seng Chua*

142. Cross-Platform Video Person ReID: A New Benchmark Dataset and Adaptation Approach; *Shizhou Zhang; Wenlong Luo; De Cheng*; Qingchun Yang; Lingyan Ran; Yinghui Xing; Yanning Zhang*

143. UniM2AE: Multi-modal Masked Autoencoders with Unified 3D Representation for 3D Perception in Autonomous Driving; *Jian Zou; Tianyu Huang; Guanglei Yang*; Zhenhua Guo; Tao Luo*; Chun-Mei Feng; Wangmeng Zuo*

144. Frontier-enhanced Topological Memory with Improved Exploration Awareness for Embodied Visual Navigation; *Xinru Cui; Qiming Liu; Zhe Liu; Hesheng Wang**

145. Dataset Enhancement with Instance-Level Augmentations; *Orest Kupyn*; Christian Rupprecht*

146. Fast Diffusion-Based Counterfactuals for Shortcut Removal and Generation; *Nina Weng*; Paraskevas Pegios; Eike Petersen; Aasa Feragen; Siavash Arjomand Bigdeli*

147. Relation DETR: Exploring Explicit Position Relation Prior for Object Detection; *Xiuquan Hou; Meiqin Liu*; Senlin Zhang; Ping Wei; Badong Chen; Xuguang Lan*

148. On Calibration of Object Detectors: Pitfalls, Evaluation and Baselines; *Selim Kuzucu*; Kemal Oksuz*; Jonathan Sadeghi; Puneet Dokania*

149. Accelerating Image Generation with Sub-path Linear Approximation Model; *Chen Xu; Tianhui Song; Weixin Feng; Xubin Li; Tiezheng Ge; Bo Zheng; Limin Wang**

150. Action2Sound: Ambient-Aware Generation of Action Sounds from Egocentric Videos; *Changan Chen*; Puyuan Peng; Ami Baid; Zihui Xue; Wei-Ning Hsu; David Harwath; Kristen Grauman*

151. EDTalk: Efficient Disentanglement for Emotional Talking Head Synthesis; *Shuai Tan*; Bin Ji; Mengxiao Bi; ye pan**

152. SphereHead: Stable 3D Full-head Synthesis with Spherical Tri-plane Representation; *Heyuan Li*; Ce Chen; Tianhao Shi; Yuda Qiu; Sizhe An; Guanying CHEN; Xiaoguang Han**

153. ConDense: Consistent 2D-3D Pre-training for Dense and Sparse Features from Multi-View Images; *Xiaoshuai Zhang*; Zhicheng Wang; Howard Zhou; Soham Ghosh; Danushen L Gnanapragasam; Varun Jampani; Hao Su; Leonidas Guibas*

154. Bridging the Gap: Studio-like Avatar Creation from a Monocular Phone Capture; *ShahRukh Athar*; Shunsuke Saito; Stanislav Pidhorskyi; Zhengyu Yang; Chen Cao*

155. LGM: Large Multi-View Gaussian Model for High-Resolution 3D Content Creation; *Jiaxiang Tang*; Zhaoxi Chen; Xiaokang Chen; Tengfei Wang; Gang Zeng; Ziwei Liu*

156. TexDreamer: Towards Zero-Shot High-Fidelity 3D Human Texture Generation; *Yufei Liu; Junwei Zhu; Junshu Tang; Shijie Zhang; Jiangning Zhang; Weijian Cao; Chengjie Wang; Yunsheng Wu; Dongjin Huang**

157. LLMGA: Multimodal Large Language Model based Generation Assistant; *bin xia*; Shiyin Wang; Yingfan Tao; Yitong Wang; Jiaya Jia*

158. TextDiffuser-2: Unleashing the Power of Language Models for Text Rendering; *Jingye Chen*; Yupan Huang; Tengchao Lv; Lei Cui; Qifeng Chen; Furu Wei*

159. FlashTex: Fast Relightable Mesh Texturing with LightControlNet; *Kangle Deng*; Timothy Omernick; Alexander B Weiss; Deva Ramanan; Jun-Yan Zhu; Tinghui Zhou; Maneesh Agrawala*

160. PatchRefiner: Leveraging Synthetic Data for Real-Domain High-Resolution Monocular Metric Depth Estimation; *Zhenyu Li*; Shariq Farooq Bhat; Peter Wonka*

161. Street Gaussians: Modeling Dynamic Urban Scenes with Gaussian Splatting; *Yunzhi Yan*; Haotong Lin; Chenxu Zhou; Weijie Wang; Haiyang Sun; Kun Zhan; Xianpeng Lang; Xiaowei Zhou; Sida Peng**

162. MeshVPR: Citywide Visual Place Recognition Using 3D Meshes; *Gabriele Berton*; Lorenz Junglas; Riccardo Zaccone; Thomas Pollok; Barbara Caputo; Carlo Masone*

163. PointNeRF++: A multi-scale, point-based Neural Radiance Field; *Weiwei Sun; Eduard Trulls; Yang-Che Tseng; Sneha Sambandam; Gopal Sharma; Andrea Tagliasacchi; Kwang Moo Yi**

164. Continuous SO(3) Equivariant Convolution for 3D Point Cloud Analysis; *Jaein Kim; HEE BIN YOO; Dong-Sig Han; Yeon-Ji Song; Byoung-Tak Zhang**

165. FrePolad: Frequency-Rectified Point Latent Diffusion for Point Cloud Generation; *Chenliang Zhou*; Fangcheng Zhong; Param Hanji; Zhilin Guo; Kyle Thomas Fogarty; Alejandro Sztrajman; Hongyun Gao; A. Cengiz Oztireli*

166. Depth on Demand: Streaming Dense Depth from a Low Frame Rate Active Sensor; *Andrea Conti*; Matteo Poggi; Valerio Cambareri; Stefano Mattoccia*

167. FRI-Net: Floorplan Reconstruction via Room-wise Implicit Representation; *Honghao Xu; Juzhan Xu; Zeyu Huang; Pengfei Xu; Hui Huang; Ruizhen Hu**

168. UniCal: Unified Neural Sensor Calibration; *Ze Yang*; George G Chen; Haowei Zhang; Kevin Ta; Ioan Andrei Bârsan; Daniel Murphy; Sivabalan Manivasagam*; Raquel Urtasun**

169. Projecting Points to Axes: Oriented Object Detection via Point-Axis Representation; *Zeyang Zhao; Qilong Xue; Yifan Bai; Yuhang He; Xing Wei*; Yihong Gong*

170. Grounding Image Matching in 3D with MASt3R; *Vincent Leroy*; Yohann Cabon; Jerome Revaud*

171. LRSLAM: Low-rank Representation of Signed Distance Fields in Dense Visual SLAM System; *Hongbeen Park; Minjeong Park; Giljoo Nam; Jinkyu Kim**

172. Zero-Shot Detection of AI-Generated Images; *Davide Cozzolino; Giovanni Poggi; Matthias Niessner; Luisa Verdoliva**

173. DiscoMatch: Fast Discrete Optimisation for Geometrically Consistent 3D Shape Matching; *Paul Roetzer*; Ahmed Abbas*; Dongliang Cao; Florian Bernard; Paul Swoboda*

174. Dense Hand-Object(HO) GraspNet with Full Grasping Taxonomy and Dynamics; *Woojin Cho; Jihyun Lee; Minjae Yi; Minje Kim; Taeyun Woo; Donghwan Kim; Taewook Ha; Hyokeun Lee; Je-Hwan Ryu; Woontack Woo; Tae-Kyun (T-K) Kim**

175. GMM-IKRS: Gaussian Mixture Models for Interpretable Keypoint Refinement and Scoring; *Emanuele Santellani*; Martin Zach; Christian Sormann; Mattia Rossi; Andreas Kuhn; Friedrich Fraundorfer*

176. SPVLoc: Semantic Panoramic Viewport Matching for 6D Camera Localization in Unseen Environments; *Niklas Gard*; Anna Hilsmann; Peter Eisert*

177. U-COPE: Taking a Further Step to Universal 9D Category-level Object Pose Estimation; *li zhang*; Weiqing Meng; Yan Zhong; Bin Kong; Mingliang Xu; Jianming Du; Xue Wang; Rujing Wang; Liu Liu*

178. EgoPoseFormer: A Simple Baseline for Stereo Egocentric 3D Human Pose Estimation; *Chenhongyi Yang*; Anastasia Tkach; Shreyas Hampali; Linguang Zhang; Elliot J Crowley; Cem Keskin*

179. Learning Cross-hand Policies of High-DOF Reaching and Grasping; *Qijin She; Shishun Zhang; Yunfan Ye; Ruizhen Hu; Kai Xu**

180. Alignist: CAD-Informed Orientation Distribution Estimation by Fusing Shape and Correspondences; *Shishir Reddy Vutukur*; Junwen Huang; Rasmus Laurvig Haugaard; Benjamin Busam; Tolga Birdal*

181. COMO: Compact Mapping and Odometry; *Eric Dexheimer*; Andrew Davison*

182. ADen: Adaptive Density Representations for Sparse-view Camera Pose Estimation; *Hao Tang; Weiyao Wang; Pierre Gleize; Matt Feiszli**

183. Six-Point Method for Multi-Camera Systems with Reduced Solution Space; *Banglei Guan; Ji Zhao*; Laurent Kneip*

184. 3D Hand Sequence Recovery from Real Blurry Images and Event Stream; *JoonKyu Park; Gyeongsik Moon; Weipeng Xu; Evan Kaseman; Takaaki Shiratori; Kyoung Mu Lee**

185. Scene Coordinate Reconstruction: Posing of Image Collections via Incremental Learning of a Relocalizer; *Eric Brachmann*; Jamie Wynn; Shuai Chen; Tommaso Cavallari; Aron Monszpart; Daniyar Turmukhambetov; Victor Adrian Prisacariu*

186. Correspondences of the Third Kind: Camera Pose Estimation from Object Reflection; *Kohei Yamashita*; Vincent Lepetit; Ko Nishino*

187. Semicalibrated Relative Pose from an Affine Correspondence and Monodepth; *Petr Hruby*; Marc Pollefeys; Daniel Barath*

188. Non-Line-of-Sight Estimation of Fast Human Motion with Slow Scanning Imagers; *Javier Grau Chopite*; Patrick Hähn; Matthias B Hullin**

189. SRPose: Two-view Relative Pose Estimation with Sparse Keypoints; *Rui Yin; Yulun Zhang; Zherong Pan; Jianjun Zhu; Cheng Wang; Biao Jia**

190. Cut out the Middleman: Revisiting Pose-based Gait Recognition; *Yang Fu; Saihui Hou*; Shibei Meng; Xuecai Hu*; Chunshui Cao; Xu Liu; Yongzhen Huang*

191. Prompting Future Driven Diffusion Model for Hand Motion Prediction; *Bowen Tang*; Kaihao Zhang*; Wenhan Luo*; Wei Liu; HONGDONG LI*

192. Synchronization of Projective Transformations; *Rakshith Madhavan*; Andrea Fusiello; Federica Arrigoni*

193. Camera Calibration using a Collimator System; *Shunkun Liang; Banglei Guan*; Zhenbao Yu; Pengju Sun; Yang Shang*

194. Binomial Self-compensation for Motion Error in Dynamic 3D Scanning; *Geyou Zhang; Ce Zhu*; Kai Liu*

195. EgoPoser: Robust Real-Time Egocentric Pose Estimation from Sparse and Intermittent Observations Everywhere; *Jiaxi Jiang*; Paul Streli; Manuel Meier; Christian Holz*

196. Light-in-Flight for a World-in-Motion; *Jongho Lee*; Ryan J Suess; Mohit Gupta*

197. Differentiable Product Quantization for Memory Efficient Camera Relocalization; *Zakaria Laskar*; Iaroslav Melekhov; Assia Benbihi; Shuzhe Wang; Juho Kannala*

198. Nymeria: A Massive Collection of Egocentric Multi-modal Human Motion in the Wild; *Lingni Ma*; Yuting Ye; Rowan Postyeni; Alexander J Gamino; Vijay Baiyya; Luis Pesqueira; Kevin M Bailey; David Soriano Fosas; Fangzhou Hong; Vladimir Guzov; Yifeng Jiang; Hyo Jin Kim; Jakob Engel; Karen Liu; Ziwei Liu; Renzo De Nardi; Richard Newcombe*

199. MVDD: Multi-View Depth Diffusion Models; *Zhen Wang*; Qiangeng Xu; Feitong Tan; Menglei Chai; Shichen Liu; Rohit Pandey; Sean Fanello; Achuta Kadambi; Yinda Zhang*

200. McGrids: Monte Carlo-Driven Adaptive Grids for Iso-Surface Extraction; *Daxuan Ren*; Hezi Shi; Jianmin Zheng; Jianfei Cai*

201. Click-Gaussian: Interactive Segmentation to Any 3D Gaussians; *Seokhun Choi; Hyeonseop Song; Jaechul Kim; Taehyeong Kim*; Hoseok Do**

202. Free-Viewpoint Video of Outdoor Sports Using a Drone; *Zhengdong Hong**

203. GGRt: Towards Generalizable 3D Gaussians without Pose Priors in Real-Time; *Hao Li; Yuanyuan Gao; Dingwen Zhang*; Chenming Wu; YALUN DAI; Chen Zhao; Haocheng Feng; Errui Ding; Jingdong Wang; Junwei Han*

204. Deep Cost Ray Fusion for Sparse Depth Video Completion; *Jungeon Kim; Soongjin Kim; Jaesik Park; Seungyong Lee**

205. High-Fidelity 3D Textured Shapes Generation by Sparse Encoding and Adversarial Decoding; *Qi Zuo*; Xiaodong Gu; Yuan Dong; Zhengyi Zhao; Weihao Yuan; Qiu Lingteng; Liefeng Bo; Zilong Dong*

206. G3R: Gradient Guided Generalizable Reconstruction; *Yun Chen*; Jingkang Wang; Ze Yang; Sivabalan Manivasagam*; Raquel Urtasun**

207. latentSplat: Autoencoding Variational Gaussians for Fast Generalizable 3D Reconstruction; *Christopher Wewer*; Kevin Raj; Eddy Ilg; Bernt Schiele; Jan E. Lenssen**

208. VEGS: View Extrapolation of Urban Scenes in 3D Gaussian Splatting using Learned Priors; *Sungwon Hwang; Min-Jung Kim; Taewoong Kang; Jayeon Kang; Jaegul Choo**

209. 3D-GOI: 3D GAN Omni-Inversion for Multifaceted and Multi-object Editing; *Haoran Li; Long Ma; Haolin Shi; Yanbin Hao; Yong Liao*; Lechao Cheng; Peng Yuan Zhou**

210. Smoothness, Synthesis, and Sampling: Re-thinking Unsupervised Multi-View Stereo with DIV Loss; *Alex Rich*; Noah Stier; Pradeep Sen; Tobias Hollerer*

211. MinD-3D: Reconstruct High-quality 3D objects in Human Brain; *Jianxiong Gao; Yuqian Fu; Yun Wang; Xuelin Qian; Jianfeng Feng; Yanwei Fu**

212. Within the Dynamic Context: Inertia-aware 3D Human Modeling with Pose Sequence; *Yutong Chen; Yifan Zhan; Zhihang Zhong*; Wei Wang; Xiao Sun*; Yu Qiao; Yinqiang Zheng*

213. UpFusion: Novel View Diffusion from Unposed Sparse View Observations; *Bharath Raj Nagoor Kani*; Hsin-Ying Lee; Sergey Tulyakov; Shubham Tulsiani*

214. Efficient NeRF Optimization - Not All Samples Remain Equally Hard; *Juuso Korhonen*; Goutham Rangu; Hamed Rezazadegan Tavakoli; Juho Kannala*

215. CPT-VR: Improving Surface Rendering via Closest Point Transform with View-Reflection Appearance; *Zhipeng Hu; Yongqiang Zhang*; Chen Liu; Lincheng Li*; Sida Peng; Xiaowei Zhou; Changjie Fan; Xin Yu*

216. Temporal Residual Jacobians for Rig-free Motion Transfer; *Sanjeev Muralikrishnan*; Niladri Shekhar Dutt; Siddhartha Chaudhuri; Noam Aigerman; Vladimir Kim; Matthew Fisher; Niloy Mitra*

217. Geometry Fidelity for Spherical Images; *Anders Christensen*; Nooshin Mojab*; Khushman Patel; Karan Ahuja; Zeynep Akata; Ole Winther; Mar Gonzalez Franco; Andrea Colaco*

218. Radiative Gaussian Splatting for Efficient X-ray Novel View Synthesis; *Yuanhao Cai*; Yixun Liang; Jiahao Wang; Angtian Wang; Yulun Zhang; Xiaokang Yang; Zongwei Zhou; Alan Yuille*

219. SlotLifter: Slot-guided Feature Lifting for Learning Object-Centric Radiance Fields; *Yu Liu; Baoxiong Jia*; Yixin Chen; Siyuan Huang*

220. GS2Mesh: Surface Reconstruction from Gaussian Splatting via Novel Stereo Views; *Yaniv Wolf*; Amit Bracha; Ron Kimmel*

221. MetaCap: Meta-learning Priors from Multi-View Imagery for Sparse-view Human Performance Capture and Rendering; *Guoxing Sun*; Rishabh Dabral; Pascal Fua; Christian Theobalt; Marc Habermann*

222. Physics-Free Spectrally Multiplexed Photometric Stereo under Unknown Spectral Composition; *Satoshi Ikehata*; Yuta Asano*

223. Non-parametric Sensor Noise Modeling and Synthesis; *Ali Mosleh*; Luxi Zhao; Atin Vikram Singh; Jaeduk Han; Abhijith Punnappurath; Marcus A Brubaker; Jihwan Choe; Michael S Brown*

224. A Compact Dynamic 3D Gaussian Representation for Real-Time Dynamic View Synthesis; *Kai Katsumata*; Duc Minh Vo; Hideki Nakayama*

225. Holodepth: Programmable Depth-Varying Projection via Computer-Generated Holography; *Dorian Chan*; Matthew O'Toole; Sizhuo Ma; Jian Wang**

226. RS-NeRF: Neural Radiance Fields from Rolling Shutter Images; *Muyao Niu; Tong Chen; Yifan Zhan; Zhuoxiao Li; Xiang Ji; Yinqiang Zheng**

227. Structured-NeRF: Hierarchical Scene Graph with Neural Representation; *Zhide Zhong; Jiakai Cao; songen gu; Sirui Xie; Liyi Luo; Hao Zhao; Guyue Zhou; Haoang Li; Zike Yan**

228. Pathformer3D: A 3D Scanpath Transformer for 360° Images; *Rong Quan; yantao Lai; Mengyu Qiu; Dong Liang**

229. BAGS: Blur Agnostic Gaussian Splatting through Multi-Scale Kernel Modeling; *Cheng Peng*; Yutao Tang; Yifan Zhou; Nengyu Wang; Xijun Liu; Deming Li; Rama Chellappa*

230. BeNeRF:Neural Radiance Fields from a Single Blurry Image and Event Stream; *Wenpu Li; Pian Wan; Peng Wang; Jinghang Li; Yi Zhou; Peidong Liu**

231. Motion Mamba: Efficient and Long Sequence Motion Generation; *Zeyu Zhang; Akide Liu; Ian Reid; RICHARD HARTLEY; Bohan Zhuang; Hao Tang**

232. Neural Metamorphosis; *Xingyi Yang*; Xinchao Wang**

233. Optimizing Illuminant Estimation in Dual-Exposure HDR Imaging; *Mahmoud Afifi*; Zhenhua Hu; Liang Liang*

234. GLARE: Low Light Image Enhancement via Generative Latent Feature based Codebook Retrieval; *Han Zhou; Wei Dong; Xiaohong Liu*; Shuaicheng Liu; Xiongkuo Min; Guangtao Zhai; Jun Chen**

235. Restoring Images in Adverse Weather Conditions via Histogram Transformer; *Shangquan Sun; Wenqi Ren*; Xinwei Gao; Rui Wang; Xiaochun Cao*

236. Towards Real-world Event-guided Low-light Video Enhancement and Deblurring; *Taewoo Kim; Jaeseok Jeong; Hoonhee Cho; Yuhwan Jeong; Kuk-Jin Yoon**

237. Osmosis: RGBD Diffusion Prior for Underwater Image Restoration; *Opher Bar Nathan*; Deborah Levy; Tali Treibitz; Dan Rosenbaum*

238. Decomposition Betters Tracking Everything Everywhere; *Rui Li; Dong Liu**

239. Handling The Non-Smooth Challenge in Tensor SVD: A Multi-Objective Tensor Recovery Framework; *Jingjing Zheng; Wanglong Lu; Wenzhe Wang; Yankai Cao*; Xiaoqin Zhang; Xianta Jiang*

240. Adaptive Multi-modal Fusion of Spatially Variant Kernel Refinement with Diffusion Model for Blind Image Super-Resolution; *Junxiong Lin*; Yan Wang; Zeng Tao; Boyang Wang; Qing Zhao; Haoran Wang; Xuan Tong; Xinji Mai; Yuxuan Lin; Wei Song; Jiawen Yu; Shaoqi Yan; Wenqiang Zhang*

241. Efficient Learning of Event-based Dense Representation using Hierarchical Memories with Adaptive Update; *Uday Kamal*; Saibal Mukhopadhyay*

242. Efficient Training with Denoised Neural Weights; *Yifan Gong*; Zheng Zhan; Yanyu Li; Yerlan Idelbayev; Andrey Zharkov; Kfir Aberman; Sergey Tulyakov; Yanzhi Wang; Jian Ren*

243. DSMix: Distortion-Induced Saliency Map Based Pre-training for No-Reference Image Quality Assessment; *Jinsong Shi; Pan Gao*; Xiaojiang Peng; Jie Qin*

244. DiffBIR: Toward Blind Image Restoration with Generative Diffusion Prior; *Xinqi Lin*; Jingwen He; Ziyan Chen; Zhaoyang Lyu; Bo Dai; Fanghua Yu; Yu Qiao; Wanli Ouyang; Chao Dong**

245. Efficient Cascaded Multiscale Adaptive Network for Image Restoration; *Yichen Zhou*; Pan Zhou*; Teck Khim Ng*

246. You Only Need One Step: Fast Super-Resolution with Stable Diffusion via Scale Distillation; *Mehdi Noroozi*; Isma Hadji*; Brais Martinez*; Adrian Bulat*; Georgios Tzimiropoulos**

247. Match-Stereo-Videos: Bidirectional Alignment for Consistent Dynamic Stereo Matching; *Junpeng Jing*; Ye Mao; Krystian Mikolajczyk**

248. Eliminating Warping Shakes for Unsupervised Online Video Stitching; *Lang Nie; Chunyu Lin*; Kang Liao; Yun Zhang; Shuaicheng Liu; Rui Ai; Yao Zhao*

249. DreamMover: Leveraging the Prior of Diffusion Models for Image Interpolation with Large Motion; *Liao Shen; Tianqi Liu; Huiqiang Sun; Xinyi Ye; Baopu Li; Jianming Zhang; Zhiguo Cao**

250. Ponymation: Learning Articulated 3D Animal Motions from Unlabeled Online Videos; *Keqiang Sun; Dor Litvak; Yunzhi Zhang; Hongsheng Li; Jiajun Wu*; Shangzhe Wu**

251. Quanta Video Restoration; *Prateek Chennuri*; Yiheng Chi; Enze Jiang; GM Dilshan Godaliyadda*; Abhiram Gnanasambandam*; Hamid R Sheikh; Istvan Gyongy; Stanley H Chan**

252. Arbitrary-Scale Video Super-Resolution with Structural and Textural Priors; *Wei Shang*; Dongwei Ren*; Wanying Zhang; Yuming Fang; Wangmeng Zuo; Kede Ma*

253. Unleashing the Potential of the Semantic Latent Space in Diffusion Models for Image Dehazing; *Zizheng Yang; Hu Yu; Bing Li; Jinghao Zhang; Jie Huang; Feng Zhao**

254. Online Video Quality Enhancement with Spatial-Temporal Look-up Tables; *Zefan Qu; Xinyang Jiang*; Yifan Yang; Dongsheng Li; Cairong Zhao**

255. Rethinking Image-to-Video Adaptation: An Object-centric Perspective; *Rui Qian*; Shuangrui Ding; Dahua Lin*

256. SIGMA: Sinkhorn-Guided Masked Video Modeling; *Mohammadreza Salehi*; Michael Dorkenwald*; Fida Mohammad Thoker; Efstratios Gavves; Cees Snoek; Yuki M Asano*

257. Hybrid Video Diffusion Models with 2D Triplane and 3D Wavelet Representation; *Kihong Kim; Haneol Lee; Jihye Park; Seyeon Kim; Kwang Hee Lee; Seungryong Kim*; Jaejun Yoo**

258. Two-Stage Video Shadow Detection via Temporal-Spatial Adaption; *Xin Duan; Yu Cao; Lei Zhu; Gang Fu; Xin Wang; Renjie ZHANG; Ping Li**

259. InsMapper: Exploring Inner-instance Information for Vectorized HD Mapping; *zhenhua xu*; Kwan-Yee K. Wong; Hengshuang Zhao*

260. Understanding Physical Dynamics with Counterfactual World Modeling; *Rahul Venkatesh*; Honglin Chen*; Kevin Feigelis; Daniel M Bear; Khaled Jedoui; Klemen Kotar; Felix J Binder; Wanhee Lee; Sherry Liu; Kevin Smith; Judith E. Fan; Daniel Yamins*

261. RICA^2: Rubric-Informed, Calibrated Assessment of Actions; *Abrar Majeedi; Viswanatha Reddy Gajjala; Satya Sai Srinath Namburi GNVV; Yin Li**

262. Training-free Video Temporal Grounding using Large-scale Pre-trained Models; *Minghang Zheng; Xinhao Cai; Qingchao Chen; Yuxin Peng; Yang Liu**

263. Pose Guided Fine-Grained Sign Language Video Generation; *Tongkai Shi; Lianyu Hu; Fanhua Shang; Jichao Feng; liu peidong; Wei Feng**

264. Learning to Localize Actions in Instructional Videos with LLM-Based Multi-Pathway Text-Video Alignment; *Yuxiao Chen*; Kai Li; Wentao Bao; Deep Patel; Yu Kong; Martin Renqiang Min; Dimitris N. Metaxas**

265. Learning Trimodal Relation for Audio-Visual Question Answering with Missing Modality; *Kyu Ri Park; Hong Joo Lee*; Jung Uk Kim**

266. EA-VTR: Event-Aware Video-Text Retrieval; *Zongyang Ma*; Ziqi Zhang; Yuxin Chen; Zhongang Qi; Chunfeng Yuan; Bing Li; Yingmin Luo; Xu LI; Xiaojuan Qi; Ying Shan; Weiming Hu*

267. Rethinking Video-Text Understanding: Retrieval from Counterfactually Augmented Data; *Wufei Ma*; Kai Li; Zhongshi Jiang; Moustafa Meshry; Qihao Liu; Huiyu Wang; Christian Haene; Alan Yuille*

268. Semantically Guided Representation Learning For Action Anticipation; *Anxhelo Diko*; Danilo Avola; Bardh Prenkaj; Federico Fontana; Luigi Cinque*

269. FunQA: Towards Surprising Video Comprehension; *Binzhu Xie; Sicheng Zhang; Zitang Zhou; Bo Li; Yuanhan Zhang; Jack Hessel; Jingkang Yang; Ziwei Liu**

270. VideoStudio: Generating Consistent-Content and Multi-Scene Videos; *Fuchen Long; Zhaofan Qiu*; Ting Yao; Tao Mei*

271. Boosting the Power of Small Multimodal Reasoning Models to Match Larger Models with Self-Consistency Training; *Cheng Tan*; Jingxuan Wei*; Zhangyang Gao; Linzhuang Sun; Siyuan Li; Ruifeng Guo; BiHui Yu; Stan Z. Li**

272. Can Textual Semantics Mitigate Sounding Object Segmentation Preference?; *Yaoting Wang; Peiwen Sun; Yuanchao Li; Honggang Zhang; Di Hu**

273. TLControl: Trajectory and Language Control for Human Motion Synthesis; *Weilin Wan*; Zhiyang Dou; Taku Komura; Wenping Wang; Dinesh Jayaraman; Lingjie Liu*

274. StoryImager: A Unified and Efficient Framework for Coherent Story Visualization and Completion; *Ming Tao*; Bingkun Bao*; Hao Tang; Yaowei Wang; Changsheng Xu*

275. DragVideo: Interactive Drag-style Video Editing; *Yufan Deng; Ruida WANG; Yuhao ZHANG; Yu-Wing Tai*; Chi-Keung Tang**

276. Animate Your Motion: Turning Still Images into Dynamic Videos; *Mingxiao Li*; Bo Wan*; Sien Moens; Tinne Tuytelaars*

277. BAMM: Bidirectional Autoregressive Motion Model; *Ekkasit Pinyoanuntapong*; Muhammad Usama Saleem; Pu Wang; Minwoo Lee; Srijan Das; Chen Chen*

278. ParCo: Part-Coordinating Text-to-Motion Synthesis; *Qiran Zou; Shangyuan Yuan; Shian Du; Yu Wang; Chang Liu; Yi Xu; Jie Chen; Xiangyang Ji**

279. MagDiff: Multi-Alignment Diffusion for High-Fidelity Video Generation and Editing; *Haoyu Zhao; Tianyi Lu; Jiaxi Gu; Xing Zhang; Qingping Zheng; Zuxuan Wu*; Hang Xu; Yu-Gang Jiang*

280. MART: MultiscAle Relational Transformer Networks for Multi-agent Trajectory Prediction; *Seongju Lee; Junseok Lee; Yeonguk Yu; Taeri Kim; Kyoobin Lee**

281. Ex2Eg-MAE: A Framework for Adaptation of Exocentric Video Masked Autoencoders for Egocentric Social Role Understanding; *Minh Tran*; Yelin Kim; Che-Chun Su; Min Sun; Cheng-Hao Kuo; Mohammad Soleymani*

282. V-Trans4Style: Visual Transition Recommendation for Video Production Style Adaptation; *Pooja Guhan*; Tsung-Wei Huang; Guan-Ming Su; Subhadra Gopalakrishnan; Dinesh Manocha*

283. EMO: Emote Portrait Alive - Generating Expressive Portrait Videos with Audio2Video Diffusion Model under Weak Conditions; *Linrui Tian*; Qi Wang*; Bang Zhang*; Liefeng Bo**

284. Taming Lookup Tables for Efficient Image Retouching; *Sidi Yang; Binxiao Huang; Mingdeng Cao; Yatai Ji; Hanzhong Guo; Ngai Wong; Yujiu Yang**

285. PreSight: Enhancing Autonomous Vehicle Perception with City-Scale NeRF Priors; *Tianyuan Yuan*; Yucheng Mao; Jiawei Yang; Yicheng LIU; Yue Wang; Hang Zhao**

286. Avatar Fingerprinting for Authorized Use of Synthetic Talking-Head Videos; *Ekta Prashnani*; Koki Nagano; Shalini De Mello; David P Luebke; Orazio Gallo*

287. EmoTalk3D: High-Fidelity Free-View Synthesis of Emotional 3D Talking Head; *Qianyun He; Xinya Ji; Yicheng Gong; Yuanxun Lu; Zhengyu Diao; Linjia Huang; Yao Yao; Siyu Zhu; Zhan Ma; Songcen Xu; Xiaofei Wu; Zixiao Zhang; Xun Cao; Hao Zhu**

288. Improving Agent Behaviors with RL Fine-tuning for Autonomous Driving; *Zhenghao Peng; Wenjie Luo; Yiren Lu*; Tianyi Shen; Cole Gulino; Ari Seff; Justin Fu*

289. Asynchronous Large Language Model Enhanced Planner for Autonomous Driving; *Yuan Chen; Zi-han Ding; Ziqin Wang; Yan Wang*; Lijun Zhang; Si Liu**

290. 3D Gaussian Parametric Head Model; *Yuelang Xu; Lizhen Wang; Zerong Zheng; Zhaoqi Su; Yebin Liu**

291. Neural graphics texture compression supporting random access; *Farzad Farhadzadeh*; Qiqi Hou; Hoang Le; Amir Said; Randall R Rauwendaal; Alex Bourd; Fatih Porikli*

292. COMPOSE: Comprehensive Portrait Shadow Editing; *Andrew Z Hou*; Zhixin Shu; Xuaner Zhang; He Zhang; Yannick Hold-Geoffroy; Jae Shin Yoon; Xiaoming Liu*

293. PhysAvatar: Learning the Physics of Dressed 3D Avatars from Visual Observations; *Yang Zheng*; Qingqing Zhao; Guandao Yang; Wang Yifan; Donglai Xiang; Florian Dubost; Dmitry Lagun; Thabo Beeler; Federico Tombari; Leonidas Guibas; Gordon Wetzstein*

294. RodinHD: High-Fidelity 3D Avatar Generation with Diffusion Models; *Bowen Zhang; Yiji Cheng; Chunyu Wang*; Ting Zhang; Jiaolong Yang; Yansong Tang; Feng Zhao; Dong Chen; Baining Guo*

295. FouriScale: A Frequency Perspective on Training-Free High-Resolution Image Synthesis; *Linjiang Huang*; Rongyao Fang; Aiping Zhang; Guanglu Song; Si Liu; Yu Liu; Hongsheng Li**

296. Texture-GS: Disentangle the Geometry and Texture for 3D Gaussian Splatting Editing; *Tianxing Xu*; Wenbo Hu; Yu-Kun Lai; Ying Shan; Song-Hai Zhang*

297. Generative End-to-End Autonomous Driving; *Wenzhao Zheng; Ruiqi Song; Xianda Guo*; Chenming Zhang; Long Chen*

298. AddMe: Zero-shot Group-photo Synthesis by Inserting People into Scenes; *Dongxu Yue; Maomao Li; Yunfei Liu; Ailing Zeng; Tianyu Yang; Qin Guo; Yu Li**

299. UniProcessor: A Text-induced Unified Low-level Image Processor; *Huiyu Duan*; Xiongkuo Min; Sijing Wu; Wei Shen; Guangtao Zhai*

300. Unified Local-Cloud Decision-Making via Reinforcement Learning; *Kathakoli Sengupta; Zhongkai Shangguan; Sandesh Bharadwaj; Sanjay Arora; Eshed Ohn-Bar*; Renato Mancuso*

301. EBDM: Exemplar-guided Image Translation with Brownian-bridge Diffusion Models; *Eungbean Lee; Somi Jeong; Kwanghoon Sohn**

302. Free-Editor: Zero-shot Text-driven 3D Scene Editing; *Nazmul Karim*; Hasan Iqbal; Umar Khalid; Chen Chen; Jing Hua*

303. FreestyleRet: Retrieving Images from Style-Diversified Queries; *Hao Li*; Yanhao Jia; Peng Jin; Zesen Cheng; Kehan Li; Jialu Sui; Chang Liu; Li Yuan**

304. BenchLMM: Benchmarking Cross-style Visual Capability of Large Multimodal Models; *Rizhao Cai*; Zirui Song; Dayan Guan*; Zhenhao Chen; Yaohang Li; Xing Luo; Chenyu Yi; Alex Kot*

305. Commonly Interesting Images; *Fitim Abdullahu*; Helmut Grabner**

306. WaSt-3D: Wasserstein-2 Distance for Scene-to-Scene Stylization on 3D Gaussians; *Dmytro Kotovenko*; Olga Grebenkova*; Nikolaos Sarafianos; Avinash Paliwal; Pingchuan Ma; Omid Poursaeed; Sreyas Mohan; Yuchen Fan; Yilei Li; Rakesh Ranjan; Bjorn Ommer*

307. VCD-Texture: Variance Alignment based 3D-2D Co-Denoising for Text-Guided Texturing; *Shang Liu*; Chaohui Yu; Chenjie Cao; Wen Qian; Fan Wang**

308. Efficient Pre-training for Localized Instruction Generation of Procedural Videos; *Anil Batra*; Davide Moltisanti; Laura Sevilla-Lara; Marcus Rohrbach; Frank Keller*

309. FlexiEdit: Frequency-Aware Latent Refinement for Enhanced Non-Rigid Editing; *Gwanhyeong Koo; Sunjae Yoon; Ji Woo Hong; Chang D. Yoo**

310. InstructGIE: Towards Generalizable Image Editing; *Zichong Meng; Changdi Yang; Jun Liu; Hao Tang*; Pu Zhao*; Yanzhi Wang**

311. Lazy Diffusion Transformer for Interactive Image Editing; *Yotam Nitzan*; Zongze Wu; Richard Zhang; Eli Shechtman; Danny Cohen-Or; Taesung Park; Michaël Gharbi*

312. MasterWeaver: Taming Editability and Face Identity for Personalized Text-to-Image Generation; *Yuxiang Wei; Zhilong Ji; Jinfeng Bai; Hongzhi Zhang; Lei Zhang*; Wangmeng Zuo**

313. Towards Reliable Advertising Image Generation Using Human Feedback; *Zhenbang Du*; Wei Feng; Haohan Wang; Yaoyu Li; Jingsen Wang; Jian Li; Zheng Zhang; Jingjing Lv; Xin Zhu; Junsheng Jin; Junjie Shen; Zhangang Lin; Jingping Shao*

314. The Lottery Ticket Hypothesis in Denoising: Towards Semantic-Driven Initialization; *Jiafeng Mao*; Xueting Wang; Kiyoharu Aizawa*

315. PreciseControl: Enhancing Text-To-Image Diffusion Models with Fine-Grained Attribute Control; *Rishubh Parihar*; Sachidanand VS; Sabariswaran Mani; Tejan Karmali; Venkatesh Babu RADHAKRISHNAN*

316. Layered Rendering Diffusion Model for Controllable Zero-Shot Image Synthesis; *Zipeng Qi; Guoxi Huang*; Chenyang Liu; Fei Ye*

317. Improving Text-guided Object Inpainting with Semantic Pre-inpainting; *Yifu Chen; Jingwen Chen; Yingwei Pan*; Yehao Li; Ting Yao; Zhineng Chen; Tao Mei*

318. DreamView: Injecting View-specific Text Guidance into Text-to-3D Generation; *Junkai Yan; Yipeng Gao; Qize Yang; Xihan Wei; Xuansong Xie; Ancong Wu*; WEI-SHI ZHENG**

319. CTRLorALTer: Conditional LoRAdapter for Efficient 0-Shot Control & Altering of T2I Models; *Nick Stracke*; Stefan Andreas Baumann; Joshua Susskind; Miguel Angel Bautista; Bjorn Ommer*

320. MobileDiffusion: Instant Text-to-Image Generation on Mobile Devices; *Yang Zhao*; Zhisheng Xiao*; Yanwu Xu; Haolin Jia; Tingbo Hou*

321. Ferret-UI: Grounded Mobile UI Understanding with Multimodal LLMs; *Keen You*; Haotian Zhang; Eldon Schoop; Floris Weers; Amanda Swearngin; Jeff Nichols; Yinfei Yang; Zhe Gan*

322. Hypernetworks for Generalizable BRDF Representation; *Fazilet Gokbudak*; Alejandro Sztrajman; Chenliang Zhou; Fangcheng Zhong; Rafal Mantiuk; A. Cengiz Oztireli*

323. UniDream: Unifying Diffusion Priors for Relightable Text-to-3D Generation; *Zexiang Liu; Yangguang Li; Youtian Lin; Xin Yu; Sida Peng; Yan-Pei Cao; Xiaojuan Qi; Xiaoshui Huang; Ding Liang*; Wanli Ouyang*

324. R3D-AD: Reconstruction via Diffusion for 3D Anomaly Detection; *Zheyuan Zhou; Le Wang; Naiyu Fang; Zili Wang; Lemiao Qiu*; Shuyou Zhang*

325. Textual-Visual Logic Challenge: Understanding and Reasoning in Text-to-Image Generation; *Peixi Xiong*; Michael A Kozuch; Nilesh Jain*

326. Zero-shot Text-guided Infinite Image Synthesis with LLM guidance; *Soyeong Kwon; Taegyeong Lee; Taehwan Kim**

327. A Task is Worth One Word: Learning with Task Prompts for High-Quality Versatile Image Inpainting; *Junhao Zhuang; Yanhong Zeng; WENRAN LIU; Chun Yuan*; Kai Chen**

328. IMMA: Immunizing text-to-image Models against Malicious Adaptation; *Amber Yijia Zheng*; Raymond A. Yeh*

329. Customized Generation Reimagined: Fidelity and Editability Harmonized; *Jian Jin; Yang Shen; Zhenyong Fu*; Jian Yang**

330. UMERegRobust – Universal Manifold Embedding Compatible Features for Robust Point Cloud Registration; *Yuval Haitman*; Amit Efraim; Joseph M Francos*

331. ColorPeel: Color Prompt Learning with Diffusion Models via Color and Shape Disentanglement; *Muhammad Atif Butt*; Kai Wang; Javier Vazquez-Corral; Joost van de Weijer*

334. ScaleDreamer: Scalable Text-to-3D Synthesis with Asynchronous Score Distillation; *Zhiyuan Ma*; Yuxiang Wei; Yabin Zhang; Xiangyu Zhu; Zhen Lei; Lei Zhang*

335. ViPer: Visual Personalization of Generative Models via Individual Preference Learning; *Sogand Salehi*; Mahdi Shafiei; Roman Bachmann; Teresa Yeo; Amir Zamir*

336. D4-VTON: Dynamic Semantics Disentangling for Differential Diffusion based Virtual Try-On; *Zhaotong Yang; Zicheng Jiang; Xinzhe Li; Huiyu Zhou; Junyu Dong; Huaidong Zhang; Yong Du**

337. PairingNet: A Learning-based Pair-searching and -matching Network for Image Fragments; *Rixin Zhou*; Ding Xia; YI ZHANG; honglin pang; Xi Yang; chuntao li*

338. PosterLlama: Bridging Design Ability of Langauge Model to Content-Aware Layout Generation; *Jaejung Seol; SeoJun Kim; Jaejun Yoo**

339. Controllable Navigation Instruction Generation with Chain of Thought Prompting; *Xianghao Kong; Jinyu Chen; Wenguan Wang*; Hang Su; Xiaolin Hu; Yi Yang; Si Liu**

340. Text to Layer-wise 3D Clothed Human Generation; *Junting Dong*; Qi Fang; Zehuan Huang; Xudong XU; Jingbo Wang; Sida Peng; Bo Dai*

341. ShoeModel: Learning to Wear on the User-specified Shoes via Diffusion Model; *Wenyu Li*; Binghui Chen; Yifeng Geng; Xuansong Xie; Wangmeng Zuo*

342. SceneTeller: Language-to-3D Scene Generation; *Basak Melis Ocal*; Maxim Tatarchenko; Sezer Karaoglu; Theo Gevers*

343. GroundUp: Rapid Sketch-Based 3D City Massing; *Gizem Esra Unlu*; Mohamed Sayed; Yulia Gryaditskaya; Gabriel Brostow*

344. Forbes: Face Obfuscation Rendering via Backpropagation Refinement Scheme; *Jintae Kim; Seungwon Yang; Seong-Gyun Jeong; Chang-Su Kim**

18:00 – 18:30
Meta Technical Session - Technical Presentation Area (Level 0)
SAM 2: Segment Anything in Images & Videos

18:30 – 19:30
Welcome Reception – Balcony Level 1

WEDNESDAY, 2ND OCTOBER

08:00 – 18:30
Registration - Badge Pickup

09:00 – 18:30
Exhibition - Level 0

09:00 – 10:30
Oral session 3A: Datasets and benchmarking - Gold Room
Chairs: Juan Carlos Niebles; Jose M Alvarez

1. PetFace: A Large-Scale Dataset and Benchmark for Animal Identification; *Risa Shinoda*; Kaede Shiohara*

2. UniIR: Training and Benchmarking Universal Multimodal Information Retrievers; *Cong Wei*; Yang Chen; Haonan Chen; Hexiang Hu; Ge Zhang; Jie Fu; Alan Ritter; Wenhu Chen*

3. Towards Model-Agnostic Dataset Condensation by Heterogeneous Models; *Jun-Yeong Moon; Jung Uk Kim*; Gyeong-Moon Park**

4. Parrot Captions Teach CLIP to Spot Text; *Yiqi Lin; Conghui He*; Alex Jinpeng Wang; Bin Wang; Weijia Li; Mike Zheng Shou*

5. Towards Open-ended Visual Quality Comparison; *Haoning Wu; Hanwei Zhu; Zicheng Zhang; Erli Zhang; Chaofeng Chen; Liang Liao; Chunyi Li; Annan Wang; Wenxiu Sun; Qiong Yan; Xiaohong Liu; Guangtao Zhai; Shiqi Wang; Weisi Lin**

6. VETRA: A Dataset for Vehicle Tracking in Aerial Imagery - New Challenges for Multi-Object Tracking; *Jens Hellekes*; Manuel Mühlhaus; Reza Bahmanyar; Seyed Majid Azimi; Franz Kurz*

7. Insect Identification in the Wild: The AMI Dataset; *Aditya Jain*; Fagner Cunha; Michael J Bunsen; Juan Sebastián Cañas; Léonard Pasi; Nathan Pinoy; Flemming Helsing; JoAnne Russo; Marc S Botham; Michael Sabourin; Jonathan Fréchette; Alexandre Anctil; Yacksecari Lopez; Eduardo Navarro; Filonila Pérez; Ana C Zamora; Jose Alejandro Ramirez-Silva; Jonathan Gagnon; Tom A August; Kim Bjerge; Alba Gomez Segura; Marc Belisle; Yves Basset; Kent P McFarland; David B Roy; Toke T Høye; Maxim Larrivee; David Rolnick*

8. MarineInst: A Foundation Model for Marine Image Analysis with Instance Visual Description; *Ziqiang Zheng*; Yiwei Chen; Huimin Zeng; Tuan-Anh Vu; Binh-Son Hua; Sai-Kit Yeung*

09:00 – 10:30 Oral session 3B: Medical and biological imaging - Auditorium
Chairs: Jose Dolz; Benjamin Busam

1. PathMMU: A Massive Multimodal Expert-Level Benchmark for Understanding and Reasoning in Pathology; *Yuxuan Sun*; Hao Wu; Chenglu Zhu; Sunyi Zheng; Qizi Chen; Kai Zhang; Yunlong Zhang; Dan Wan; Xiaoxiao Lan; Mengyue Zheng; Jingxiong Li; Xinheng Lyu; Tao Lin*; Lin Yang** BEST PAPER CANDIDATE

2. Self-Supervised Video Desmoking for Laparoscopic Surgery; *Renlong Wu; Zhilu Zhang*; Shuohao Zhang; Longfei Gou; Haobin Chen; Lei Zhang; Hao Chen*; Wangmeng Zuo*

3. CardiacNet: Learning to Reconstruct Abnormalities for Cardiac Disease Assessment from Echocardiogram Videos; *Jiewen Yang*; Yiqun Lin; Bin Pu; Jiarong GUO; Xiaowei Xu*; Xiaomeng Li**

4. Rethinking Deep Unrolled Model for Accelerated MRI Reconstruction; *Bingyu Xin*; Meng Ye; Leon Axel; Dimitris N. Metaxas*

5. Adaptive Correspondence Scoring for Unsupervised Medical Image Registration; *Xiaoran Zhang*; John C. Stendahl; Lawrence H. Staib; Albert J. Sinusas; Alex Wong; James S. Duncan*

6. Revisiting Adaptive Cellular Recognition Under Domain Shifts: A Contextual Correspondence View; *Jianan Fan*; Dongnan Liu; Canran Li; Hang Chang; Heng Huang; Filip Braet; Mei Chen; Weidong Cai**

7. SparseSSP: 3D Subcellular Structure Prediction from Sparse-View Transmitted Light Images; *Jintu Zheng; Yi Ding; Qizhe Liu; Yuehui Chen; Yi Cao; Ying Hu; Zenan Wang**

8. Knowledge-enhanced Visual-Language Pretraining for Computational Pathology; *Xiao Zhou; Xiaoman Zhang; Chaoyi Wu; Ya Zhang; Weidi Xie; Yan-Feng Wang**

09:00 – 10:30
Oral session 3C: Point clouds - Silver Room
Chairs: Yiming Wang; Yuchao Dai

1. HGL: Hierarchical Geometry Learning for Test-time Adaptation in 3D Point Cloud Segmentation; *Tianpei Zou; Sanqing Qu; Zhijun Li; Alois C. Knoll; 何 良; Guang Chen*; Changjun Jiang*

2. PointLLM: Empowering Large Language Models to Understand Point Clouds; *Runsen Xu*; Xiaolong Wang; Tai Wang*; Yilun Chen; Jiangmiao Pang*; Dahua Lin* BEST PAPER CANDIDATE

3. RISurConv: Rotation Invariant Surface Attention-Augmented Convolutions for 3D Point Cloud Classification and Segmentation; *Zhiyuan Zhang*; Licheng Yang; Zhiyu Xiang*

4. DVLO: Deep Visual-LiDAR Odometry with Local-to-Global Feature Fusion and Bi-Directional Structure Alignment; *Jiuming Liu; Dong Zhuo; Zhiheng Feng; Siting Zhu; Chensheng Peng; Zhe Liu; Hesheng Wang**

5. KeypointDETR: An End-to-End 3D Keypoint Detector; *Hairong Jin; Yuefan Shen; Jianwen Lou; Kun Zhou; Youyi Zheng**

6. Rethinking Data Augmentation for Robust LiDAR Semantic Segmentation in Adverse Weather; *Junsung Park; Kyungmin Kim; Hyunjung Shim**

7. RAPiD-Seg: Range-Aware Pointwise Distance Distribution Networks for 3D LiDAR Segmentation; *Li Li*; Hubert P. H. Shum; Toby P Breckon*

8. Equi-GSPR: Equivariant SE(3) Graph Network Model for Sparse Point Cloud Registration; *Xueyang Kang*; Zhaoliang Luan; Kourosh Khoshelham; Bing WANG**

09:00 – 12:30
Demo session 3 - Level 0
1. SLAM with Stereo Event Cameras; *Suman Ghosh, Valentina Cavinato, Guillermo Gallego - Technische Universität Berlin*
2. PROCEDO: *Alessandro Flaborea, Luca Franco, Alessandro Raimondi, Fabio Galasso, Luca Rigazio - Sapienza University of Rome*

3. Better Call SAL: Segment Anything in Lidar; *Aljoša Ošep, Tim Meinhardt, Francesco Ferroni, Neehar Peri, Deva Ramanan, Laura Leal-Taixé1 - NVIDIA*

4. AI3D Sculpt - Create 3D by rough sculpting 3D to 3D in 3D; *Yosun Chang - AI3D.foundation*

5. Real-time Multi-Person Whole-Body Human Mesh Recovery with Multi-HMR; *Fabien Baradel, Matthieu Armando, Romain Brégier, Thomas Lucas, Philippe Weinzaepfel, Grégory Rogez - Naver Labs Europe*

10:30 – 11:00
Bending Spoons Technical Session - Technical Presentation Area (Level 0)
Scaling Generative AI at Bending Spoons

10:30 – 11:00
Coffee Break - Exhibition Area (Level 0)

10:30 – 12:30
Poster session 3
1. Rethinking Normalization Layers for Domain Generalizable Person Re-identification; *Ren Nie; Jin Ding; Xue Zhou*; Xi Li*
2. De-Confusing Pseudo-Labels in Source-Free Domain Adaptation; *Idit Diamant*; Amir Rosenfeld; Idan Achituve; Jacob Goldberger; Arnon Netzer*

3. Hierarchical Unsupervised Relation Distillation for Source Free Domain Adaptation; *Bowei Xing*; Xianghua Ying; Ruibin Wang; Ruohao Guo; Ji Shi; Wenzhen Yue*

4. Distribution-Aware Robust Learning from Long-Tailed Data with Noisy Labels; *Jae Soon Baik*; In Young Yoon; Kun Hoon Kim; Jun Won Choi**

5. Reshaping the Online Data Buffering and Organizing Mechanism for Continual Test-Time Adaptation; *Zhilin Zhu*; Xiaopeng Hong*; Zhiheng Ma; Weijun Zhuang; YaoHui Ma; Yong Dai; Yaowei Wang*

6. Improving Unsupervised Domain Adaptation: A Pseudo-Candidate Set Approach; *Aveen Dayal*; Rishabh Lalla; Linga Reddy Cenkeramaddi; C. Krishna Mohan; Abhinav Kumar; Vineeth N Balasubramanian*

7. Learning to Complement and to Defer to Multiple Users; *Zheng Zhang; Wenjie Ai; Kevin Wells; David M Rosewarne; Thanh-Toan Do; Gustavo Carneiro**

8. PFedEdit: Personalized Federated Learning via Automated Model Editing; *Haolin Yuan*; William Paul; John Aucott; Philippe Burlina; Yinzhi Cao**

9. Personalized Federated Domain-Incremental Learning based on Adaptive Knowledge Matching; *Yichen Li; Wenchao Xu; Haozhao Wang*; Yining Qi*; Jingcai Guo; Ruixuan Li**

10. Feature Diversification and Adaptation for Federated Domain Generalization; *Seunghan Yang*; Seokeon Choi; Hyunsin Park; Sungha Choi; Simyung Chang; Sungrack Yun*

11. Adapting to Shifting Correlations with Unlabeled Data Calibration; *Minh Nguyen*; Alan Q Wang; Heejong Kim; Mert Sabuncu*

12. An Information Theoretical View for Out-Of-Distribution Detection; *Hu Jinjing; Wenrui Liu; Hong Chang*; Bingpeng MA; Shiguang Shan; Xilin Chen*

13. Revisiting Supervision for Continual Representation Learning; *Daniel Marczak*; Sebastian Cygert*; Tomasz Trzcinski*; Bartlomiej Twardowski**

14. Source-Free Domain-Invariant Performance Prediction; *Ekaterina Khramtsova*; Mahsa Baktashmotlagh; Guido Zuccon; Xi Wang; Mathieu Salzmann*

15. Overcome Modal Bias in Multi-modal Federated Learning via Balanced Modality Selection; *Yunfeng FAN*; Wenchao Xu*; Haozhao Wang; Fushuo Huo; Jinyu Chen; Song Guo*

16. Contrastive Learning with Synthetic Positives; *Dewen Zeng*; Xinrong Hu; Yawen Wu; Xiaowei Xu; Yiyu Shi*

17. On Pretraining Data Diversity for Self-Supervised Learning; *Hasan Abed Al Kader Hammoud*; Tuhin Das; Fabio Pizzati*; Philip Torr; Adel Bibi; Bernard Ghanem*

18. ProSub: Probabilistic Open-Set Semi-Supervised Learning with Subspace-Based Out-of-Distribution Detection; *Erik Wallin*; Lennart Svensson; Fredrik Kahl; Lars Hammarstrand*

19. Harmonizing knowledge Transfer in Neural Network with Unified Distillation; *yaomin huang; Faming Fang; Zaoming Yan; Chaomin Shen; Guixu Zhang**

20. Training A Secure Model against Data-Free Model Extraction; *Zhenyi Wang*; Li Shen*; junfeng guo; Tiehang Duan; Siyu Luan; Tongliang Liu; Mingchen Gao*

21. Learning Scalable Model Soup on a Single GPU: An Efficient Subspace Training Strategy; *Tao Li*; Weisen Jiang; Fanghui Liu; Xiaolin Huang; James Kwok*

22. Operational Open-Set Recognition and PostMax Refinement; *Steve Cruz*; Ryan Rabinowitz; Manuel Günther; Terrance E. Boult*

23. Challenging Forgets: Unveiling the Worst-Case Forget Sets in Machine Unlearning; *Chongyu Fan; Jiancheng Liu*; Alfred Hero; Sijia Liu*

24. Benchmarking Spurious Bias in Few-Shot Image Classifiers; *Guangtao Zheng*; Wenqian Ye; Aidong Zhang*

25. FroSSL: Frobenius Norm Minimization for Efficient Multiview Self-Supervised Learning; *Oscar Skean*; Aayush Dhakal; Nathan Jacobs; Luis G Sanchez Giraldo*

26. Deep Companion Learning: Enhancing Generalization Through Historical Consistency; *Ruizhao Zhu*; Venkatesh Saligrama**

27. Tight and Efficient Upper Bound on Spectral Norm of Convolutional Layers; *Ekaterina Grishina*; Mikhail Gorbunov; Maxim Rakhuba*

28. Deciphering the Role of Representation Disentanglement: Investigating Compositional Generalization in CLIP Models; *Reza Abbasi; Mohammad Rohban; Mahdieh Soleymani Baghshah**

29. Reinforcement Learning via Auxillary Task Distillation; *Abhinav N Harish*; Larry Heck; Josiah P Hanna; Zsolt Kira; Andrew Szot*

30. Dependency-aware Differentiable Neural Architecture Search; *Buang Zhang*; Xinle Wu; Hao Miao; Bin Yang; Chenjuan Guo*

31. Multimodal Cross-Domain Few-Shot Learning for Egocentric Action Recognition; *Masashi Hatano*; Ryo Hachiuma; Ryo Fujii; Hideo Saito*

32. Towards Model-Agnostic Dataset Condensation by Heterogeneous Models; *Jun-Yeong Moon; Jung Uk Kim*; Gyeong-Moon Park**

33. Enhanced Sparsification via Stimulative Training; *Shengji Tang; Weihao Lin; Hancheng Ye; Peng Ye; Chong Yu; Baopu Li; Tao Chen**

34. Interleaving One-Class and Weakly-Supervised Models with Adaptive Thresholding for Unsupervised Video Anomaly Detection; *Yongwei Nie; Hao Huang; Chengjiang Long; Qing Zhang; Pradipta Maji; Hongmin Cai**

35. Layer-Wise Relevance Propagation with Conservation Property for ResNet; *Seitaro Otsuki*; Tsumugi Iida*; Félix Doublet*; Tsubasa Hirakawa*; Takayoshi Yamashita*; Hironobu Fujiyoshi*; Komei Sugiura**

36. CLIP-Guided Generative Networks for Transferable Targeted Adversarial Attacks; *Hao Fang; Jiawei Kong; Bin Chen*; Tao Dai; Hao Wu; Shu-Tao Xia*

37. Leveraging Imperfect Restoration for Data Availability Attack; *YI HUANG*; Jeremy Styborski*; Mingzhi Lyu*; Fan Wang*; Wai-Kin Adams Kong**

38. Any Target Can be Offense: Adversarial Example Generation via Generalized Latent Infection; *Youheng Sun; Shengming Yuan; Xuanhan Wang*; Lianli Gao; Jingkuan Song*

39. Data-to-Model Distillation: Data-Efficient Learning Framework; *Ahmad Sajedi*; Samir Khaki; Lucy Z. Liu; Ehsan Amjadian; Yuri A. Lawryshyn; Konstantinos N. Plataniotis*

40. Active Generation for Image Classification; *Tao Huang; Jiaqi Liu; Shan You*; Chang Xu*

41. Augmented Neural Fine-tuning for Efficient Backdoor Purification; *Nazmul Karim*; Abdullah Al Arafat; Umar Khalid; Zhishan Guo; Nazanin Rahnavard*

42. DIFFender: Diffusion-Based Adversarial Defense against Patch Attacks; *Caixin Kang*; Yinpeng Dong; Zhengyi Wang; Shouwei Ruan; Yubo Chen; Hang Su*; Xingxing Wei**

43. GenQ: Quantization in Low Data Regimes with Generative Synthetic Data; *Yuhang Li*; Youngeun Kim; Donghyun Lee; Souvik Kundu; Priyadarshini Panda*

44. FYI: Flip Your Images for Dataset Distillation; *Byunggwan Son*; Youngmin Oh; Donghyeon Baek; Bumsub Ham**

45. Veil Privacy on Visual Data: Concealing Privacy for Humans, Unveiling for DNNs; *Shuchao Pang*; Ruhao Ma; Bing Li*; Yongbin Zhou; Yazhou Yao*

46. Efficient Image Pre-Training with Siamese Cropped Masked Autoencoders; *Alexandre Eymaël; Renaud Vandeghen*; Anthony Cioppa; Silvio Giancola; Bernard Ghanem; Marc Van Droogenbroeck*

47. Learning to Detect Multi-class Anomalies with Just One Normal Image Prompt; *Bin-Bin Gao**

48. Get Your Embedding Space in Order: Domain-Adaptive Regression for Forest Monitoring; *Sizhuo Li; Dimitri Gominski*; Martin Brandt; Xiaoye Tong; Philippe Ciais*

49. DecentNeRFs: Decentralized Neural Radiance Fields from Crowdsourced Images; *Zaid Tasneem*; Akshat Dave; Abhishek Singh; Kushagra Tiwary; Praneeth Vepakomma; Ashok Veeraraghavan; Ramesh Raskar*

50. Learning Representations of Satellite Images From Metadata Supervision; *Jules Bourcier*; Gohar Dashyan; Karteek Alahari; Jocelyn Chanussot*

51. Distributed Semantic Segmentation with Efficient Joint Source and Task Decoding; *Danish Nazir*; Timo Bartels; Jan Piewek; Thorsten Bagdonat; Tim Fingscheidt*

52. Learning with Unmasked Tokens Drives Stronger Vision Learners; *Taekyung Kim*; Sanghyuk Chun; Byeongho Heo; Dongyoon Han**

53. InfMAE: A Foundation Model in The Infrared Modality; *Fangcen Liu; Chenqiang Gao*; Yaming Zhang; Junjie Guo; Jinghao Wang; Deyu Meng*

54. Image Manipulation Detection With Implicit Neural Representation and Limited Supervision; *Zhenfei Zhang*; Mingyang Li; Xin Li; Ming-Ching Chang; Jun-Wei Hsieh*

55. Towards Latent Masked Image Modeling for Self-Supervised Visual Representation Learning; *Yibing Wei*; Abhinav Gupta; Pedro Morgado**

56. TransFusion -- A Transparency-Based Diffusion Model for Anomaly Detection; *Matic Fučka*; Vitjan Zavrtanik; Danijel Skočaj*

57. AFreeCA: Annotation-Free Counting for All; *Adriano D'Alessandro*; Ali Mahdavi-Amiri; Ghassan Hamarneh*

58. SAIR: Learning Semantic-aware Implicit Representation; *Canyu Zhang*; Xiaoguang Li*; Qing Guo*; Song Wang**

59. Dual-stage Hyperspectral Image Classification Model with Spectral Supertoken; *Peifu Liu; Tingfa Xu*; Jie Wang; Huan Chen; Huiyan Bai; Jianan Li**

60. The Role of Masking for Efficient Supervised Knowledge Distillation of Vision Transformers; *Seungwoo Son*; Jegwang Ryu; Namhoon Lee; Jaeho Lee**

61. Look Around and Learn: Self-Training Object Detection by Exploration; *Gianluca Scarpellini*; Stefano Rosa*; Pietro Morerio; Lorenzo Natale; Alessio Del Bue*

62. Distilling Knowledge from Large-Scale Image Models for Object Detection; *Gang Li*; Wenhai Wang; Xiang Li; Ziheng Li; Jian Yang; Jifeng Dai; Yu Qiao; Shanshan Zhang**

63. SHERL: Synthesizing High Accuracy and Efficient Memory for Resource-Limited Transfer Learning; *Haiwen Diao*; Bo Wan; Xu Jia; Yunzhi Zhuge; Ying Zhang; Huchuan Lu*; Long Chen*

64. OpenPSG: Open-set Panoptic Scene Graph Generation via Large Multimodal Models; *Zijian Zhou*; Zheng Zhu; Holger Caesar; Miaojing Shi**

65. Knowledge Transfer with Simulated Inter-Image Erasing for Weakly Supervised Semantic Segmentation; *Tao Chen*; Xiruo Jiang; Gensheng Pei; Zeren Sun; Yucheng Wang; Yazhou Yao*

66. ProMerge: Prompt and Merge for Unsupervised Instance Segmentation; *Dylan J Li; Gyungin Shin**

67. Adaptive Multi-task Learning for Few-shot Object Detection; *Yan Ren*; Yanling Li; Adams Wai-Kin Kong*

68. Crowd-SAM:SAM as a smart annotator for object detection in crowded scenes; *Zhi Cai; Yingjie Gao; Yaoyan Zheng; Nan Zhou; Di Huang**

69. Revisiting Domain-Adaptive Object Detection in Adverse Weather by the Generation and Composition of High-Quality Pseudo-Labels; *Rui Zhao; Huibin Yan; Shuoyao Wang**

70. VCP-CLIP: A visual context prompting model for zero-shot anomaly segmentation; *Zhen Qu; Xian Tao*; Mukesh Prasad; Fei Shen; Zhengtao Zhang; Xinyi Gong; Guiguang Ding*

71. UniFS: Universal Few-shot Instance Perception with Point Representations; *Sheng Jin*; Ruijie Yao; Lumin Xu; Wentao Liu*; Chen Qian; Ji Wu; Ping Luo**

72. Mind the Interference: Retaining Pre-trained Knowledge in Parameter Efficient Continual Learning of Vision-Language Models; *Longxiang Tang*; Zhuotao Tian; Kai Li; Chunming He; Hantao Zhou; Hengshuang Zhao; Xiu Li; Jiaya Jia*

73. Boosting Gaze Object Prediction via Pixel-level Supervision from Vision Foundation Model; *Yang Jin; Lei Zhang; Shi Yan; Bin Fan; Binglu Wang**

74. MarvelOVD: Marrying Object Recognition and Vision-Language Models for Robust Open-Vocabulary Object Detection; *Kuo Wang; Lechao Cheng*; Weikai Chen; Pingping Zhang; Liang Lin; Fan Zhou; Guanbin Li**

75. Prioritized Semantic Learning for Zero-shot Instance Navigation; *xinyu sun*; Lizhao Liu; Hongyan Zhi; Ronghe Qiu; Junwei Liang**

76. Diffusion for Out-of-Distribution Detection on Road Scenes and Beyond; *Silvio Galesso*; Philipp Schröppel*; Hssan Driss; Thomas Brox*

77. DIAL: Dense Image-text ALignment for Weakly Supervised Semantic Segmentation; *Soojin Jang; JungMin Yun; JuneHyoung Kwon; Eunju Lee; YoungBin Kim**

78. Unified Embedding Alignment for Open-Vocabulary Video Instance Segmentation; *Hao Fang; Peng Wu; Yawei Li; Xinxin Zhang; Xiankai Lu**

79. Robust Calibration of Large Vision-Language Adapters; *Balamurali Murugesan*; Julio Silva-Rodríguez; Ismail Ben Ayed; Jose Dolz*

80. Explore the Potential of CLIP for Training-Free Open Vocabulary Semantic Segmentation; *Tong Shao; Zhuotao Tian*; Hang Zhao; Jingyong Su**

81. Emerging Property of Masked Token for Effective Pre-training; *Hyesong Choi; Hunsang Lee; Seyoung Joung; Hyejin Park; Jiyeong Kim; Dongbo Min**

82. SemiVL: Semi-Supervised Semantic Segmentation with Vision-Language Guidance; *Lukas Hoyer*; David Joseph Tan; Muhammad Ferjad Naeem; Luc Van Gool; Federico Tombari*

83. Cascade Prompt Learning for Visual-Language Model Adaptation; *Ge Wu; Xin Zhang; Zheng Li; Zhaowei Chen; Jiajun Liang; Jian Yang; Xiang Li**

84. Removing Rows and Columns of Tokens in Vision Transformer enables Faster Dense Prediction without Retraining; *Diwei Su; cheng fei; Jianxu Luo**

85. Dense Multimodal Alignment for Open-Vocabulary 3D Scene Understanding; *Ruihuang Li*; Zhengqiang ZHANG; Chenhang He; Zhiyuan Ma; Vishal Patel; Lei Zhang*

86. ClearCLIP: Decomposing CLIP Representations for Dense Vision-Language Inference; *Mengcheng Lan; Chaofeng Chen; Yiping Ke; Xinjiang Wang; Litong Feng*; Wayne Zhang*

87. Towards Open-Ended Visual Recognition with Large Language Models; *Qihang Yu*; Xiaohui Shen; Liang-Chieh Chen*

88. EventBind: Learning a Unified Representation to Bind Them All for Event-based Open-world Understanding; *jiazhou zhou*; Xu Zheng; Yuanhuiyi Lyu; Lin Wang*

89. TIP: Tabular-Image Pre-training for Multimodal Classification with Incomplete Data; *Siyi Du*; Shaoming Zheng; Yinsong Wang; Wenjia Bai; Declan P. O'Regan; Chen Qin**

90. Self-Adapting Large Visual-Language Models to Edge Devices across Visual Modalities; *Kaiwen Cai; ZheKai Duan; Gaowen Liu; Charles Fleming; Chris Xiaoxuan Lu**

91. ArtVLM: Attribute Recognition Through Vision-Based Prefix Language Modeling; *William Yicheng Zhu*; Keren Ye*; Junjie Ke; Jiahui Yu; Leonidas Guibas; Peyman Milanfar; Feng Yang**

92. Unified Medical Image Pre-training in Language-Guided Common Semantic Space; *Xiaoxuan He; Yifan Yang; Xinyang Jiang; Xufang Luo*; Haoji Hu; Siyun Zhao; Dongsheng Li; Yuqing Yang; Lili Qiu*

93. Enhancing Recipe Retrieval with Foundation Models: A Data Augmentation Perspective; *Fangzhou Song; Bin Zhu; Yanbin Hao*; Shuo Wang*

94. Object-Aware Query Perturbation for Cross-Modal Image-Text Retrieval; *Naoya Sogi*; Takashi Shibata*; Makoto Terao**

95. Parrot Captions Teach CLIP to Spot Text; *Yiqi Lin; Conghui He*; Alex Jinpeng Wang; Bin Wang; Weijia Li; Mike Zheng Shou*

96. IG Captioner: Information Gain Captioners are Strong Zero-shot Classifiers; *Chenglin Yang*; Siyuan Qiao; Yuan Cao; Yu Zhang; Tao Zhu; Alan Yuille; Jiahui Yu*

97. Language-Image Pre-training with Long Captions; *Kecheng Zheng*; Yifei Zhang; Wei Wu; Fan Lu; Shuailei Ma; Xin Jin; Wei Chen; Yujun Shen*

98. CIC-BART-SSA: : Controllable Image Captioning with Structured Semantic Augmentation; *Kalliopi Basioti*; Mohamed A Abdelsalam*; Federico Fancellu*; Vladimir Pavlovic*; Afsaneh Fazly**

99. X-Former: Unifying Contrastive and Reconstruction Learning for MLLMs; *Sirnam Swetha*; Jinyu Yang; Tal Neiman; Mamshad Nayeem Rizve; Son Tran; Benjamin Yao; Trishul A Chilimbi; Mubarak Shah*

100. The Hard Positive Truth about Vision-Language Compositionality; *Amita Kamath*; Cheng-Yu Hsieh; Kai-Wei Chang; Ranjay Krishna*

101. HiFi-Score: Fine-grained Image Description Evaluation with Hierarchical Parsing Graphs; *Ziwei Yao; Ruiping Wang*; Xilin Chen*

102. UniCode : Learning a Unified Codebook for Multimodal Large Language Models; *Sipeng Zheng*; Bohan Zhou; Yicheng Feng; Ye Wang; Zongqing Lu**

103. SAM4MLLM: Enhance Multi-Modal Large Language Model for Referring Expression Segmentation; *Yi-Chia Chen; Wei-Hua Li; Cheng Sun; Yu-Chiang Frank Wang; Chu-Song Chen**

104. LLMCO4MR: LLMs-aided Neural Combinatorial Optimization for Ancient Manuscript Restoration from Fragments with Case Studies on Dunhuang; *Yuqing Zhang; Hangqi Li; Shengyu Zhang*; Runzhong Wang; Baoyi He; Huaiyong Dou; Junchi Yan*; Yongquan Zhang; Fei Wu*

105. PointLLM: Empowering Large Language Models to Understand Point Clouds; *Runsen Xu*; Xiaolong Wang; Tai Wang*; Yilun Chen; Jiangmiao Pang*; Dahua Lin* BEST PAPER CANDIDATE

106. Towards Open-ended Visual Quality Comparison; *Haoning Wu; Hanwei Zhu; Zicheng Zhang; Erli Zhang; Chaofeng Chen; Liang Liao; Chunyi Li; Annan Wang; Wenxiu Sun; Qiong Yan; Xiaohong Liu; Guangtao Zhai; Shiqi Wang; Weisi Lin**

107. A Comprehensive Study of Multimodal Large Language Models for Image Quality Assessment; *Tianhe Wu; Kede Ma*; Jie Liang; Yujiu Yang*; Lei Zhang*

108. PathMMU: A Massive Multimodal Expert-Level Benchmark for Understanding and Reasoning in Pathology; *Yuxuan Sun*; Hao Wu; Chenglu Zhu; Sunyi Zheng; Qizi Chen; Kai Zhang; Yunlong Zhang; Dan Wan; Xiaoxiao Lan; Mengyue Zheng; Jingxiong Li; Xinheng Lyu; Tao Lin*; Lin Yang** BEST PAPER CANDIDATE

109. Groma: Localized Visual Tokenization for Grounding Multimodal Large Language Models; *Chuofan Ma*; Yi Jiang*; Jiannan Wu; Zehuan Yuan; Xiaojuan Qi**

110. Knowledge-enhanced Visual-Language Pretraining for Computational Pathology; *Xiao Zhou; Xiaoman Zhang; Chaoyi Wu; Ya Zhang; Weidi Xie; Yan-Feng Wang**

111. The First to Know: How Token Distributions Reveal Hidden Knowledge in Large Vision-Language Models?; *Qinyu Zhao*; Ming Xu; Kartik Gupta; Akshay Asthana; Liang Zheng; Stephen Gould*

112. Unifying 3D Vision-Language Understanding via Promptable Queries; *ziyu zhu*; Zhuofan Zhang; Xiaojian Ma; Xuesong Niu; Yixin Chen; Baoxiong Jia; Zhidong Deng*; Siyuan Huang*; Qing Li**

113. Grounding Language Models for Visual Entity Recognition; *Zilin Xiao*; Ming Gong; Paola Cascante-Bonilla; Xingyao Zhang; Jie Wu; Vicente Ordonez**

114. AdaShield: Safeguarding Multimodal Large Language Models from Structure-based Attack via Adaptive Shield Prompting; *Yu Wang*; Xiaogeng Liu*; Yu Li*; Muhao Chen; Chaowei Xiao**

115. CoReS: Orchestrating the Dance of Reasoning and Segmentation; *Xiaoyi Bao; Siyang Sun; Shuailei Ma; Kecheng Zheng; Yuxin Guo; Guosheng Zhao; Yun Zheng; Xingang Wang**

116. UniIR: Training and Benchmarking Universal Multimodal Information Retrievers; *Cong Wei*; Yang Chen; Haonan Chen; Hexiang Hu; Ge Zhang; Jie Fu; Alan Ritter; Wenhu Chen*

117. PARIS3D: Reasoning-based 3D Part Segmentation Using Large Multimodal Model; *Amrin Kareem*; Jean Lahoud; Hisham Cholakkal**

118. M3DBench: Towards Omni 3D Assistant with Interleaved Multi-modal Instructions; *Mingsheng Li; Xin Chen; Chi Zhang; Sijin Chen; Hongyuan Zhu; Fukun Yin; Zhuoyuan Li; Gang Yu; Tao Chen**

119. UMBRAE: Unified Multimodal Brain Decoding; *Weihao Xia*; Raoul de Charette; A. Cengiz Oztireli; Jing-Hao Xue*

120. QUAR-VLA: Vision-Language-Action Model for Quadruped Robots; *Pengxiang Ding; Han Zhao; Wenjie Zhang; Wenxuan Song; Min Zhang; Siteng Huang; Ningxi Yang; Donglin Wang**

121. Navigation Instruction Generation with BEV Perception and Large Language Models; *Sheng Fan; Rui Liu; Wenguan Wang*; Yi Yang*

122. OmniACT: A Dataset and Benchmark for Enabling Multimodal Generalist Autonomous Agents for Desktop and Web; *Raghav Kapoor*; Yash Parag Butala*; Melisa A Russak; Jing Yu Koh; Kiran Kamble; Waseem AlShikh; Ruslan Salakhutdinov*

123. Towards Multimodal Sentiment Analysis Debiasing via Bias Purification; *Dingkang Yang; Mingcheng Li; Dongling Xiao; Yang Liu; Kun Yang; Zhaoyu Chen; Yuzheng Wang; Peng Zhai*; Ke Li; Lihua Zhang**

124. V-IRL: Grounding Virtual Intelligence in Real Life; *Jihan Yang*; Runyu Ding; Ellis L Brown; Xiaojuan Qi; Saining Xie*

125. MotionChain: Conversational Motion Controllers via Multimodal Prompts; *Biao Jiang; Xin Chen; Chi Zhang; Fukun Yin; Zhuoyuan Li; Gang Yu; Jiayuan Fan**

126. BI-MDRG: Bridging Image History in Multimodal Dialogue Response Generation; *Hee Suk Yoon; Eunseop Yoon; Joshua Tian Jin Tee; Kang Zhang; Yu-Jung Heo; Du-Seong Chang; Chang D. Yoo**

127. AutoEval-Video: An Automatic Benchmark for Assessing Large Vision Language Models in Open-Ended Video Question Answering; *Weiran Huang*; Xiuyuan Chen*; Yuan Lin*; Yuchen Zhang**

128. Learning Video Context as Interleaved Multimodal Sequences; *Kevin Qinghong Lin; Pengchuan Zhang; Difei Gao; Xide Xia; Joya Chen; Ziteng Gao; Jinheng Xie; Xuhong Xiao; Mike Zheng Shou**

130. Multi-Modal Video Dialog State Tracking in the Wild; *Adnen Abdessaied*; Lei Shi; Andreas Bulling*

131. VideoAgent: Long-form Video Understanding with Large Language Model as Agent; *Xiaohan Wang*; Yuhui Zhang; Orr Zohar; Serena Yeung-Levy*

132. Elysium: Exploring Object-level Perception in Videos through Semantic Integration Using MLLMs; *Han Wang*; Yanjie Wang; Ye Yongjie; Yuxiang Nie; Can Huang*

133. VITATECS: A Diagnostic Dataset for Temporal Concept Understanding of Video-Language Models; *Shicheng Li; Lei Li; Yi Liu; Shuhuai Ren; Yuanxin Liu; Rundong Gao; Xu Sun*; Lu Hou*

134. Referring Atomic Video Action Recognition; *Kunyu Peng*; Jia Fu; Kailun Yang; Di Wen; Yufan Chen; Ruiping Liu; Junwei Zheng; Jiaming Zhang; Saquib Sarfraz; Rainer Stiefelhagen; Alina Roitberg*

135. N2F2: Hierarchical Scene Understanding with Nested Neural Feature Fields; *Yash Bhalgat*; Iro Laina; Joao F Henriques; Andrew Zisserman; Andrea Vedaldi*

136. Snuffy: Efficient Whole Slide Image Classifier; *Hossein Jafarinia*; Alireza Alipanah; Saeed Razavi; Nahal Mirzaie; Mohammad Hossein Rohban**

137. Enhancing Cross-Subject fMRI-to-Video Decoding with Global-Local Functional Alignment; *Chong Li*; Xuelin Qian; Yun Wang; Jingyang Huo; Xiangyang Xue*; Yanwei Fu*; Jianfeng Feng*

138. AdaGlimpse: Active Visual Exploration with Arbitrary Glimpse Position and Scale; *Adam Pardyl*; Michał Wronka; Maciej Wołczyk; Kamil Adamczewski; Tomasz Trzcinski; Bartosz Zieliński**

139. Asynchronous Bioplausible Neuron for Spiking Neural Networks for Event-Based Vision; *Hussain Sajwani; Dimitrios Makris; Yahya Prof. Zweiri; Fariborz Baghaei Naeini; Sanket Mr Kachole**

140. Wavelet Convolutions for Large Receptive Fields; *Shahaf E Finder*; Roy Amoyal; Eran Treister; Oren Freifeld**

141. AUFormer: Vision Transformers are Parameter-Efficient Facial Action Unit Detectors; *Kaishen Yuan; Zitong Yu*; Xin Liu*; Weicheng Xie; Huanjing Yue; Jingyu Yang*

142. Implicit Neural Models to Extract Heart Rate from Video; *Pradyumna Chari*; Anirudh Bindiganavale Harish; Adnan Armouti; Alexander Vilesov; Sanjit Sarda; Laleh Jalilian; Achuta Kadambi*

143. Learning Natural Consistency Representation for Face Forgery Video Detection; *Daichi Zhang*; Zihao Xiao; Shikun Li; Fanzhao Lin; Jianmin Li; Shiming Ge**

144. Effective Lymph Nodes Detection in CT Scans Using Location Debiased Query Selection and Contrastive Query Representation in Transformer; *Qinji Yu*; Yirui Wang*; Ke Yan; Haoshen Li; Dazhou Guo; Li Zhang; Na Shen; Qifeng Wang; Xiaowei Ding; Le Lu; Xianghua Ye*; Dakai Jin**

145. Revisiting Adaptive Cellular Recognition Under Domain Shifts: A Contextual Correspondence View; *Jianan Fan*; Dongnan Liu; Canran Li; Hang Chang; Heng Huang; Filip Braet; Mei Chen; Weidong Cai**

146. MarineInst: A Foundation Model for Marine Image Analysis with Instance Visual Description; *Ziqiang Zheng*; Yiwei Chen; Huimin Zeng; Tuan-Anh Vu; Binh-Son Hua; Sai-Kit Yeung*

147. Insect Identification in the Wild: The AMI Dataset; *Aditya Jain*; Fagner Cunha; Michael J Bunsen; Juan Sebastián Cañas; Léonard Pasi; Nathan Pinoy; Flemming Helsing; JoAnne Russo; Marc S Botham; Michael Sabourin; Jonathan Fréchette; Alexandre Anctil; Yacksecari Lopez; Eduardo Navarro; Filonila Pérez; Ana C Zamora; Jose Alejandro Ramirez-Silva; Jonathan Gagnon; Tom A August; Kim Bjerge; Alba Gomez Segura; Marc Belisle; Yves Basset; Kent P McFarland; David B Roy; Toke T Høye; Maxim Larrivee; David Rolnick*

148. VETRA: A Dataset for Vehicle Tracking in Aerial Imagery - New Challenges for Multi-Object Tracking; *Jens Hellekes*; Manuel Mühlhaus; Reza Bahmanyar; Seyed Majid Azimi; Franz Kurz*

149. Rethinking Data Augmentation for Robust LiDAR Semantic Segmentation in Adverse Weather; *Junsung Park; Kyungmin Kim; Hyunjung Shim**

150. Rethinking Deep Unrolled Model for Accelerated MRI Reconstruction; *Bingyu Xin*; Meng Ye; Leon Axel; Dimitris N. Metaxas*

151. RAPiD-Seg: Range-Aware Pointwise Distance Distribution Networks for 3D LiDAR Segmentation; *Li Li*; Hubert P. H. Shum; Toby P Breckon*

152. Adaptive Correspondence Scoring for Unsupervised Medical Image Registration; *Xiaoran Zhang*; John C. Stendahl; Lawrence H. Staib; Albert J. Sinusas; Alex Wong; James S. Duncan*

153. DVLO: Deep Visual-LiDAR Odometry with Local-to-Global Feature Fusion and Bi-Directional Structure Alignment; *Jiuming Liu; Dong Zhuo; Zhiheng Feng; Siting Zhu; Chensheng Peng; Zhe Liu; Hesheng Wang**

154. CardiacNet: Learning to Reconstruct Abnormalities for Cardiac Disease Assessment from Echocardiogram Videos; *Jiewen Yang*; Yiqun Lin; Bin Pu; Jiarong GUO; Xiaowei Xu*; Xiaomeng Li**

155. Equi-GSPR: Equivariant SE(3) Graph Network Model for Sparse Point Cloud Registration; *Xueyang Kang*; Zhaoliang Luan; Kourosh Khoshelham; Bing WANG**

156. SparseSSP: 3D Subcellular Structure Prediction from Sparse-View Transmitted Light Images; *Jintu Zheng; Yi Ding; Qizhe Liu; Yuehui Chen; Yi Cao; Ying Hu; Zenan Wang**

157. HGL: Hierarchical Geometry Learning for Test-time Adaptation in 3D Point Cloud Segmentation; *Tianpei Zou; Sanqing Qu; Zhijun Li; Alois C. Knoll; 何 良 ; Guang Chen*; Changjun Jiang*

158. RISurConv: Rotation Invariant Surface Attention-Augmented Convolutions for 3D Point Cloud Classification and Segmentation; *Zhiyuan Zhang*; Licheng Yang; Zhiyu Xiang*

159. KeypointDETR: An End-to-End 3D Keypoint Detector; *Hairong Jin; Yuefan Shen; Jianwen Lou; Kun Zhou; Youyi Zheng**

160. Learning Exhaustive Correlation for Spectral Super-Resolution: Where Spatial-Spectral Attention Meets Linear Dependence; *Hongyuan Wang; Lizhi Wang*; Jiang Xu; Chang Chen; Xue Hu; Fenglong Song; Youliang Yan*

161. LightenDiffusion: Unsupervised Low-Light Image Enhancement with Latent-Retinex Diffusion Models; *Hai Jiang; Ao Luo; Xiaohong Liu; Songchen Han; Shuaicheng Liu**

162. Kalman-Inspired Feature Propagation for Video Face Super-Resolution; *Ruicheng Feng; Chongyi Li; Chen Change Loy**

163. Kernel Diffusion: An Alternate Approach to Blind Deconvolution; *Yash Sanghvi*; Yiheng Chi; Stanley Chan*

164. Zero-Shot Adaptation for Approximate Posterior Sampling of Diffusion Models in Inverse Problems; *Yasar U Alcalar*; Mehmet Akcakaya*

165. A Comparative Study of Image Restoration Networks for General Backbone Network Design; *Xiangyu Chen*; Zheyuan Li; Yuandong Pu; Yihao Liu; Jiantao Zhou*; Yu Qiao; Chao Dong**

166. Compensation Sampling for Improved Convergence in Diffusion Models; *Hui Lu*; Albert Ali Salah; Ronald Poppe*

167. Unmasking Bias in Diffusion Model Training; *Hu Yu; Li Shen; Jie Huang; Hongsheng Li; Feng Zhao**

168. DiffiT: Diffusion Vision Transformers for Image Generation; *Ali Hatamizadeh*; Jiaming Song; Guilin Liu; Jan Kautz; Arash Vahdat*

169. DC-Solver: Improving Predictor-Corrector Diffusion Sampler via Dynamic Compensation; *Wenliang Zhao; Haolin Wang; Jie Zhou; Jiwen Lu**

170. RealViformer: Investigating Attention for Real-World Video Super-Resolution; *Yuehan Zhang*; Angela Yao*

171. OAPT: Offset-Aware Partition Transformer for Double JPEG Artifacts Removal; *Qiao Mo; Yukang Ding; Jinhua Hao*; Qiang Zhu; Ming Sun; Chao Zhou; Feiyu Chen; Shuyuan Zhu**

172. Enhancing Perceptual Quality in Video Super-Resolution through Temporally-Consistent Detail Synthesis using Diffusion Models; *Claudio Rota*; Marco Buzzelli; Joost van de Weijer*

173. ∞-Brush: Controllable Large Image Synthesis with Diffusion Models in Infinite Dimensions; *Minh-Quan Le*; Alexandros Graikos; Srikar Yellapragada; Rajarsi Gupta; Joel Saltz; Dimitris Samaras*

174. Domain-adaptive Video Deblurring via Test-time Blurring; *Jin-Ting He*; Fu-Jen Tsai; Jia-Hao Wu; Yan-Tsung Peng; Chung-Chi Tsai; Chia-Wen Lin; Yen-Yu Lin*

175. Gaze Target Detection Based on Head-Local-Global Coordination; *Yaokun Yang; Feng Lu**

176. A Watermark-Conditioned Diffusion Model for IP Protection; *Rui Min*; Sen Li*; Hongyang Chen*; Minhao Cheng**

177. Dual-Rain: Video Rain Removal using Assertive and Gentle Teachers; *Tingting Chen*; Beibei Lin; Yeying Jin; Wending Yan; WEI YE; Yuan Yuan; Robby T. Tan*

178. Self-Supervised Video Desmoking for Laparoscopic Surgery; *Renlong Wu; Zhilu Zhang*; Shuohao Zhang; Longfei Gou; Haobin Chen; Lei Zhang; Hao Chen*; Wangmeng Zuo*

179. PEA-Diffusion: Parameter-Efficient Adapter with Knowledge Distillation in non-English Text-to-Image Generation; *jian ma; Chen Chen*; Qingsong Xie; Haonan Lu**

180. Lost in Translation: Latent Concept Misalignment in Text-to-Image Diffusion Models; *Juntu Zhao; Junyu Deng; Yixin Ye; Chongxuan Li; Zhijie Deng*; Dequan Wang**

181. Distilling Diffusion Models into Conditional GANs; *MinGuk Kang*; Richard Zhang; Connelly Barnes; Sylvain Paris; Suha Kwak; Jaesik Park; Eli Shechtman; Jun-Yan Zhu; Taesung Park**

182. Reliable and Efficient Concept Erasure of Text-to-Image Diffusion Models; *Chao Gong*; Kai Chen; Zhipeng Wei; Jingjing Chen*; Yu-Gang Jiang*

183. Post-training Quantization with Progressive Calibration and Activation Relaxing for Text-to-Image Diffusion Models; *Siao Tang; Xin Wang*; Hong Chen; Chaoyu Guan; Zewen Wu; Yansong Tang; Wenwu Zhu**

184. MaxFusion: Plug&Play Multi-Modal Generation in Text-to-Image Diffusion Models; *Nithin Gopalakrishnan Nair*; Jeya Maria Jose Valanarasu; Vishal Patel*

185. DCDM: Diffusion-Conditioned-Diffusion Model for Scene Text Image Super-Resolution; *Shrey Singh*; Prateek Keserwani; Masakazu Iwamura*; Partha Pratim Roy*

186. Teaching Tailored to Talent: Adverse Weather Restoration via Prompt Pool and Depth-Anything Constraint; *Sixiang Chen; Tian Ye; Kai Zhang; Zhaohu Xing; Yunlong Lin; Lei Zhu**

187. AccDiffusion: An Accurate Method for Higher-Resolution Image Generation; *Zhihang Lin; Mingbao Lin; Meng Zhao; Rongrong Ji**

188. ComFusion: Enhancing Personalized Generation by Instance-Scene Compositing and Fusion; *Yan Hong*; Yuxuan Duan; Bo Zhang; Haoxing Chen; Jun Lan; Huijia Zhu; Weiqiang Wang; Jianfu Zhang**

189. BeyondScene: Higher-Resolution Human-Centric Scene Generation With Pretrained Diffusion; *Gwanghyun Kim; Hayeon Kim; Hoigi Seo; Dong Un Kang; Se Young Chun**

190. TP2O: Creative Text Pair-to-Object Generation using Balance Swap-Sampling; *Jun Li*; Zedong Zhang; Jian Yang*

191. StyleTokenizer: Defining Image Style by a Single Instance for Controlling Diffusion Models; *Wen Li*; Muyuan Fang; Cheng Zou; Biao Gong; Ruobing Zheng; Meng Wang; Jingdong Chen; Ming Yang*

192. Glyph-ByT5: A Customized Text Encoder for Accurate Visual Text Rendering; *Zeyu Liu; Weicong Liang; Zhanhao Liang; Chong Luo; Ji Li; Gao Huang; Yuhui Yuan**

193. The Fabrication of Reality and Fantasy: Scene Generation with LLM-Assisted Prompt Interpretation; *Yi Yao; Chan-Feng Hsu*; Jhe-Hao Lin; Hongxia Xie; Terence Lin; Yi-Ning Huang; Hong-Han Shuai*; Wen-Huang Cheng**

194. DreamDiffusion: High-Quality EEG-to-Image Generation with Temporal Masked Signal Modeling and CLIP Alignment; *Yunpeng Bai*; Xintao Wang; Yan-Pei Cao; Yixiao Ge; Chun Yuan; Ying Shan*

195. Tuning-Free Image Customization with Image and Text Guidance; *Pengzhi Li; Qiang Nie; Ying Chen; Xi Jiang; Kai Wu; Yuhuan Lin; Yong Liu; Jinlong Peng; Chengjie Wang; Feng Zheng**

196. Curved Diffusion: A Generative Model With Optical Geometry Control; *Andrey Voynov*; Amir Hertz; Moab Arar; Shlomi Fruchter; Daniel Cohen-Or*

197. AID-AppEAL: Automatic Image Dataset and Algorithm for Content Appeal Enhancement and Assessment Labeling; *Sherry X. Chen*; Yaron Vaxman; Elad Ben Baruch; David Asulin; Aviad Moreshet; Misha Sra; Pradeep Sen*

198. HiEI: A Universal Framework for Generating High-quality Emerging Images from Natural Images; *Jingmeng Li; Lukang Fu; Surun Yang; Hui Wei**

199. MagicEraser: Erasing Any Objects via Semantics-Aware Control; *Fan Li*; Zixiao Zhang; Yi Huang; Jianzhuang Liu; Renjing Pei; Bin Shao; Songcen Xu*

200. Improving Diffusion Models for Authentic Virtual Try-on in the Wild; *Yisol Choi*; Sangkyung Kwak; Kyungmin Lee; Hyungwon Choi; Jinwoo Shin**

201. Kinetic Typography Diffusion Model; *Seonmi Park; Inhwan Bae; Seunghyun Shin; Hae-Gon Jeon**

202. SAVE: Protagonist Diversification with Structure Agnostic Video Editing; *Yeji Song*; Wonsik Shin; Junsoo Lee; Jeesoo Kim; Nojun Kwak**

203. Responsible Visual Editing; *Minheng Ni; Yeli Shen; Lei Zhang*; Wangmeng Zuo**

204. SMooDi: Stylized Motion Diffusion Model; *Lei Zhong; Yiming Xie; Varun Jampani; Deqing Sun; Huaizu Jiang**

205. Eta Inversion: Designing an Optimal Eta Function for Diffusion-based Real Image Editing; *Wonjun Kang; Kevin Galim; Hyung Il Koo**

206. Mutual Learning for Acoustic Matching and Dereverberation via Visual Scene-driven Diffusion; *Jian Ma; Wenguan Wang*; Yi Yang; Feng Zheng*

207. Chains of Diffusion Models; *Yanheng Wei*; Lianghua Huang*; Zhi-Fan Wu; Wei Wang; Yu Liu; Mingda Jia; Shuailei Ma*

208. M2D2M: Multi-Motion Generation from Text with Discrete Diffusion Models; *Seunggeun Chi*; Hyung-gun Chi; Hengbo Ma; Nakul Agarwal; Faizan Siddiqui; Karthik Ramani*; Kwonjoon Lee**

209. VFusion3D: Learning Scalable 3D Generative Models from Video Diffusion Models; *Junlin Han*; Filippos Kokkinos; Philip Torr*

210. All You Need is Your Voice: Emotional Face Representation with Audio Perspective for Emotional Talking Face Generation; *Seongho Kim; Byung Cheol Song**

211. GIVT: Generative Infinite-Vocabulary Transformers; *Michael Tschannen*; Cian Eastwood; Fabian Mentzer*

212. Scene-Conditional 3D Object Stylization and Composition; *Jinghao Zhou*; Tomas Jakab; Philip Torr; Christian Rupprecht*

213. HeadGaS: Real-Time Animatable Head Avatars via 3D Gaussian Splatting; *Helisa Dhamo*; Yinyu Nie; Arthur Moreau; Jifei Song; Richard Shaw; Yiren Zhou; Eduardo Pérez-Pellitero**

214. Stable Video Portraits; *Mirela Ostrek*; Justus Thies*

215. TalkingGaussian: Structure-Persistent 3D Talking Head Synthesis via Gaussian Splatting; *Jiahe Li; Jiawei Zhang; Xiao Bai*; Jin Zheng*; Xin Ning; Jun Zhou; Lin Gu*

216. DreamScene: 3D Gaussian-based Text-to-3D Scene Generation via Formation Pattern Sampling; *Haoran Li; Haolin Shi; Wenli Zhang; Wenjun Wu; Yong Liao*; Lin Wang; Lik-Hang Lee; Peng Yuan Zhou**

217. PhysGen: Rigid-Body Physics-Grounded Image-to-Video Generation; *Shaowei Liu; Zhongzheng Ren; Saurabh Gupta; Shenlong Wang**

218. RoomTex: Texturing Compositional Indoor Scenes via Iterative Inpainting; *Qi WANG*; Ruijie Lu; Xudong XU; Jingbo Wang; Michael Yu Wang; Bo Dai; Gang Zeng; Dan Xu*

219. POCA: Post-training Quantization with Temporal Alignment for Codec Avatars; *Jian Meng*; Yuecheng Li*; Leo (Chenghui) Li; Syed Shakib Sarwar; Dilin Wang; Jae-sun Seo**

220. FocusDiffuser: Perceiving Local Disparities for Camouflaged Object Detection; *Jianwei Zhao*; Xin Li; Fan Yang; Qiang Zhai*; Ao Luo; Zhicheng Jiao; Hong Cheng*

221. LayoutDETR: Detection Transformer Is a Good Multimodal Layout Designer; *Ning Yu*; Chia-chih Chen; Zeyuan Chen; Rui Meng; Gang Wu; Paul W Josel; Juan Carlos Niebles; Caiming Xiong; Ran Xu*

222. Real-time 3D-aware Portrait Editing from a Single Image; *Qingyan Bai*; Zifan Shi; Yinghao Xu; Hao Ouyang; Qiuyu Wang; Ceyuan Yang; Xuan Wang; Gordon Wetzstein; Yujun Shen*; Qifeng Chen**

223. iHuman: Instant Animatable Digital Humans From Monocular Videos; *Pramish Paudel*; Anubhav Khanal; Danda Pani Paudel; Jyoti Tandukar; Ajad Chhatkuli*

224. Topology-Preserving Downsampling of Binary Images; *Chia-Chia Chen*; Chi-Han Peng**

225. PanoFree: Tuning-Free Holistic Multi-view Image Generation with Cross-view Self-Guidance; *Aoming Liu*; Zhong Li*; Zhang Chen*; Nannan Li; Yi Xu; Bryan Plummer*

226. An Optimization Framework to Enforce Multi-View Consistency for Texturing 3D Meshes; *Zhengyi Zhao; Chen Song; Xiaodong Gu; Yuan Dong; Qi Zuo; Weihao Yuan; Zilong Dong*; Liefeng Bo; Qixing Huang**

227. EMDM: Efficient Motion Diffusion Model for Fast, High-Quality Human Motion Generation; *Wenyang Zhou; Zhiyang Dou*; Zeyu Cao; Zhouyingcheng Liao; Jingbo Wang; Wenjia Wang; Yuan Liu; Taku Komura; Wenping Wang; Lingjie Liu*

228. StableDrag: Stable Dragging for Point-based Image Editing; *Yutao Cui; Xiaotong Zhao; Guozhen Zhang; Shengming Cao; Kai Ma; Limin Wang**

229. DeCo: Decoupled Human-Centered Diffusion Video Editing with Motion Consistency; *Xiaojing Zhong; Xinyi Huang; Xiaofeng Yang; Guosheng Lin*; Qingyao Wu**

230. Occlusion-Aware Seamless Segmentation; *Yihong Cao; Jiaming Zhang; Hao Shi; Kunyu Peng; Yuhongxuan Zhang; Hui Zhang*; Rainer Stiefelhagen; Kailun Yang**

231. IDOL: Unified Dual-Modal Latent Diffusion for Human-Centric Joint Video-Depth Generation; *Yuanhao Zhai*; Kevin Lin; Linjie Li; Chung-Ching Lin; Jianfeng Wang; Zhengyuan Yang; David Doermann; Junsong Yuan; Zicheng Liu; Lijuan Wang*

232. General and Task-Oriented Video Segmentation; *Mu Chen; Liulei Li; Wenguan Wang; Ruijie Quan; Yi Yang**

233. COIN: Control-Inpainting Diffusion Prior for Human and Camera Motion Estimation; *Jiefeng Li*; Ye Yuan; Davis Rempe; Haotian Zhang; Pavlo Molchanov; Cewu Lu; Jan Kautz; Umar Iqbal**

234. Long-term Temporal Context Gathering for Neural Video Compression; *Linfeng Qi; Zhaoyang Jia; Jiahao Li; Bin Li; Houqiang Li; Yan Lu**

235. Towards High-Quality 3D Motion Transfer with Realistic Apparel Animation; *Rong Wang*; Wei Mao; Changsheng Lu; HONGDONG LI*

236. S-JEPA: A Joint Embedding Predictive Architecture for Skeletal Action Recognition; *Mohamed Abdelfattah*; Alexandre Alahi*

237. Event-Based Motion Magnification; *Yutian Chen; Shi Guo*; Yu Fangzheng; Feng Zhang; Jinwei Gu; Tianfan Xue*

238. SkateFormer: Skeletal-Temporal Transformer for Human Action Recognition; *Jeonghyeok Do; Munchurl Kim**

239. Self-Supervised Video Copy Localization with Regional Token Representation; *Minlong Lu*; Yichen Lu; Siwei Nie; Xudong Yang; Xiaobo Zhang*

240. Long-Tail Temporal Action Segmentation with Group-wise Temporal Logit Adjustment; *Zhanzhong Pang*; Fadime Sener; Shrinivas Ramasubramanian; Angela Yao*

241. RGNet: A Unified Clip Retrieval and Grounding Network for Long Videos; *Tanveer Hannan*; Md Mohaiminul Islam; Thomas Seidl; Gedas Bertasius*

242. ZigMa: A DiT-style Zigzag Mamba Diffusion Model; *Vincent Tao Hu*; Stefan A Baumann; Ming Gui; Olga Grebenkova; Pingchuan Ma; Johannes S Fischer; Bjorn Ommer*

243. Temporal-Mapping Photography for Event Cameras; *Yuhan Bao; Lei Sun*; Yuqin Ma; Kaiwei Wang**

244. Motion Aware Event Representation-driven Image Deblurring; *Zhijing Sun; Xueyang Fu; Longzhuo Huang; Aiping Liu; Zheng-Jun Zha**

245. EDformer: Transformer-Based Event Denoising Across Varied Noise Levels; *Bin Jiang; Bo Xiong; Bohan Qu; M. Salman Asif; You Zhou*; Zhan Ma**

246. Free Lunch for Gait Recognition: A Novel Relation Descriptor; *Jilong Wang*; Saihui Hou; Yan Huang; Chunshui Cao; Xu Liu; Yongzhen Huang; Tianzhu Zhang; Liang Wang**

247. Interaction-centric Spatio-Temporal Context Reasoning for Multi-Person Video HOI Recognition; *Yisong Wang; Nan Xi*; Jingjing Meng; Junsong Yuan*

248. Spatio-Temporal Proximity-Aware Dual-Path Model for Panoramic Activity Recognition; *Sumin Lee*; Yooseung Wang; Sangmin Woo; Changick Kim*

249. Bidirectional Progressive Transformer for Interaction Intention Anticipation; *Zichen Zhang*; Hongchen Luo; Wei Zhai*; Yu Kang; Yang Cao*

250. Revisit Human-Scene Interaction via Space Occupancy; *Xinpeng Liu; Haowen Hou; Yanchao Yang; Yong-Lu Li*; Cewu Lu*

251. BlinkVision: A Benchmark for Optical Flow, Scene Flow and Point Tracking Estimation using RGB Frames and Events; *Yijin Li; Yichen Shen; Zhaoyang Huang; Shuo Chen; Weikang Bian; Xiaoyu Shi; Fu-Yun Wang; Keqiang Sun; Hujun Bao; Zhaopeng Cui; Guofeng Zhang*; Hongsheng Li**

252. LiDAR-Event Stereo Fusion with Hallucinations; *Luca Bartolomei*; Matteo Poggi; Andrea Conti; Stefano Mattoccia**

253. CoTracker: It is Better to Track Together; *Nikita Karaev*; Ignacio Rocco; Ben Graham; Natalia Neverova; Andrea Vedaldi; Christian Rupprecht*

254. RecurrentBEV: A Long-term Temporal Fusion Framework for Multi-view 3D Detection; *Ming Chang; Xishan Zhang*; Rui Zhang; Zhipeng Zhao; Guanhua He; Shaoli Liu*

255. JDT3D: Addressing the Gaps in LiDAR-Based Tracking-by-Attention; *Brian Cheong*; Jiachen Zhou*; Steven L Waslander**

256. PetFace: A Large-Scale Dataset and Benchmark for Animal Identification; *Risa Shinoda*; Kaede Shiohara*

257. Keypoint Promptable Re-Identification; *Vladimir Somers*; Alexandre Alahi; Christophe De Vleeschouwer*

258. Ray Denoising: Depth-aware Hard Negative Sampling for Multi-view 3D Object Detection; *Feng Liu*; Tengteng Huang; Qianjing Zhang; Haotian Yao; Chi Zhang; Fang Wan; Qixiang Ye; Yanzhao Zhou**

259. ARoFace: Alignment Robustness to Improve Low-quality Face Recognition; *Mohammad Saeed Ebrahimi Saadabadi*; Sahar Rahimi Malakshan; Ali Dabouei; Nasser Nasrabadi*

260. Close, But Not There: Boosting Geographic Distance Sensitivity in Visual Place Recognition; *Sergio Izquierdo*; Javier Civera**

261. Domesticating SAM for Breast Ultrasound Image Segmentation via Spatial-frequency Fusion and Uncertainty Correction; *Wanting Zhang; Huisi Wu*; Jing Qin*

262. VisionTrap: Vision-Augmented Trajectory Prediction Guided by Textual Descriptions; *Seokha Moon; Hyun Woo; Hongbeen Park; Haeji Jung; Reza Mahjourian; Hyung-gun Chi; Hyerin Lim; Sangpil Kim; Jinkyu Kim**

263. DriveDreamer: Towards Real-world-driven World Models for Autonomous Driving; *Xiaofeng Wang*; Zheng Zhu; Guan Huang; Chen Xinze; Jiagang Zhu; Jiwen Lu*

264. SLEDGE: Synthesizing Driving Environments with Generative Models and Rule-Based Traffic; *Kashyap Chitta*; Daniel Dauner; Andreas Geiger*

265. V2X-Real: a Largs-Scale Dataset for Vehicle-to-Everything Cooperative Perception; *Hao Xiang; Xin Xia; Zhaoliang Zheng; Runsheng Xu; Letian Gao; Zewei Zhou; xu han; Xinkai Ji; Mingxi Li; Zonglin Meng; Li Jin; Mingyue Lei; Zhaoyang Ma; Zihang He; Haoxuan Ma; Yunshuang Yuan; Yingqian Zhao; Jiaqi Ma**

266. Enhancing Vectorized Map Perception with Historical Rasterized Maps; *Xiaoyu Zhang; Guangwei Liu; Zihao Liu; Ningyi Xu; Yunhui Liu*; Ji Zhao*

267. Caltech Aerial RGB-Thermal Dataset in the Wild; *Connor Lee*; Matthew Anderson; Nikhil Ranganathan; Xingxing Zuo; Kevin T Do; Georgia Gkioxari; Soon-Jo Chung*

268. UAV First-Person Viewers Are Radiance Field Learners; *Liqi Yan*; Qifan Wang; Junhan Zhao; Qiang Guan; Zheng Tang; Jianhui Zhang; Dongfang Liu**

269. RoadPainter: Points Are Ideal Navigators for Topology transformER; *Zhongxing Ma; Liang Shuang; Yongkun Wen; Weixin Lu; Guowei Wan**

270. Adapting Fine-Grained Cross-View Localization to Areas without Fine Ground Truth; *Zimin Xia*; Yujiao Shi; Hongdong Li; Julian F. P. Kooij*

271. DA-BEV: Unsupervised Domain Adaptation for Bird's Eye View Perception; *Kai Jiang*; Jiaxing Huang; Weiying Xie; Jie Lei; Yunsong Li; Ling Shao; Shijian Lu*

272. CVT-Occ: Cost Volume Temporal Fusion for 3D Occupancy Prediction; *Zhangchen Ye; Tao Jiang; Chenfeng Xu; Yiming Li; Hang Zhao**

273. Detecting As Labeling: Rethinking LiDAR-camera Fusion in 3D Object Detection; *Junjie Huang*; Yun Ye; Zhujin Liang; Yi Shan; Dalong Du*

274. Camera Height Doesn't Change: Unsupervised Training for Metric Monocular Road-Scene Depth Estimation; *Genki Kinoshita*; Ko Nishino*

275. LabelDistill: Label-guided Cross-modal Knowledge Distillation for Camera-based 3D Object Detection; *Sanmin Kim; Youngseok Kim; Sihwan Hwang; Hyeonjun Jeong; Dongsuk Kum**

276. Diffusion Model is a Good Pose Estimator from 3D RF-Vision; *Junqiao Fan; Jianfei Yang*; Yuecong Xu; Lihua Xie*

277. MMVR: Millimeter-wave Multi-View Radar Dataset and Benchmark for Indoor Perception; *Mohammad Mahbubur Rahman; Ryoma Yataka; Sorachi Kato; Pu Wang*; Peizhao Li; Adriano Cardace; Petros Boufounos*

278. Dual-level Adaptive Self-Labeling for Novel Class Discovery in Point Cloud Segmentation; *Ruijie Xu*; CHUYU ZHANG; Hui Ren; Xuming He*

279. Brain-ID: Learning Contrast-agnostic Anatomical Representations for Brain Imaging; *Peirong Liu*; Oula Puonti; Xiaoling Hu; Daniel C. Alexander; Juan E. Iglesias*

280. Domain Generalization of 3D Object Detection by Density-Resampling; *Shuangzhi Li; Lei Ma; Xingyu Li**

281. Progressive Classifier and Feature Extractor Adaptation for Unsupervised Domain Adaptation on Point Clouds; *Zicheng Wang; Zhen Zhao; Yiming Wu; Luping Zhou*; Dong Xu**

282. Part2Object: Hierarchical Unsupervised 3D Instance Segmentation; *Cheng Shi; Yulin Zhang; Bin Yang; Jiajin Tang; Yuexin Ma; Sibei Yang**

283. AnatoMask: Enhancing Medical Image Segmentation with Reconstruction-guided Self-masking; *Yuheng Li; Tianyu Luan; Yizhou Wu; Shaoyan Pan; Yenho Chen; Xiaofeng Yang**

284. Representing Topological Self-Similarity Using Fractal Feature Maps for Accurate Segmentation of Tubular Structures; *Jiaxing Huang; Yanfeng Zhou; Yaoru Luo; Guole Liu; Heng Guo; Ge Yang**

285. Transferable 3D Adversarial Shape Completion using Diffusion Models; *Xuelong Dai*; Bin Xiao*

286. Fast Training of Diffusion Transformer with Extreme Masking for 3D Point Clouds Generation; *Shentong Mo; Enze Xie*; Yue Wu; Junsong Chen; Matthias Niessner; Zhenguo Li*

287. SpaceJAM: a Lightweight and Regularization-free Method for Fast Joint Alignment of Images; *Nir Barel*; Ron A Shapira Weber*; Nir Mualem; Shahaf E Finder; Oren Freifeld**

288. Task-Driven Uncertainty Quantification in Inverse Problems via Conformal Prediction; *Jeffrey Wen*; Rizwan Ahmad; Phillip Schniter*

289. Deep Diffusion Image Prior for Efficient OOD Adaptation in 3D Inverse Problems; *Hyungjin Chung; Jong Chul Ye**

290. Heterogeneous Graph Learning for Scene Graph Prediction in 3D Point Clouds; *Yanni Ma; Hao Liu; Yun Pei; Yulan Guo**

291. PointRegGPT: Boosting 3D Point Cloud Registration using Generative Point-Cloud Pairs for Training; *Suyi Chen; Hao Xu; Haipeng Li; Kunming Luo; Guanghui Liu; Chi-Wing Fu; Ping Tan; Shuaicheng Liu**

292. Reprojection Errors as Prompts for Efficient Scene Coordinate Regression; *Ting-Ru Liu*; Hsuan-Kung Yang; Jou-Min Liu; Chun-Wei Huang; Tsung-Chih Chiang; Quan Kong; Norimasa Kobori; Chun-Yi Lee*

293. Sparse Beats Dense: Rethinking Supervision in Radar-Camera Depth Completion; *Huadong Li; Minhao Jing; Jin Wang; Shichao Dong; Jiajun Liang; Haoqiang Fan; Renhe Ji**

294. VF-NeRF: Viewshed Fields for Rigid NeRF Registration; *Leo Segre*; Shai Avidan*

295. Improving 2D Feature Representations by 3D-Aware Fine-Tuning; *Yuanwen Yue*; Anurag Das; Francis Engelmann; Siyu Tang; Jan Eric Lenssen*

296. Omni6D: Large-Vocabulary 3D Object Dataset for Category-Level 6D Object Pose Estimation; *Mengchen Zhang*; Tong Wu; Tai Wang; Tengfei Wang; Ziwei Liu; Dahua Lin**

297. Unsupervised Variational Translator for Bridging Image Restoration and High-Level Vision Tasks; *Jiawei Wu; Zhi Jin**

298. 3D Congealing: 3D-Aware Image Alignment in the Wild; *Yunzhi Zhang*; Zizhang Li; Amit Raj; Andreas Engelhardt; Yuanzhen Li; Tingbo Hou; Jiajun Wu; Varun Jampani*

299. GraspXL: Generating Grasping Motions for Diverse Objects at Scale; *Hui Zhang*; Sammy Christen; Zicong Fan; Otmar Hilliges; Jie Song*

300. GS-Pose: Category-Level Object Pose Estimation via Geometric and Semantic Correspondence; *Pengyuan Wang*; Takuya Ikeda; Robert Lee; Koichi Nishiwaki*

301. 3D Reconstruction of Objects in Hands without Real World 3D Supervision; *Aditya Prakash*; Matthew Chang; Matthew Jin; Ruisen Tu; Saurabh Gupta*

302. Local Occupancy-Enhanced Object Grasping with Multiple Triplanar Projection; *Kangqi Ma*; Hao Dong; Yadong Mu*

303. Coarse-to-Fine Implicit Representation Learning for 3D Hand-Object Reconstruction from a Single RGB-D Image; *Xingyu Liu; Pengfei Ren; Jingyu Wang*; Qi Qi; Haifeng Sun; Zirui Zhuang*; Jianxin Liao*

304. Rotated Orthographic Projection for Self-Supervised 3D Human Pose Estimation; *YAO YAO; Yixuan Pan; Wenjun Shi; Dongchen Zhu; Lei Wang; Jiamao Li**

305. Weakly-Supervised 3D Hand Reconstruction with Knowledge Prior and Uncertainty Guidance; *Yufei Zhang*; Jeffrey Kephart; Qiang Ji**

306. MANIKIN: Biomechanically Accurate Neural Inverse Kinematics for Human Motion Estimation; *Jiaxi Jiang*; Paul Streli; Xuejing Luo; Christoph Gebhardt; Christian Holz*

307. Occlusion Handling in 3D Human Pose Estimation with Perturbed Positional Encoding; *Niloofar Azizi*; Mohsen Fayyaz; Horst Bischof*

308. COSMU: Complete 3D human shape from monocular unconstrained images; *Marco Pesavento*; Marco Volino; Adrian Hilton*

309. FastCAD: Real-Time CAD Retrieval and Alignment from Scans and Videos; *Florian Maximilian Langer*; Jihong Ju; Georgi Dikov; Gerhard Reitmayr; Mohsen Ghafoorian*

310. RGBD GS-ICP SLAM; *Seongbo Ha; Jiung Yeon; Hyeonwoo Yu**

311. Revisiting Calibration of Wide-Angle Radially Symmetric Cameras; *Andrea Porfiri Dal Cin*; Francesco Azzoni; Giacomo Boracchi; Luca Magri**

312. Neural Surface Detection for Unsigned Distance Fields; *Federico Stella*; Nicolas Talabot; Hieu Le; Pascal Fua*

313. Zero-Shot Multi-Object Scene Completion; *Shun Iwase*; Katherine Liu; Vitor Guizilini; Adrien Gaidon; Kris Kitani; Rareș A Ambruș; Sergey Zakharov*

314. Decomposition of Neural Discrete Representations for Large-Scale 3D Mapping; *Minseong Park; Suhan Woo; Euntai Kim**

315. HSR: Holistic 3D Human-Scene Reconstruction from Monocular Videos; *Lixin Xue*; Chen Guo; Chengwei Zheng; Fangjinhua Wang; Tianjian Jiang; Hsuan-I Ho; Manuel Kaufmann; Jie Song; Otmar Hilliges*

316. Learning Neural Deformation Representation for 4D Dynamic Shape Generation; *Gyojin Han*; Jiwan Hur; Jaehyun Choi; Junmo Kim**

317. NeuSDFusion: A Spatial-Aware Generative Model for 3D Shape Completion, Reconstruction, and Generation; *Ruikai Cui; Weizhe Liu*; Weixuan Sun; Senbo Wang; Taizhang Shang; Yang Li; Xibin Song; Han Yan; ZHENNAN WU; Shenzhou Chen; HONGDONG LI; Pan Ji*

318. MeshFeat: Multi-Resolution Features for Neural Fields on Meshes; *Mihir Mahajan*; Florian Hofherr*; Daniel Cremers*

319. CG-SLAM: Efficient Dense RGB-D SLAM in a Consistent Uncertainty-aware 3D Gaussian Field; *Jiarui Hu; Xianhao Chen; Boyin Feng; Guanglin Li; Liangjing Yang; Hujun Bao; Guofeng Zhang; Zhaopeng Cui**

320. Reconstruction and Simulation of Elastic Objects with Spring-Mass 3D Gaussians; *Licheng Zhong; Hong-Xing Yu; Jiajun Wu; Yunzhu Li**

321. Object-Aware NIR-to-Visible Translation; *Yunyi Gao; Lin Gu; Qiankun Liu; Ying Fu**

322. Physics-informed Knowledge Transfer for Underwater Monocular Depth Estimation; *Jinghe Yang*; Mingming Gong; Ye Pu*

323. SEDiff: Structure Extraction for Domain Adaptive Depth Estimation via Denoising Diffusion Models; *Dongseok Shim*; Hyoun Jin Kim**

324. Learning Representations from Foundation Models for Domain Generalized Stereo Matching; *Yongjian Zhang; Longguang Wang; Kunhong Li; WANG Yun; Yulan Guo**

325. MesonGS: Post-training Compression of 3D Gaussians via Efficient Attribute Transformation; *Shuzhao Xie*; Weixiang Zhang; Chen Tang; Yunpeng Bai; Rongwei Lu; Shjia Ge; Zhi Wang*

326. Towards Image Ambient Lighting Normalization; *Florin-Alexandru Vasluianu*; Tim Seizinger; Zongwei WU*; Rakesh Ranjan; Radu Timofte*

327. RANRAC: Robust Neural Scene Representations via Random Ray Consensus; *Benno Buschmann*; Andreea Dogaru; Elmar Eisemann; Michael Weinmann; Bernhard Egger*

328. SplatFields: Neural Gaussian Splats for Sparse 3D and 4D Reconstruction; *Marko Mihajlovic*; Sergey Prokudin; Siyu Tang; Robert Maier; Federica Bogo; Tony Tung; Edmond Boyer*

329. On the Error Analysis of 3D Gaussian Splatting and an Optimal Projection Strategy; *Letian Huang; Jiayang Bai; Jie Guo*; Yuanqi Li; Yanwen Guo*

330. CoR-GS: Sparse-View 3D Gaussian Splatting via Co-Regularization; *Jiawei Zhang; Jiahe Li; Xiaohan Yu; Lei Huang; Lin Gu; Jin Zheng*; Xiao Bai**

331. Efficient Snapshot Spectral Imaging: Calibration-Free Parallel Structure with Aperture Diffraction Fusion; *Tao Lv*; Lihao Hu; Shiqiao Li; Chenglong Huang; Xun Cao*

334. Revising Densification in Gaussian Splatting; *Samuel Rota Bulò*; Lorenzo Porzi; Peter Kontschieder*

335. Analysis-by-Synthesis Transformer for Single-View 3D Reconstruction; *Dian Jia; Xiaoqian Ruan; Kun Xia; Zhiming Zou; Le Wang; Wei Tang**

336. SparseCraft: Few-Shot Neural Reconstruction through Stereopsis Guided Geometric Linearization; *Mae Younes*; Amine Ouasfi; Adnane Boukhayma*

337. LaRa: Efficient Large-Baseline Radiance Fields; *Anpei Chen*; Haofei Xu; Stefano Esposito; Siyu Tang; Andreas Geiger*

338. Depth-guided NeRF Training via Earth Mover's Distance; *Anita Rau*; Josiah Aklilu; Floyd C Holsinger; Serena Yeung-Levy*

339. RoGUENeRF: A Robust Geometry-Consistent Universal Enhancer for NeRF; *Sibi Catley-Chandar*; Richard Shaw; Gregory Slabaugh; Eduardo Pérez Pellitero*

340. Physically Plausible Color Correction for Neural Radiance Fields; *Qi Zhang*; Ying Feng; HONGDONG LI**

341. Distractor-Free Novel View Synthesis via Exploiting Memorization Effect in Optimization; *Yukun Wang*; Kunhong Li; Minglin Chen; Longguang Wang; Shunbo Zhou; Kaiwen Xue; Yulan Guo**

342. Deblurring 3D Gaussian Splatting; *Byeonghyeon Lee*; Howoong Lee; Xiangyu Sun; Usman Ali; Eunbyung Park**

343. TriNeRFLet: A Wavelet Based Triplane NeRF Representation; *Rajaei Khatib*; Raja Giryes**

344. Volumetric Rendering with Baked Quadrature Fields; *Gopal Sharma*; Daniel Rebain; Kwang Moo Yi; Andrea Tagliasacchi*

12:00 – 13:30
Speed Mentoring - Space 4

12:30 – 13:30
Lunch - Exhibition Area (Level 0) & Balcony Level 1

13:30 – 15:30
Oral session 4A: Neural 3D rendering - Gold Room
Chairs: Martin R. Oswald; Gim Hee Lee
1. Generative Camera Dolly: Extreme Monocular Dynamic Novel View Synthesis; *Basile Van Hoorick*; Rundi Wu; Ege Ozguroglu; Kyle Sargent; Ruoshi Liu; Pavel Tokmakov; Achal Dave; Changxi Zheng; Carl Vondrick*

2. Gaussian Frosting: Editable Complex Radiance Fields with Real-Time Rendering; *Antoine Guédon*; Vincent Lepetit*

3. Analytic-Splatting: Anti-Aliased 3D Gaussian Splatting via Analytic Integration; *Zhihao Liang*; Qi Zhang*; Wenbo Hu; Ying Feng; Lei ZHU; Kui Jia**

4. FisherRF: Active View Selection and Mapping with Radiance Fields using Fisher Information; *Wen Jiang*; BOSHU LEI; Kostas Daniilidis**

5. RaFE: Generative Radiance Fields Restoration; *Zhongkai Wu; Ziyu Wan; Jing Zhang*; Jing Liao; Dong Xu*

6. Watch Your Steps: Local Image and Scene Editing by Text Instructions; *Ashkan Mirzaei*; Tristan T Aumentado-Armstrong; Marcus A Brubaker; Jonathan Kelly; Alex Levinshtein; Konstantinos G Derpanis; Igor Gilitschenski*

7. MVSplat: Efficient 3D Gaussian Splatting from Sparse Multi-View Images; *Yuedong Chen*; Haofei Xu; Chuanxia Zheng; Bohan Zhuang; Marc Pollefeys; Andreas Geiger; Tat-Jen Cham; Jianfei Cai*

8. RPBG: Towards Robust Neural Point-based Graphics in the Wild; *Qingtian Zhu; Zizhuang Wei; Zhongtian Zheng; Yifan Zhan; Zhuyu Yao; Jiawang Zhang; Kejian Wu; Yinqiang Zheng**

9. Omni-Recon: Harnessing Image-based Rendering for General-Purpose Neural Radiance Fields; *Yonggan Fu; Huaizhi Qu; Zhifan Ye; Chaojian Li; Kevin Zhao; Yingyan (Celine) Lin**

10. Learning 3D-aware GANs from Unposed Images with Template Feature Field; *Xinya Chen; Hanlei Guo; Yanrui Bin; Shangzhan Zhang; Yuanbo Yang; Yujun Shen; Yue Wang; Yiyi Liao**

11. MIGS: Multi-Identity Gaussian Splatting via Tensor Decomposition; *Aggelina Chatziagapi*; Grigorios Chrysos; Dimitris Samaras*

13:30 – 15:30
Oral session 4B: Video generation / editing / prediction - Auditorium
Chairs: Richard Zhang; Saining Xie
1. LEGO: Learning EGOcentric Action Frame Generation via Visual Instruction Tuning; *Bolin Lai*; Xiaoliang Dai; Lawrence Chen; Guan Pang; James M Rehg; Miao Liu* BEST PAPER CANDIDATE

2. SV3D: Novel Multi-view Synthesis and 3D Generation from a Single Image using Latent Video Diffusion; *Vikram Voleti*; Chun-Han Yao; Mark Boss; Adam Letts; David Pankratz; Dmitrii Tochilkin; Christian Laforte; Robin Rombach; Varun Jampani**

3. Efficient Neural Video Representation with Temporally Coherent Modulation; *Seungjun Shin*; Suji Kim*; Dokwan Oh*

4. Clearer Frames, Anytime: Resolving Velocity Ambiguity in Video Frame Interpolation; *Zhihang Zhong; Gurunandan Krishnan; Xiao Sun; Yu Qiao; Sizhuo Ma*; Jian Wang**

5. Video Editing via Factorized Diffusion Distillation; *Uriel Singer*; Amit Zohar*; Yuval Kirstain; Shelly Sheynin; Adam Polyak; Devi Parikh; Yaniv Taigman*

6. ReSyncer: Rewiring Style-based Generator for Unified Audio-Visually Synced Facial Performer; *Jiazhi Guan*; Zhiliang Xu; Hang Zhou; Kaisiyuan Wang; Shengyi He; Zhanwang Zhang; Borong Liang; Haocheng Feng; Errui Ding; Jingtuo Liu; Jingdong Wang; Youjian Zhao; Ziwei Liu*

7. Audio-Synchronized Visual Animation; *Lin Zhang; Shentong Mo; Yijing Zhang; Pedro Morgado**

8. DynamiCrafter: Animating Open-domain Images with Video Diffusion Priors; *Jinbo Xing*; Menghan Xia; Yong Zhang; Haoxin Chen; Wangbo Yu; Hanyuan Liu; Gongye Liu; Xintao Wang; Ying Shan; Tien-Tsin Wong*

9. MotionDirector: Motion Customization of Text-to-Video Diffusion Models; *Rui Zhao; Yuchao Gu; Jay Zhangjie Wu; David Junhao Zhang; Jia-Wei Liu; weijia wu; Jussi Keppo; Mike Zheng Shou**

10. ZoLA: Zero-Shot Creative Long Animation Generation with Short Video Model; *Fu-Yun Wang*; Zhaoyang Huang*; Qiang Ma; Guanglu Song; Xudong LU; Weikang Bian; Yijin Li; Yu Liu; Hongsheng Li**

11. Temporal Residual Guided Diffusion Framework for Event-Driven Video Reconstruction; *Lin Zhu*; Yunlong Zheng; Yijun Zhang; Xiao Wang; Lizhi Wang; Hua Huang*

13:30 – 15:30
Oral session 4C: Humans: Biometrics, pose and motion - Silver Room
Chairs: Georgios Pavlakos; Federica Bogo
1. AttentionHand: Text-driven Controllable Hand Image Generation for 3D Hand Reconstruction in the Wild; *Junho Park; Kyeongbo Kong; Suk-Ju Kang**

2. Sapiens: Foundation for Human Vision Models; *Rawal Khirodkar*; Timur Bagautdinov; jltmtzc@gmail.com Martinez; Zhaoen Su; Austin T James; Peter Selednik; Stuart Anderson; Shunsuke Saito* BEST PAPER CANDIDATE

3. POET: Prompt Offset Tuning for Continual Human Action Adaptation; *Prachi Garg*; Joseph K J; Vineeth N Balasubramanian; Necati Cihan Camgoz; Chengde Wan; Kenrick Kin; Weiguang Si; Shugao Ma; Fernando de la Torre*

4. Harnessing Text-to-Image Diffusion Models for Category-Agnostic Pose Estimation; *Duo Peng; Zhengbo Zhang; Ping Hu; Qiuhong Ke; David Yau; Jun Liu**

5. SemGrasp: Semantic Grasp Generation via Language Aligned Discretization; *Kailin Li*; Jingbo Wang; Lixin Yang; Cewu Lu*; Bo Dai*

6. UGG: Unified Generative Grasping; *Jiaxin Lu; Hao Kang; Haoxiang Li; Bo Liu; Yiding Yang; Qixing Huang; Gang Hua**

7. NL2Contact: Natural Language Guided 3D Hand-Object Contact Modeling with Diffusion Model; *Zhongqun Zhang*; Hengfei Wang; Ziwei Yu; Yihua Cheng*; Angela Yao; Hyung Jin Chang*

8. Beyond the Contact: Discovering Comprehensive Affordance for 3D Objects from Pre-trained 2D Diffusion Models; *Hyeonwoo Kim; Sookwan Han; Patrick Kwon; Hanbyul Joo**

9. LiveHPS++: Robust and Coherent Motion Capture in Dynamic Free Environment; *Yiming Ren; Xiao Han; Yichen Yao; Xiaoxiao Long; Yujing Sun*; Yuexin Ma**

10. Controllable Human-Object Interaction Synthesis; *Jiaman Li*; Alexander Clegg; Roozbeh Mottaghi; Jiajun Wu; Xavier Puig; C. Karen Liu*

11. NeRMo: Learning Implicit Neural Representations for 3D Human Motion Prediction; *Dong Wei; Huaijiang Sun; Xiaoning Sun*; Shengxiang Hu*

14:30 – 18:00
Demo session 4 - Level 0
1. Live Demo of Matching and Dense 3D Reconstruction with MASt3R; *Vincent Leroy, Yohann Cabon, Jerome Revaud - Naver Labs Europe*

2. ON-STAGE 3D: Link-based Investigation into Spatial Iconographic Heritage; *Emile Blettery, Valérie Gouet-Brunet, Livio de Luca - Institut national de l'information géographique et forestière - LaSTIG / CNRS - MAP*

3. Controllable Neural Reconstruction for Autonomous Driving; *Máté Tóth, Péter Kovács, Zoltán Bendefy, Zoltán Hortsin, Tamás Matuszka - AImotive*

4. Spiky DVS Piano; *Muhammad Aitsam, Gaurvi Goyal, Chiara Bartolozzi, Alessandro di Nuovo - Sheffield Hallam University*

5. Automating Parasite Egg Detection: Artificial Intelligence based Kubic FLOTAC microscope (KFM); *Antonio Bosco, Salvatore Capuozzo, Giuseppe Cringoli, Michela Gravina, Stefano Marrone, Maria Paola Maurelli, Laura Rinaldi, Carlo Sansone - University of Napoli Federico II*

15:30 – 16:30
Keynote Lecture - Gold Room (live), Auditorium (broadcast), Silver Room (broadcast)
Fair, transparent, and accountable AI: What is legally required, what is ethically desired, and what is technically feasible?; *Sandra Wachter*

16:30 – 17:00
Weight & Biases Technical Session - Technical Presentation Area (Level 0)
Reproducible ML: Tooling is your friend

16:30 – 17:00
Coffee Break - Exhibition Area (Level 0)

16:30 – 18:30
Poster session 4
1. A Secure Image Watermarking Framework with Statistical Guarantees via Adversarial Attacks on Secret Key Networks; *Feiyu CHEN*; Wei Lin; Ziquan Liu; Antoni Chan*

2. Efficient Training of Spiking Neural Networks with Multi-Parallel Implicit Stream Architecture; *Zhigao Cao; Meng Li; Xiashuang Wang; Haoyu Wang; Fan Wang; Youjun Li; Zigang Huang**

3. On the Vulnerability of Skip Connections to Model Inversion Attacks; *Jun Hao Koh*; Sy-Tuyen Ho; Ngoc-Bao Nguyen; Ngai-Man Cheung*

4. Clean & Compact: Efficient Data-Free Backdoor Defense with Model Compactness; *Huy Phan*; Jinqi Xiao; Yang Sui; Tianfang Zhang; Zijie Tang; Cong Shi; Yan Wang; Yingying Chen; Bo Yuan*

5. Non-transferable Pruning; *Ruyi Ding*; Lili Su; A. Adam Ding; Yunsi Fei*

6. Cross-Input Certified Training for Universal Perturbations; *Changming Xu*; Gagandeep Singh*

7. Interpretability-Guided Test-Time Adversarial Defense; *Akshay Kulkarni*; Tsui-Wei Weng*

8. Idling Neurons, Appropriately Lenient Workload During Fine-tuning Leads to Better Generalization; *Hongjing Niu*; Hanting Li; Bin Li; Feng Zhao**

9. DεpS: Delayed ε-Shrinking for Faster Once-For-All Training; *Aditya Annavajjala*; Alind Khare*; Animesh Agrawal; Igor Fedorov; Hugo M Latapie; Myungjin Lee; Alexey Tumanov*

10. Straightforward Layer-wise Pruning for More Efficient Visual Adaptation; *Ruizi Han*; Jinglei Tang**

11. Dataset Quantization with Active Learning based Adaptive Sampling; *Zhenghao Zhao*; Yuzhang Shang; Junyi Wu; Yan Yan*

12. Auto-DAS: Automated Proxy Discovery for Training-free Distillation-aware Architecture Search; *Haosen Sun; Lujun Li*; Peijie Dong; Zimian Wei; Shitong Shao*

13. Local and Global Flatness for Federated Domain Generalization; *Hao Yan; Yuhong Guo**

14. Beta-Tuned Timestep Diffusion Model; *Tianyi Zheng*; Peng-Tao Jiang; Ben Wan; Hao Zhang; Jinwei Chen; Jia Wang*; Bo Li**

15. PILoRA: Prototype Guided Incremental LoRA for Federated Class-Incremental Learning; *Haiyang Guo*; Fei Zhu; Wenzhuo Liu; Xu-Yao Zhang*; Cheng-Lin Liu*

16. Exploring Guided Sampling of Conditional GANs; *Yifei Zhang*; Mengfei Xia; Yujun Shen; Jiapeng Zhu; Ceyuan Yang; Kecheng Zheng; Lianghua Huang; Yu Liu; Fan Cheng**

17. Self-Rectifying Diffusion Sampling with Perturbed-Attention Guidance; *Donghoon Ahn; Hyoungwon Cho; Jaewon Min; Jungwoo Kim; Wooseok Jang; SeonHwa Kim; Hyun Hee Park; Kyong Hwan Jin*; Seungryong Kim**

18. DiffClass: Diffusion-Based Class Incremental Learning; *Zichong Meng; Jie Zhang; Changdi Yang; Zheng Zhan; Pu Zhao*; Yanzhi Wang**

19. Direct Distillation between Different Domains; *Jialiang Tang; Shuo Chen*; Gang Niu; Hongyuan Zhu; Joey Tianyi Zhou; Chen Gong*; Masashi Sugiyama*

20. How to Train the Teacher Model for Effective Knowledge Distillation; *Shayan Mohajer Hamidi*; Xizhen Deng; Renhao Tan; Linfeng Ye; Ahmed Hussein Salamah*

21. Is Retain Set All You Need in Machine Unlearning? Restoring Performance of Unlearned Models with Out-Of-Distribution Images; *Jacopo Bonato*; Marco Cotogni; Luigi Sabetta**

22. Is user feedback always informative? Retrieval Latent Defending for Semi-Supervised Domain Adaptation without Source Data; *Junha Song*; Tae Soo Kim; Junha Kim; Gunhee Nam; Thijs Kooi; Jaegul Choo**

23. MemBN: Robust Test-Time Adaptation via Batch Norm with Statistics Memory; *Juwon Kang*; Nayeong Kim; Jungseul Ok; Suha Kwak**

24. PromptFusion: Decoupling Stability and Plasticity for Continual Learning; *Haoran Chen; Zuxuan Wu*; Xintong Han; Menglin Jia; Yu-Gang Jiang*

25. Dual-Path Adversarial Lifting for Domain Shift Correction in Online Test-time Adaptation; *Yushun Tang; Shuoshuo Chen; Zhihe Lu; Xinchao Wang; Zhihai He**

26. Cs2K: Class-specific and Class-shared Knowledge Guidance for Incremental Semantic Segmentation; *Wei Cong*; Yang Cong; Yuyang Liu; Gan Sun*

27. Strike a Balance in Continual Panoptic Segmentation; *Jinpeng Chen; Runmin Cong*; Yuxuan Luo; Horace Ho Shing Ip; Sam Kwong**

28. HVCLIP: High-dimensional Vector in CLIP for Unsupervised Domain Adaptation; *Noranart Vesdapunt*; Kah Kuen Fu; Yue Wu; Xu Zhang; Pradeep Natarajan*

29. Learning from the Web: Language Drives Weakly-Supervised Incremental Learning for Semantic Segmentation; *Chang Liu; Giulia Rizzoli; Pietro Zanuttigh; Fu Li; Yi Niu**

30. Select and Distill: Selective Dual-Teacher Knowledge Transfer for Continual Learning on Vision-Language Models; *Yu-Chu Yu*; Chi-Pin Huang; Jr-Jen Chen; Kai-Po Chang; Yung-Hsuan Lai; Fu-En Yang; Yu-Chiang Frank Wang*

31. SAFT: Towards Out-of-Distribution Generalization in Fine-Tuning; *Bac Nguyen*; Stefan Uhlich; Fabien Cardinaux; Lukas Mauch; Marzieh Edraki; Aaron Courville*

32. Gradient-Aware for Class-Imbalanced Semi-supervised Medical Image Segmentation; *Wenbo Qi; Jiafei Wu*; S. C. Chan**

33. Tendency-driven Mutual Exclusivity for Weakly Supervised Incremental Semantic Segmentation; *Chongjie Si; Xuehui Wang; Xiaokang Yang; Wei Shen**

34. Dual-Decoupling Learning and Metric-Adaptive Thresholding for Semi-Supervised Multi-Label Learning; *Jia-Hao Xiao; Ming-Kun Xie; Heng-Bo Fan; Gang Niu; Masashi Sugiyama; Sheng-Jun Huang**

35. Robust Multimodal Learning via Representation Decoupling; *Shicai Wei; Yang Luo; Yuji Wang; Chunbo Luo**

36. Adapt without Forgetting: Distill Proximity from Dual Teachers in Vision-Language Models; *Mengyu Zheng*; Yehui Tang; Zhiwei Hao; Kai Han; Yunhe Wang; Chang Xu**

37. IGNORE: Information Gap-based False Negative Loss Rejection for Single Positive Multi-Label Learning; *Gyeong Ryeol Song; Noo-ri Kim; Jin-Seop Lee; Jee-Hyong Lee**

38. Instance-dependent Noisy-label Learning with Graphical Model Based Noise-rate Estimation; *Arpit Garg*; Cuong Cao Nguyen; RAFAEL FELIX; Thanh-Toan Do; Gustavo Carneiro*

39. Image-Feature Weak-to-Strong Consistency: An Enhanced Paradigm for Semi-Supervised Learning; *Zhiyu Wu*; Jinshi Cui**

40. Understanding and Mitigating Human-Labelling Errors in Supervised Contrastive Learning; *Zijun Long*; Lipeng Zhuang; George W Killick; Richard Mccreadie; Gerardo Aragon-Camarasa; Paul Henderson*

41. Learning to Distinguish Samples for Generalized Category Discovery; *Fengxiang Yang; Nan Pu; Wenjing Li; Zhiming Luo*; Shaozi Li; Nicu Sebe; Zhun Zhong**

42. SUMix: Mixup with Semantic and Uncertain Information; *Huafeng Qin; Xin Jin*; Hongyu Zhu; Hongchao Liao; Mounim A. El Yacoubi; Xinbo Gao*

43. MetaAT: Active Testing for Label-Efficient Evaluation of Dense Recognition Tasks; *Sanbao Su; Xin Li*; Thang Doan; Sima Behpour; Wenbin He; Liang Gou; Fei Miao; Liu Ren*

44. Simplifying Source-Free Domain Adaptation for Object Detection: Effective Self-Training Strategies and Performance Insights; *Yan Hao; Florent Forest*; Olga Fink*

45. CamoTeacher: Dual-Rotation Consistency Learning for Semi-Supervised Camouflaged Object Detection; *xunfa lai; Zhiyu Yang; Jie Hu; ShengChuan Zhang*; Liujuan Cao; Guannan Jiang; Songan Zhang; zhiyu wang; Rongrong Ji*

46. MetaAug: Meta-Data Augmentation for Post-Training Quantization; *Cuong Van Pham*; Hoang Anh Dung; Cuong Cao Nguyen; Trung Le; Dinh Q Phung; Gustavo Carneiro; Thanh-Toan Do*

47. HyTAS: A Hyperspectral Image Transformer Architecture Search Benchmark and Analysis; *Fangqin Zhou*; Mert Kilickaya; Joaquin Vanschoren; Ran Piao*

48. Stitched ViTs are Flexible Vision Backbones; *Zizheng Pan*; Jing Liu; Haoyu He; Jianfei Cai; Bohan Zhuang**

49. SpecFormer: Guarding Vision Transformer Robustness via Maximum Singular Value Penalization; *Xixu Hu; Runkai Zheng; Jindong Wang*; Cheuk Hang Leung; Qi Wu*; Xing Xie*

50. One-stage Prompt-based Continual Learning; *Youngeun Kim*; Yuhang Li; Priyadarshini Panda*

51. AMD: Automatic Multi-step Distillation of Large-scale Vision Models; *Cheng Han; Qifan Wang; Sohail A Dianat; Majid Rabbani; Raghuveer Rao; Yi Fang; Qiang Guan; Lifu Huang; Dongfang Liu**

52. Enhancing Tracking Robustness with Auxiliary Adversarial Defense Networks; *Zhewei Wu; Ruilong Yu; Qihe Liu*; Shuying Cheng; Shilin Qiu; Shijie Zhou*

53. Self-Supervised Representation Learning for Adversarial Attack Detection; *Yi Li*; Plamen Angelov; Neeraj Suri*

54. SeiT++: Masked Token Modeling Improves Storage-efficient Training; *Minhyun Lee; Song Park; Byeongho Heo; Dongyoon Han; Hyunjung Shim**

55. Real Appearance Modeling for More General Deepfake Detection; *Jiahe Tian; Cai Yu; Xi Wang; Peng Chen; Zihao Xiao; Jiao Dai; Yesheng Chai*; Jizhong Han*

56. DECIDER: Leveraging Foundation Model Priors for Improved Model Failure Detection and Explanation; *Rakshith Subramanyam*; Kowshik Thopalli*; Vivek Sivaraman Narayanaswamy; Jayaraman J. Thiagarajan*

57. Unlocking Attributes' Contribution to Successful Camouflage: A Combined Textual and Visual Analysis Strategy; *Hong Zhang; Yixuan Lyu; Qian Yu; Hanyang Liu; Huimin Ma; Yuan Ding; Yifan Yang**

58. LAPT: Label-driven Automated Prompt Tuning for OOD Detection with Vision-Language Models; *Yabin Zhang*; Wenjie Zhu; Chenhang He; Lei Zhang**

59. Self-supervised Feature Adaptation for 3D Industrial Anomaly Detection; *Yuanpeng Tu; Boshen Zhang; Liang Liu; YUXI LI; Jiangning Zhang; Yabiao Wang*; Chengjie Wang; cairong zhao**

60. MC-PanDA: Mask Confidence for Panoptic Domain Adaptation; *Ivan Martinović*; Josip Šarić; Siniša Šegvić*

61. Global Counterfactual Directions; *Bartłomiej Sobieski*; Przemyslaw Biecek**

62. Linking in Style: Understanding learned features in deep learning models; *Maren Wehrheim*; Pamela Osuna Vargas; Matthias Kaschube*

63. On Spectral Properties of Gradient-based Explanation Methods; *Amir Mehrpanah*; Erik Englesson; Hossein Azizpour*

64. Quantized Prompt for Efficient Generalization of Vision-Language Models; *Tianxiang Hao; Xiaohan Ding*; Juexiao Feng; Yuhong Yang; Hui Chen; Guiguang Ding**

65. Bottom-Up Domain Prompt Tuning for Generalized Face Anti-Spoofing; *Siqi Liu*; Qirui Wang; Pong C. Yuen*

66. ItTakesTwo: Leveraging Peer Representations for Semi-supervised LiDAR Semantic Segmentation; *Yuyuan Liu*; Yuanhong Chen; Hu Wang; Vasileios Belagiannis; Ian Reid; Gustavo Carneiro*

67. Constructing Concept-based Models to Mitigate Spurious Correlations with Minimal Human Effort; *Jeeyung Kim*; Ze Wang; Qiang Qiu*

68. On-the-fly Category Discovery for LiDAR Semantic Segmentation; *Hyeonseong Kim; Sung-Hoon Yoon; Minseok Kim; Kuk-Jin Yoon**

69. Open Vocabulary 3D Scene Understanding via Geometry Guided Self-Distillation; *Pengfei Wang*; Yuxi Wang; Shuai Li; Zhaoxiang Zhang; Zhen Lei; Lei Zhang*

70. Bridging Synthetic and Real Worlds for Pre-training Scene Text Detectors; *Tongkun Guan; Wei Shen*; Xue Yang; Xuehui Wang; Xiaokang Yang*

71. Global-Local Collaborative Inference with LLM for Lidar-Based Open-Vocabulary Detection; *Xingyu Peng; Yan Bai; Chen Gao; Lirong Yang; Fei Xia; Beipeng Mu; Xiaofei Wang; Si Liu**

72. CLIP-DINOiser: Teaching CLIP a few DINO tricks for open-vocabulary semantic segmentation; *Monika Wysoczańska*; Oriane Siméoni; Michaël Ramamonjisoa; Andrei Bursuc; Tomasz Trzciński; Patrick Pérez*

73. Harnessing Text-to-Image Diffusion Models for Category-Agnostic Pose Estimation; *Duo Peng; Zhengbo Zhang; Ping Hu; Qiuhong Ke; David Yau; Jun Liu**

74. VeCLIP: Improving CLIP Training via Visual-enriched Captions; *Zhengfeng Lai*; Haotian Zhang; Bowen Zhang; Wentao Wu; Haoping Bai; Aleksei Timofeev; Xianzhi Du; Zhe Gan; Jiulong Shan; Chen-Nee Chuah; Yinfei Yang; Meng Cao*

75. Unveiling Typographic Deceptions: Insights of the Typographic Vulnerability in Large Vision-Language Models; *Hao Cheng; Erjia Xiao; Jindong Gu; Le Yang; Jinhao Duan; Jize Zhang; Jiahang Cao; Kaidi Xu; Renjing Xu**

76. Exploring Conditional Multi-Modal Prompts for Zero-shot HOI Detection; *Ting Lei; Shaofeng Yin; Yuxin Peng; Yang Liu**

77. SCLIP: Rethinking Self-Attention for Dense Vision-Language Inference; *Feng Wang*; Jieru Mei; Alan Yuille*

78. Eyes Closed, Safety On: Protecting Multimodal LLMs via Image-to-Text Transformation; *Yunhao Gou*; Kai Chen; Zhili LIU; Lanqing Hong; Hang Xu; Zhenguo Li; Dit-Yan Yeung; James Kwok; Yu Zhang**

79. ControlCap: Controllable Region-level Captioning; *Yuzhong Zhao; Liu Yue; Zonghao Guo; weijia wu; Chen Gong; Qixiang Ye; Fang Wan**

80. OLAF: A Plug-and-Play Framework for Enhanced Multi-object Multi-part Scene Parsing; *Pranav Gupta*; Rishubh Singh; Pradeep Shenoy; Ravi Kiran Sarvadevabhatla**

81. Scene-Graph ViT: End-to-End Open-Vocabulary Visual Relationship Detection; *Tim Salzmann; Markus Ryll; Alex Bewley; Matthias Minderer**

82. Textual Query-Driven Mask Transformer for Domain Generalized Segmentation; *Byeonghyun Pak; Byeongju Woo; Sunghwan Kim; Dae-hwan Kim; Hoseong Kim**

83. Multi-Granularity Sparse Relationship Matrix Prediction Network for End-to-End Scene Graph Generation; *lei wang; Zejian Yuan; Badong Chen**

84. ViGoR: Improving Visual Grounding of Large Vision Language Models with Fine-Grained Reward Modeling; *Siming Yan*; Min Bai; Weifeng Chen; Xiong Zhou; Qixing Huang; Li Erran Li*

85. MoAI: Mixture of All Intelligence for Large Language and Vision Models; *Byung-Kwan Lee; Beomchan Park; Chae Won Kim; Yong Man Ro**

86. LoA-Trans: Enhancing Visual Grounding by Location-Aware Transformers; *Ziling Huang*; Shin'ichi Satoh*

87. Uni3DL: A Unified Model for 3D Vision-Language Understanding; *Xiang Li*; Jian Ding; Zhaoyang Chen; Mohamed Elhoseiny*

88. CONDA: Condensed Deep Association Learning for Co-Salient Object Detection.; *Long Li; Nian Liu*; Dingwen Zhang; Zhongyu Li; Salman Khan; Rao Anwer; Hisham Cholakkal; Junwei Han*; Fahad Shahbaz Khan*

89. VisFocus: Prompt-Guided Vision Encoders for OCR-Free Dense Document Understanding; *Ofir Abramovich*; Niv Nayman*; Sharon Fogel; Inbal Lavi; Ron Litman; Shahar Tsiper; Royee Tichauer; Srikar Appalaraju; Shai Mazor; R. Manmatha*

90. ChEX: Interactive Localization and Region Description in Chest X-rays; *Philip Müller*; Georgios Kaissis; Daniel Rueckert*

91. WSI-VQA: Interpreting Whole Slide Images by Generative Visual Question Answering; *Pingyi Chen*; Chenglu Zhu; Sunyi Zheng; Honglin Li; Lin Yang**

92. X-InstructBLIP: A Framework for Aligning Image, 3D, Audio, Video to LLMs and its Emergent Cross-modal Reasoning; *Artemis Panagopoulou*; Le Xue; Ning Yu; LI JUNNAN; DONGXU LI; Shafiq Joty; Ran Xu; Silvio Savarese; Caiming Xiong; Juan Carlos Niebles*

93. ScanReason: Empowering 3D Visual Grounding with Reasoning Capabilities; *Chenming Zhu; Tai Wang; Wenwei Zhang; Kai Chen; Xihui Liu**

94. Training A Small Emotional Vision Language Model for Visual Art Comprehension; *Jing Zhang; Liang Zheng*; Meng Wang; Dan Guo**

95. Attention Decomposition for Cross-Domain Semantic Segmentation; *Liqiang He*; Sinisa Todorovic*

96. SpatialFormer: Towards Generalizable Vision Transformers with Explicit Spatial Understanding; *Han Xiao; Wenzhao Zheng; Sicheng Zuo; Peng Gao; Jie Zhou; Jiwen Lu**

97. Compositional Substitutivity of Visual Reasoning for Visual Question Answering; *Chuanhao Li; Zhen Li; Chenchen Jing*; Yuwei Wu*; Mingliang Zhai; Yunde Jia*

98. The All-Seeing Project V2: Towards General Relation Comprehension of the Open World; *Weiyun Wang; yiming ren; Haowen Luo; Tiantong Li; Chenxiang Yan; Zhe Chen; Wenhai Wang; Qingyun Li; Lewei Lu; Xizhou Zhu; Yu Qiao; Jifeng Dai**

99. Finding Visual Task Vectors; *Alberto Hojel*; Yutong Bai; Trevor Darrell; Amir Globerson; Amir Bar**

100. A Simple Latent Diffusion Approach for Panoptic Segmentation and Mask Inpainting; *Wouter Van Gansbeke*; Bert De Brabandere*

101. ControlLLM: Augment Language Models with Tools by Searching on Graphs; *Zhaoyang Liu; Zeqiang Lai; Zhangwei Gao; erfei cui; Ziheng Li; Xizhou Zhu; Lewei Lu; Qifeng Chen*; Yu Qiao; Jifeng Dai; Wenhai Wang**

102. Watching it in Dark: A Target-aware Representation Learning Framework for High-Level Vision Tasks in Low Illumination; *Yunan Li*; Yihao Zhang; Shoude Li; Long Tian; DOU QUAN; Chaoneng Li; Qiguang Miao**

103. Causality-inspired Discriminative Feature Learning in Triple Domains for Gait Recognition; *Haijun Xiong; Bin Feng*; Xinggang Wang; Wenyu Liu*

104. Rethinking Features-Fused-Pyramid-Neck for Object Detection; *Hulin Li**

105. SCAPE: A Simple and Strong Category-Agnostic Pose Estimator; *Yujia Liang; Zixuan Ye; Wenze Liu; Hao Lu**

106. Open-Set Biometrics: Beyond Good Closed-Set Models; *Yiyang Su; Minchul Kim; Feng Liu; Anil Jain; Xiaoming Liu**

107. SLAck: Semantic, Location, and Appearance Aware Open-Vocabulary Tracking; *Siyuan Li*; Lei Ke; Yung-Hsu Yang; Luigi Piccinelli; Mattia Segù; Martin Danelljan; Luc Van Gool*

108. General Geometry-aware Weakly Supervised 3D Object Detection; *Guowen Zhang*; Junsong Fan; Liyi Chen; Zhaoxiang Zhang; Zhen Lei; Lei Zhang*

109. Domain Shifting: A Generalized Solution for Heterogeneous Cross-Modality Person Re-Identification; *Yan Jiang; Xu Cheng*; Hao Yu; Xingyu Liu; Haoyu Chen; Guoying Zhao*

110. 3D Small Object Detection with Dynamic Spatial Pruning; *Zhihao Sun; Ziwei Wang; Hongmin Liu; Jie Zhou; Jiwen Lu*; Xiuwei Xu**

111. VLAD-BuFF: Burst-aware Fast Feature Aggregation for Visual Place Recognition; *Ahmad Khaliq; Ming Xu; Stephen Hausler; Michael J Milford; Sourav Garg**

112. PCF-Lift: Panoptic Lifting by Probabilistic Contrastive Fusion; *Runsong Zhu*; Shi Qiu*; Qianyi Wu; Ka-Hei Hui; Pheng-Ann Heng; Chi-Wing Fu*

113. Sapiens: Foundation for Human Vision Models; *Rawal Khirodkar*; Timur Bagautdinov; jltmtzc@gmail.com Martinez; Zhaoen Su; Austin T James; Peter Selednik; Stuart Anderson; Shunsuke Saito* BEST PAPER CANDIDATE

114. LaPose: Laplacian Mixture Shape Modeling for RGB-Based Category-Level Object Pose Estimation; *Ruida Zhang; Ziqin Huang; Gu Wang; Chenyangguang Zhang; Yan Di; Xingxing Zuo; Jiwen Tang; Xiangyang Ji**

115. Camera-LiDAR Cross-modality Gait Recognition; *Wenxuan Guo*; Yingping Liang; Zhiyu Pan; Ziheng Xi; Jianjiang Feng; Jie Zhou*

116. SceneGraphLoc: Cross-Modal Coarse Visual Localization on 3D Scene Graphs; *Yang Miao; Francis Engelmann; Olga Vysotska; Federico Tombari; Marc Pollefeys; Daniel Barath**

117. Mask as Supervision: Leveraging Unified Mask Information for Unsupervised 3D Pose Estimation; *Yuchen Yang; Yu Qiao; Xiao Sun**

118. WorldPose: A World Cup Dataset for Global 3D Human Pose Estimation; *Tianjian Jiang*; Johsan Billingham; Sebastian Müksch; Juan J Zarate; Nicolas Evans; Martin R. Oswald; Marc Pollefeys; Otmar Hilliges; Manuel Kaufmann; Jie Song*

119. Shape2Scene: 3D Scene Representation Learning Through Pre-training on Shape Data; *Tuo Feng; Wenguan Wang; Ruijie Quan; Yi Yang**

120. GPSFormer: A Global Perception and Local Structure Fitting-based Transformer for Point Cloud Understanding; *Changshuo Wang*; Meiqing Wu; Siew-Kei Lam; Xin Ning; Shangshu Yu; Ruiping Wang; Weijun Li; Thambipillai Srikanthan*

121. SCPNet: Unsupervised Cross-modal Homography Estimation via Intra-modal Self-supervised Learning; *Runmin Zhang*; Jun Ma; Lun Luo; Beinan Yu; Shu-Jie Chen; Junwei Li; Hui-Liang Shen; Si-Yuan Cao**

122. Frugal 3D Point Cloud Model Training via Progressive Near Point Filtering and Fused Aggregation; *Donghyun Lee; Yejin Lee; Jae W. Lee*; Hongil Yoon**

123. DynoSurf: Neural Deformation-based Temporally Consistent Dynamic Surface Reconstruction; *Yuxin Yao; Siyu Ren; Junhui Hou*; Zhi Deng; Juyong Zhang; Wenping Wang*

124. FLAT: Flux-aware Imperceptible Adversarial Attacks on 3D Point Clouds; *Keke Tang; Lujie Huang; Weilong Peng*; Daizong Liu; Xiaofei Wang; Yang Ma; Ligang Liu; Zhihong Tian*

125. Skeleton Recall Loss for Connectivity Conserving and Resource Efficient Segmentation of Thin Tubular Structures; *Yannick Kirchhoff*; Maximilian R Rokuss*; Saikat Roy*; Balint Kovacs; Constantin Ulrich; Tassilo Wald; Maximilian Zenk; Philipp Vollmuth; Jens Kleesiek; Fabian Isensee; Klaus H. Maier-Hein*

126. PFGS: High Fidelity Point Cloud Rendering via Feature Splatting; *Jiaxu Wang; Zhang Ziyi; Junhao He; Renjing Xu**

127. Masked Motion Prediction with Semantic Contrast for Point Cloud Sequence Learning; *yuehui han*; Can Xu; Rui Xu; Jianjun Qian; Jin Xie*

128. SemReg: Semantics Constrained Point Cloud Registration; *Sheldon Fung; Xuequan Lu*; Dasith de Silva Edirimuni; Wei Pan; Xiao Liu; HONGDONG LI*

129. Fast Point Cloud Geometry Compression with Context-based Residual Coding and INR-based Refinement; *Hao Xu*; Xi Zhang; Xiaolin Wu**

130. 3D Single-object Tracking in Point Clouds with High Temporal Variation; *Qiao Wu; Kun Sun; Pei An; Mathieu Salzmann; Yanning Zhang; Jiaqi Yang**

131. RangeLDM: Fast Realistic LiDAR Point Cloud Generation; *Qianjiang Hu; Zhimin Zhang; Wei Hu**

132. LISO: Lidar-only Self-Supervised 3D Object Detection; *Stefan Andreas Baur*; Frank Moosmann; Andreas Geiger*

133. UPose3D: Uncertainty-Aware 3D Human Pose Estimation with Cross-View and Temporal Cues; *Vandad Davoodnia*; Saeed Ghorbani; Marc-André Carbonneau; Alexandre Messier; Ali Etemad*

134. Real-time Holistic Robot Pose Estimation with Unknown States; *Shikun Ban; Juling Fan; Xiaoxuan Ma; Wentao Zhu*; Yu QIAO*; Yizhou Wang*

135. MLPHand: Real Time Multi-View 3D Hand Reconstruction via MLP Modeling; *Jian Yang; Jiakun Li; Guoming Li; Huaiyu Wu; Zhen Shen; Zhaoxin Fan**

136. Multi-RoI Human Mesh Recovery with Camera Consistency and Contrastive Losses; *Yongwei Nie; Changzhen Liu; Chengjiang Long; Qing Zhang; Guiqing Li; Hongmin Cai**

137. ProDepth: Boosting Self-Supervised Multi-Frame Monocular Depth with Probabilistic Fusion; *Sungmin Woo*; Wonjoon Lee; Woo Jin Kim; Dogyoon Lee; Sangyoun Lee**

138. TCLC-GS: Tightly Coupled LiDAR-Camera Gaussian Splatting for Autonomous Driving; *Cheng Zhao*; su sun; Ruoyu Wang; Yuliang Guo; Jun-Jun Wan; Zhou Huang; Xinyu Huang; Yingjie Victor Chen; Liu Ren*

139. Mahalanobis Distance-based Multi-view Optimal Transport for Multi-view Crowd Localization; *Qi Zhang; Kaiyi Zhang; Antoni B. Chan; Hui Huang**

140. Hierarchical Temporal Context Learning for Camera-based Semantic Scene Completion; *Bohan Li*; Jiajun Deng; Wenyao Zhang; Zhujin Liang; Dalong Du; Xin Jin; Wenjun Zeng*

141. FAFA: Frequency-Aware Flow-Aided Self-Supervision for Underwater Object Pose Estimation; *Jingyi Tang*; Gu Wang; Zeyu Chen; Shengquan Li; Xiu Li*; Xiangyang Ji*

142. Robust Incremental Structure-from-Motion with Hybrid Features; *Shaohui Liu*; Yidan Gao; Tianyi Zhang; Rémi Pautrat; Johannes L Schönberger; Viktor Larsson; Marc Pollefeys*

143. IFTR: An Instance-Level Fusion Transformer for Visual Collaborative Perception; *Shaohong Wang; Lu Bin; Xinyu Xiao; Zhiyu Xiang; Hangguan Shan; Eryun Liu**

144. MonoWAD: Weather-Adaptive Diffusion Model for Robust Monocular 3D Object Detection; *Youngmin Oh; Hyung-Il Kim; Seong Tae Kim*; Jung Uk Kim**

145. LiveHPS++: Robust and Coherent Motion Capture in Dynamic Free Environment; *Yiming Ren; Xiao Han; Yichen Yao; Xiaoxiao Long; Yujing Sun*; Yuexin Ma**

146. Learning 3D-aware GANs from Unposed Images with Template Feature Field; *Xinya Chen; Hanlei Guo; Yanrui Bin; Shangzhan Zhang; Yuanbo Yang; Yujun Shen; Yue Wang; Yiyi Liao**

147. Watch Your Steps: Local Image and Scene Editing by Text Instructions; *Ashkan Mirzaei*; Tristan T Aumentado-Armstrong; Marcus A Brubaker; Jonathan Kelly; Alex Levinshtein; Konstantinos G Derpanis; Igor Gilitschenski*

148. Beyond the Contact: Discovering Comprehensive Affordance for 3D Objects from Pre-trained 2D Diffusion Models; *Hyeonwoo Kim; Sookwan Han; Patrick Kwon; Hanbyul Joo**

149. Gaussian Frosting: Editable Complex Radiance Fields with Real-Time Rendering; *Antoine Guédon*; Vincent Lepetit*

150. Analytic-Splatting: Anti-Aliased 3D Gaussian Splatting via Analytic Integration; *Zhihao Liang*; Qi Zhang*; Wenbo Hu; Ying Feng; Lei ZHU; Kui Jia**

151. RaFE: Generative Radiance Fields Restoration; *Zhongkai Wu; Ziyu Wan; Jing Zhang*; Jing Liao; Dong Xu*

152. RPBG: Towards Robust Neural Point-based Graphics in the Wild; *Qingtian Zhu; Zizhuang Wei; Zhongtian Zheng; Yifan Zhan; Zhuyu Yao; Jiawang Zhang; Kejian Wu; Yinqiang Zheng**

153. Omni-Recon: Harnessing Image-based Rendering for General-Purpose Neural Radiance Fields; *Yonggan Fu; Huaizhi Qu; Zhifan Ye; Chaojian Li; Kevin Zhao; Yingyan (Celine) Lin**

154. SV3D: Novel Multi-view Synthesis and 3D Generation from a Single Image using Latent Video Diffusion; *Vikram Voleti*; Chun-Han Yao; Mark Boss; Adam Letts; David Pankratz; Dmitrii Tochilkin; Christian Laforte; Robin Rombach; Varun Jampani**

155. FisherRF: Active View Selection and Mapping with Radiance Fields using Fisher Information; *Wen Jiang*; BOSHU LEI; Kostas Daniilidis**

156. Generative Camera Dolly: Extreme Monocular Dynamic Novel View Synthesis; *Basile Van Hoorick*; Rundi Wu; Ege Ozguroglu; Kyle Sargent; Ruoshi Liu; Pavel Tokmakov; Achal Dave; Changxi Zheng; Carl Vondrick*

157. MVSplat: Efficient 3D Gaussian Splatting from Sparse Multi-View Images; *Yuedong Chen*; Haofei Xu; Chuanxia Zheng; Bohan Zhuang; Marc Pollefeys; Andreas Geiger; Tat-Jen Cham; Jianfei Cai*

158. EMIE-MAP: Large-Scale Road Surface Reconstruction Based on Explicit Mesh and Implicit Encoding; *Wenhua Wu; Qi Wang; Guangming Wang; Junping Wang; Tiankun Zhao; Yang Liu; Dongchao Gao; Zhe Liu*; Hesheng Wang**

159. Fully Sparse 3D Occupancy Prediction; *Haisong Liu; Yang Chen; Haiguang Wang; Zetong Yang; Tianyu Li; Jia Zeng; Li Chen; Hongyang Li; Limin Wang**

160. Embodied Understanding of Driving Scenarios; *Yunsong Zhou*; Linyan Huang; Qingwen Bu; Jia Zeng; Tianyu Li; Hang Qiu; Hongzi Zhu; Minyi Guo; Yu Qiao; Hongyang Li*

161. Think2Drive: Efficient Reinforcement Learning by Thinking with Latent World Model for Autonomous Driving (in CARLA-v2); *Qifeng Li*; Xiaosong Jia; Shaobo Wang; Junchi Yan*

162. Solving Motion Planning Tasks with a Scalable Generative Model; *Yihan Hu*; Siqi Chai; Zhening Yang; Jingyu Qian; Kun Li; Wenxin Shao; Haichao Zhang; Wei Xu; Qiang Liu**

163. FipTR: A Simple yet Effective Transformer Framework for Future Instance Prediction in Autonomous Driving; *Xingtai Gui*; Tengteng Huang; Haonan Shao; Haotian Yao; Chi Zhang*

164. OGNI-DC: Robust Depth Completion with Optimization-Guided Neural Iterations; *Yiming Zuo*; Jia Deng*

165. Octopus: Embodied Vision-Language Programmer from Environmental Feedback; *Jingkang Yang; Yuhao Dong; Shuai Liu; Bo Li; Ziyue Wang; ChenCheng Jiang; Haoran Tan; Jiamu Kang; Yuanhan Zhang; Kaiyang Zhou; Ziwei Liu**

166. Weight Conditioning for Smooth Optimization of Neural Networks; *Hemanth Saratchandran*; Thomas X Wang; Simon Lucey*

167. Reliability in Semantic Segmentation: Can We Use Synthetic Data?; *Thibaut Loiseau; Tuan-Hung Vu*; Mickael Chen; Patrick Pérez; Matthieu Cord*

168. SegGen: Supercharging Segmentation Models with Text2Mask and Mask2Img Synthesis; *Hanrong Ye*; Jason Kuen; Qing Liu; Zhe Lin; Brian Price; Dan Xu**

169. DGInStyle: Domain-Generalizable Semantic Segmentation with Image Diffusion Models and Stylized Semantic Control; *Yuru Jia; Lukas Hoyer; Shengyu Huang; Tianfu Wang; Luc Van Gool; Konrad Schindler; Anton Obukhov**

170. AFF-ttention! Affordances and Attention models for Short-Term Object Interaction Anticipation; *Lorenzo Mur-Labadia*; Ruben Martinez-Cantin; Jose J Guerrero; Giovanni Maria Farinella; Antonino Furnari*

171. Look Hear: Gaze Prediction for Speech-directed Human Attention; *Sounak Mondal*; Seoyoung Ahn; Zhibo Yang; Niranjan Balasubramanian; Dimitris Samaras; Gregory Zelinsky; Minh Hoai*

172. Event-Aided Time-To-Collision Estimation for Autonomous Driving; *Jinghang Li; Bangyan Liao; Xiuyuan Lu; Peidong Liu; Shaojie Shen; Yi Zhou**

173. Exploring Vulnerabilities in Spiking Neural Networks: Direct Adversarial Attacks on Raw Event Data; *Yanmeng Yao; Xiaohan Zhao; Bin Gu**

174. Motion Keyframe Interpolation for Any Human Skeleton using Point Cloud-based Human Motion Data Homogenisation; *Clinton A Mo; Kun Hu*; Chengjiang Long; Dong Yuan; Zhiyong Wang*

175. Random Walk on Pixel Manifolds for Anomaly Segmentation of Complex Driving Scenes; *Zelong Zeng*; Kaname Tomite*

176. NeRMo: Learning Implicit Neural Representations for 3D Human Motion Prediction; *Dong Wei; Huaijiang Sun; Xiaoning Sun*; Shengxiang Hu*

177. Toward INT4 Fixed-Point Training via Exploring Quantization Error for Gradients; *Dohyung Kim; Junghyup Lee; Jeimin Jeon; JAEHYEON MOON; Bumsub Ham**

178. MRSP: Learn Multi-Representations of Single Primitive for Compositional Zero-Shot Learning; *Dongyao Jiang; Hui Chen; Haodong Jing; Yongqiang Ma; Nanning Zheng**

179. Event Camera Data Dense Pre-training; *Yan Yang; Liyuan Pan*; Liu liu*

180. KDProR: A Knowledge-Decoupling Probabilistic Framework for Video-Text Retrieval; *Xianwei Zhuang*; Hongxiang Li; Xuxin Cheng; Zhihong Zhu; Yuxin Xie; Yuexian Zou*

181. Beyond MOT: Semantic Multi-Object Tracking; *Yunhao Li; Qin Li; Hao Wang; Xue Ma; Jiali Yao; Shaohua Dong; Heng Fan; Libo Zhang**

182. VideoAgent: A Memory-augmented Multimodal Agent for Video Understanding; *Yue Fan; Xiaojian Ma*; Rujie Wu; yuntao du; Jiaqi Li; Zhi Gao; Qing Li**

183. Uncertainty-aware sign language video retrieval with probability distribution modeling; *Xuan Wu*; Hongxiang Li; yuanjiang luo; Xuxin Cheng; Xianwei Zhuang; Meng Cao; Keren Fu**

184. Leveraging temporal contextualization for video action recognition; *Minji Kim; Dongyoon Han; Taekyung Kim*; Bohyung Han**

185. Unsupervised Moving Object Segmentation with Atmospheric Turbulence; *Dehao Qin*; Ripon k Saha; Woojeh Chung; Suren Jayasuriya; Jinwei Ye; Nianyi Li*

186. FedVAD: Enhancing Federated Video Anomaly Detection with GPT-Driven Semantic Distillation; *Fan Qi*; Ruijie Pan; Huaiwen Zhang; Changsheng Xu**

187. Open Vocabulary Multi-Label Video Classification; *Rohit Gupta*; Mamshad Nayeem Rizve; Jayakrishnan Unnikrishnan; Ashish Tawari; Son Tran; Mubarak Shah; Benjamin Yao; Trishul A Chilimbi*

188. Bayesian Evidential Deep Learning for Online Action Detection; *Hongji Guo; Hanjing Wang; Qiang Ji**

189. HowToCaption: Prompting LLMs to Transform Video Annotations at Scale; *Nina Shvetsova*; Anna Kukleva; Xudong Hong; Christian Rupprecht; Bernt Schiele; Hilde Kuehne*

190. InternVideo2: Scaling Foundation Models for Multimodal Video Understanding; *Yi Wang*; Kunchang Li; Xinhao Li; Jiashuo Yu; Yinan He; Guo Chen; Baoqi Pei; Rongkun Zheng; Jilan Xu; Zun Wang; Yansong Shi; Tianxiang Jiang; SongZe Li; hongjie Zhang; Yifei Huang; Yu Qiao*; Yali Wang*; Limin Wang**

191. OphNet: A Large-Scale Video Benchmark for Ophthalmic Surgical Workflow Understanding; *Ming Hu*; Peng Xia; Lin Wang; Siyuan Yan; Feilong Tang; zhongxing xu; Yimin Luo; Kaimin Song; Jurgen Leitner; Xuelian Cheng; Jun Cheng; Chi Liu; Kaijing Zhou*; Zongyuan Ge**

192. Brain Netflix: Scaling Data to Reconstruct Videos from Brain Signals; *Camilo L Fosco*; Benjamin Lahner; Bowen Pan; Alex Andonian; Emilie L Josephs; Alex Lascelles; Aude Oliva*

193. Label-anticipated Event Disentanglement for Audio-Visual Video Parsing; *Jinxing Zhou*; Dan Guo*; Yuxin Mao; Yiran Zhong; Xiaojun Chang; Meng Wang**

194. VSViG: Real-time Video-based Seizure Detection via Skeleton-based Spatiotemporal ViG; *Yankun Xu*; Junzhe Wang; Yun-Hsuan Chen; Jie Yang; Wenjie Ming; Shuang Wang; Mohamad Sawan**

195. Stepping Stones: A Progressive Training Strategy for Audio-Visual Semantic Segmentation; *Juncheng Ma; Peiwen Sun; Yaoting Wang; Di Hu**

196. FinePseudo: Improving Pseudo-Labelling through Temporal-Alignablity for Semi-Supervised Fine-Grained Action Recognition; *Ishan Rajendrakumar Dave*; Mamshad Nayeem Rizve*; Mubarak Shah*

197. Spiking Wavelet Transformer; *Yuetong Fang; Ziqing Wang; Lingfeng Zhang; Jiahang Cao; Honglei Chen; Renjing Xu**

198. Language-Assisted Skeleton Action Understanding for Skeleton-Based Temporal Action Segmentation; *Haoyu Ji; Bowen Chen; Xinglong Xu; Weihong Ren; Zhiyong Wang*; Honghai Liu*

199. Optimizing Factorized Encoder Models: Time and Memory Reduction for Scalable and Efficient Action Recognition; *Shreyank N Gowda*; Anurag Arnab; Jonathan Huang*

200. R^2-Tuning: Efficient Image-to-Video Transfer Learning for Video Temporal Grounding; *Ye Liu; Jixuan He; Wanhua Li*; Junsik Kim; Donglai Wei; Hanspeter Pfister; Chang Wen Chen**

201. Revisit Event Generation Model: Self-Supervised Learning of Event-to-Video Reconstruction with Implicit Neural Representations; *Zipeng Wang*; yunfan lu; Lin Wang**

202. Temporal Residual Guided Diffusion Framework for Event-Driven Video Reconstruction; *Lin Zhu*; Yunlong Zheng; Yijun Zhang; Xiao Wang; Lizhi Wang; Hua Huang*

203. Spike-Temporal Latent Representation for Energy-Efficient Event-to-Video Reconstruction; *Jianxiong Tang*; Jian-Huang Lai*; Lingxiao Yang; Xiaohua Xie*

204. NAMER: Non-Autoregressive Modeling for Handwritten Mathematical Expression Recognition; *Chenyu Liu; Jia Pan; Jinshui Hu; Baocai Yin; Bing Yin; Mingjun Chen; Cong Liu; Jun Du*; Qingfeng Liu*

205. Efficient Neural Video Representation with Temporally Coherent Modulation; *Seungjun Shin*; Suji Kim*; Dokwan Oh*

206. Temporal As a Plugin: Unsupervised Video Denoising with Pre-Trained Image Denoisers; *Zixuan Fu*; Lanqing Guo; Chong Wang; Yufei Wang; Zhihao Li; Bihan Wen*

207. Learned Rate Control for Frame-Level Adaptive Neural Video Compression via Dynamic Neural Network; *Chenhao Zhang; Wei Gao**

208. Clearer Frames, Anytime: Resolving Velocity Ambiguity in Video Frame Interpolation; *Zhihang Zhong; Gurunandan Krishnan; Xiao Sun; Yu Qiao; Sizhuo Ma*; Jian Wang**

209. Perceptual Evaluation of Audio-Visual Synchrony Grounded in Viewers' Opinion Scores; *Lucas Goncalves; Prashant Mathur*; Chandrashekhar Lavania; Metehan Cekic; Marcello Federico; Kyu Han*

210. TimeLens-XL: Real-time Event-based Video Frame Interpolation with Large Motion; *Shi Guo; Yutian Chen; Tianfan Xue; Jinwei Gu; Yongrui Ma**

211. POET: Prompt Offset Tuning for Continual Human Action Adaptation; *Prachi Garg*; Joseph K J; Vineeth N Balasubramanian; Necati Cihan Camgoz; Chengde Wan; Kenrick Kin; Weiguang Si; Shugao Ma; Fernando de la Torre*

212. Elucidating the Hierarchical Nature of Behavior with Masked Autoencoders; *Lucas Stoffl; Andy Bonnetto; Stéphane D'Ascoli; Alexander Mathis**

213. Getting it Right: Improving Spatial Consistency in Text-to-Image Models; *Agneet Chatterjee*; Gabriela Ben Melech Stan; Estelle Guez Aflalo; Sayak Paul; Dhruba Ghosh; Tejas Gokhale; Ludwig Schmidt; Hanna Hajishirzi; Vasudev Lal; Chitta R Baral; Yezhou Yang*

214. Generating Physically Realistic and Directable Human Motions from Multi-Modal Inputs; *Aayam Shrestha; Pan Liu*; German Ros; Kai Yuan*; Alan Fern*

215. Learning-based Axial Video Motion Magnification; *Kwon Byung-Ki; Oh Hyun-Bin; Kim Jun-Seong; Hyunwoo Ha; Tae-Hyun Oh**

216. An Economic Framework for 6-DoF Grasp Detection; *Xiao-Ming Wu*; Jia-Feng Cai; Jian-Jian Jiang; Dian Zheng; Yi-Lin Wei; Wei-Shi Zheng**

217. UGG: Unified Generative Grasping; *Jiaxin Lu; Hao Kang; Haoxiang Li; Bo Liu; Yiding Yang; Qixing Huang; Gang Hua**

218. SemGrasp: Semantic Grasp Generation via Language Aligned Discretization; *Kailin Li*; Jingbo Wang; Lixin Yang; Cewu Lu*; Bo Dai*

219. NL2Contact: Natural Language Guided 3D Hand-Object Contact Modeling with Diffusion Model; *Zhongqun Zhang*; Hengfei Wang; Ziwei Yu; Yihua Cheng*; Angela Yao; Hyung Jin Chang*

220. MIGS: Multi-Identity Gaussian Splatting via Tensor Decomposition; *Aggelina Chatziagapi*; Grigorios Chrysos; Dimitris Samaras*

221. Norface: Improving Facial Expression Analysis by Identity Normalization; *Hanwei Liu*; Rudong An; Zhimeng Zhang; Bowen Ma; Wei Zhang; Yan Song; Yujing Hu; Chen Wei; Yu Ding**

222. Scalable Group Choreography via Variational Phase Manifold Learning; *Nhat Le; Khoa Do; Xuan Bui; Tuong Do; Erman Tjiputra; Quang D.Tran; Anh Nguyen**

223. FreeMotion: MoCap-Free Human Motion Synthesis with Multimodal Large Language Models; *Zhikai Zhang; Yitang Li; Haofeng Huang; Mingxian Lin; Li Yi**

224. Controllable Human-Object Interaction Synthesis; *Jiaman Li*; Alexander Clegg; Roozbeh Mottaghi; Jiajun Wu; Xavier Puig; C. Karen Liu*

225. Plan, Posture and Go: Towards Open-vocabulary Text-to-Motion Generation; *Jinpeng Liu; Wenxun Dai; Chunyu Wang; Yiji Cheng; Yansong Tang*; Xin Tong*

226. KMTalk: Speech-Driven 3D Facial Animation with Key Motion Embedding; *Zhihao Xu; Shengjie Gong; Jiapeng Tang; Lingyu Liang; Yining Huang; Haojie Li; Shuangping Huang**

227. E.T. the Exceptional Trajectory: Text-to-camera-trajectory generation with character awareness; *Robin Courant*; Nicolas Dufour; Xi WANG; Marc Christie; Vicky Kalogeiton*

228. Drag Anything: Motion Control for Anything using Entity Representation; *Weijia Wu ; Zhuang Li; Yuchao Gu; Rui Zhao; Yefei He; David Junhao Zhang; Mike Zheng Shou*; Yan Li; Tingting Gao; Zhang Di*

229. ZoLA: Zero-Shot Creative Long Animation Generation with Short Video Model; *Fu-Yun Wang*; Zhaoyang Huang*; Qiang Ma; Guanglu Song; Xudong LU; Weikang Bian; Yijin Li; Yu Liu; Hongsheng Li**

230. Audio-Synchronized Visual Animation; *Lin Zhang; Shentong Mo; Yijing Zhang; Pedro Morgado**

231. Champ: Controllable and Consistent Human Image Animation with 3D Parametric Guidance; *Shenhao Zhu; Junming Leo Chen; Zuozhuo Dai; Zilong Dong; Yinghui Xu; Xun Cao; Yao Yao; Hao Zhu*; Siyu Zhu**

232. GroupDiff: Diffusion-based Group Portrait Editing; *Yuming Jiang; Nanxuan Zhao*; Qing Liu; Krishna Kumar Singh; Shuai Yang; Chen Change Loy; Ziwei Liu*

233. ReSyncer: Rewiring Style-based Generator for Unified Audio-Visually Synced Facial Performer; *Jiazhi Guan*; Zhiliang Xu; Hang Zhou; Kaisiyuan Wang; Shengyi He; Zhanwang Zhang; Borong Liang; Haocheng Feng; Errui Ding; Jingtuo Liu; Jingdong Wang; Youjian Zhao; Ziwei Liu*

234. Object-Centric Diffusion for Efficient Video Editing; *Kumara Kahatapitiya*; Adil Karjauv; Davide Abati*; Fatih Porikli; Yuki M Asano; Amirhossein Habibian*

235. MotionDirector: Motion Customization of Text-to-Video Diffusion Models; *Rui Zhao; Yuchao Gu; Jay Zhangjie Wu; David Junhao Zhang; Jia-Wei Liu; weijia wu; Jussi Keppo; Mike Zheng Shou**

236. DynamiCrafter: Animating Open-domain Images with Video Diffusion Priors; *Jinbo Xing*; Menghan Xia; Yong Zhang; Haoxin Chen; Wangbo Yu; Hanyuan Liu; Gongye Liu; Xintao Wang; Ying Shan; Tien-Tsin Wong*

237. Let the Avatar Talk using Texts without Paired Training Data; *Xiuzhe Wu; Yang-Tian Sun; Handi Chen; Hang Zhou; Jingdong Wang; Zhengzhe Liu; Xiaojuan Qi**

238. Modeling and Driving Human Body Soundfields through Acoustic Primitives; *Chao Huang*; Dejan Markovic*; Chenliang Xu*; Alexander Richard**

239. SparseCtrl: Adding Sparse Controls to Text-to-Video Diffusion Models; *Yuwei Guo; Ceyuan Yang*; Anyi Rao; Maneesh Agrawala; Dahua Lin*; Bo Dai**

240. LEGO: Learning EGOcentric Action Frame Generation via Visual Instruction Tuning; *Bolin Lai*; Xiaoliang Dai; Lawrence Chen; Guan Pang; James M Rehg; Miao Liu* BEST PAPER CANDIDATE

241. EAFormer: Scene Text Segmentation with Edge-Aware Transformers; *Haiyang Yu; Teng Fu; Bin Li*; Xiangyang Xue*

242. Video Editing via Factorized Diffusion Distillation; *Uriel Singer*; Amit Zohar*; Yuval Kirstain; Shelly Sheynin; Adam Polyak; Devi Parikh; Yaniv Taigman*

243. MultiGen: Zero-shot Image Generation from Multi-modal Prompts; *Zhi-Fan Wu*; Lianghua Huang; Wei Wang; Yanheng Wei; Yu Liu*

244. InterFusion: Text-Driven Generation of 3D Human-Object Interaction; *Sisi Dai; Wenhao Li; Haowen Sun; Haibin Huang; Chongyang Ma; Hui Huang; Kai Xu*; Ruizhen Hu**

245. When and How do negative prompts take effect?; *Yuanhao Ban; Ruochen Wang; Tianyi Zhou; Minhao Cheng; Boqing Gong; Cho-Jui Hsieh**

246. LogoSticker: Inserting Logos into Diffusion Models for Customized Generation; *Mingkang Zhu; Xi CHEN; Zhongdao Wang; Hengshuang Zhao*; Jiaya Jia**

247. Bridging Different Language Models and Generative Vision Models for Text-to-Image Generation; *Shihao Zhao*; Shaozhe Hao; Bojia Zi; Huaizhe Xu; Kwan-Yee K. Wong**

248. Text-Anchored Score Composition: Tackling Condition Misalignment in Text-to-Image Diffusion Models; *Luozhou Wang*; Guibao Shen; Wenhang Ge; Guangyong Chen; Yijun Li; Yingcong Chen**

249. AttentionHand: Text-driven Controllable Hand Image Generation for 3D Hand Reconstruction in the Wild; *Junho Park; Kyeongbo Kong; Suk-Ju Kang**

250. Lego: Learning to Disentangle and Invert Personalized Concepts Beyond Object Appearance in Text-to-Image Diffusion Models; *Saman Motamed*; Danda Pani Paudel; Luc Van Gool*

251. Implicit Concept Removal of Diffusion Models; *Zhili Liu*; Kai Chen; Yifan Zhang; Jianhua Han; Lanqing Hong; Hang Xu; Zhenguo Li; Dit-Yan Yeung; James Kwok*

252. EditShield: Protecting Unauthorized Image Editing by Instruction-guided Diffusion Models; *Ruoxi Chen; Haibo Jin; Yixin Liu; Jinyin Chen*; Haohan Wang; Lichao Sun*

253. SwapAnything: Enabling Arbitrary Object Swapping in Personalized Image Editing; *Jing Gu*; Nanxuan Zhao; Wei Xiong; Qing Liu; Zhifei Zhang; He Zhang; Jianming Zhang; HyunJoon Jung; Yilin Wang*; Xin Eric Wang**

254. UDiffText: A Unified Framework for High-quality Text Synthesis in Arbitrary Images via Character-aware Diffusion Models; *Yiming Zhao*; Zhouhui Lian**

255. InstaStyle: Inversion Noise of a Stylized Image is Secretly a Style Adviser; *Xing Cui; Zekun Li; Peipei Li*; Huaibo Huang; Xuannan Liu; Zhaofeng He*

256. Enhancing Semantic Fidelity in Text-to-Image Synthesis: Attention Regulation in Diffusion Models; *Yang Zhang*; Tze Tzun Teoh; Wei Hern Lim; Kenji Kawaguchi*

257. Enhancing Diffusion Models with Text-Encoder Reinforcement Learning; *Chaofeng Chen*; Annan Wang; Haoning Wu; Liang Liao; Wenxiu Sun; Qiong Yan; Weisi Lin**

258. Towards compact reversible image representations for neural style transfer; *Xiyao Liu; Siyu Yang; Jian Zhang*; Gerald Schaefer; Jiya Li; Xunli FAN; Songtao Wu; Hui Fang**

259. LayerDiff: Exploring Text-guided Multi-layered Composable Image Synthesis via Layer-Collaborative Diffusion Model; *Runhui Huang; Kaixin Cai; Jianhua Han; Xiaodan Liang*; Renjing Pei; Guansong Lu; Songcen Xu; Wei Zhang; Hang Xu*

260. SwiftBrush v2: Make Your One-step Diffusion Model Better Than Its Teacher; *Trung Tuan Dao*; Thuan Hoang Nguyen; Thanh Van Le; Duc H Vu; Khoi Nguyen; Cuong Pham; Anh T Tran**

261. Relightable Neural Actor with Intrinsic Decomposition and Pose Control; *Diogo Carbonera Luvizon*; Vladislav Golyanik; Adam Kortylewski; Marc Habermann; Christian Theobalt*

262. SPIRE: Semantic Prompt-Driven Image Restoration; *Chenyang QI*; Zhengzhong Tu; Keren Ye; Mauricio Delbracio; Peyman Milanfar; Qifeng Chen; Hossein Talebi*

263. Source Prompt Disentangled Inversion for Boosting Image Editability with Diffusion Models; *Ruibin Li*; Ruihuang Li; Song Guo; Lei Zhang*

264. Pixel-Aware Stable Diffusion for Realistic Image Super-Resolution and Personalized Stylization; *Tao Yang*; Rongyuan Wu; Peiran Ren; Xuansong Xie; Lei Zhang*

265. AdaNAT: Exploring Adaptive Policy for Token-Based Image Generation; *Zanlin Ni; Yulin Wang; Renping Zhou; Rui Lu; Jiayi Guo; Jinyi Hu; Zhiyuan Liu; Yuan Yao*; Gao Huang**

266. Learning Neural Volumetric Pose Features for Camera Localization; *Jingyu Lin; Jiaqi Gu; Bojian Wu; Lubin Fan*; Renjie Chen*; Ligang Liu; Jieping Ye*

267. InstructIR: High-Quality Image Restoration Following Human Instructions; *Marcos V. Conde*; Gregor Geigle; Radu Timofte*

268. BrushNet: A Plug-and-Play Image Inpainting Model with Decomposed Dual-Branch Diffusion; *Xuan Ju*; Xian Liu; Xintao Wang*; Yuxuan Bian; Ying Shan; Qiang Xu**

269. Test-Time Stain Adaptation with Diffusion Models for Histopathology Image Classification; *Cheng-Chang Tsai*; Yuan-Chih Chen; Chun-Shien Lu**

270. Efficient Diffusion-Driven Corruption Editor for Test-Time Adaptation; *Yeongtak Oh; Jonghyun Lee; Jooyoung Choi; Dahuin Jung; Uiwon Hwang*; Sungroh Yoon**

271. UCIP: A Universal Framework for Compressed Image Super-Resolution using Dynamic Prompt; *Xin Li*; Bingchen Li; Yeying Jin; Cuiling Lan; Hanxin Zhu; Yulin Ren; Zhibo Chen*

272. Guide-and-Rescale: Self-Guidance Mechanism for Effective Tuning-Free Real Image Editing; *Vadim Titov*; Madina Khalmatova*; Alexandra Ivanova*; Dmitry P Vetrov; Aibek Alanov**

273. Improving Virtual Try-On with Garment-focused Diffusion Models; *Siqi Wan; Yehao Li; Jingwen Chen; Yingwei Pan*; Ting Yao; Yang Cao; Tao Mei*

274. DreamDissector: Learning Disentangled Text-to-3D Generation from 2D Diffusion Priors; *Zizheng Yan*; Jiapeng Zhou; Fanpeng Meng; Yushuang Wu; Lingteng Qiu; Zisheng Ye; Shuguang Cui; Guanying CHEN; Xiaoguang Han**

275. LatentEditor: Text Driven Local Editing of 3D Scenes; *Umar Khalid*; Hasan Iqbal; Muhammad Tayyab; Md Nazmul Karim; Jing Hua; Chen Chen*

276. Retargeting Visual Data with Deformation Fields; *Tim Elsner*; Julia Berger; Tong Wu; Victor Czech; Lin Gao; Leif Kobbelt*

277. CanonicalFusion: Generating Drivable 3D Human Avatars from Multiple Images; *Jisu Shin; Junmyeong Lee; Seongmin Lee; Min-Gyu Park; Jumi Kang; Ju Hong Yoon; Hae-Gon Jeon**

278. SelfSwapper: Self-Supervised Face Swapping via Shape Agnostic Masked AutoEncoder; *Jaeseong Lee*; Junha Hyung*; Sohyun Jeong; Jaegul Choo**

279. LayoutFlow: Flow Matching for Layout Generation; *Julian Jorge Andrade Guerreiro*; Naoto Inoue*; Kento Masui; Mayu Otani; Hideki Nakayama*

280. A Unified Anomaly Synthesis Strategy with Gradient Ascent for Industrial Anomaly Detection and Localization; *Qiyu Chen; Huiyuan Luo; Chengkan Lv*; Zhengtao Zhang*

281. Gravity-aligned Rotation Averaging with Circular Regression; *Linfei Pan*; Marc Pollefeys; Daniel Barath*

282. FAMOUS: High-Fidelity Monocular 3D Human Digitization Using View Synthesis; *Vishnu Mani Hema*; Shubhra Aich; Christian Haene; Jean-Charles Bazin; Fernando de la Torre*

283. GarmentCodeData: A Dataset of 3D Made-to-Measure Garments With Sewing Patterns; *Maria Korosteleva*; Timur Levent Kesdogan; Fabian Kemper; Stephan Wenninger; Jasmin Koller; Yuhan Zhang; Mario Botsch; Olga Sorkine-Hornung*

284. Synchronous Diffusion for Unsupervised Smooth Non-Rigid 3D Shape Matching; *Dongliang Cao*; Zorah Laehner; Florian Bernard*

285. Surf-D: Generating High-Quality Surfaces of Arbitrary Topologies Using Diffusion Models; *Zhengming Yu*; Zhiyang Dou; Xiaoxiao Long; Cheng Lin; Zekun Li; Yuan Liu; Norman Müller; Taku Komura; Marc Habermann; Christian Theobalt; Xin Li; Wenping Wang**

286. DoughNet: A Visual Predictive Model for Topological Manipulation of Deformable Objects; *Dominik Bauer*; Zhenjia Xu; Shuran Song*

287. AWOL: Analysis WithOut synthesis using Language; *Silvia Zuffi*; Michael J. Black*

288. Generating 3D House Wireframes with Semantics; *Xueqi Ma; Yilin Liu; Wenjun Zhou; Ruowei Wang; Hui Huang**

289. Connecting Consistency Distillation to Score Distillation for Text-to-3D Generation; *Zongrui Li*; Minghui Hu; Qian Zheng*; Xudong Jiang*

290. Mesh2NeRF: Direct Mesh Supervision for Neural Radiance Field Representation and Generation; *Yujin Chen*; Yinyu Nie; Benjamin Ummenhofer; Reiner Birkl; Michael Paulitsch; Matthias Müller; Matthias Niessner*

291. Pairwise Distance Distillation for Unsupervised Real-World Image Super-Resolution; *Yuehan Zhang*; Seungjun Lee; Angela Yao*

292. Make a Cheap Scaling: A Self-Cascade Diffusion Model for Higher-Resolution Adaptation; *Lanqing Guo; Yingqing HE; Haoxin Chen; Menghan Xia; Xiaodong Cun; Yufei Wang; Siyu Huang; Yong Zhang; Xintao Wang; Qifeng Chen; Ying Shan; Bihan Wen**

293. OneRestore: A Universal Restoration Framework for Composite Degradation; *Yu Guo*; Yuan Gao; Yuxu Lu; Huilin Zhu; Wen Liu; Shengfeng He*

294. SCP-Diff: Spatial-Categorical Joint Prior for Diffusion Based Semantic Image Synthesis; *Huan-ang Gao; Mingju Gao; Jiaju Li; Wenyi Li; Rong Zhi; Hao Tang; Hao Zhao**

295. Towards Real-World Adverse Weather Image Restoration: Enhancing Clearness and Semantics with Vision-Language Models; *Jiaqi Xu*; Mengyang Wu; Xiaowei Hu*; Chi-Wing Fu; Qi Dou; Pheng-Ann Heng*

296. NVS-Adapter: Plug-and-Play Novel View Synthesis from a Single Image; *Yoonwoo Jeong; Jinwoo Lee; Chiheon Kim; Minsu Cho*; Doyup Lee**

297. CrossScore: A Multi-View Approach to Image Evaluation and Scoring; *Zirui Wang*; Wenjing Bian; Victor Adrian Prisacariu*

298. SuperGaussian: Repurposing Video Models for 3D Super Resolution; *Yuan Shen*; Duygu Ceylan*; Paul Guerrero; Zexiang Xu; Niloy J. Mitra; Shenlong Wang; Anna Fruehstueck**

299. Leveraging Representations from Intermediate Encoder-blocks for Synthetic Image Detection; *Christos Koutlis*; Symeon Papadopoulos*

300. GOEmbed: Gradient Origin Embeddings for Representation Agnostic 3D Feature Learning; *Animesh Karnewar*; Roman Shapovalov; Tom Monnier; Andrea Vedaldi; Niloy J. Mitra*; David Novotny**

301. On Learning Discriminative Features from Synthesized Data for Self-Supervised Fine-Grained Visual Recognition; *Zihu Wang*; Lingqiao Liu; Scott Ricardo Figueroa Weston; Samuel Tian; Peng Li*

302. STAG4D: Spatial-Temporal Anchored Generative 4D Gaussians; *Yifei Zeng; Yanqin Jiang; Siyu Zhu; Yuanxun Lu; Youtian Lin; Hao Zhu; Weiming Hu; Xun Cao; Yao Yao**

303. Every Pixel Has its Moments: Ultra-High-Resolution Unpaired Image-to-Image Translation via Dense Normalization; *Ming-Yang Ho; Che-Ming Wu; Min-Sheng Wu;*

304. When Fast Fourier Transform Meets Transformer for Image Restoration; *Xingyu Jiang; Xiuhui Zhang; Ning Gao; Yue Deng**

305. StyleCity: Large-Scale 3D Urban Scenes Stylization; *Yingshu Chen; Huajian Huang*; Tuan-Anh Vu; Ka Chun Shum; Sai-Kit Yeung*

306. Learning 3D Geometry and Feature Consistent Gaussian Splatting for Object Removal; *Yuxin Wang; Qianyi Wu; Guofeng Zhang; Dan Xu**

307. Content-Aware Radiance Fields: Aligning Model Complexity with Scene Intricacy Through Learned Bitwidth Quantization; *Weihang Liu; Xue Xian Zheng; Jingyi Yu; Xin Lou**

308. Unveiling Advanced Frequency Disentanglement Paradigm for Low-Light Image Enhancement; *Kun Zhou*; Xinyu Lin; Wenbo Li; Xiaogang Xu; Yuanhao Cai; Zhonghang Liu; Xiaoguang Han; Jiangbo Lu*

309. Collaborative Control for Geometry-Conditioned PBR Image Generation; *Shimon Vainer; Mark Boss; Mathias Parger; Konstantin Kutsy; Dante De Nigris; Ciara Rowles; Nicolas Perony; Simon Donné**

310. Intrinsic Single-Image HDR Reconstruction; *Sebastian Dille*; Chris Careaga*; Yagiz Aksoy*

311. 3R-INN: How to be climate friendly while consuming/delivering videos?; *ZOUBIDA AMEUR*; Claire-Helene Demarty; Olivier LE MEUR; Daniel Menard*

312. Gaussian Grouping: Segment and Edit Anything in 3D Scenes; *Mingqiao Ye; Martin Danelljan; Fisher Yu; Lei Ke**

313. Repaint123: Fast and High-quality One Image to 3D Generation with Progressive Controllable Repainting; *Junwu Zhang*; Zhenyu Tang; Yatian Pang; Xinhua Cheng; Peng Jin; Yida Wei; xing zhou; munan ning; Li Yuan**

314. SAGS: Structure-Aware 3D Gaussian Splatting; *Evangelos Ververas; Rolandos Alexandros Potamias*; Jifei Song; Jiankang Deng; Stefanos Zafeiriou*

315. Photorealistic Object Insertion with Diffusion-Guided Inverse Rendering; *Ruofan Liang; Zan Gojcic; Merlin Nimier-David; David Acuna; Nandita Vijaykumar; Sanja Fidler; Zian Wang**

316. HAC: Hash-grid Assisted Context for 3D Gaussian Splatting Compression; *Yihang Chen*; Qianyi Wu; Weiyao Lin*; Mehrtash Harandi; Jianfei Cai*

317. Compact 3D Scene Representation via Self-Organizing Gaussian Grids; *Wieland Morgenstern*; Florian Barthel; Anna Hilsmann; Peter Eisert*

318. Consistent 3D Line Mapping; *Xulong Bai; Hainan Cui*; Shuhan Shen**

319. GSD: View-Guided Gaussian Splatting Diffusion for 3D Reconstruction; *Yuxuan Mu*; Xinxin Zuo; Chuan Guo; Yilin Wang; Juwei Lu; Xiaofei Wu; Songcen Xu; Peng Dai; Youliang Yan; Li Cheng*

320. Scalar Function Topology Divergence: Comparing Topology of 3D Objects; *Ilya Trofimov*; Daria Voronkova; Eduard Tulchinskii; Evgeny Burnaev; Serguei Barannikov*

321. Parameterization-driven Neural Surface Reconstruction for Object-oriented Editing in Neural Rendering; *Baixin Xu; Jiangbei Hu; Fei Hou; Kwan-Yee Lin; Wayne Wu; Chen Qian; Ying He**

322. Imaging Interiors: An Implicit Solution to Electromagnetic Inverse Scattering Problems; *Ziyuan Luo; Boxin Shi; Haoliang Li; Renjie Wan**

323. Synthesizing Time-varying BRDFs via Latent Space; *Takuto Narumoto*; Hiroaki Santo; Fumio Okura*

324. MERLiN: Single-Shot Material Estimation and Relighting for Photometric Stereo; *Ashish Tiwari*; Satoshi Ikehata; Shanmuganathan Raman*

325. GAURA: Generalizable Approach for Unified Restoration and Rendering of Arbitrary Views; *Vinayak Gupta*; Rongali Simhachala Venkata Girish; Mukund Varma T; Ayush Tewari; Kaushik Mitra*

326. Adaptive Annealing for Robust Averaging; *Sidhartha Chitturi*; Venu Madhav Govindu*

327. GeoCalib: Learning Single-image Calibration with Geometric Optimization; *Alexander Veicht*; Paul-Edouard Sarlin*; Philipp Lindenberger; Marc Pollefeys*

328. RePOSE: 3D Human Pose Estimation via Spatio-Temporal Depth Relational Consistency; *Ziming Sun; Yuan Liang; Zejun Ma; Tianle Zhang; Linchao Bao; Guiqing Li; Shengfeng He**

329. MUSES: The Multi-Sensor Semantic Perception Dataset for Driving under Uncertainty; *Tim Broedermann*; David Brüggemann; Christos Sakaridis; Kevin Ta; Odysseas Liagouris; Jason Corkill; Luc Van Gool*

330. Reinforcement Learning Meets Visual Odometry; *Nico Messikommer*; Giovanni Cioffi; Mathias Gehrig; Davide Scaramuzza*

331. FSGS: Real-Time Few-shot View Synthesis using Gaussian Splatting; *Zehao Zhu; Zhiwen Fan*; Yifan Jiang; Zhangyang Wang**

333. Enhanced Motion Forecasting with Visual Relation Reasoning; *Sungjune Kim; Hadam Baek; Seunggwan Lee; Hyung-gun Chi; Hyerin Lim; Jinkyu Kim*; Sangpil Kim**

334. Continuity Preserving Online CenterLine Graph Learning; *Yunhui Han; Kun Yu; Zhiwei Li**

335. KFD-NeRF: Rethinking Dynamic NeRF with Kalman Filter; *Yifan Zhan; Zhuoxiao Li; Muyao Niu; Zhihang Zhong; Shohei Nobuhara; Ko Nishino; Yinqiang Zheng**

336. Concise Plane Arrangements for Low-Poly Surface and Volume Modelling; *Raphael Sulzer; Florent Lafarge**

337. Instant Uncertainty Calibration of NeRFs Using a Meta-Calibrator; *Niki Amini-Naieni*; Tomas Jakab; Andrea Vedaldi; Ronald Clark*

338. DoubleTake: Geometry Guided Depth Estimation; *Mohamed Sayed*; Filippo Aleotti; Jamie Watson; Zawar Qureshi; Guillermo Garcia-Hernando; Gabriel Brostow; Sara Vicente; Michael Firman*

339. Domain Reduction Strategy for Non-Line-of-Sight Imaging; *Hyunbo Shim; In Cho; Daekyu Kwon; Seon Joo Kim**

340. TrackNeRF: Bundle Adjusting NeRF from Sparse and Noisy Views via Feature Tracks; *Jinjie Mai*; Wenxuan Zhu; Sara Rojas; Jesus Zarzar; Abdullah Hamdi; Guocheng Qian; Bing Li; Silvio Giancola; Bernard Ghanem*

341. Gaussian Splatting on the Move: Blur and Rolling Shutter Compensation for Natural Camera Motion; *Otto Seiskari*; Jerry Ylilammi; Valtteri Kaatrasalo; Pekka Rantalankila; Matias Turkulainen; Juho Kannala; Esa Rahtu; Arno Solin*

342. URS-NeRF: Unordered Rolling Shutter Bundle Adjustment for Neural Radiance Fields; *Bo Xu*; Liu Ziao; Mengqi Guo; jiancheng Li; Gim Hee Lee*

343. Resolving Scale Ambiguity in Multi-view 3D Reconstruction using Dual-Pixel Sensors; *Kohei Ashida*; Hiroaki Santo; Fumio Okura; Yasuyuki Matsushita*

344. Event-based Mosaicing Bundle Adjustment; *Shuang Guo*; Guillermo Gallego*

THURSDAY, 3RD OCTOBER

08:00 – 18:30
Registration - Badge Pickup

09:00 – 18:30
Exhibition - Level 0

09:00 – 10:30
Oral session 5A: Segmentation - Gold Room
Chairs: Jiri Matas; Jing Zhang
1. WPS-SAM: Towards Weakly-Supervised Part Segmentation with Foundation Models; *Xin-Jian Wu*; Ruisong Zhang; Jie Qin; Shijie Ma; Cheng-Lin Liu**

2. AlignDiff: Aligning Diffusion Models for General Few-Shot Segmentation; *Ri-Zhao Qiu*; Yu-Xiong Wang; Kris Hauser*

3. CAT-SAM: Conditional Tuning for Few-Shot Adaptation of Segment Anything Model; *Aoran Xiao; Weihao Xuan; Heli Qi; Yun Xing; Ruijie Ren; Xiaoqin Zhang; Ling Shao; Shijian Lu**

4. Collaborative Vision-Text Representation Optimizing for Open-Vocabulary Segmentation; *Siyu Jiao*; hongguang Zhu; Yunchao Wei; Yao Zhao*; Jiannan Huang; Humphrey Shi*

5. Efficient Active Domain Adaptation for Semantic Segmentation by Selecting Information-rich Superpixels; *Yuan Gao; Zilei Wang*; Yixin Zhang; Bohai Tu*

6. ActionVOS: Actions as Prompts for Video Object Segmentation; *Liangyang Ouyang*; Ruicong Liu; Yifei Huang*; Ryosuke Furuta; Yoichi Sato**

7. Learning Modality-agnostic Representation for Semantic Segmentation from Any Modalities; *Xu Zheng*; Yuanhuiyi Lyu; Lin Wang**

8. Diffusion Models for Open-Vocabulary Segmentation; *Laurynas Karazija*; Iro Laina; Andrea Vedaldi; Christian Rupprecht*

09:00 – 10:30 Oral session 5B: Vision applications - Auditorium
Chairs: Nicoletta Noceti; Joachim Denzler
1. Robust Fitting on a Gate Quantum Computer; *Frances F Yang*; Michele Sasdelli; Tat-Jun Chin* BEST PAPER CANDIDATE

2. Geospecific View Generation - Geometry-Context Aware High-resolution Ground View Inference from Satellite Views; *Ningli Xu; Rongjun Qin**

3. Language-Driven 6-DoF Grasp Detection Using Negative Prompt Guidance; *Toan Nguyen; Minh Nhat Nhat Vu; Baoru Huang; An Dinh Vuong; Quan Vuong; Ngan Le; Thieu Vo; Anh Nguyen**

4. MaxMI: A Maximal Mutual Information Criterion for Manipulation Concept Discovery; *Pei Zhou; Yanchao Yang**

5. Align before Collaborate: Mitigating Feature Misalignment for Robust Multi-Agent Perception; *Dingkang Yang; Dingkang Yang; Ke Li; Dongling Xiao; Zedian Shao; Peng Sun; Liang Song**

6. Faceptor: A Generalist Model for Face Perception; *Lixiong Qin*; Mei Wang; Xuannan Liu; Yuhang Zhang; Wei Deng; Xiaoshuai Song; Weiran Xu*; Weihong Deng*

7. A Geometric Distortion Immunized Deep Watermarking Framework with Robustness Generalizability; *Linfeng Ma; Han Fang*; Tianyi Wei; Zijin Yang; Zehua Ma*; Weiming Zhang; Nenghai Yu*

8. COHO: Context-Sensitive City-Scale Hierarchical Urban Layout Generation; *Liu He*; Daniel Aliaga*

09:00 – 10:30
Oral session 5C: Representation learning - Silver Room
Chairs: Yuki Asano; Stella Yu
1. PiTe: Pixel-Temporal Alignment for Large Video-Language Model; *Yang Liu*; Pengxiang Ding; Siteng Huang; Min Zhang; Han Zhao; Donglin Wang*

2. Pose-Aware Self-Supervised Learning with Viewpoint Trajectory Regularization; *Jiayun Wang*; Yubei Chen; Stella X. Yu*

3. Emergent Visual-Semantic Hierarchies in Image-Text Representations; *Morris Alper*; Hadar Averbuch-Elor*

4. Learning Multimodal Latent Generative Models with Energy-Based Prior; *Shiyu Yuan*; Jiali Cui; Hanao Li; Tian Han*

5. Decoupling Common and Unique Representations for Multimodal Self-supervised Learning; *Yi Wang*; Conrad M Albrecht; Nassim Ait Ali Braham; Chenying Liu; Zhitong Xiong; Xiao Xiang Zhu*

6. SINDER: Repairing the Singular Defects of DINOv2; *Haoqi Wang; Tong Zhang; Mathieu Salzmann**

7. Denoising Vision Transformers; *Jiawei Yang*; Katie Z Luo; Jiefeng Li; Congyue Deng; Leonidas Guibas; Dilip Krishnan; Kilian Weinberger; Yonglong Tian; Yue Wang*

8. Exploring the Feature Extraction and Relation Modeling For Light-Weight Transformer Tracking; *Jikai Zheng; Mingjiang Liang; Shaoli Huang; Jifeng Ning**

09:00 – 12:30
Demo session 5 - Level 0
1. Multi-Setup Depth Perception through Virtual Image Hallucination; *Luca Bartolomei, Matteo Poggi, Fabio Tosi, Andrea Conti, Stefano Mattoccia - University of Bologna*

2. COMO: Compact Mapping and Odometry; *Eric Dexheimer, Andrew J. Davison - Imperial College London*

3. H-Unique: 3D Hand Reconstruction and Automated Mapping of Anatomical Detail for Forensic Identification; *Bryan M. Williams, Hossein Rahmani, Sue Black, Xinyu Yang, Zheheng Jiang, Andrei Banica - Lancaster University*

4. ScribblePrompt: Fast and Flexible Interactive Segmentation for Any Biomedical Image; *Hallee E. Wong, Marianne Rakic, John Guttag, Adrian V. Dalca - MIT CSAIL*

5. Showcase: Contrasting Deepfakes Embeddings; *Lorenzo Baraldi, Federico Cocchi, Stefano Savian, Marcella Cornia, Lorenzo Baraldi, Alessandro Nicolosi, Marina Geymonat, Rita Cucchiara - University of Modena and Reggio Emilia*

10:30 – 11:00
Baidu Technical Session - Technical Presentation Area (Level 0)

10:30 – 11:00
Coffee Break - Exhibition Area (Level 0)

10:30 – 12:30
Poster session 5
1. Adaptive Selection of Sampling-Reconstruction in Fourier Compressed Sensing; *Seongmin Hong; Jaehyeok Bae; Jongho Lee*; Se Young Chun**
2. Uncertainty-Driven Spectral Compressive Imaging with Spatial-Frequency Transformer; *Lintao Peng; Siyu Xie; Liheng Bian**

3. Masked Angle-Aware Autoencoder for Remote Sensing Images; *Zhihao Li*; Biao Hou; Siteng Ma; zitong wu; Xianpeng Guo; bo ren; Licheng Jiao*

4. Data Overfitting for On-Device Super-Resolution with Dynamic Algorithm and Compiler Co-Design; *Gen Li*; zhihao shu; Jie Ji; Minghai Qin; Fatemeh Afghah; Wei Niu; Xiaolong Ma**

5. Accelerating Image Super-Resolution Networks with Pixel-Level Classification; *Jinho Jeong; Jinwoo Kim; Younghyun Jo; Seon Joo Kim**

6. Bidirectional Stereo Image Compression with Cross-Dimensional Entropy Model; *Zhening Liu; Xinjie Zhang; Jiawei Shao; Zehong Lin*; Jun Zhang*

7. Overcoming Distribution Mismatch in Quantizing Image Super-Resolution Networks; *Cheeun Hong; Kyoung Mu Lee**

8. Rate-Distortion-Cognition Controllable Versatile Neural Image Compression; *Jinming Liu*; Ruoyu Feng; Yunpeng Qi; Qiuyu Chen; Zhibo Chen; Wenjun Zeng; Xin Jin*

9. Rethinking Image Super Resolution from Training Data Perspectives; *Go Ohtani*; Ryu Tadokoro; Ryosuke Yamada; Yuki M Asano; Iro Laina; Christian Rupprecht; Nakamasa Inoue; Rio Yokota; Hirokatsu Kataoka; Yoshimitsu Aoki*

10. Confidence-Based Iterative Generation for Real-World Image Super-Resolution; *Jialun Peng; Xin Luo; Jingjing Fu*; Dong Liu**

11. Learned HDR Image Compression for Perceptually Optimal Storage and Display; *Peibei Cao; HAOYU CHEN; Jingzhe Ma; Yu-Chieh Yuan; Zhiyong Xie; Xin Xie; Haiqing Bai; Kede Ma**

12. MoE-DiffIR: Task-customized Diffusion Priors for Universal Compressed Image Restoration; *Yulin Ren; Xin Li*; Bingchen Li; Xingrui Wang; Mengxi China Guo; Shijie Zhao; Li Zhang; Zhibo Chen**

13. Efficient Frequency-Domain Image Deraining with Contrastive Regularization; *Ning Gao; Xingyu Jiang; Xiuhui Zhang; Yue Deng**

14. Restore Anything with Masks: Leveraging Mask Image Modeling for Blind All-in-One Image Restoration; *Chujie Qin; Ruiqi Wu; Zikun Liu; Xin Lin; Chun-Le Guo; Hyun Hee Park; Chongyi Li**

15. Implicit Steganography Beyond the Constraints of Modality; *Sojeong Song*; Seoyun Yang*; Chang D. Yoo*; Junmo Kim**

16. A Geometric Distortion Immunized Deep Watermarking Framework with Robustness Generalizability; *Linfeng Ma; Han Fang*; Tianyi Wei; Zijin Yang; Zehua Ma*; Weiming Zhang; Nenghai Yu*

17. EGIC: Enhanced Low-Bit-Rate Generative Image Compression Guided by Semantic Segmentation; *Nikolai Körber*; Eduard Kromer; Andreas Siebert; Sascha Hauke; Daniel Mueller-Gritschneder; Björn Schuller*

18. Diffusion for Natural Image Matting; *Yihan Hu*; Yiheng Lin; Wei Wang; Yao Zhao; Yunchao Wei*; Humphrey Shi*

19. Blind Image Deconvolution by Generative-based Kernel Prior and Initializer via Latent Encoding; *Jiangtao Zhang; Zongsheng Yue*; Hui Wang; Qian Zhao*; Deyu Meng*

20. TTT-MIM: Test-Time Training with Masked Image Modeling for Denoising Distribution Shifts; *Youssef Mansour*; Xuyang Zhong; Serdar Caglar; Reinhard Heckel*

21. AdaIFL: Adaptive Image Forgery Localization via a Dynamic and Importance-aware Transformer Network; *Yuxi Li*; Fuyuan Cheng; Wangbo Yu; Guangshuo Wang; Guibo Luo*; Yuesheng Zhu**

22. Switch Diffusion Transformer: Synergizing Denoising Tasks with Sparse Mixture-of-Experts; *Byeongjun Park; Hyojun Go; Jin-Young Kim; Sangmin Woo; Seokil Ham; Changick Kim**

23. DiffFAS: Face Anti-Spoofing via Generative Diffusion Models; *Xinxu Ge; Xin Liu*; Zitong Yu*; Jingang Shi; Chun Qi; Jie Li; Heikki Kälviäinen*

24. Debiasing surgeon: fantastic weights and how to find them; *Remi Nahon; Ivan Luiz De Moura Matos; Van-Tam Nguyen; Enzo Tartaglione**

25. Face Reconstruction Transfer Attack as Out-of-Distribution Generalization; *Yoon Gyo Jung*; Jaewoo Park; Xingbo Dong; Hojin Park; Andrew Beng Jin Teoh; Octavia Camps**

26. Improving Robustness to Model Inversion Attacks via Sparse Coding Architectures; *Sayanton V. Dibbo*; Adam Breuer; Juston Moore; Michael Teti*

27. UNIT: Backdoor Mitigation via Automated Neural Distribution Tightening; *Siyuan Cheng*; Guangyu Shen; Kaiyuan Zhang; Guanhong Tao; Shengwei An; Hanxi Guo; Shiqing Ma; Xiangyu Zhang*

28. BAFFLE: A Baseline of Backpropagation-Free Federated Learning; *Haozhe Feng*; Tianyu Pang*; Chao Du; Wei Chen*; Shuicheng Yan; Min Lin*

29. Trainable Highly-expressive Activation Functions; *Irit Chelly*; Shahaf E. Finder; Shira Ifergane; Oren Freifeld*

30. Toward Tiny and High-quality Facial Makeup with Data Amplify Learning; *Qiaoqiao Jin; Xuanhong Chen; Meiguang Jin; Ying Chen; Rui Shi; Yucheng Zheng; Yupeng Zhu; Bingbing Ni**

31. Improving Adversarial Transferability via Model Alignment; *Avery Ma*; Amir-massoud Farahmand; Yangchen Pan; Philip Torr; Jindong Gu*

32. Faceptor: A Generalist Model for Face Perception; *Lixiong Qin*; Mei Wang; Xuannan Liu; Yuhang Zhang; Wei Deng; Xiaoshuai Song; Weiran Xu*; Weihong Deng*

33. HPFF: Hierarchical Locally Supervised Learning with Patch Feature Fusion; *Junhao Su; Chenghao He; Feiyu Zhu; Xiaojie Xu; Dongzhi Guan; Chenyang Si**

34. AdaDistill: Adaptive Knowledge Distillation for Deep Face Recognition; *Fadi Boutros*; Vitomir Struc; Naser Damer*

35. To Supervise or Not to Supervise: Understanding and Addressing the Key Challenges of Point Cloud Transfer Learning; *Souhail Hadgi*; Lei Li; Maks Ovsjanikov*

36. Linearly Controllable GAN: Unsupervised Feature Categorization and Decomposition for Image Generation and Manipulation; *sehyung lee*; Mijung Kim; Yeongnam Chae; Bjorn Stenger*

37. SeA: Semantic Adversarial Augmentation for Last Layer Features from Unsupervised Representation Learning; *Qi Qian*; Yuanhong Xu; Juhua Hu*

38. On the Evaluation Consistency of Attribution-based Explanations; *Jiarui Duan; Haoling Li; Haofei Zhang; Hao Jiang; Mengqi Xue; Li Sun; Mingli Song; Jie Song**

39. MO-EMT-NAS: Multi-Objective Continuous Transfer of Architectural Knowledge Between Tasks from Different Datasets; *PENG LIAO*; Xilu Wang*; Yaochu Jin*; Wenli Du**

40. Facial Affective Behavior Analysis with Instruction Tuning; *Yifan Li*; Anh Dao; Wentao Bao; Zhen Tan; Tianlong Chen; Huan Liu; Yu Kong*

41. Learning Unified Reference Representation for Unsupervised Multi-class Anomaly Detection; *Liren He; Zhengkai Jiang; Jinlong Peng; Wenbing Zhu; Liang Liu; Qiangang Du; Xiaobin Hu; Mingmin Chi*; Yabiao Wang*; Chengjie Wang**

42. Model Breadcrumbs: Scaling Multi-Task Model Merging with Sparse Masks; *MohammadReza Davari*; Eugene Belilovsky*

43. Unsupervised Representation Learning by Balanced Self Attention Matching; *Daniel Shalam*; Simon Korman**

44. Learning Diffusion Models for Multi-View Anomaly Detection; *Chieh Liu*; Yu-Min Chu*; Ting-I Hsieh*; Hwann-Tzong Chen*; Tyng-Luh Liu**

45. SLIM: Spuriousness Mitigation with Minimal Human Annotations; *Xiwei Xuan*; Ziquan Deng; Hsuan-Tien Lin; Kwan-Liu Ma*

46. CipherDM: Secure Three-Party Inference for Diffusion Model Sampling; *Xin Zhao; Xiaojun Chen*; Xudong Chen; He Li; Tingyu Fan; Zhendong Zhao*

47. Auto-GAS: Automated Proxy Discovery for Training-free Generative Architecture Search; *Lujun Li; Haosen Sun; Shiwen Li; Peijie Dong; Wenhan Luo; Wei Xue; Qifeng Liu*; Yike Guo**

48. Simple Unsupervised Knowledge Distillation With Space Similarity; *Aditya Singh*; Haohan Wang*

49. Gradient-based Out-of-Distribution Detection; *Taha Entesari*; Sina Sharifi*; Bardia Safaei*; Vishal Patel; Mahyar Fazlyab*

50. Learning Differentially Private Diffusion Models via Stochastic Adversarial Distillation; *Bochao Liu; Pengju Wang; Shiming Ge**

51. Leveraging Hierarchical Feature Sharing for Efficient Dataset Condensation; *Haizhong Zheng*; Jiachen Sun; Shutong Wu; Bhavya Kailkhura; Zhuoqing Morley Mao; Chaowei Xiao*; Atul Prakash**

52. SlimFlow: Training Smaller One-Step Diffusion Models with Rectified Flow; *Yuanzhi Zhu*; Xingchao Liu; Qiang Liu**

53. Self-Guided Generation of Minority Samples Using Diffusion Models; *Soobin Um; Jong Chul Ye**

54. A Semantic Space is Worth 256 Language Descriptions: Make Stronger Segmentation Models with Descriptive Properties; *Junfei Xiao; Ziqi Zhou; Wenxuan Li; Shiyi Lan; Jieru Mei; Zhiding Yu; Bingchen Zhao; Alan Yuille; Yuyin Zhou; Cihang Xie**

55. Memory-Efficient Fine-Tuning for Quantized Diffusion Model; *Hyogon Ryu; Seohyun Lim; Hyunjung Shim**

56. Certifiably Robust Image Watermark; *Zhengyuan Jiang*; Moyang Guo; Yuepeng Hu; Jinyuan Jia; Neil Zhenqiang Gong*

57. DomainFusion: Generalizing To Unseen Domains with Latent Diffusion Models; *Yuyang Huang; Yabo Chen; Yuchen Liu; xiaopeng zhang*; Wenrui Dai*; Hongkai Xiong; Qi Tian*

58. Distribution Alignment for Fully Test-Time Adaptation with Dynamic Online Data Streams; *Ziqiang Wang*; Zhixiang Chi; Yanan Wu; Li Gu; Zhi Liu*; Konstantinos N Plataniotis*; Yang Wang**

59. Idempotent Unsupervised Representation Learning for Skeleton-Based Action Recognition; *Lilang Lin; Lehong Wu; Jiahang Zhang; Jiaying Liu**

60. MONTAGE: Monitoring Training for Attribution of Generative Diffusion Models; *Jonathan Brokman*; Omer Hofman; Roman Vainshtein; Amit Giloni; Toshiya Shimizu; Inderjeet Singh; Oren Rachmil; Alon Zolfi; Asaf Shabtai; Yuki Unno; Hisashi Kojima*

61. Dataset Growth; *Ziheng Qin*; zhaopan xu; YuKun Zhou; Kai Wang*; Zangwei Zheng; Zebang Cheng; Hao Tang; Lei Shang; Baigui Sun; Radu Timofte; Xiaojiang Peng; Hongxun Yao*; Yang You**

62. Self-Cooperation Knowledge Distillation for Novel Class Discovery; *Yuzheng Wang*; Zhaoyu Chen; Dingkang Yang; Yunquan Sun; Lizhe Qi**

63. To Generate or Not? Safety-Driven Unlearned Diffusion Models Are Still Easy To Generate Unsafe Images ... For Now; *Yimeng Zhang*; jinghan jia; Xin Chen; Aochuan Chen; Yihua Zhang; Jiancheng Liu; Ke Ding; Sijia Liu*

64. Foster Adaptivity and Balance in Learning with Noisy Labels; *Mengmeng Sheng; Zeren Sun*; Tao Chen; Shuchao Pang; yucheng wang; Yazhou Yao**

65. Can OOD Object Detectors Learn from Foundation Models?; *Jiahui Liu*; Xin Wen; Shizhen Zhao; Yingxian Chen; Xiaojuan Qi**

66. Optimal Transport of Diverse Unsupervised Tasks for Robust Learning from Noisy Few-Shot Data; *Xiaofan Que; Qi Yu**

67. The Gaussian Discriminant Variational Autoencoder (GdVAE): A Self-Explainable Model with Counterfactual Explanations; *Anselm Haselhoff*; Kevin Trelenberg; Fabian Küppers; Jonas Schneider*

68. Exemplar-free Continual Representation Learning via Learnable Drift Compensation; *Alex Gomez-Villa*; Dipam Goswami; Kai Wang; Andy Bagdanov; Bartlomiej Twardowski; Joost van de Weijer*

69. Few-shot Class Incremental Learning with Attention-Aware Self-Adaptive Prompt; *Chenxi Liu*; Zhenyi Wang; Tianyi Xiong; Ruibo Chen; Yihan Wu; junfeng guo; Heng Huang**

70. Rebalancing Using Estimated Class Distribution for Imbalanced Semi-Supervised Learning under Class Distribution Mismatch; *Taemin Park; Hyuck Lee; Heeyoung Kim**

71. Diagnosing and Re-learning for Balanced Multimodal Learning; *Yake Wei; Siwei Li; Ruoxuan Feng; Di Hu**

72. Non-Exemplar Domain Incremental Learning via Cross-Domain Concept Integration; *Qiang Wang*; Yuhang He; Songlin Dong; Xinyuan Gao; Shaokun Wang; Yihong Gong*

73. Which Model Generated This Image? A Model-Agnostic Approach for Origin Attribution; *Fengyuan Liu; Haochen Luo; Yiming Li; Philip Torr; Jindong Gu**

74. Open-World Dynamic Prompt and Continual Visual Representation Learning; *Youngeun Kim; Jun Fang*; Qin Zhang; Zhaowei Cai; Yantao Shen; Rahul Duggal; Dripta S. Raychaudhuri; Zhuowen Tu; Yifan Xing; Onkar Dabeer*

75. Continual Learning and Unknown Object Discovery in 3D Scenes via Self-Distillation; *Mohamed El Amine Boudjoghra*; Jean Lahoud; Salman Khan; Hisham Cholakkal; Rao M Anwer; Fahad Shahbaz Khan*

76. Scaling Up Personalized Image Aesthetic Assessment via Task Vector Customization; *Jooyeol Yun*; Jaegul Choo*

77. AlignDiff: Aligning Diffusion Models for General Few-Shot Segmentation; *Ri-Zhao Qiu*; Yu-Xiong Wang; Kris Hauser*

78. Factorized Diffusion: Perceptual Illusions by Noise Decomposition; *Daniel Geng*; Inbum Park; Andrew Owens*

79. Efficient 3D-Aware Facial Image Editing via Attribute-Specific Prompt Learning; *Amandeep Kumar*; Muhammad Awais; Sanath Narayan; Hisham Cholakkal; Salman Khan; Rao Muhammad Anwer*

80. ZipLoRA: Any Subject in Any Style by Effectively Merging LoRAs; *Viraj Shah; Nataniel Ruiz; Forrester Cole; Erika Lu; Svetlana Lazebnik; Yuanzhen Li; Varun Jampani**

81. OMG: Occlusion-friendly Personalized Multi-concept Generation in Diffusion Models; *Zhe Kong*; Yong Zhang*; Tianyu Yang; Tao Wang; Kaihao Zhang; Bizhu Wu; Guanying Chen; Wei Liu; Wenhan Luo**

82. JointDreamer: Ensuring Geometry Consistency and Text Congruence in Text-to-3D Generation via Joint Score Distillation; *ChenHan Jiang*; Yihan Zeng; Tianyang Hu; Songcen Xu; Wei Zhang; Hang Xu; Dit-Yan Yeung*

83. WAVE: Warping DDIM Inversion Features for Zero-shot Text-to-Video Editing; *Yutang Feng; Sicheng Gao*; Yuxiang Bao; Xiaodi Wang; Shumin Han*; Juan Zhang*; Baochang Zhang; Angela Yao*

84. DiffusionPen: Towards Controlling the Style of Handwritten Text Generation; *Konstantina Nikolaidou*; George Retsinas; Giorgos Sfikas; Marcus Liwicki*

85. RegionDrag: Fast Region-Based Image Editing with Diffusion Models; *Jingyi Lu; Xinghui Li; Kai Han**

86. TurboEdit: Real-time text-based disentangled real image editing; *Zongze Wu*; Nicholas I Kolkin; Jonathan Brandt; Richard Zhang; Eli Shechtman*

87. HiDiffusion: Unlocking Higher-Resolution Creativity and Efficiency in Pretrained Diffusion Models; *Shen Zhang; Zhaowei CHEN; Zhenyu Zhao; Yuhao Chen; Yao Tang; Jiajun Liang**

88. Few-shot Defect Image Generation based on Consistency Modeling; *Qingfeng Shi; Jing Wei; Fei Shen*; Zhengtao Zhang*

89. Be Yourself: Bounded Attention for Multi-Subject Text-to-Image Generation; *Omer Dahary*; Or Patashnik; Kfir Aberman; Danny Cohen-Or*

90. Merging and Splitting Diffusion Paths for Semantically Coherent Panoramas; *Fabio Quattrini*; Vittorio Pippi; Silvia Cascianelli*; Rita Cucchiara*

91. Object-Conditioned Energy-Based Attention Map Alignment in Text-to-Image Diffusion Models; *Yasi Zhang*; Peiyu Yu; Ying Nian Wu*

92. FontStudio: Shape-Adaptive Diffusion Model for Coherent and Consistent Font Effect Generation; *Xinzhi Mu*; Li Chen; Bohan CHEN; Shuyang Gu; Jianmin Bao; Dong Chen; Ji Li; Yuhui Yuan*

93. Viewpoint textual inversion: discovering scene representations and 3D view control in 2D diffusion models; *James Burgess*; Kuan-Chieh Wang; Serena Yeung-Levy*

94. AnyControl: Create Your Artwork with Versatile Control on Text-to-Image Generation; *Yanan Sun*; Yanchen Liu; Yinhao Tang; Wenjie Pei; Kai Chen*

95. BK-SDM: A Lightweight, Fast, and Cheap Version of Stable Diffusion; *Bo-Kyeong Kim*; Hyoung-Kyu Song; Thibault Castells; Shinkook Choi*

96. ProTIP: Probabilistic Robustness Verification on Text-to-Image Diffusion Models against Stochastic Perturbation; *Yi Zhang; Yun Tang; Wenjie Ruan; Xiaowei Huang; Siddartha Khastgir; Paul A Jennings; Xingyu Zhao**

97. Deep Reward Supervisions for Tuning Text-to-Image Diffusion Models; *Xiaoshi Wu; Yiming Hao; Manyuan Zhang*; Keqiang Sun; Zhaoyang Huang; Guanglu Song; Yu Liu; Hongsheng Li**

98. HARIVO: Harnessing Text-to-Image Models for Video Generation; *Mingi Kwon; Seoung Wug Oh; Yang Zhou; Joon-Young Lee; Difan Liu; Haoran Cai; Baqiao Liu; Feng Liu; Youngjung Uh**

99. Training-free Composite Scene Generation for Layout-to-Image Synthesis; *Jiaqi Liu*; Tao Huang; Chang Xu*

100. Diverse Text-to-3D Synthesis with Augmented Text Embedding; *Uy Dieu Tran*; Minh N. Hoang Luu*; Phong Ha Nguyen*; Khoi Nguyen*; Binh-Son Hua**

101. Skews in the Phenomenon Space Hinder Generalization in Text-to-Image Generation; *Yingshan Chang*; Yasi Zhang; Zhiyuan Fang; Ying Nian Wu; Yonatan Bisk; Feng Gao*

102. Unleashing Text-to-Image Diffusion Prior for Zero-Shot Image Captioning; *Jianjie Luo; Jingwen Chen; Yehao Li; Yingwei Pan*; Jianlin Feng; Hongyang Chao; Ting Yao*

103. RAVE: Residual Vector Embedding for CLIP-Guided Backlit Image Enhancement; *Tatiana Gaintseva*; Martin Benning; Gregory Slabaugh**

104. ReGround: Improving Textual and Spatial Grounding at No Cost; *Phillip Y. Lee; Minhyuk Sung**

105. Enriching Information and Preserving Semantic Congruence in Expanding Curvilinear Object Segmentation Datasets; *Qin Lei*; Jiang Zhong; Qizhu Dai*

106. COHO: Context-Sensitive City-Scale Hierarchical Urban Layout Generation; *Liu He*; Daniel Aliaga*

107. cDP-MIL: Robust Multiple Instance Learning via Cascaded Dirichlet Process; *Yihang Chen; Tsai Hor Chan; Guosheng Yin; Yuming Jiang; Lequan Yu**

108. Learning with Counterfactual Explanations for Radiology Report Generation; *Mingjie Li*; Haokun Lin; Liang Qiu; Xiaodan Liang*; Ling Chen; Abdulmotaleb Elsaddik; Xiaojun Chang*

109. Pathology-knowledge Enhanced Multi-instance Prompt Learning for Few-shot Whole Slide Image Classification; *Linhao Qu*; Dingkang Yang; Dan Huang; Qinhao Guo; rongkui luo; Shaoting Zhang; Xiaosong Wang**

110. Alternate Diverse Teaching for Semi-supervised Medical Image Segmentation; *Zhen Zhao*; Zicheng Wang; Dian Yu; Longyue Wang*; Yixuan Yuan; Luping Zhou*

111. Efficient Active Domain Adaptation for Semantic Segmentation by Selecting Information-rich Superpixels; *Yuan Gao; Zilei Wang*; Yixin Zhang; Bohai Tu*

112. Unleashing the Power of Prompt-driven Nucleus Instance Segmentation; *Zhongyi Shui*; Yunlong Zhang; Kai Yao; Chenglu Zhu; Sunyi Zheng; Jingxiong Li; Honglin Li; YUXUAN SUN; Ruizhe Guo; Lin Yang**

113. Beyond Pixels: Semi-Supervised Semantic Segmentation with a Multi-scale Patch-based Multi-Label Classifier; *Prantik Howlader*; Srijan Das; Hieu Le; Dimitris Samaras*

114. Weakly Supervised Co-training with Swapping Assignments for Semantic Segmentation; *Xinyu Yang*; Hossein Rahmani; Dame S Black; Bryan M Williams*

115. Improving Medical Multi-modal Contrastive Learning with Expert Annotations; *Yogesh Kumar*; Pekka Marttinen*

116. CAT-SAM: Conditional Tuning for Few-Shot Adaptation of Segment Anything Model; *Aoran Xiao; Weihao Xuan; Heli Qi; Yun Xing; Ruijie Ren; Xiaoqin Zhang; Ling Shao; Shijian Lu**

117. Pseudo-Labelling Should Be Aware of Disguising Channel Activations; *Changrui Chen; Kurt Debattista; Jungong Han**

118. Modeling Label Correlations with Latent Context for Multi-Label Recognition; *Zhaomin Chen*; Quan Cui; Ruoxi Deng; Jie Hu; Guodao Zhang**

119. Pro2SAM: Mask Prompt to SAM with Grid Points for Weakly Supervised Object Localization; *Xi Yang; Songsong Duan*; Nannan Wang; Xinbo Gao*

120. SINDER: Repairing the Singular Defects of DINOv2; *Haoqi Wang; Tong Zhang; Mathieu Salzmann**

121. Placing Objects in Context via Inpainting for Out-of-distribution Segmentation; *Pau de Jorge Aranda*; Riccardo Volpi; Puneet Dokania; Philip Torr; Gregory Rogez*

122. Enhancing Optimization Robustness in 1-bit Neural Networks through Stochastic Sign Descent; *NianHui Guo*; Hong Guo; Christoph Meinel; Haojin Yang*

123. SAM-COD: SAM-guided Unified Framework for Weakly-Supervised Camouflaged Object Detection; *Huafeng Chen; Pengxu Wei; Guangqian Guo; Shan Gao**

124. Finding NeMo: Negative-mined Mosaic Augmentation for Referring Image Segmentation; *Seongsu Ha; Chaeyun Kim; Donghwa Kim; Junho Lee; Sangho Lee; Joonseok Lee**

125. Preventing Catastrophic Forgetting through Memory Networks in Continuous Detection; *Gaurav Bhatt*; Leonid Sigal; James Ross*

126. Leveraging Text Localization for Scene Text Removal via Text-aware Masked Image Modeling; *Zixiao Wang*; Hongtao Xie; YuXin Wang; Yadong Qu; Fengjun Guo; Pengwei Liu*

127. SAFARI: Adaptive Sequence Transformer for Weakly Supervised Referring Expression Segmentation; *Sayan Nag*; Koustava Goswami; Srikrishna Karanam*

128. Long-CLIP: Unlocking the Long-Text Capability of CLIP; *Beichen Zhang*; Pan Zhang; Xiaoyi Dong*; Yuhang Zang; Jiaqi Wang**

130. DSA: Discriminative Scatter Analysis for Early Smoke Segmentation; *Lujian Yao*; Haitao Zhao*; Jingchao Peng; Zhongze Wang; Kaijie Zhao*

131. CSOT: Cross-Scan Object Transfer for Semi-Supervised LiDAR Object Detection; *Jinglin Zhan; Tiejun Liu; Rengang Li; Zhaoxiang Zhang; Yuntao Chen**

132. Platypus: A Generalized Specialist Model for Reading Text in Various Forms; *Peng Wang; Zhaohai Li; Jun Tang; Humen Zhong; Fei Huang; Zhibo Yang*; Cong Yao**

133. Localization and Expansion: A Decoupled Framework for Point Cloud Few-shot Semantic Segmentation; *Zhaoyang Li*; Yuan Wang; Wangkai Li; Rui Sun; Tianzhu Zhang*

134. Segment3D: Learning Fine-Grained Class-Agnostic 3D Segmentation without Manual Labels; *Rui Huang; Songyou Peng; Ayca Takmaz; Federico Tombari; Marc Pollefeys; Shiji Song; Gao Huang*; Francis Engelmann*

135. Bayesian Self-Training for Semi-Supervised 3D Segmentation; *Ozan Unal*; Christos Sakaridis; Luc Van Gool*

136. Context-Guided Spatial Feature Reconstruction for Efficient Semantic Segmentation; *Zhenliang Ni; Xinghao Chen*; Yingjie Zhai; Yehui Tang; Yunhe Wang**

137. Tokenize Anything via Prompting; *Ting Pan*; Lulu Tang; Xinlong Wang*; Shiguang Shan*

138. ReMamber: Referring Image Segmentation with Mamba Twister; *Yuhuan Yang; Chaofan Ma; Jiangchao Yao; Zhun Zhong*; Ya Zhang; Yanfeng Wang**

139. Salience-Based Adaptive Masking: Revisiting Token Dynamics for Enhanced Pre-training; *Hyesong Choi; Hyejin Park; Kwang Moo Yi; Sungmin Cha; Dongbo Min**

140. Benchmarking Object Detectors with COCO: A New Path Forward; *Shweta Singh; Aayan Yadav; Jitesh Jain; Humphrey Shi; Justin Johnson; Karan Desai**

141. Visual Prompting via Partial Optimal Transport; *Mengyu Zheng*; Zhiwei Hao; Yehui Tang; Chang Xu**

142. PartGLEE: A Foundation Model for Recognizing and Parsing Any Objects; *Junyi Li; Junfeng Wu; Weizhi Zhao; Song Bai; Xiang Bai**

143. Efficient and Versatile Robust Fine-Tuning of Zero-shot Models; *Sungyeon Kim*; Boseung Jeong; Donghyun Kim; Suha Kwak**

144. Multi-Memory Matching for Unsupervised Visible-Infrared Person Re-Identification; *Jiangming Shi; Xiangbo Yin; Yeyun Chen; Yachao Zhang; Zhizhong Zhang; Yuan Xie*; Yanyun Qu**

145. Learning Multimodal Latent Generative Models with Energy-Based Prior; *Shiyu Yuan*; Jiali Cui; Hanao Li; Tian Han*

146. WPS-SAM: Towards Weakly-Supervised Part Segmentation with Foundation Models; *Xin-Jian Wu*; Ruisong Zhang; Jie Qin; Shijie Ma; Cheng-Lin Liu**

147. Diffusion Models for Open-Vocabulary Segmentation; *Laurynas Karazija*; Iro Laina; Andrea Vedaldi; Christian Rupprecht*

148. Emergent Visual-Semantic Hierarchies in Image-Text Representations; *Morris Alper*; Hadar Averbuch-Elor*

149. Collaborative Vision-Text Representation Optimizing for Open-Vocabulary Segmentation; *Siyu Jiao*; hongguang Zhu; Yunchao Wei; Yao Zhao*; Jiannan Huang; Humphrey Shi*

150. Denoising Vision Transformers; *Jiawei Yang*; Katie Z Luo; Jiefeng Li; Congyue Deng; Leonidas Guibas; Dilip Krishnan; Kilian Weinberger; Yonglong Tian; Yue Wang*

151. Robust Fitting on a Gate Quantum Computer; *Frances F Yang*; Michele Sasdelli; Tat-Jun Chin* BEST PAPER CANDIDATE

152. Learning Modality-agnostic Representation for Semantic Segmentation from Any Modalities; *Xu Zheng*; Yuanhuiyi Lyu; Lin Wang**

153. Geospecific View Generation - Geometry-Context Aware High-resolution Ground View Inference from Satellite Views; *Ningli Xu; Rongjun Qin**

154. Introducing Routing Functions to Vision-Language Parameter-Efficient Fine-Tuning with Low-Rank Bottlenecks; *Tingyu Qu*; Tinne Tuytelaars; Marie-Francine Moens*

155. OpenKD: Opening Prompt Diversity for Zero- and Few-shot Keypoint Detection; *Changsheng Lu*; Zheyuan Liu; Piotr Koniusz**

156. MM-SafetyBench: A Benchmark for Safety Evaluation of Multimodal Large Language Models; *Xin Liu*; Yichen Zhu; Jindong Gu; Yunshi Lan; Chao Yang; Yu Qiao*

157. Griffon: Spelling out All Object Locations at Any Granularity with Large Language Models; *Yufei Zhan; Yousong Zhu*; Zhiyang Chen; Fan Yang; Ming Tang; Jinqiao Wang*

158. Exploiting Semantic Reconstruction to Mitigate Hallucinations in Vision-Language Models; *Minchan Kim; Minyeong Kim; Junik Bae; Suhwan Choi; Sungkyung Kim; Buru Chang**

159. CoLA: Conditional Dropout and Language-driven Robust Dual-modal Salient Object Detection; *Shuang Hao; Chunlin Zhong; He Tang**

160. UMG-CLIP: A Unified Multi-Granularity Vision Generalist for Open-World Understanding; *Bowen Shi; Peisen Zhao; Zichen Wang; Yuhang Zhang; Yaoming Wang; Jin Li; Wenrui Dai; Junni Zou; Hongkai Xiong; Qi Tian; Xiaopeng Zhang**

161. Decoupling Common and Unique Representations for Multimodal Self-supervised Learning; *Yi Wang*; Conrad M Albrecht; Nassim Ait Ali Braham; Chenying Liu; Zhitong Xiong; Xiao Xiang Zhu*

162. Boosting Transferability in Vision-Language Attacks via Diversification along the Intersection Region of Adversarial Trajectory; *Sensen Gao; Xiaojun Jia*; Xuhong Ren; Ivor Tsang; Qing Guo**

163. Multi-Task Domain Adaptation for Language Grounding with 3D Objects; *Penglei Sun; Yaoxian Song; Xinglin Pan; Peijie Dong; Xiaofei Yang; Qiang Wang*; Zhixu Li; Tiefeng Li; Xiaowen Chu**

164. How Many Unicorns Are in This Image? A Safety Evaluation Benchmark for Vision LLMs; *Haoqin Tu*; Chenhang Cui; Zijun Wang; Yiyang Zhou; Bingchen Zhao; Junlin Han; Wangchunshu Zhou; Huaxiu Yao; Cihang Xie**

165. LLaVA-Grounding: Grounded Visual Chat with Large Multimodal Models; *Hao Zhang*; Hongyang Li; Feng Li; Tianhe Ren; Xueyan Zou; Shilong Liu; Shijia Huang; Jianfeng Gao; Lei Zhang; Chunyuan Li; Jianwei Yang*

166. Q&A Prompts: Discovering Rich Visual Clues through Mining Question-Answer Prompts for VQA requiring Diverse World Knowledge; *Haibo Wang*; Weifeng Ge**

167. An Efficient and Effective Transformer Decoder-Based Framework for Multi-Task Visual Grounding; *Wei Chen; Long Chen; Yu Wu**

168. DQ-DETR: DETR with Dynamic Query for Tiny Object Detection; *Yi-Xin Huang*; Hou-I Liu; Hong-Han Shuai; Wen-Huang Cheng*

169. Siamese Vision Transformers are Scalable Audio-visual Learners; *Yan-Bo Lin*; Gedas Bertasius*

170. Take A Step Back: Rethinking the Two Stages in Visual Reasoning; *Mingyu Zhang; Jiting Cai; Mingyu Liu; Yue Xu; Cewu Lu; Yong-Lu Li**

171. Token Compensator: Altering Inference Cost of Vision Transformer without Re-Tuning; *Shibo Jie; Yehui Tang; Jianyuan Guo; Zhi-Hong Deng*; Kai Han*; Yunhe Wang**

172. MathVerse: Does Your Multi-modal LLM Truly See the Diagrams in Visual Math Problems?; *Renrui Zhang; Dongzhi Jiang; Yichi Zhang; Haokun Lin; Ziyu Guo; Pengshuo Qiu; Aojun Zhou; Pan Lu; Kai-Wei Chang; Peng Gao; Hongsheng Li**

173. Tiny Models are the Computational Saver for Large Models; *Qingyuan Wang*; Barry Cardiff; Antoine Frappé; Benoit Larras; Deepu John**

174. Efficient Vision Transformers with Partial Attention; *Xuan-Thuy Vo*; Duy-Linh Nguyen; Adri Priadana; Kang-Hyun Jo**

175. SNP: Structured Neuron-level Pruning to Preserve Attention Scores; *KyungHwan Shim; Jaewoong Yun; Shinkook Choi**

176. GRA: Detecting Oriented Objects through Group-wise Rotating and Attention; *Jiangshan Wang*; Yifan Pu; Yizeng Han; Jiayi Guo; Yiru Wang; Xiu Li*; Gao Huang**

177. HYDRA: A Hyper Agent for Dynamic Compositional Visual Reasoning; *Fucai Ke*; Zhixi Cai; Simindokht Jahangard; Weiqing Wang; Pari Delir Haghighi; Hamid Rezatofighi*

178. AMES: Asymmetric and Memory-Efficient Similarity Estimation for Instance-level Retrieval; *Pavel Suma*; Giorgos Kordopatis-Zilos; Ahmet Iscen; Giorgos Tolias*

179. Fast Encoding and Decoding for Implicit Video Representation; *Hao Chen*; Saining Xie; Ser-Nam Lim; Abhinav Shrivastava*

180. Language-Driven 6-DoF Grasp Detection Using Negative Prompt Guidance; *Toan Nguyen; Minh Nhat Nhat Vu; Baoru Huang; An Dinh Vuong; Quan Vuong; Ngan Le; Thieu Vo; Anh Nguyen**

181. PiTe: Pixel-Temporal Alignment for Large Video-Language Model; *Yang Liu*; Pengxiang Ding; Siteng Huang; Min Zhang; Han Zhao; Donglin Wang*

182. CoLeaF: A Contrastive-Collaborative Learning Framework for Weakly Supervised Audio-Visual Video Parsing; *Faegheh Sardari*; Armin Mustafa; Philip JB Jackson; Adrian Hilton*

183. MEVG : Multi-event Video Generation with Text-to-Video Models; *Gyeongrok Oh*; Jaehwan Jeong; Sieun Kim; Wonmin Byeon; Jinkyu Kim; Sungwoong Kim; Sangpil Kim**

184. Rethinking Weakly-supervised Video Temporal Grounding From a Game Perspective; *Xiang Fang; Zeyu Xiong; Wanlong Fang; Xiaoye Qu; Chen Chen; Jianfeng Dong; Keke Tang; Pan Zhou*; Yu Cheng; Daizong Liu**

185. Contextual Correspondence Matters: Bidirectional Graph Matching for Video Summarization; *Yunzuo Zhang*; Yameng Liu*

186. Weakly-Supervised Spatio-Temporal Video Grounding with Variational Cross-Modal Alignment; *Yang Jin*; Yadong Mu**

187. Delving Deep into Engagement Prediction of Short Videos; *dasong Li; Wenjie Li; Baili Lu; Hongsheng Li; Sizhuo Ma; Gurunandan Krishnan; Jian Wang**

188. LITA: Language Instructed Temporal-Localization Assistant; *De-An Huang*; Shijia Liao; Subhashree Radhakrishnan; Hongxu Yin; Pavlo Molchanov; Zhiding Yu; Jan Kautz*

189. Frequency-Spatial Entanglement Learning for Camouflaged Object Detection; *Yanguang Sun; Chunyan Xu; Jian Yang; Hanyu Xuan*; Lei Luo**

190. Reinforcement Learning Friendly Vision-Language Model for Minecraft; *Haobin Jiang; Junpeng Yue; Hao Luo; Ziluo Ding; Zongqing Lu**

191. UniMD: Towards Unifying Moment Retrieval and Temporal Action Detection; *Yingsen Zeng; Yujie Zhong*; Chengjian Feng; Lin Ma*

192. TimeCraft: Navigate Weakly-Supervised Temporal Grounded Video Question Answering via Bi-directional Reasoning; *Huabin Liu; Xiao Ma; Cheng Zhong; Yang Zhang; Weiyao Lin**

193. AMEGO: Active Memory from long EGOcentric videos; *Gabriele Goletto*; Tushar Nagarajan; Giuseppe Averta; Dima Damen*

194. See and Think: Embodied Agent in Virtual Environment; *Zhonghan Zhao; Xuan Wang; Wenhao Chai; Boyi Li; Shengyu Hao; Shidong Cao; Tian Ye; Gaoang Wang**

195. STSP: Spatial-Temporal Subspace Projection for Video Class-incremental Learning; *Hao Cheng; SIYUAN YANG; Chong Wang; Joey Tianyi Zhou; Alex Kot; Bihan Wen**

196. VideoClusterNet: Self-Supervised and Adaptive Face Clustering for Videos; *Devesh Walawalkar*; Pablo Garrido*

197. EvSign: Sign Language Recognition and Translation with Streaming Events; *Pengyu Zhang*; Hao Yin; Zeren Wang; Wenyue Chen; Sheng Ming Li; Dong Wang; Huchuan Lu; Xu Jia*

198. Follow the Rules: Reasoning for Video Anomaly Detection with Large Language Models; *Yuchen Yang*; Kwonjoon Lee; Behzad Dariush; Yinzhi Cao*; Shao-Yuan Lo**

199. Data Collection-free Masked Video Modeling; *Yuchi Ishikawa*; Masayoshi Kondo; Yoshimitsu Aoki*

200. Discovering Novel Actions from Open World Egocentric Videos with Object-Grounded Visual Commonsense Reasoning; *Sanjoy Kundu; Shubham Trehan; Sathyanarayanan N Aakur**

201. VISAGE: Video Instance Segmentation with Appearance-Guided Enhancement; *Hanjung Kim; Jaehyun Kang; Miran Heo; Sukjun Hwang; Seoung Wug Oh; Seon Joo Kim**

202. ActionSwitch: Class-agnostic Detection of Simultaneous Actions in Streaming Videos; *Hyolim Kang; Jeongseok Hyun; Joungbin An; Youngjae Yu; Seon Joo Kim**

203. Trajectory-aligned Space-time Tokens for Few-shot Action Recognition; *Pulkit Kumar*; Namitha Padmanabhan; Luke Luo; Sai Saketh Rambhatla; Abhinav Shrivastava*

204. Interactive 3D Object Detection with Prompts; *Ruifei Zhang; Xiangru Lin; Wei Zhang; Jincheng Lu; Xuekuan Wang; Xiao Tan; Yingying Li; Errui Ding; Jingdong Wang; Guanbin Li**

205. ActionVOS: Actions as Prompts for Video Object Segmentation; *Liangyang Ouyang*; Ruicong Liu; Yifei Huang*; Ryosuke Furuta; Yoichi Sato**

206. Empowering Embodied Visual Tracking with Visual Foundation Models and Offline RL; *Fangwei Zhong*; Kui Wu; Hai Ci; Chu-ran Wang; Hao Chen*

207. Improving Video Segmentation via Dynamic Anchor Queries; *Yikang Zhou; Tao Zhang*; Xiangtai Li*; Shunping Ji*; Shuicheng Yan*

208. MaxMI: A Maximal Mutual Information Criterion for Manipulation Concept Discovery; *Pei Zhou; Yanchao Yang**

209. Robo-ABC: Affordance Generalization Beyond Categories via Semantic Correspondence for Robot Manipulation; *Yuanchen Ju; Kaizhe Hu; Guowei Zhang; Gu Zhang; Mingrun Jiang; Huazhe Xu**

210. DISCO: Embodied Navigation and Interaction via Differentiable Scene Semantics and Dual-level Control; *Xinyu Xu*; Shengcheng Luo; Yanchao Yang; Yong-Lu Li*; Cewu Lu**

211. F-HOI: Toward Fine-grained Semantic-Aligned 3D Human-Object Interactions; *Jie Yang; Xuesong Niu; Nan Jiang; Ruimao Zhang*; Siyuan Huang**

212. PoseEmbroider: Towards a 3D, Visual, Semantic-aware Human Pose Representation; *Ginger Delmas*; Philippe Weinzaepfel; Francesc Moreno-Noguer; Gregory Rogez*

213. Bridging the Gap Between Human Motion and Action Semantics via Kinematics Phrases; *Xinpeng Liu; Yong-Lu Li*; Ailing Zeng; Zizheng Zhou; Yang You; Cewu Lu**

214. Self-Supervised Any-Point Tracking by Contrastive Random Walks; *Ayush Shrivastava*; Andrew Owens*

215. Local Action-Guided Motion Diffusion Model for Text-to-Motion Generation; *Peng Jin*; Hao Li; Zesen Cheng; Kehan Li; Runyi Yu; Chang Liu*; Xiangyang Ji; Li Yuan*; Jie Chen*

216. CoMo: Controllable Motion Generation through Language Guided Pose Code Editing; *Yiming Huang*; Weilin Wan; Yue Yang; Chris Callison-Burch; Mark Yatskar; Lingjie Liu*

217. CadVLM: Bridging Language and Vision in the Generation of Parametric CAD Sketches; *Sifan Wu*; Amir Hosein Khasahmadi; Mor Katz; Pradeep Kumar Jayaraman; Yewen Pu; Karl D.D. Willis; Bang Liu**

218. Exploring the Feature Extraction and Relation Modeling For Light-Weight Transformer Tracking; *Jikai Zheng; Mingjiang Liang; Shaoli Huang; Jifeng Ning**

219. PoseSOR: Human Pose Can Guide Our Attention; *Huankang Guan; Rynson W.H. Lau**

220. PapMOT: Exploring Adversarial Patch Attack against Multiple Object Tracking; *Jiahuan Long*; Tingsong Jiang*; Wen Yao*; Shuai Jia*; Weijia Zhang*; Weien Zhou*; Chao Ma*; Xiaoqian Chen**

221. EgoExo-Fitness: Towards Egocentric and Exocentric Full-Body Action Understanding; *Yuan-Ming Li; Wei-Jin Huang; An-Lan Wang; Ling-An Zeng; Jing-Ke Meng*; Wei-Shi Zheng**

222. Merlin: Empowering Multimodal LLMs with Foresight Minds; *En Yu; Liang Zhao; YANA WEI; Jinrong Yang; Dongming Wu; Lingyu Kong; Haoran Wei; Tiancai Wang; Zheng Ge; Xiangyu Zhang; Wenbing Tao**

223. FARSE-CNN: Fully Asynchronous, Recurrent and Sparse Event-Based CNN; *Riccardo Santambrogio*; Marco Cannici; Matteo Matteucci*

224. OMR: Occlusion-Aware Memory-Based Refinement for Video Lane Detection; *Dongkwon Jin; Chang-Su Kim**

225. WTS: A Pedestrian-Centric Traffic Video Dataset for Fine-grained Spatial-Temporal Understanding; *Quan Kong*; Yuki Kawana; Rajat Saini; Ashutosh Kumar; Jingjing Pan; Ta Gu; Yohei Ozao; Balazs Opra; Yoichi Sato; Norimasa Kobori*

226. Event-Adapted Video Super-Resolution; *Zeyu Xiao; Dachun Kai; Yueyi Zhang; Zheng-Jun Zha; Xiaoyan Sun; Zhiwei Xiong**

227. Temporal Event Stereo via Joint Learning with Stereoscopic Flow; *Hoonhee Cho; Jae-Young Kang; Kuk-Jin Yoon**

228. WiMANS: A Benchmark Dataset for WiFi-based Multi-user Activity Sensing; *Shuokang Huang*; Kaihan Li; Di You; Yichong Chen; Arvin Lin; Siying Liu; Xiaohui Li; Julie A. McCann**

229. Motion-prior Contrast Maximization for Dense Continuous-Time Motion Estimation; *Friedhelm Hamann*; Ziyun Wang; Ioannis Asmanis; Kenneth Chaney; Guillermo Gallego; Kostas Daniilidis*

230. Probabilistic Weather Forecasting with Deterministic Guidance-based Diffusion Model; *Donggeun Yoon; Minseok Seo; Doyi Kim; Yeji Choi; Donghyeon Cho**

231. Adaptive Human Trajectory Prediction via Latent Corridors; *Neerja Thakkar*; Karttikeya Mangalam; Andrea Bajcsy; Jitendra Malik*

232. Modelling Competitive Behaviors in Autonomous Driving Under Generative World Model; *Guanren Qiao; Guiliang Liu*; Guorui Quan; Rongxiao Qu*

233. Diffusion Models as Optimizers for Efficient Planning in Offline RL; *Renming Huang; Yunqiang Pei; Guoqing Wang*; Yangming Zhang; Yang Yang; Peng Wang; Heng Tao Shen*

234. Early Anticipation of Driving Maneuvers; *Abdul Wasi Lone; Shankar Gangisetty*; Shyam Nandan Rai; C. V. Jawahar*

235. Accelerating Online Mapping and Behavior Prediction via Direct BEV Feature Attention; *Xunjiang Gu; Guanyu Song; Igor Gilitschenski; Marco Pavone; Boris Ivanovic**

236. SimPB: A Single Model for 2D and 3D Object Detection from Multiple Cameras; *Yingqi Tang; Zhaotie Meng; Guoliang Chen; Erkang Cheng**

237. UniTraj: A Unified Framework for Scalable Vehicle Trajectory Prediction; *Lan Feng; Mohammadhossein Bahari*; Kaouther Messaoud; Eloi Zablocki; Matthieu Cord; Alexandre Alahi*

238. VEON: Vocabulary-Enhanced Occupancy Prediction; *Jilai Zheng; Pin Tang; Zhongdao Wang; Guoqing Wang; Xiangxuan Ren; Bailan Feng; Chao Ma**

239. OccGen: Generative Multi-modal 3D Occupancy Prediction for Autonomous Driving; *Guoqing Wang; Zhongdao Wang; Pin Tang; Jilai Zheng; Xiangxuan Ren; Bailan Feng; Chao Ma**

240. Wear-Any-Way: Manipulable Virtual Try-on via Sparse Correspondence Alignment; *Mengting Chen*; Xi Chen; Zhonghua Zhai; Chen Ju; Xuewen Hong; Jinsong Lan; Shuai Xiao*

241. SkyScenes: A Synthetic Dataset for Aerial Scene Understanding; *Sahil S Khose*; Anisha Pal; Aayushi Agarwal; . Deepanshi; Judy Hoffman; Prithvijit Chattopadhyay*

242. CoMusion: Towards Consistent Stochastic Human Motion Prediction via Motion Diffusion; *Jiarui Sun*; Girish Chowdhary**

243. HENet: Hybrid Encoding for End-to-end Multi-task 3D Perception from Multi-view Cameras; *Zhongyu Xia; ZhiWei Lin; Xinhao Wang; Yongtao Wang*; Yun Xing; Shengxiang Qi; Nan Dong; Ming-Hsuan Yang*

244. Revisit Anything: Visual Place Recognition via Image Segment Retrieval; *Kartik Garg; Sai Shubodh; Shishir N Y Kolathaya; Madhava Krishna; Sourav Garg**

245. Weakly Supervised 3D Object Detection via Multi-Level Visual Guidance; *Kuan-Chih Huang*; Yi-Hsuan Tsai; Ming-Hsuan Yang*

246. Equivariant Spatio-Temporal Self-Supervision for LiDAR Object Detection; *Deepti Hegde; Suhas Lohit*; Kuan-Chuan Peng*; Michael J. Jones; Vishal M. Patel*

247. Align before Collaborate: Mitigating Feature Misalignment for Robust Multi-Agent Perception; *Dingkang Yang; Dingkang Yang; Ke Li; Dongling Xiao; Zedian Shao; Peng Sun; Liang Song**

248. Formula-Supervised Visual-Geometric Pre-training; *Ryosuke Yamada*; Kensho Hara*; Hirokatsu Kataoka; Koshi Makihara; Nakamasa Inoue; Rio Yokota; Yutaka Satoh*

249. SAMFusion: Sensor-Adaptive Multimodal Fusion for 3D Object Detection in Adverse Weather; *Edoardo Palladin*; Roland Dietze*; Praveen Narayanan; Mario Bijelic; Felix Heide*

250. LiDAR-based All-weather 3D Object Detection via Prompting and Distilling 4D Radar; *Yujeong Chae; Hyeonseong Kim; Changgyoon Oh; Minseok Kim; Kuk-Jin Yoon**

251. FutureDepth: Learning to Predict the Future Improves Video Depth Estimation; *Rajeev Yasarla*; Manish Kumar Singh; Hong Cai; Yunxiao Shi; Jisoo Jeong; Yinhao Zhu; Shizhong Han; Risheek Garrepalli; Fatih Porikli*

252. Scene-aware Human Motion Forecasting via Mutual Distance Prediction; *Chaoyue Xing*; Wei Mao; Miaomiao Liu*

253. 3D Human Pose Estimation via Non-Causal Retentive Networks; *Kaili Zheng; Feixiang Lu; Yihao Lv; Liangjun Zhang; Chenyi Guo*; Ji Wu**

254. De-confounded Gaze Estimation; *Ziyang Liang; Yiwei Bao; Feng Lu**

255. EgoBody3M: Egocentric Body Tracking on a VR Headset using a Diverse Dataset; *Amy Zhao; Chengcheng Tang; Lezi Wang; Yijing Li; Mihika Dave; Lingling Tao*; Christopher D. Twigg; Robert Y. Wang*

256. Pose-Aware Self-Supervised Learning with Viewpoint Trajectory Regularization; *Jiayun Wang*; Yubei Chen; Stella X. Yu*

257. HPE-Li: WiFi-enabled Lightweight Dual Selective Kernel Convolution for Human Pose Estimation; *Toan D. Gian; Tien Dac Lai; Thien Van Luong; Kok-Seng Wong; Van-Dinh Nguyen**

258. A Graph-Based Approach for Category-Agnostic Pose Estimation; *Or Hirschorn*; Shai Avidan*

259. HandDAGT: A Denoising Adaptive Graph Transformer for 3D Hand Pose Estimation; *WENCAN CHENG; Eunji Kim; Jong Hwan Ko**

260. 3DSA:Multi-View 3D Human Pose Estimation With 3D Space Attention Mechanisms; *Po Han Chen; Chia-Chi Tsai**

261. WHAC: World-grounded Humans and Cameras; *Wanqi Yin; Zhongang Cai; Chen Wei; Fanzhou Wang; Ruisi Wang; Haiyi Mei; Weiye Xiao; Zhitao Yang; Qingping Sun; Atsushi Yamashita; Ziwei Liu; Lei Yang**

262. Mono-ViFI: A Unified Learning Framework for Self-supervised Single- and Multi-frame Monocular Depth Estimation; *Jinfeng Liu*; Lingtong Kong; Bo Li; Zerong Wang; Hong Gu; Jinwei Chen*

263. View-Consistent Hierarchical 3D Segmentation Using Ultrametric Feature Fields; *Haodi He; Colton Stearns; Adam Harley; Leonidas Guibas**

264. Open Panoramic Segmentation; *Junwei Zheng; Ruiping Liu; Yufan Chen; Kunyu Peng; Chengzhi Wu; Kailun Yang; Jiaming Zhang*; Rainer Stiefelhagen*

265. R3DS: Reality-linked 3D Scenes for Panoramic Scene Understanding; *Qirui Wu*; Sonia Raychaudhuri; Daniel Ritchie; Manolis Savva; Angel X Chang*

266. When Do We Not Need Larger Vision Models?; *Baifeng Shi*; Ziyang Wu; Maolin Mao; Xin Wang; Trevor Darrell*

267. Canonical Shape Projection is All You Need for 3D Few-shot Class Incremental Learning; *Ali Cheraghian*; Zeeshan Hayder; Sameeea Ramasinghe; Shafin Rahman; Javad Jafaryahya; Lars Petersson; Mehrtash Harandi*

268. FuseTeacher: Modality-fused Encoders are Strong Vision Supervisors; *Chen-Wei Xie*; Siyang Sun; Liming Zhao; Pandeng Li; Shuailei Ma; Yun Zheng*

269. CMD: A Cross Mechanism Domain Adaptation Dataset for 3D Object Detection; *Jinhao Deng; Wei Ye; Hai Wu; Qiming Xia; Xun Huang; Xin Li; Jin Fang; Wei Li*; Chenglu Wen*; Cheng Wang*

270. ML-SemReg: Boosting Point Cloud Registration with Multi-level Semantic Consistency; *Shaocheng Yan; Pengcheng Shi; Jiayuan Li**

271. UL-VIO: Ultra-lightweight Visual-Inertial Odometry with Noise Robust Test-time Adaptation; *Jinho Park*; Se Young Chun; Mingoo Seok*

272. Stream Query Denoising for Vectorized HD-Map Construction; *Shuo Wang*; Fan Jia; Weixin Mao; Yingfei Liu; Yucheng Zhao; Zehui Chen; Tiancai Wang; Chi Zhang; Xiangyu Zhang; Feng Zhao**

273. Rawformer: Unpaired Raw-to-Raw Translation for Learnable Camera ISPs; *Georgy Perevozchikov*; Nancy Mehta*; Mahmoud Afifi*; Radu Timofte**

274. Correspondence-Free SE(3) Point Cloud Registration in RKHS via Unsupervised Equivariant Learning; *Ray Zhang*; Zheming Zhou; Min Sun; Omid Ghasemalizadeh; Cheng-Hao Kuo; Ryan M. Eustice; Maani Ghaffari Jadidi; Arnie Sen*

275. EINet: Point Cloud Completion via Extrapolation and Interpolation; *Pingping Cai*; Canyu Zhang; LINGJIA SHI; Lili Wang; Nasrin Imanpour; Song Wang*

276. DiffPMAE: Diffusion Masked Autoencoders for Point Cloud Reconstruction; *Yanlong LI*; Chamara Madarasingha; Kanchana Thilakarathna*

277. DrivingDiffusion: Layout-Guided Multi-View Driving Scenarios Video Generation with Latent Diffusion Model; *Li Xiaofan*; Zhang Yifu*; Ye Xiaoqing**

278. VQA-Diff: Exploiting VQA and Diffusion for Zero-Shot Image-to-3D Vehicle Asset Generation in Autonomous Driving; *YIBO LIU*; Zheyuan Yang; Guile Wu; Yuan Ren; Kejian Lin; Liu Bingbing; Yang Liu; JINJUN SHAN*

279. TransCAD: A Hierarchical Transformer for CAD Sequence Inference from Point Clouds; *Elona Dupont*; Kseniya Cherenkova; Dimitrios Mallis; Gleb A Gusev; Anis Kacem; Djamila Aouada*

280. FoundPose: Unseen Object Pose Estimation with Foundation Features; *Evin Pınar Örnek*; Yann Labbé; Bugra Tekin; Lingni Ma; Cem Keskin; Christian Forster; Tomas Hodan*

281. MOFA-Video: Controllable Image Animation via Generative Motion Field Adaptions in Frozen Image-to-Video Diffusion Model; *Muyao Niu; Xiaodong Cun*; Xintao Wang; Yong Zhang; Ying Shan; Yinqiang Zheng**

282. Global Structure-from-Motion Revisited; *Linfei Pan*; Daniel Barath; Marc Pollefeys; Johannes L Schönberger*

283. GalLop: Learning global and local prompts for vision-language models; *Marc Lafon*; Elias Ramzi*; Clément Rambour; Nicolas Audebert; Nicolas Thome*

284. StereoGlue: Joint Feature Matching and Robust Estimation; *Daniel Barath*; Dmytro Mishkin; Luca Cavalli; Paul-Edouard Sarlin; Petr Hruby; Marc Pollefeys*

285. Raising the Ceiling: Conflict-Free Local Feature Matching with Dynamic View Switching; *Xiaoyong Lu*; Songlin Du**

286. Möbius Transform for Mitigating Perspective Distortions in Representation Learning; *Prakash Chandra Chhipa*; Meenakshi Subhash Chippa; Kanjar De; Rajkumar Saini; Marcus Liwicki; Mubarak Shah*

287. DualBEV: Unifying Dual View Transformation with Probabilistic Correspondences; *Peidong Li*; Wancheng Shen; Qihao Huang; Dixiao Cui**

288. Diff-Reg: Diffusion Model in Doubly Stochastic Matrix Space for Registration Problem; *Qianliang Wu*; Haobo Jiang*; Lei Luo; Jun Li; Yaqing Ding*; Jin Xie*; Jian Yang**

289. LineFit: A Geometric Approach for Fitting Line Segments in Images; *Marion Boyer; David Youssefi; Florent Lafarge**

290. The Nerfect Match: Exploring NeRF Features for Visual Localization; *Qunjie Zhou*; Maxim Maximov; Or Litany; Laura Leal-Taixé*

291. Learned Neural Physics Simulation for Articulated 3D Human Pose Reconstruction; *Misha Andriluka*; Baruch Tabanpour; Daniel Freeman; Cristian Sminchisescu*

292. DGD: Dynamic 3D Gaussians Distillation; *Isaac Labe; Noam Issachar; Itai Lang; Sagie Benaim**

293. NICP: Neural ICP for 3D Human Registration at Scale; *Riccardo Marin*; Enric Corona; Gerard Pons-Moll*

294. A Cephalometric Landmark Regression Method based on Dual-encoder for High-resolution X-ray Image; *Chao Dai; yang wang*; Chaolin Huang; zhou jiakai; Qilin Xu; Minpeng Xu*

295. SHIC: Shape-Image Correspondences with no Keypoint Supervision; *Aleksandar Shtedritski*; Christian Rupprecht; Andrea Vedaldi*

296. External Knowledge Enhanced 3D Scene Generation from Sketch; *Zijie Wu; Mingtao Feng*; Yaonan Wang; He Xie; Weisheng Dong; Bo Miao; Ajmal Mian*

297. G2fR: Frequency Regularization in Grid-based Feature Encoding Neural Radiance Fields; *Shuxiang Xie*; Shuyi Zhou; Ken Sakurada; Ryoichi Ishikawa; Masaki Onishi; Takeshi Oishi*

298. Vista3D: unravel the 3d darkside of a single image; *Qiuhong Shen; Xingyi Yang; Michael Bi Mi; Xinchao Wang**

299. LEIA: Latent View-invariant Embeddings for Implicit 3D Articulation; *Archana Swaminathan*; Anubhav Gupta; Kamal Gupta; Shishira R Maiya; Vatsal Agarwal; Abhinav Shrivastava*

300. RadEdit: stress-testing biomedical vision models via diffusion image editing; *Fernando Pérez-García; Sam Bond-Taylor; Pedro Sanchez; Boris van Breugel; Daniel Coelho de Castro; Harshita Sharma; Valentina Salvatelli; Maria Teodora A Wetscherek; Hannah CM Richardson; Lungren Matthew; Aditya Nori; Javier Alvarez-Valle; Ozan Oktay; Maximilian Ilse**

301. An Adaptive Screen-Space Meshing Approach for Normal Integration; *Moritz Heep*; Eduard Zell*

302. ComboVerse: Compositional 3D Assets Creation Using Spatially-Aware Diffusion Guidance; *Yongwei Chen; Tengfei Wang; Tong Wu; Xingang Pan; Kui Jia*; Ziwei Liu*

303. GeoWizard: Unleashing the Diffusion Priors for 3D Geometry Estimation from a Single Image; *Xiao Fu*; Wei Yin; Mu Hu; Kaixuan Wang; Yuexin Ma; Ping Tan; Shaojie Shen; Dahua Lin; Xiaoxiao Long*

304. Object-Oriented Anchoring and Modal Alignment in Multimodal Learning; *Shibin Mei; Bingbing Ni*; Hang Wang; Chenglong Zhao; fengfa hu; Zhiming Pi; BiLian Ke*

305. Learning Pseudo 3D Guidance for View-consistent Texturing with 2D Diffusion; *Kehan Li; Yanbo Fan*; Yang Wu; Zhongqian Sun; Wei Yang; Xiangyang Ji; Li Yuan; Jie Chen**

306. Test-time Model Adaptation for Image Reconstruction Using Self-supervised Adaptive Layers; *Yutian Zhao; Tianjing Zhang; Hui Ji**

307. 4Diff: 3D-Aware Diffusion Model for Third-to-First Viewpoint Translation; *Feng Cheng*; Mi Luo*; Huiyu Wang; Alex Dimakis; Lorenzo Torresani; Gedas Bertasius; Kristen Grauman*

308. SweepNet: Unsupervised Learning Shape Abstraction via Neural Sweepers; *Mingrui Zhao*; Yizhi Wang; Fenggen Yu; Changqing Zou; Ali Mahdavi-Amiri*

309. EchoScene: Indoor Scene Generation via Information Echo over Scene Graph Diffusion; *Guangyao Zhai*; Evin Pınar Örnek; Dave Zhenyu Chen; Ruotong Liao; Yan Di; Nassir Navab; Federico Tombari; Benjamin Busam*

310. DiffSurf: A Transformer-based Diffusion Model for Generating and Reconstructing 3D Surfaces in Pose; *Yusuke Yoshiyasu*; Leyuan Sun*

311. Surface Reconstruction for 3D Gaussian Splatting via Local Structural Hints; *Qianyi Wu*; Jianmin Zheng; Jianfei Cai*

312. Disentangled Generation and Aggregation for Robust Radiance Fields; *Shihe Shen; Huachen Gao; Wangze Xu; Rui Peng; Luyang Tang; Kaiqiang Xiong; Jianbo Jiao; Ronggang Wang**

313. SAFNet: Selective Alignment Fusion Network for Efficient HDR Imaging; *Lingtong Kong*; Bo Li; Yike Xiong; Hao Zhang; Hong Gu; Jinwei Chen*

314. InfoNorm: Mutual Information Shaping of Normals for Sparse-View Reconstruction; *Xulong Wang; Siyan Dong*; Youyi Zheng; Yanchao Yang**

315. 3DEgo: 3D Editing on the Go!; *Umar Khalid*; Hasan Iqbal*; Azib Farooq; Jing Hua; Chen Chen**

316. Learning to Enhance Aperture Phasor Field for Non-Line-of-Sight Imaging; *In Cho; Hyunbo Shim; Seon Joo Kim**

317. Pixel-GS Density Control with Pixel-aware Gradient for 3D Gaussian Splatting; *Zheng Zhang; Wenbo Hu*; Yixing Lao; Tong He; Hengshuang Zhao**

318. SWAG: Splatting in the Wild images with Appearance-conditioned Gaussians; *Hiba Dahmani*; Moussab Bennehar; Nathan Piasco; Luis G Roldao Jimenez; Dzmitry Tsishkou*

319. CoherentGS: Sparse Novel View Synthesis with Coherent 3D Gaussians; *Avinash Paliwal*; Wei Ye; Jinhui Xiong; Dmytro Kotovenko; Rakesh Ranjan; Vikas Chandra; Nima Khademi Kalantari*

320. GS-LRM: Large Reconstruction Model for 3D Gaussian Splatting; *Kai Zhang*; Sai Bi; Hao Tan; Yuanbo Xiangli; Nanxuan Zhao; Kalyan Sunkavalli; Zexiang Xu*

321. SWinGS: Sliding Windows for Dynamic 3D Gaussian Splatting; *Richard Shaw*; Michal Nazarczuk; Jifei Song; Arthur Moreau; Sibi Catley-Chandar; Helisa Dhamo; Eduardo Pérez Pellitero*

322. MirrorGaussian: Reflecting 3D Gaussians for Reconstructing Mirror Reflections; *Jiayue Liu; Xiao Tang; Freeman Cheng; Zihao Yang; Zhihao Li*; Jianzhuang Liu; Yi Huang; Jiaqi Lin; Shiyong Liu; Xiaofei Wu; Songcen Xu; Chun Yuan**

323. A Probability-guided Sampler for Neural Implicit Surface Rendering; *Gonçalo José Dias Pais; Valter André Piedade; Moitreya Chatterjee; Marcus Greiff; Pedro Miraldo**

324. LN3Diff: Scalable Latent Neural Fields Diffusion for Speedy 3D Generation; *Yushi Lan; Fangzhou Hong; Shuai Yang; Shangchen Zhou; Xuyi Meng; Bo Dai; Xingang Pan; Chen Change Loy**

325. Fast View Synthesis of Casual Videos with Soup-of-Planes; *Yao-Chih Lee*; Zhoutong Zhang; Kevin Blackburn-Matzen; Simon Niklaus; Jianming Zhang; Jia-Bin Huang; Feng Liu**

326. The Sky's the Limit: Relightable Outdoor Scenes via a Sky-pixel Constrained Illumination Prior and Outside-In Visibility; *James A D Gardner*; Evgenii Kashin; Bernhard Egger; William Smith*

327. Boost Your NeRF: A Model-Agnostic Mixture of Experts Framework for High Quality and Efficient Rendering; *Francesco Di Sario*; Riccardo Renzulli; Marco Grangetto; Enzo Tartaglione*

328. Leveraging Thermal Modality to Enhance Reconstruction in Low-Light Conditions; *Jiacong Xu*; Mingqian Liao; Ram Prabhakar Kathirvel; Vishal Patel*

329. DMiT: Deformable Mipmapped Tri-Plane Representation for Dynamic Scenes; *Jing-Wen Yang; Jia-Mu Sun; Yong-Liang Yang; Jie Yang; Ying Shan; Yan-Pei Cao; Lin Gao**

330. VersatileGaussian: Real-time Neural Rendering for Versatile Tasks using Gaussian Splatting; *Renjie Li; Zhiwen Fan*; Bohua Wang; Peihao Wang; Zhangyang Wang; Xi Wu*

331. Dynamic Neural Radiance Field From Defocused Monocular Video; *Xianrui Luo; Huiqiang Sun; Juewen Peng; Zhiguo Cao**

332. Learning to Robustly Reconstruct Dynamic Scenes from Low-light Spike Streams; *Liwen Hu*; Ziluo Ding; Mianzhi Liu; Lei Ma*; Tiejun Huang*

334. NeRF-XL: NeRF at Any Scale with Multi-GPU; *Ruilong Li*; Sanja Fidler; Angjoo Kanazawa; Francis Williams*

335. Wavelength-Embedding-guided Filter-Array Transformer for Spectral Demosaicing; *Haijin Zeng*; Hiep Luong; Wilfried Philips*

336. REFRAME: Reflective Surface Real-Time Rendering for Mobile Devices; *Chaojie Ji*; Yufeng Li; Yiyi Liao*

337. MeshAvatar: Learning High-quality Triangular Human Avatars from Multi-view Videos; *Yushuo Chen*; Zerong Zheng; Zhe Li; Chao Xu; Yebin Liu*

338. Motion-Oriented Compositional Neural Radiance Fields for Monocular Dynamic Human Modeling; *Jaehyeok Kim; Dongyoon Wee; Dan Xu**

339. Portrait4D-v2: Pseudo Multi-View Data Creates Better 4D Head Synthesizer; *Yu Deng*; Duomin Wang; Baoyuan Wang*

340. Learning to Generate Conditional Tri-plane for 3D-aware Expression Controllable Portrait Animation; *Taekyung Ki*; Dongchan Min; Gyeongsu Chae**

341. Fast Registration of Photorealistic Avatars for VR Facial Animation; *Chaitanya Patel*; Shaojie Bai; Te-Li Wang; Jason Saragih; Shih-En Wei*

342. ScanTalk: 3D Talking Heads from Unregistered Scans; *Federico Nocentini*; Thomas Besnier; Claudio Ferrari; Sylvain Arguillere; Stefano Berretti; Mohamed Daoudi*

343. Co-speech Gesture Video Generation with 3D Human Meshes; *Aniruddha Mahapatra*; Richa Mishra*; Ziyi Chen; Boyang Ding; Renda Li; Shoulei Wang; Jun-Yan Zhu; Peng Chang; Mei Han; Jing Xiao*

344. Audio-driven Talking Face Generation with Stabilized Synchronization Loss; *Dogucan Yaman*; Fevziye Irem Eyiokur; Leonard Bärmann; HAZIM KEMAL EKENEL; Alexander Waibel*

12:00 – 13:30
Speed Mentoring - Space 4

12:30 – 13:30
Lunch – Balcony Level 1

13:30 – 15:30
Oral session 6A: Generative models II - Gold Room
Chairs: Nicu Sebe; Vicky Kalogeiton
1. Controlling the World by Sleight of Hand; *Sruthi Sudhakar*; Ruoshi Liu; Basile Van Hoorick; Carl Vondrick; Richard Zemel* BEST PAPER CANDIDATE

2. Pyramid Diffusion for Fine 3D Large Scene Generation; *Yuheng Liu*; Xinke Li; Xueting Li; Lu Qi*; Chongshou Li; Ming-Hsuan Yang*

3. FMBoost: Boosting Latent Diffusion with Flow Matching; *Johannes S Fischer*; Ming Gui; Pingchuan Ma; Nick Stracke; Stefan Andreas Baumann; Vincent Tao Hu; Björn Ommer*

4. ConceptExpress: Harnessing Diffusion Models for Single-image Unsupervised Concept Extraction; *Shaozhe Hao*; Kai Han*; Zhengyao Lv; Shihao Zhao; Kwan-Yee K. Wong**

5. Exact Diffusion Inversion via Bidirectional Integration Approximation; *Guoqiang Zhang*; j.p. lewis; W. Bastiaan Kleijn*

6. Tackling Structural Hallucination in Image Translation with Local Diffusion; *Seunghoi Kim*; Chen Jin; Tom Diethe; Matteo Figini; Henry FJ Tregidgo; Asher Mullokandov; Philip A Teare; Daniel Alexander*

7. Diffusion Prior-Based Amortized Variational Inference for Noisy Inverse Problems; *Sojin Lee; Dogyun Park; Inho Kong; Hyunwoo J. Kim**

8. Adversarial Diffusion Distillation; *Axel Sauer*; Dominik Lorenz; Andreas Blattmann; Robin Rombach*

9. Arc2Face: A Foundation Model for ID-Consistent Human Faces; *Foivos Paraperas Papantoniou*; Alexandros Lattas; Stylianos Moschoglou; Jiankang Deng; Bernhard Kainz; Stefanos Zafeiriou*

10. Diffusion-Driven Data Replay: A Novel Approach to Combat Forgetting in Federated Class Continual Learning; *Jinglin Liang; Jin Zhong; Hanlin Gu; Zhongqi Lu; Xingxing Tang; Gang Dai; Shuangping Huang*; Lixin Fan; Qiang Yang*

11. OmniSSR: Zero-shot Omnidirectional Image Super-Resolution using Stable Diffusion Model; *Runyi Li*; Xuhan Sheng; Weiqi Li; Jian Zhang**

13:30 – 15:30
Oral session 6B: Video understanding - Auditorium
Chairs: Hazel Doughty; Lamberto Ballan
1. E3M: Zero-Shot Spatio-Temporal Video Grounding with Expectation-Maximization Multimodal Modulation; *Peijun Bao*; Zihao Shao; Wenhan Yang; Boon Poh Ng; Alex Kot*

2. Animal Avatars: Reconstructing Animatable 3D Animals from Casual Videos; *Remy Sabathier*; David Novotny; Niloy Mitra*

3. Made to Order: Discovering monotonic temporal changes via self-supervised video ordering; *Charig Yang*; Weidi Xie; Andrew Zisserman*

4. MAGR: Manifold-Aligned Graph Regularization for Continual Action Quality Assessment; *Kanglei Zhou; Liyuan Wang; Xingxing Zhang; Hubert P. H. Shum; Frederick W. B. Li; Jianguo Li; Xiaohui Liang**

5. C2C: Component-to-Composition Learning for Zero-Shot Compositional Action Recognition; *Rongchang Li; Zhenhua Feng; Tianyang Xu; Linze Li; Xiao-Jun Wu*; Muhammad Awais; Sara Atito; Josef Kittler*

6. LongVLM: Efficient Long Video Understanding via Large Language Models; *Yuetian Weng; Mingfei Han; Haoyu He; Xiaojun Chang; Bohan Zhuang**

7. Propose, Assess, Search: Harnessing LLMs for Goal-Oriented Planning in Instructional Videos; *Md Mohaiminul Islam*; Tushar Nagarajan; Huiyu Wang; FU-JEN CHU; Kris Kitani; Gedas Bertasius; Xitong Yang*

8. Towards Neuro-Symbolic Video Understanding; *Minkyu Choi*; Harsh Goel; Mohammad Omama; Yunhao Yang; Sahil Shah; Sandeep Chinchali*

9. Classification Matters: Improving Video Action Detection with Class-Specific Attention; *Jinsung Lee; Taeoh Kim; Inwoong Lee; Minho Shim; Dongyoon Wee; Minsu Cho; Suha Kwak**

10. DEVIAS: Learning Disentangled Video Representations of Action and Scene; *Kyungho Bae; Youngrae Kim; Geo Ahn; Jinwoo Choi**

11. Sync from the Sea: Retrieving Alignable Videos from Large-Scale Datasets; *Ishan Rajendrakumar Dave*; Fabian Caba; Mubarak Shah; Simon Jenni**

13:30 – 15:30
Oral session 6C: Vision and other modalities - Silver Room
Chairs: Shizhe Chen; Vicente Ordonez
1. GiT: Towards Generalist Vision Transformer through Universal Language Interface; *Haiyang Wang*; Hao Tang; Li Jiang; Shaoshuai Shi; Muhammad Ferjad Naeem; Hongsheng Li; Bernt Schiele; Liwei Wang*

2. Omniview-Tuning: Boosting Viewpoint Invariance of Vision-Language Pre-training Models; *Shouwei Ruan*; Yinpeng Dong; Liu Hanqing; Yao Huang; Hang Su; Xingxing Wei**

3. Turbo: Informativity-Driven Acceleration Plug-In for Vision-Language Large Models; *Chen Ju*; Haicheng Wang; Haozhe Cheng; Xu Chen; Zhonghua Zhai; Weilin Huang; Jinsong Lan; Shuai Xiao*; Bo Zheng*

4. MMBENCH: Is Your Multi-Modal Model an All-around Player?; *Yuan Liu*; Haodong Duan*; Yuanhan Zhang; Bo Li; Songyang Zhang; Wangbo Zhao; Yike Yuan; Jiaqi Wang; Conghui He; Ziwei Liu; Kai Chen; Dahua Lin*

5. Strengthening Multimodal Large Language Model with Bootstrapped Preference Optimization; *Renjie Pi*; Tianyang Han; Wei Xiong; Jipeng ZHANG; Runtao Liu; Rui Pan; Tong Zhang*

6. Beat-It: Beat-Synchronized Multi-Condition 3D Dance Generation; *Zikai Huang; Xuemiao Xu*; Cheng Xu*; Huaidong Zhang; Chenxi Zheng; Jing Qin; Shengfeng He*

7. A Simple Baseline for Spoken Language to Sign Language Translation with 3D Avatars; *Ronglai Zuo; Fangyun Wei*; Zenggui Chen; Brian Mak; Jiaolong Yang; Xin Tong*

8. HYPE: Hyperbolic Entailment Filtering for Underspecified Images and Texts; *Wonjae Kim*; Sanghyuk Chun; Taekyung Kim; Dongyoon Han; Sangdoo Yun*

9. An Image is Worth 1/2 Tokens After Layer 2: Plug-and-Play Inference Acceleration for Large Vision-Language Models; *Liang Chen; Haozhe Zhao; Tianyu Liu; Shuai Bai; Junyang Lin; Chang Zhou; Baobao Chang**

10. uCAP: An Unsupervised Prompting Method for Vision-Language Models; *A. Tuan Nguyen*; Kai Sheng Tai; Bor-Chun Chen; Satya Narayan Shukla; Hanchao Yu; Philip Torr; Tai-Peng Tian; Ser-Nam Lim*

11. BRAVE: Broadening the visual encoding of vision-language models; *Oğuzhan Fatih Kar*; Alessio Tonioni*; Petra Poklukar; Achin Kulshrestha; Amir Zamir; Federico Tombari*

14:30 – 18:00
Demo session 6 - Level 0
1. Fruit Ninja with an Event Camera; *Gaurvi Goyal, Massimiliano Iacano, Arren Glover, Chiara Bartolozzi - Istituto Italiano di Tecnologia*

2. ViPer: Visual Personalization of Generative Models via Individual Preference Learning; *Sogand Salehi, Mahdi Shafiei, Teresa Yeo, Roman Bachmann, Amir Zamir - EPFL*

3. Open-Vocabulary Interactive 3D Scenes with Spot; *Tim Engelbracht, Zuria Bauer, Herman Blum, Francis Engelmann - ETH Zurich*

4. R^2-Tuning: Efficient Image-to-Video Transfer Learning for Video Temporal Grounding; *Ye Liu, Jixuan He, Wanhua Li, Junsik Kim, Donglai Wei, Hanspeter Pfister, Chang Wen Chen - Hong Kong Polytechnic University*

5. Automatic Data Curation for Self-Supervised Learning of Visual Features; *Vasil Khalidov, Huy Vo, Claire Roberts, Piotr Bojanowski, Patrick Labatut - Meta* FAIR

15:30 – 16:30
Keynote Lecture - Gold Room (live), Auditorium (broadcast), Silver Room (broadcast)
Is distribution shift still an AI problem?; *Sanmi Koyejo*

16:30 – 17:00
Istituto Italiano di Tecnologia Technical Session - Technical Presentation Area (Level 0)
The Italian Institute of Technology: our research of di embodiment of intelligence

16:30 – 17:00
Coffee Break - Exhibition Area (Level 0)

16:30 – 18:30
Poster session 6
1. Missing Modality Prediction for Unpaired Multimodal Learning via Joint Embedding of Unimodal Models; *Taesup Kim*; Donggeun Kim*

2. Towards Multimodal Open-Set Domain Generalization and Adaptation through Self-supervision; *Hao Dong*; Eleni Chatzi*; Olga Fink**

3. Pick-a-back: Selective Device-to-Device Knowledge Transfer in Federated Continual Learning; *HyungJune Lee*; JinYi Yoon*

4. Diffusion-Driven Data Replay: A Novel Approach to Combat Forgetting in Federated Class Continual Learning; *Jinglin Liang; Jin Zhong; Hanlin Gu; Zhongqi Lu; Xingxing Tang; Gang Dai; Shuangping Huang*; Lixin Fan; Qiang Yang*

5. FedTSA: A Cluster-based Two-Stage Aggregation Method for Model-heterogeneous Federated Learning; *Boyu Fan*; Chenrui Wu; Xiang Su; Pan HUI*

6. MagMax: Leveraging Model Merging for Seamless Continual Learning; *Daniel Marczak*; Bartlomiej Twardowski*; Tomasz Trzcinski*; Sebastian Cygert**

7. Exploring Active Learning in Meta-Learning: Enhancing Context Set Labeling; *Wonho Bae; Jing Wang; Danica J. Sutherland**

8. Forget More to Learn More: Domain-specific Feature Unlearning for Semi-supervised and Unsupervised Domain Adaptation; *Hritam Basak*; Zhaozheng Yin*

9. Learning to Unlearn for Robust Machine Unlearning; *Mark He Huang*; Lin Geng Foo; Jun Liu**

10. UNIC: Universal Classification Models via Multi-teacher Distillation; *Yannis Kalantidis; Diane Larlus; Mert Bulent Sariyildiz*; Philippe Weinzaepfel; Thomas LUCAS*

11. CLOSER: Towards Better Representation Learning for Few-Shot Class-Incremental Learning; *Junghun Oh; Sungyong Baik; Kyoung Mu Lee**

12. PromptCCD: Learning Gaussian Mixture Prompt Pool for Continual Category Discovery; *Fernando Julio Cendra; Bingchen Zhao; Kai Han**

13. Distributed Active Client Selection With Noisy Clients Using Model Association Scores; *Kwang In Kim**

14. SCOD: From Heuristics to Theory; *Vojtech Franc*; Jakub Paplham*; Daniel Prusa**

15. Regulating Model Reliance on Non-Robust Features by Smoothing Input Marginal Density; *Peiyu Yang*; Naveed Akhtar; Mubarak Shah; Ajmal Mian*

16. LNL+K: Enhancing Learning with Noisy Labels Through Noise Source Knowledge Integration; *Siqi Wang*; Bryan Plummer*

17. Rethinking Fast Adversarial Training: A Splitting Technique To Overcome Catastrophic Overfitting; *Masoumeh Zareapoor; Pourya Shamsolmoali**

18. SCOMatch: Alleviating Overtrusting in Open-set Semi-supervised Learning; *Zerun Wang*; Liuyu Xiang; Lang Huang; Jiafeng Mao; Ling Xiao; Toshihiko Yamasaki*

19. A high-quality robust diffusion framework for corrupted dataset; *Quan Dao*; Binh Ta; Tung Pham; Anh Tran*

20. Dynamic Guidance Adversarial Distillation with Enhanced Teacher Knowledge; *Hyejin Park; Dongbo Min**

21. Learning to Obstruct Few-Shot Image Classification over Restricted Classes; *Amber Yijia Zheng*; Chiao-An Yang*; Raymond A. Yeh*

22. Cross-Domain Learning for Video Anomaly Detection with Limited Supervision; *Yashika Jain; Ali Dabouei*; Min Xu**

23. Labeled Data Selection for Category Discovery; *Bingchen Zhao*; Nico Lang; Serge Belongie; Oisin Mac Aodha**

24. Unsqueeze [CLS] Bottleneck to Learn Rich Representations; *Qing Su*; Shihao Ji*

25. PartImageNet++ Dataset: Scaling up Part-based Models for Robust Recognition; *Xiao Li*; Yining Liu; Na Dong; Sitian Qin; Xiaolin Hu*

26. HyperSpaceX: Radial and Angular Exploration of HyperSpherical Dimensions; *Chiranjeev Chiranjeev; Muskan Dosi; Kartik Thakral; Mayank Vatsa*; Richa Singh*

27. Norma: A Noise Robust Memory-Augmented Framework for Whole Slide Image Classification; *Yu Bai; Bo Zhang*; Zheng Zhang; Shuo Yan; Zibo Ma; Wu Liu; Xiuzhuang Zhou; Xiangyang Gong; Wendong Wang*

28. Improving Hyperbolic Representations via Gromov-Wasserstein Regularization; *Yifei Yang; Wonjun Lee; Dongmian Zou*; Gilad Lerman*

29. GLAD: Towards Better Reconstruction with Global and Local Adaptive Diffusion Models for Unsupervised Anomaly Detection; *Hang Yao; Ming Liu*; Zhicun Yin; Zifei Yan; Xiaopeng Hong; Wangmeng Zuo*

30. Lagrangian Hashing for Compressed Neural Field Representations; *Shrisudhan Govindarajan*; Zeno Sambugaro; Akhmedkhan Shabanov; Towaki Takikawa; Weiwei Sun; Daniel Rebain; Nicola Conci; Kwang Moo Yi; Andrea Tagliasacchi*

31. Unsupervised, Online and On-The-Fly Anomaly Detection For Non-Stationary Image Distributions; *Declan GD McIntosh*; Alexandra Branzan Albu*

32. AD3: Introducing a score for Anomaly Detection Dataset Difficulty assessment using VIADUCT dataset; *Jan D Lehr*; Jan H Philipps; Alik Sargsyan; Martin Pape; Jörg Krüger*

33. Weighting Pseudo-Labels via High-Activation Feature Index Similarity and Object Detection for Semi-Supervised Segmentation; *Prantik Howlader*; Hieu Le; Dimitris Samaras*

34. Rectify the Regression Bias in Long-Tailed Object Detection; *Ke Zhu; Minghao Fu; Jie Shao; Tianyu Liu; Jianxin Wu**

35. AlignZeg: Mitigating Objective Misalignment for Zero-shot Semantic Segmentation; *Jiannan Ge*; Lingxi Xie; Hongtao Xie; Pandeng Li; Xiaopeng Zhang; Yongdong Zhang; Qi Tian*

36. Just a Hint: Point-Supervised Camouflaged Object Detection; *Huafeng Chen; Dian SHAO*; Guangqian Guo; shan gao**

37. Learning Camouflaged Object Detection from Noisy Pseudo Label; *Jin Zhang*; Ruiheng Zhang*; Yanjiao Shi; Zhe Cao; Nian Liu; Fahad Shahbaz Khan*

38. Click Prompt Learning with Optimal Transport for Interactive Segmentation; *Jie Liu*; Haochen wang; Wenzhe Yin; Jan-Jakob Sonke; Efstratios Gavves*

39. SOS: Segment Object System for Open-World Instance Segmentation With Object Priors; *Christian Wilms*; Tim Rolff; Maris N Hillemann; Robert Johanson; Simone Frintrop*

40. Segment and Recognize Anything at Any Granularity; *Feng Li*; Hao Zhang; Peize Sun; Xueyan Zou; Shilong Liu; Chunyuan Li; Jianwei Yang; Lei Zhang*; Jianfeng Gao**

41. Active Coarse-to-Fine Segmentation of Moveable Parts from Real Images; *Ruiqi Wang*; Akshay Gadi Patil; Fenggen Yu; Hao Zhang*

42. SEGIC: Unleashing the Emergent Correspondence for In-Context Segmentation; *Lingchen Meng; Shiyi Lan; Hengduo Li; Jose M Alvarez; Zuxuan Wu*; Yu-Gang Jiang*

43. Phase Concentration and Shortcut Suppression for Weakly Supervised Semantic Segmentation; *Hoyong Kwon; Jaeseok Jeong; Sung-Hoon Yoon; Kuk-Jin Yoon**

44. BugNIST - a Large Volumetric Dataset for Detection under Domain Shift; *Patrick M Jensen; Vedrana A Dahl; Rebecca Engberg; Carsten Gundlach; Hans Martin Kjer; Anders B Dahl**

45. In Defense of Lazy Visual Grounding for Open-Vocabulary Semantic Segmentation; *Dahyun Kang; Minsu Cho**

46. LaMI-DETR: Open-Vocabulary Detection with Language Model Instruction; *Penghui Du; Yu Wang; Yifan Sun; Luting Wang; Yue Liao; gang zhang; Errui Ding; Yan Wang*; Jingdong Wang; Si Liu**

47. Shifted Autoencoders for Point Annotation Restoration in Object Counting; *Yuda Zou; Xin Xiao; Peilin Zhou; Zhichao Sun; Bo Du; Yongchao Xu**

48. DetToolChain: A New Prompting Paradigm to Unleash Detection Ability of MLLM; *Yixuan Wu*; Yizhou Wang; Shixiang Tang; Wenhao Wu; Tong He; Wanli Ouyang; Philip Torr; Jian Wu*

49. Toward Open Vocabulary Aerial Object Detection with CLIP-Activated Student-Teacher Learning; *Yan Li; Weiwei Guo*; Xue Yang; Ning Liao; Dunyun He; Jiaqi Zhou; Wenxian Yu**

50. 3D Open-Vocabulary Panoptic Segmentation with 2D-3D Vision-Language Distillation; *Zihao Xiao*; Longlong Jing; Shangxuan Wu; Alex Zihao Zhu; Jingwei Ji; Chiyu Max Jiang; Wei-Chih Hung; Thomas Funkhouser; Weicheng Kuo; Anelia Angelova; Yin Zhou; Shiwei Sheng*

51. Human-in-the-Loop Visual Re-ID for Population Size Estimation; *Gustavo Perez*; Daniel Sheldon; Grant Van Horn; Subhransu Maji*

52. OV-Uni3DETR: Towards Unified Open-Vocabulary 3D Object Detection via Cycle-Modality Propagation; *Zhenyu Wang*; Ya-Li Li; TAICHI LIU; Hengshuang Zhao; Shengjin Wang*

53. Diff3DETR: Agent-based Diffusion Model for Semi-supervised 3D Object Detection; *Jiacheng Deng*; Jiahao Lu; Tianzhu Zhang*

54. Fine-Grained Scene Graph Generation via Sample-Level Bias Prediction; *Yansheng Li; Tingzhu Wang*; Kang Wu; Linlin Wang; Xin Guo; Wenbin Wang*

55. SAM-guided Graph Cut for 3D Instance Segmentation; *Haoyu Guo*; He Zhu; Sida Peng; Yuang Wang; Yujun Shen; Ruizhen Hu*; Xiaowei Zhou**

56. Subspace Prototype Guidance for Mitigating Class Imbalance in Point Cloud Semantic Segmentation; *Jiawei Han; Kaiqi Liu*; Wei Li; Guangzhi Chen*

57. Learning Local Pattern Modularization for Point Cloud Reconstruction from Unseen Classes; *Chao Chen; Yu-Shen Liu*; Zhizhong Han*

58. T-CorresNet: Template Guided 3D Point Cloud Completion with Correspondence Pooling Query Generation Strategy; *Fan Duan; Jiahao Yu; Li Chen**

59. Implicit Filtering for Learning Neural Signed Distance Functions from 3D Point Clouds; *Shengtao Li*; Ge Gao; Yudong Liu; Ming Gu; Yu-Shen Liu*

60. SEED: A Simple and Effective 3D DETR in Point Clouds; *Zhe Liu; Jinghua Hou; Xiaoqing Ye; Tong Wang; Jingdong Wang; Xiang Bai**

61. ProtoComp: Diverse Point Cloud Completion with Controllable Prototype; *Xumin Yu; Yanbo Wang; Jie Zhou; Jiwen Lu**

62. CloudFixer: Test-Time Adaptation for 3D Point Clouds via Diffusion-Guided Geometric Transformation; *Hajin Shim; Changhun Kim; Eunho Yang**

63. FastPCI: Motion-Structure Guided Fast Point Cloud Frame Interpolation; *tianyu zhang; Guocheng Qian; Jin Xie*; Jian Yang*

64. Multi-modal Relation Distillation for Unified 3D Representation Learning; *Huiqun Wang; Yiping Bao; Panwang Pan; Zeming Li; Xiao Liu; Ruijie Yang; Di Huang**

65. Rethinking LiDAR Domain Generalization: Single Source as Multiple Density Domains; *Jaeyeul Kim; Jungwan Woo; Jeonghoon Kim; Sunghoon Im**

66. Visible and Clear: Finding Tiny Objects in Difference Map; *Bing Cao; Haiyu Yao; Pengfei Zhu*; Qinghua Hu*

67. LEROjD: Lidar Extended Radar-Only Object Detection; *Patrick Palmer*; Martin Krüger; Stefan Schütte; Richard Altendorfer; Ganesh Adam; Torsten Bertram*

68. Improving Point-based Crowd Counting and Localization Based on Auxiliary Point Guidance; *I-HSIANG CHEN; Wei-Ting Chen; Yu-Wei Liu; Ming-Hsuan Yang; Sy-Yen Kuo**

69. WildRefer: 3D Object Localization in Large-scale Dynamic Scenes with Multi-modal Visual Data and Natural Language; *Zhenxiang Lin; Xidong Peng; Peishan Cong; Ge Zheng; Yujing Sun; Yuenan HOU; Xinge Zhu; Sibei Yang; Yuexin Ma**

70. Visual Relationship Transformation; *Xiaoyu Xu*; Jiayan Qiu; Baosheng Yu; Zhou Wang*

71. GaussianFormer: Scene as Gaussians for Vision-Based 3D Semantic Occupancy Prediction; *Yuanhui Huang; Wenzhao Zheng; Yunpeng Zhang; Jie Zhou; Jiwen Lu**

72. Benchmarking the Robustness of Cross-view Geo-localization Models; *Qingwang Zhang; Yingying Zhu**

73. Learning High-resolution Vector Representation from Multi-Camera Images for 3D Object Detection; *Zhili Chen; Shuangjie Xu; Maosheng Ye; Zian Qian; Xiaoyi Zou; Dit-Yan Yeung; Qifeng Chen**

74. GraphBEV: Towards Robust BEV Feature Alignment for Multi-Modal 3D Object Detection; *Ziying Song; Lei Yang; Shaoqing Xu; Lin Liu; Dongyang Xu; Caiyan Jia*; Feiyang Jia; Li Wang*

75. Boosting 3D Single Object Tracking with 2D Matching Distillation and 3D Pre-training; *Qiangqiang Wu; Yan Xia*; Jia Wan; Antoni Chan*

76. ViewFormer: Exploring Spatiotemporal Modeling for Multi-View 3D Occupancy Perception via View-Guided Transformers; *Jinke Li*; Xiao He*; Chonghua Zhou; Xiaoqiang Cheng; Yang Wen; Dan Zhang**

77. Towards Stable 3D Object Detection; *Jiabao Wang; Qiang Meng; Guochao Liu; Liujiang Yan; Ke Wang; Ming-Ming Cheng; Qibin Hou**

78. ADMap: Anti-disturbance Framework for Vectorized HD Map Construction; *Haotian Hu; Fanyi Wang*; Yaonong Wang; Laifeng Hu; Jingwei Xu; Zhiwang Zhang**

79. Lane Graph as Path: Continuity-preserving Path-wise Modeling for Online Lane Graph Construction; *Bencheng Liao; Shaoyu Chen; Bo Jiang; Tianheng Cheng; Qian Zhang; Wenyu Liu; Chang Huang; Xinggang Wang**

80. Single-Photon 3D Imaging with Equi-Depth Photon Histograms; *Kaustubh Sadekar*; David Maier; Atul Ingle*

81. Beyond the Data Imbalance: Employing the Heterogeneous Datasets for Vehicle Maneuver Prediction; *Hyeongseok Jeon; Sanmin Kim; Abi Rahman Syamil; Junsoo Kim; Dongsuk Kum**

82. DySeT: a Dynamic Masked Self-distillation Approach for Robust Trajectory Prediction; *Mozghan Pourkeshavarz*; Arielle Zhang; Amir Rasouli*

83. CarFormer: Self-Driving with Learned Object-Centric Representations; *Shadi Hamdan*; Fatma Guney*

84. Beyond Viewpoint: Robust 3D Object Recognition under Arbitrary Views through Joint Multi-Part Representation; *Linlong Fan; Ye Huang*; Yanqi Ge; Wen Li; Lixin Duan*

85. NeuroNCAP: Photorealistic Closed-loop Safety Testing for Autonomous Driving; *William Ljungbergh*; Adam Tonderski; Joakim Johnander; Holger Caesar; Kalle Åström; Michael Felsberg; Christoffer Petersson*

86. Leveraging scale- and orientation-covariant features for planar motion estimation; *Marcus Valtonen Örnhag*; Alberto Jaenal*

87. Weakly-supervised Camera Localization by Ground-to-satellite Image Registration; *Yujiao Shi*; HONGDONG LI; Akhil Perincherry; Ankit Vora*

88. TreeSBA: Tree-Transformer for Self-Supervised Sequential Brick Assembly; *Mengqi Guo*; Chen Li; Yuyang Zhao; Gim Hee Lee*

89. Learning to Build by Building Your Own Instructions; *Aaron T Walsman*; Muru Zhang; Adam Fishman; Ali Farhadi; Dieter Fox*

90. Learn to Memorize and to Forget: A Continual Learning Perspective of Dynamic SLAM; *Baicheng Li*; Zike Yan*; Dong Wu; Hanqing Jiang; Hongbin Zha**

91. High-Precision Self-Supervised Monocular Depth Estimation with Rich-Resource Prior; *Jianbing Shen*; Wencheng Han*

92. EgoPet: Egomotion and Interaction Data from an Animal's Perspective; *Amir Bar*; Arya Bakhtiar; Danny L Tran; Antonio Loquercio; Jathushan Rajasegaran; yann lecun; Amir Globerson; Trevor Darrell*

93. Revisit Self-supervision with Local Structure-from-Motion; *Shengjie Zhu*; Xiaoming Liu*

94. NeRF-MAE: Masked AutoEncoders for Self-Supervised 3D Representation Learning for Neural Radiance Fields; *Muhammad Zubair Irshad*; Sergey Zakharov; Vitor Guizilini; Adrien Gaidon; Zsolt Kira; Rares Ambrus*

95. Local All-Pair Correspondence for Point Tracking; *Seokju Cho; Jiahui Huang; Jisu Nam; Honggyu An; Seungryong Kim*; Joon-Young Lee**

96. AugUndo: Scaling Up Augmentations for Monocular Depth Completion and Estimation; *Yangchao Wu*; Tian Yu Liu; Hyoungseob Park; Stefano Soatto; Dong Lao; Alex Wong*

97. Power Variable Projection for Initialization-Free Large-Scale Bundle Adjustment; *Simon Weber*; Je Hyeong Hong; Daniel Cremers*

98. Sketch2Vox: Learning 3D Reconstruction from a Single Monocular Sketch Image; *Fei Wang**

99. SUP-NeRF: A Streamlined Unification of Pose Estimation and NeRF for Monocular 3D Object Reconstruction; *Yuliang Guo*; Abhinav Kumar; Cheng Zhao; Ruoyu Wang; Xinyu Huang; Liu Ren*

100. 3DGazeNet: Generalizing Gaze Estimation with Weak Supervision from Synthetic Views; *Evangelos Ververas*; Polydefkis Gkagkos; Jiankang Deng; Michail C Doukas; Jia Guo; Stefanos Zafeiriou*

101. SelfGeo: Self-supervised and Geodesic-consistent Estimation of Keypoints on Deformable Shapes; *Mohammad Zohaib*; Luca Cosmo; Alessio Del Bue*

102. VQ-HPS: Human Pose and Shape Estimation in a Vector-Quantized Latent Space; *Guénolé Fiche*; Simon Leglaive; Xavier Alameda-Pineda; Antonio Agudo; Francesc Moreno*

103. OneTrack: Demystifying the Conflict Between Detection and Tracking in End-to-End 3D Trackers; *Qitai Wang; Jiawei He; Yuntao Chen; Zhaoxiang Zhang**

104. GRAPE: Generalizable and Robust Multi-view Facial Capture; *Jing Li; Di Kang; Zhenyu He**

105. HandDGP: Camera-Space Hand Mesh Prediction with Differentiable Global Positioning; *Eugene Valassakis; Guillermo Garcia-Hernando**

106. Human Pose Recognition via Occlusion-Preserving Abstract Images; *Saad Manzur*; Wayne B Hayes**

107. 6DoF Head Pose Estimation through Explicit Bidirectional Interaction with Face Geometry; *Sungho Chun; Ju Yong Chang**

108. Bones Can't Be Triangles: Accurate and Efficient Vertebrae Keypoint Estimation through Collaborative Error Revision; *Jinhee Kim; Taesung Kim; Jaegul Choo**

109. RT-Pose: A 4D Radar-Tensor based 3D Human Pose Estimation and Localization Benchmark; *Yuan-Hao Ho; Jen-Hao Cheng; Sheng Yao Kuan; Zhongyu Jiang; Wenhao Chai; Hsiang-Wei Huang; Chih-Lung Lin; Jenq-Neng Hwang**

110. DINO-Tracker: Taming DINO for Self-Supervised Point Tracking in a Single Video; *Narek Tumanyan*; Assaf Singer; Shai Bagon; Tali Dekel*

111. Un-EVIMO: Unsupervised Event-based Independent Motion Segmentation; *Ziyun Wang*; Jinyuan Guo; Kostas Daniilidis*

112. Betrayed by Attention: A Simple yet Effective Approach for Self-supervised Video Object Segmentation; *Shuangrui Ding*; Rui Qian; Haohang Xu; Dahua Lin; Hongkai Xiong*

113. Appearance-based Refinement for Object-Centric Motion Segmentation; *Junyu Xie*; Weidi Xie; Andrew Zisserman*

114. Track2Act: Predicting Point Tracks from Internet Videos enables Generalizable Robot Manipulation; *Homanga Bharadhwaj*; Roozbeh Mottaghi; Abhinav Gupta; Shubham Tulsiani*

115. Ref-AVS: Refer and Segment Objects in Audio-Visual Scenes; *Yaoting Wang; Peiwen Sun; Dongzhan Zhou; Guangyao Li; Honggang Zhang; Di Hu**

116. Fine-grained Dynamic Network for Generic Event Boundary Detection; *Ziwei Zheng; Lijun He; Le Yang; Fan Li**

117. Made to Order: Discovering monotonic temporal changes via self-supervised video ordering; *Charig Yang*; Weidi Xie; Andrew Zisserman*

118. ManiGaussian: Dynamic Gaussian Splatting for Multi-task Robotic Manipulation; *Guanxing Lu; Shiyi Zhang; Ziwei Wang*; Changliu Liu; Jiwen Lu; Yansong Tang*

119. Controlling the World by Sleight of Hand; *Sruthi Sudhakar*; Ruoshi Liu; Basile Van Hoorick; Carl Vondrick; Richard Zemel* BEST PAPER CANDIDATE

120. Multi-Person Pose Forecasting with Individual Interaction Perceptron and Prior Learning; *Peng Xiao; Yi Xie; Xuemiao Xu*; Weihong Chen; Huaidong Zhang**

121. Upper-body Hierarchical Graph for Skeleton Based Emotion Recognition in Assistive Driving; *Jiehui Wu; Jiansheng Chen*; Qifeng Luo; Siqi Liu; Youze Xue; Huimin Ma*

122. On the Utility of 3D Hand Poses for Action Recognition; *Md Salman Shamil*; Dibyadip Chatterjee; Fadime Sener; Shugao Ma; Angela Yao**

123. DragAPart: Learning a Part-Level Motion Prior for Articulated Objects; *Ruining Li*; Chuanxia Zheng; Christian Rupprecht; Andrea Vedaldi*

124. HIMO: A New Benchmark for Full-Body Human Interacting with Multiple Objects; *Xintao Lv; Liang Xu; Yichao Yan*; Xin Jin; Congsheng Xu; Wu Shuwen; Yifan Liu; Lincheng Li; Mengxiao Bi; Wenjun Zeng; Xiaokang Yang*

125. ReMoS: 3D Motion-Conditioned Reaction Synthesis for Two-Person Interactions; *Anindita Ghosh*; Rishabh Dabral; Vladislav Golyanik; Christian Theobalt; Philipp Slusallek*

126. Aligning Neuronal Coding of Dynamic Visual Scenes with Foundation Vision Models; *Rining Wu*; Feixiang Zhou; Ziwei Yin; Jian Liu**

127. Learning Semantic Latent Directions for Accurate and Controllable Human Motion Prediction; *Guowei Xu; Jiale Tao; Wen Li*; Lixin Duan*

128. Context-Aware Action Recognition: Introducing a Comprehensive Dataset for Behavior Contrast; *Tatsuya Sasaki*; Yoshiki Ito; Satoshi Kondo*

130. Semi-Supervised Teacher-Reference-Student Architecture for Action Quality Assessment; *Wulian Yun; Mengshi Qi; Fei Peng; Huadong Ma**

131. Free-VSC: Free Semantics from Visual Foundation Models for Unsupervised Video Semantic Compression; *Yuan Tian*; Guo Lu*; Guangtao Zhai**

132. MotionLCM: Real-time Controllable Motion Generation via Latent Consistency Model; *Wenxun Dai; Ling-Hao Chen; Jingbo Wang*; Jinpeng Liu; Bo Dai*; Yansong Tang*

133. Nonverbal Interaction Detection; *Jianan Wei; Tianfei Zhou; Yi Yang; Wenguan Wang**

134. Chronologically Accurate Retrieval for Temporal Grounding of Motion-Language Models; *Kent Fujiwara*; Mikihiro Tanaka; Qing Yu*

135. SA-DVAE: Improving Zero-Shot Skeleton-Based Action Recognition by Disentangled Variational Autoencoders; *Sheng-Wei Li; Zi-Xiang Wei; Wei-Jie Chen; Yi-Hsin Yu; Chih-Yuan Yang*; Jane Yung-jen Hsu**

136. Flow-Assisted Motion Learning Network for Weakly-Supervised Group Activity Recognition; *Muhammad Adi Nugroho*; Sangmin Woo; Sumin Lee; Jinyoung Park; Yooseung Wang; Donguk Kim; Changick Kim*

137. EgoCVR: An Egocentric Benchmark for Fine-Grained Composed Video Retrieval; *Thomas Hummel*; Shyamgopal Karthik; Mariana-Iuliana Georgescu; Zeynep Akata*

138. RAP: Retrieval-Augmented Planner for Adaptive Procedure Planning in Instructional Videos; *Ali Zare*; Yulei Niu; Hammad Ayyubi; Shih-Fu Chang*

139. Video Question Answering with Procedural Programs; *Rohan Choudhury*; Koichiro Niinuma; Kris Kitani; Laszlo A Jeni*

140. Self-supervised visual learning from interactions with objects; *Arthur Aubret*; Céline Teulière; Jochen Triesch*

141. HAT: History-Augmented Anchor Transformer for Online Temporal Action Localization; *Sakib Reza; Yuexi Zhang; Mohsen Moghaddam; Octavia Camps**

142. PreLAR: World Model Pre-training with Learnable Action Representation; *Lixuan Zhang; Meina Kan; Shiguang Shan; Xilin Chen**

143. Efficient Few-Shot Action Recognition via Multi-Level Post-Reasoning; *Cong Wu; Xiao-Jun Wu*; Linze Li; Tianyang Xu; Zhenhua Feng; Josef Kittler*

144. Sequential Representation Learning via Static-Dynamic Conditional Disentanglement; *Mathieu Cyrille Simon*; Pascal Frossard; Christophe De Vleeschouwer*

145. Towards Neuro-Symbolic Video Understanding; *Minkyu Choi*; Harsh Goel; Mohammad Omama; Yunhao Yang; Sahil Shah; Sandeep Chinchali*

146. Beat-It: Beat-Synchronized Multi-Condition 3D Dance Generation; *Zikai Huang; Xuemiao Xu*; Cheng Xu*; Huaidong Zhang; Chenxi Zheng; Jing Qin; Shengfeng He*

147. Sync from the Sea: Retrieving Alignable Videos from Large-Scale Datasets; *Ishan Rajendrakumar Dave*; Fabian Caba; Mubarak Shah; Simon Jenni**

148. MAGR: Manifold-Aligned Graph Regularization for Continual Action Quality Assessment; *Kanglei Zhou; Liyuan Wang; Xingxing Zhang; Hubert P. H. Shum; Frederick W. B. Li; Jianguo Li; Xiaohui Liang**

149. E3M: Zero-Shot Spatio-Temporal Video Grounding with Expectation-Maximization Multimodal Modulation; *Peijun Bao*; Zihao Shao; Wenhan Yang; Boon Poh Ng; Alex Kot*

150. Propose, Assess, Search: Harnessing LLMs for Goal-Oriented Planning in Instructional Videos; *Md Mohaiminul Islam*; Tushar Nagarajan; Huiyu Wang; FU-JEN CHU; Kris Kitani; Gedas Bertasius; Xitong Yang*

151. C2C: Component-to-Composition Learning for Zero-Shot Compositional Action Recognition; *Rongchang Li; Zhenhua Feng; Tianyang Xu; Linze Li; Xiao-Jun Wu*; Muhammad Awais; Sara Atito; Josef Kittler*

152. Classification Matters: Improving Video Action Detection with Class-Specific Attention; *Jinsung Lee; Taeoh Kim; Inwoong Lee; Minho Shim; Dongyoon Wee; Minsu Cho; Suha Kwak**

153. DEVIAS: Learning Disentangled Video Representations of Action and Scene; *Kyungho Bae; Youngrae Kim; Geo Ahn; Jinwoo Choi**

154. LongVLM: Efficient Long Video Understanding via Large Language Models; *Yuetian Weng; Mingfei Han; Haoyu He; Xiaojun Chang; Bohan Zhuang**

155. Adversarial Diffusion Distillation; *Axel Sauer*; Dominik Lorenz; Andreas Blattmann; Robin Rombach*

156. FMBoost: Boosting Latent Diffusion with Flow Matching; *Johannes S Fischer*; Ming Gui; Pingchuan Ma; Nick Stracke; Stefan Andreas Baumann; Vincent Tao Hu; Björn Ommer*

157. Animal Avatars: Reconstructing Animatable 3D Animals from Casual Videos; *Remy Sabathier*; David Novotny; Niloy Mitra*

158. Pyramid Diffusion for Fine 3D Large Scene Generation; *Yuheng Liu*; Xinke Li; Xueting Li; Lu Qi*; Chongshou Li; Ming-Hsuan Yang*

159. OmniSSR: Zero-shot Omnidirectional Image Super-Resolution using Stable Diffusion Model; *Runyi Li*; Xuhan Sheng; Weiqi Li; Jian Zhang**

160. Tackling Structural Hallucination in Image Translation with Local Diffusion; *Seunghoi Kim*; Chen Jin; Tom Diethe; Matteo Figini; Henry FJ Tregidgo; Asher Mullokandov; Philip A Teare; Daniel Alexander*

161. Exact Diffusion Inversion via Bidirectional Integration Approximation; *Guoqiang Zhang*; j.p. lewis; W. Bastiaan Kleijn*

162. Diffusion Prior-Based Amortized Variational Inference for Noisy Inverse Problems; *Sojin Lee; Dogyun Park; Inho Kong; Hyunwoo J. Kim**

163. WBP: Training-time Backdoor Attacks through Hardware-based Weight Bit Poisoning; *Kunbei Cai*; Zhenkai Zhang; Qian Lou; Fan Yao**

164. Not Just Change the Labels, Learn the Features: Watermarking Deep Neural Networks with Multi-View Data; *Yuxuan Li; Sarthak Kumar Maharana; Yunhui Guo**

165. Rotary Position Embedding for Vision Transformer; *Byeongho Heo*; Song Park; Dongyoon Han; Sangdoo Yun*

166. OvSW: Overcoming Silent Weights for Accurate Binary Neural Networks; *jingyang xiang*; Zuohui Chen; Siqi Li; Qing Wu; Yong Liu*

167. Similarity of Neural Architectures using Adversarial Attack Transferability; *Jaehui Hwang; Dongyoon Han; Byeongho Heo; Song Park; Sanghyuk Chun*; Jong-Seok Lee*

168. An Image is Worth 1/2 Tokens After Layer 2: Plug-and-Play Inference Acceleration for Large Vision-Language Models; *Liang Chen; Haozhe Zhao; Tianyu Liu; Shuai Bai; Junyang Lin; Chang Zhou; Baobao Chang**

169. GiT: Towards Generalist Vision Transformer through Universal Language Interface; *Haiyang Wang*; Hao Tang; Li Jiang; Shaoshuai Shi; Muhammad Ferjad Naeem; Hongsheng Li; Bernt Schiele; Liwei Wang*

170. Reflective Instruction Tuning: Mitigating Hallucinations in Large Vision-Language Models; *Jinrui Zhang; Teng Wang; Haigang Zhang; Ping Lu; Feng Zheng**

171. Multi-branch Collaborative Learning Network for 3D Visual Grounding; *Zhipeng Qian; Yiwei Ma; Zhekai Lin; Jiayi Ji; Xiawu Zheng; Xiaoshuai Sun*; Rongrong Ji*

172. Turbo: Informativity-Driven Acceleration Plug-In for Vision-Language Large Models; *Chen Ju*; Haicheng Wang; Haozhe Cheng; Xu Chen; Zhonghua Zhai; Weilin Huang; Jinsong Lan; Shuai Xiao*; Bo Zheng*

173. SegPoint: Segment Any Point Cloud via Large Language Model; *Shuting He; Henghui Ding; Xudong Jiang; Bihan Wen**

174. Situated Instruction Following; *So Yeon Min*; Xavier Puig; Devendra Singh Chaplot; Tsung-Yen Yang; Priyam Parashar; Akshara Rai; Ruslan Salakhutdinov; Yonatan Bisk; Roozbeh Mottaghi*

175. Vary: Scaling up the Vision Vocabulary for Large Vision-Language Models; *Haoran Wei*; Lingyu Kong; Jinyue Chen; Liang Zhao; Zheng Ge; Jinrong Yang; Jianjian Sun; Chunrui Han; Xiangyu Zhang*

176. Common Sense Reasoning for Deep Fake Detection; *Yue Zhang*; Ben Colman; Xiao Guo; Ali Shahriyari; Gaurav Bharaj**

177. Dissecting Dissonance: Benchmarking Large Multimodal Models Against Self-Contradictory Instructions; *Jin Gao; Lei Gan; Yuankai Li; Yixin Ye; Dequan Wang**

178. GRACE: Graph-Based Contextual Debiasing for Fair Visual Question Answering; *Yifeng Zhang; Ming Jiang; Qi Zhao**

179. Resilience of Entropy Model in Distributed Neural Networks; *Milin Zhang*; Mohammad Abdi; Shahriar Rifat; Francesco Restuccia*

180. PosFormer: Recognizing Complex Handwritten Mathematical Expression with Position Forest Transformer; *Tongkun Guan; Chengyu Lin; Wei Shen*; Xiaokang Yang*

181. Efficient Inference of Vision Instruction-Following Models with Elastic Cache; *Zuyan Liu; Benlin Liu; Jiahui Wang; Yuhao Dong; Guangyi Chen; Yongming Rao; Ranjay Krishna; Jiwen Lu**

182. MMBENCH: Is Your Multi-Modal Model an All-around Player?; *Yuan Liu*; Haodong Duan*; Yuanhan Zhang; Bo Li; Songyang Zhang; Wangbo Zhao; Yike Yuan; Jiaqi Wang; Conghui He; Ziwei Liu; Kai Chen; Dahua Lin*

183. A Simple Baseline for Spoken Language to Sign Language Translation with 3D Avatars; *Ronglai Zuo; Fangyun Wei*; Zenggui Chen; Brian Mak; Jiaolong Yang; Xin Tong*

184. LiteSAM is Actually what you Need for segment Everything; *Jianhai Fu; Yuanjie Yu; Ningchuan Li*; Yi Zhang; Qichao Chen; Jianping Xiong; Jun Yin; Zhiyu Xiang**

185. ShapeLLM: Universal 3D Object Understanding for Embodied Interaction; *Zekun Qi; Runpei Dong; Shaochen Zhang; Haoran Geng; Chunrui Han; Zheng Ge; Li Yi*; Kaisheng Ma**

186. Affective Visual Dialog: A Large-Scale Benchmark for Emotional Reasoning Based on Visually Grounded Conversations; *Kilichbek Haydarov*; Xiaoqian Shen; Avinash Madasu; Mahmoud Salem; Li-Jia Li; Gamaleldin F Elsayed; Mohamed Elhoseiny*

187. CLIP-DPO: Vision-Language Models as a Source of Preference for Fixing Hallucinations in LVLMs; *Yassine Ouali*; Adrian Bulat*; Brais Martinez; Georgios Tzimiropoulos*

188. Discovering Unwritten Visual Classifiers with Large Language Models; *Mia Chiquier*; Utkarsh Mall; Carl Vondrick*

189. BLINK: Multimodal Large Language Models Can See but Not Perceive; *Xingyu Fu*; Yushi Hu*; Bangzheng Li; Yu Feng; Haoyu Wang; Xudong Lin; Dan Roth; Noah A Smith; Wei-Chiu Ma; Ranjay Krishna*

190. BRAVE: Broadening the visual encoding of vision-language models; *Oğuzhan Fatih Kar*; Alessio Tonioni*; Petra Poklukar; Achin Kulshrestha; Amir Zamir; Federico Tombari*

191. Omniview-Tuning: Boosting Viewpoint Invariance of Vision-Language Pre-training Models; *Shouwei Ruan*; Yinpeng Dong; Liu Hanqing; Yao Huang; Hang Su; Xingxing Wei**

192. Conceptual Codebook Learning for Vision-Language Models; *Yi Zhang*; Ke Yu; Siqi Wu; Zhihai He**

193. Improving Vision and Language Concepts Understanding with Multimodal Counterfactual Samples; *Chengen Lai; Shengli Song*; Sitong Yan; Guangneng Hu*

194. Meta-Prompting for Automating Zero-shot Visual Recognition with LLMs; *Muhammad Jehanzeb Mirza*; Leonid Karlinsky; Wei Lin; Sivan Doveh; Jakub Micorek; Mateusz Kozinski; Hilde Kuehne; Horst Possegger*

195. Teach CLIP to Develop a Number Sense for Ordinal Regression; *Yao DU*; Qiang Zhai; Weihang Dai; Xiaomeng Li**

196. Where am I? Scene Retrieval with Language; *Jiaqi Chen*; Daniel Barath; Iro Armeni; Marc Pollefeys; Hermann Blum*

197. Do Generalised Classifiers really work on Human Drawn Sketches?; *Hmrishav Bandyopadhyay*; Pinaki Nath Chowdhury; Aneeshan Sain; Subhadeep Koley; Tao Xiang; Ayan Kumar Bhunia; Yi-Zhe Song*

198. Seeing Faces in Things: A Model and Dataset for Pareidolia; *Mark T Hamilton*; Simon Stent; Vasha G DuTell; Anne Harrington; Jennifer E Corbett; Ruth Rosenholtz; William T. Freeman*

199. Strengthening Multimodal Large Language Model with Bootstrapped Preference Optimization; *Renjie Pi*; Tianyang Han; Wei Xiong; Jipeng ZHANG; Runtao Liu; Rui Pan; Tong Zhang*

200. SDPT: Synchronous Dual Prompt Tuning for Fusion-based Visual-Language Pre-trained Models; *Yang Zhou*; Yongjian Wu; Jiya Saiyin; Bingzheng Wei; Maode Lai; Eric I Chang; Yan Xu**

201. Improving Zero-Shot Generalization for CLIP with Variational Adapter; *Ziqian Lu; Fengli Shen; Mushui Liu; Yunlong Yu*; Xi Li*

202. uCAP: An Unsupervised Prompting Method for Vision-Language Models; *A. Tuan Nguyen*; Kai Sheng Tai; Bor-Chun Chen; Satya Narayan Shukla; Hanchao Yu; Philip Torr; Tai-Peng Tian; Ser-Nam Lim*

203. LLM as Dataset Analyst: Subpopulation Structure Discovery with Large Language Model; *Yulin Luo; Ruichuan An; Bocheng Zou; Yiming Tang; Jiaming Liu; Shanghang Zhang**

204. SILC: Improving Vision Language Pretraining with Self-Distillation; *Muhammad Ferjad Naeem*; Yongqin Xian; Xiaohua Zhai; Lukas Hoyer; Luc Van Gool; Federico Tombari*

205. Evaluating Text-to-Visual Generation with Image-to-Text Generation; *Zhiqiu Lin*; Deepak Pathak; Baiqi Li; Jiayao Li; Xide Xia; Graham Neubig; Pengchuan Zhang; Deva Ramanan*

206. LLaVA-UHD: an LMM Perceiving any Aspect Ratio and High-Resolution Images; *Zonghao Guo; Ruyi Xu; Yuan Yao*; Junbo Cui; Zanlin Ni; Chunjiang Ge; Tat-Seng Chua; Zhiyuan Liu; Gao Huang**

207. Removing Distributional Discrepancies in Captions Improves Image-Text Alignment; *Mu Cai; Haotian Liu; Yuheng Li*; Yijun Li; Eli Shechtman; Zhe Lin; Yong Jae Lee; Krishna Kumar Singh*

208. DOCCI: Descriptions of Connected and Contrasting Images; *Yasumasa Onoe*; Sunayana Rane; Zachary E Berger; Yonatan Bitton; Jaemin Cho; Roopal Garg; Alexander Ku; Zarana Parekh; Jordi Pont-Tuset; Garrett Tanzer; Su Wang; Jason M Baldridge*

209. Image Compression for Machine and Human Vision With Spatial-Frequency Adaptation; *Han Li*; Shaohui Li*; Shuangrui Ding; Wenrui Dai*; Maida Cao; Chenglin Li; Junni Zou; Hongkai Xiong*

210. WeConvene: Learned Image Compression with Wavelet-Domain Convolution and Entropy Model; *Haisheng Fu*; Jie Liang; Zhenman Fang; Jingning Han; Feng Liang; Guohe Zhang*

211. HYPE: Hyperbolic Entailment Filtering for Underspecified Images and Texts; *Wonjae Kim*; Sanghyuk Chun; Taekyung Kim; Dongyoon Han; Sangdoo Yun*

212. CoPT: Unsupervised Domain Adaptive Segmentation using Domain-Agnostic Text Embeddings; *Cristina Mata*; Kanchana N Ranasinghe; Michael S Ryoo*

213. PLOT: Text-based Person Search with Part Slot Attention for Corresponding Part Discovery; *Jicheol Park; Dongwon Kim; Boseung Jeong; Suha Kwak**

214. TTD: Text-Tag Self-Distillation Enhancing Image-Text Alignment in CLIP to Alleviate Single Tag Bias; *Sanghyun Jo; Soohyun Ryu; Sungyub Kim; Eunho Yang; Kyungsu Kim**

215. Distractors-Immune Representation Learning with Cross-modal Contrastive Regularization for Change Captioning; *Yunbin Tu*; Liang Li; Li Su; Chenggang Yan; Qingming Huang*

216. IRGen: Generative Modeling for Image Retrieval; *Yidan Zhang*; Ting Zhang*; Dong Chen; Yujing Wang; Qi Chen; Xing Xie; Hao Sun; Weiwei Deng; Qi Zhang; Fan Yang; Mao Yang; Qingmin Liao; Jingdong Wang; Baining Guo*

217. ConceptExpress: Harnessing Diffusion Models for Single-image Unsupervised Concept Extraction; *Shaozhe Hao*; Kai Han*; Zhengyao Lv; Shihao Zhao; Kwan-Yee K. Wong**

218. Teddy: Efficient Large-Scale Dataset Distillation via Taylor-Approximated Matching; *Ruonan Yu; Songhua Liu; Jingwen Ye; Xinchao Wang**

219. Neural Spectral Decomposition for Dataset Distillation; *Shaolei Yang; Shen Cheng; Mingbo Hong; Haoqiang Fan; Xing Wei; Shuaicheng Liu**

220. DECap: Towards Generalized Explicit Caption Editing via Diffusion Mechanism; *Zhen Wang; Xinyun Jiang; Jun Xiao; Tao Chen; Long Chen**

221. Towards Unified Representation of Invariant-Specific Features in Missing Modality Face Anti-Spoofing; *Guanghao Zheng; Yuchen Liu; Wenrui Dai*; Chenglin Li; Junni Zou; Hongkai Xiong*

222. Personalized Privacy Protection Mask Against Unauthorized Facial Recognition; *Ka-Ho Chow*; Sihao Hu; Tiansheng Huang; Ling Liu*

223. Rethinking Tree-Ring Watermarking for Enhanced Multi-Key Identification; *Hai Ci*; Pei Yang; Yiren Song; Mike Zheng Shou**

224. T2IShield: Defending Against Backdoors on Text-to-Image Diffusion Models; *Zhongqi Wang; Jie Zhang*; Shiguang Shan; Xilin Chen*

225. Enhancing Tampered Text Detection through Frequency Feature Fusion and Decomposition; *Zhongxi Chen; Shen Chen; Taiping Yao*; Ke Sun; Shouhong Ding; Xianming Lin*; Liujuan Cao; Rongrong Ji*

226. GAMMA-FACE: GAussian Mixture Models Amend Diffusion Models for Bias Mitigation in Face Images; *Basudha Pal*; Arunkumar Kannan*; Ram Prabhakar Kathirvel; Alice O'Toole; Rama Chellappa*

227. An Empirical Study and Analysis of Text-to-Image Generation Using Large Language Model-Powered Textual Representation; *Zhiyu Tan; Mengping Yang; Luozheng Qin ; Hao Yang; Ye Qian ; Qiang Zhou; Cheng Zhang; Hao Li**

228. Latent Guard: a Safety Framework for Text-to-image Generation; *Runtao Liu*; Ashkan Khakzar; Jindong Gu; Qifeng Chen*; Philip Torr; Fabio Pizzati**

229. Arc2Face: A Foundation Model for ID-Consistent Human Faces; *Foivos Paraperas Papantoniou*; Alexandros Lattas; Stylianos Moschoglou; Jiankang Deng; Bernhard Kainz; Stefanos Zafeiriou*

230. SpeedUpNet: A Plug-and-Play Adapter Network for Accelerating Text-to-Image Diffusion Models; *Weilong Chai*; Dandan Zheng; Jiajiong Cao; Zhiquan Chen; Changbao Wang; Chenguang Ma*

231. CogView3: Finer and Faster Text-to-Image Generation via Relay Diffusion; *Wendi Zheng*; Jiayan Teng; Zhuoyi Yang; Weihan Wang; Jidong Chen; Xiaotao Gu; Yuxiao Dong*; Ming Ding*; Jie Tang**

232. Efficient Diffusion Transformer with Step-wise Dynamic Attention Mediators; *Yifan Pu*; Zhuofan Xia; Jiayi Guo; Dongchen Han; Qixiu Li; Duo Li; Yuhui Yuan; Ji Li; Yizeng Han; Shiji Song; Gao Huang*; Xiu Li**

233. Large-scale Reinforcement Learning for Diffusion Models; *Yinan Zhang*; Eric Tzeng; Yilun Du; Dmitry Kislyuk**

234. Stable Preference: Redefining training paradigm of human preference model for Text-to-Image Synthesis; *Hanting Li; Hongjing Niu; Feng Zhao**

235. Instant 3D Human Avatar Generation using Image Diffusion Models; *Nikos Kolotouros*; Thiemo Alldieck; Enric Corona; Eduard Gabriel Bazavan; Cristian Sminchisescu*

236. Photorealistic Video Generation with Diffusion Models; *Agrim Gupta*; Lijun Yu; Kihyuk Sohn; Xiuye Gu; Meera Hahn; Li Fei-Fei; Irfan Essa; Lu Jiang; Jose Lezama*

237. Closed-Loop Unsupervised Representation Disentanglement with β-VAE Distillation and Diffusion Probabilistic Feedback; *Xin Jin*; Bohan Li*; Baao Xie; Wenyao Zhang; Jinming Liu; Ziqiang Li; Tao Yang; Wenjun Zeng*

238. Customize-A-Video: One-Shot Motion Customization of Text-to-Video Diffusion Models; *Yixuan Ren*; Yang Zhou; Jimei Yang; Jing Shi; Difan Liu; Feng Liu; Mingi Kwon; Abhinav Shrivastava*

239. Learn to Optimize Denoising Scores: A Unified and Improved Diffusion Prior for 3D Generation; *Xiaofeng Yang*; Yiwen Chen; Cheng Chen; Chi Zhang; Yi Xu; Xulei Yang; Fayao Liu; Guosheng Lin*

240. LivePhoto: Real Image Animation with Text-guided Motion Control; *Xi Chen; Zhiheng Liu; Mengting Chen; Yutong Feng; Yu Liu; Yujun Shen; Hengshuang Zhao**

241. Text-to-Sticker: Style Tailoring Latent Diffusion Models for Human Expression; *Animesh Sinha*; Bo Sun; Anmol Kalia; Arantxa Casanova; Elliot Blanchard; David Yan; Winnie Zhang; Tony Nelli; Jiahui Chen; Hardik Shah; Licheng Yu; Mitesh Kumar Singh; Ankit Ramchandani; Maziar Sanjabi; Sonal Gupta; Amy L Bearman; Dhruv Mahajan*

242. Self-Supervised Audio-Visual Soundscape Stylization; *Tingle Li*; Renhao Wang; Po-Yao Huang; Andrew Owens; Gopala Krishna Anumanchipalli*

243. ProCreate, Don't Reproduce! Propulsive Energy Diffusion for Creative Generation; *Jack Lu*; Ryan Teehan*; Mengye Ren**

244. Implicit Style-Content Separation using B-LoRA; *Yarden Frenkel*; Yael Vinker; Ariel Shamir; Danny Cohen-Or*

245. TC4D: Trajectory-Conditioned Text-to-4D Generation; *Sherwin Bahmani*; Xian Liu; Wang Yifan; Ivan Skorokhodov; Victor Rong; Ziwei Liu; Xihui Liu; Jeong Joon Park; Sergey Tulyakov; Gordon Wetzstein; Andrea Tagliasacchi; David B Lindell*

246. ColorMAE: Exploring data-independent masking strategies in Masked AutoEncoders; *Carlos Hinojosa*; Shuming Liu; Bernard Ghanem*

247. Diffusion-Based Image-to-Image Translation by Noise Correction via Prompt Interpolation; *Junsung Lee; Minsoo Kang; Bohyung Han**

248. Put Myself in Your Shoes: Lifting the Egocentric Perspective from Exocentric Videos; *Mi Luo*; Zihui Xue; Alex Dimakis; Kristen Grauman*

249. Context Diffusion: In-Context Aware Image Generation; *Ivona Najdenkoska*; Animesh Sinha; Abhimanyu Dubey; Dhruv Mahajan; Vignesh Ramanathan; Filip Radenovic*

250. Label-free Neural Semantic Image Synthesis; *Jiayi Wang*; Kevin A Laube; Yumeng Li; Jan Hendrik Metzen; Shin-I Cheng; Julio Borges; Anna Khoreva*

251. Dolfin: Diffusion Layout Transformers without Autoencoder; *Yilin Wang; Zeyuan Chen; Liangjun Zhong; Zheng Ding; Zhuowen Tu**

252. ByteEdit: Boost, Comply and Accelerate Generative Image Editing; *Yuxi Ren; Jie Wu*; Yanzuo Lu; Huafeng Kuang; Xin Xia; Xionghui Wang; Qianqian Wang; Yixing Zhu; Pan Xie; Shiyin Wang; Xuefeng Xiao; Yitong Wang; Min Zheng; Lean FU*

253. Revisiting Feature Disentanglement Strategy in Diffusion Training and Breaking Conditional Independence Assumption in Sampling; *Wonwoong Cho*; Hareesh Ravi*; Midhun Harikumar; Vinh Khuc; Krishna Kumar Singh; Jingwan Lu; David Iseri Inouye*; Ajinkya Kale**

254. Layout-Corrector: Alleviating Layout Sticking Phenomenon in Discrete Diffusion Model; *Shoma Iwai*; Atsuki Osanai; Shunsuke Kitada; Shinichiro Omachi*

255. DreamSampler: Unifying Diffusion Sampling and Score Distillation for Image Manipulation; *Jeongsol Kim; Geon Yeong Park; Jong Chul Ye**

256. TexGen: Text-Guided 3D Texture Generation with Multi-view Sampling and Resampling; *Dong Huo*; Zixin Guo; Xinxin Zuo; Zhihao Shi; Juwei Lu; Peng Dai; Songcen Xu; Li Cheng; Yee-Hong Yang*

257. EraseDraw : Learning to Insert Objects by Erasing Them from Images; *Alper Canberk*; Maksym Bondarenko; Ege Ozguroglu; Ruoshi Liu; Carl Vondrick*

258. Make-Your-3D: Fast and Consistent Subject-Driven 3D Content Generation; *Fangfu Liu; Hanyang Wang; Weiliang Chen; Haowen Sun; Yueqi Duan**

259. Time-Efficient and Identity-Consistent Virtual Try-On Using A Variant of Altered Diffusion Models; *Phuong*

Hoang Dam; Jihoon Jeong*; Anh T Tran*; Daeyoung Kim**

260. Text2Place: Affordance-aware Text Guided Human Placement; *Rishubh Parihar*; Harsh Gupta; Sachidanand VS; Venkatesh Babu RADHAKRISHNAN*

261. GaussCtrl: Multi-View Consistent Text-Driven 3D Gaussian Splatting Editing; *Jing Wu*; Jia-Wang Bian; Xinghui Li; Guangrun Wang; Ian Reid; Philip Torr; Victor Adrian Prisacariu**

262. Editable Image Elements for Controllable Synthesis; *Jiteng Mu*; Michaël Gharbi; Richard Zhang; Eli Shechtman; Nuno Vasconcelos; Xiaolong Wang; Taesung Park**

263. WebRPG: Automatic Web Rendering Parameters Generation for Visual Presentation; *Zirui Shao; Feiyu Gao; Hangdi Xing; Zepeng Zhu; Zhi Yu*; Jiajun Bu; Qi Zheng; Cong Yao*

264. GeometrySticker: Enabling Ownership Claim of Recolorized Neural Radiance Fields; *Xiufeng HUANG*; Ka Chun Cheung; Simon See; Renjie Wan**

265. Chat-Edit-3D: Interactive 3D Scene Editing via Text Prompts; *shuangkang fang*; Yufeng Wang*; Yi-Hsuan Tsai; Yi Yang; Wenrui Ding; Shuchang Zhou; Ming-Hsuan Yang*

266. Synthesizing Environment-Specific People in Photographs; *Mirela Ostrek*; Carol O'Sullivan; Michael J. Black; Justus Thies*

267. Be-Your-Outpainter: Mastering Video Outpainting through Input-Specific Adaptation; *Fu-Yun Wang*; Xiaoshi Wu; Zhaoyang Huang; Xiaoyu Shi; Dazhong Shen; Guanglu Song; Yu Liu; Hongsheng Li**

268. Head360: Learning a Parametric 3D Full-Head for Free-View Synthesis in 360°; *Yuxiao He; Yiyu Zhuang; Yanwen Wang; Yao Yao; Siyu Zhu; Xiaoyu Li; Qi Zhang; Xun Cao; Hao Zhu**

269. Hierarchical Conditioning of Diffusion Models Using Tree-of-Life for Studying Species Evolution; *Mridul Khurana*; Arka Daw; M. Maruf; Josef C. Uyeda; Wasila Dahdul; Caleb Charpentier; Yasin Bakış; Henry L. Bart Jr.; Paula M. Mabee; Hilmar Lapp; James P. Balhoff; Wei-Lun Chao; Charles Stewart; Tanya Berger-Wolf; Anuj Karpatne**

270. AnimateMe: 4D Facial Expressions via Diffusion Models; *Dimitrios Gerogiannis*; Foivos Paraperas Papantoniou; Rolandos Alexandros Potamias; Alexandros Lattas; Stylianos Moschoglou; Stylianos Ploumpis; Stefanos Zafeiriou*

271. Tri^{2}-plane: Thinking Head Avatar via Feature Pyramid; *Luchuan Song*; Pinxin Liu; Lele Chen; Guojun Yin; Chenliang Xu*

272. Shapefusion: 3D localized human diffusion models; *Rolandos Alexandros Potamias*; Michael Tarasiou; Stylianos Ploumpis; Stefanos Zafeiriou*

273. Few-Shot Image Generation by Conditional Relaxing Diffusion Inversion; *Yu Cao*; Shaogang Gong*

274. ViLA: Efficient Video-Language Alignment for Video Question Answering; *Xijun Wang*; Junbang Liang; Chun-Kai Wang; Kenan Deng; Yu Lou; Ming C Lin; Shan Yang*

275. ST-LLM: Large Language Models Are Effective Temporal Learners; *Ruyang Liu; Chen Li; Haoran Tang; Yixiao Ge; Ying Shan; Ge Li**

276. Attention Beats Linear for Fast Implicit Neural Representation Generation; *Shuyi Zhang; Ke Liu; Jingjun Gu; Xiaoxu Cai; Zhihua Wang; Jiajun Bu; Haishuai Wang**

277. AvatarPose: Avatar-guided 3D Pose Estimation of Close Human Interaction from Sparse Multi-view Videos; *Feichi Lu*; Zijian Dong*; Jie Song; Otmar Hilliges*

278. Human Hair Reconstruction with Strand-Aligned 3D Gaussians; *Egor Zakharov*; Vanessa Sklyarova; Michael J. Black; Giljoo Nam; Justus Thies; Otmar Hilliges*

279. Rejection Sampling IMLE: Designing Priors for Better Few-Shot Image Synthesis; *Chirag Vashist*; Shichong Peng; Ke Li*

280. LATTE3D: Large-scale Amortized Text-To-Enhanced3D Synthesis; *Kevin Xie*; Tianshi Cao; Jonathan P Lorraine; Jun Gao; James R Lucas; Antonio Torralba; Sanja Fidler; Xiaohui Zeng*

281. TetraDiffusion: Tetrahedral Diffusion Models for 3D Shape Generation; *Nikolai Kalischek*; Torben Peters; Jan Dirk Wegner; Konrad Schindler*

282. Fast Sprite Decomposition from Animated Graphics; *Tomoyuki Suzuki*; Kotaro Kikuchi; Kota Yamaguchi*

283. MVDiffHD: A Dense High-resolution Multi-view Diffusion Model for Single or Sparse-view 3D Object Reconstruction; *Shitao Tang*; Jiacheng Chen; Dilin Wang; Chengzhou Tang; Fuyang Zhang; Yuchen Fan; Vikas*

Chandra; Yasutaka Furukawa; Rakesh Ranjan

284. MSD: A Benchmark Dataset for Floor Plan Generation of Building Complexes; *Casper van Engelenburg*; Fatemeh Mostafavi; Emanuel Kuhn; Yuntae Jeon; Michael Franzen; Matthias Standfest; Jan van Gemert; Seyran Khademi*

285. SC4D: Sparse-Controlled Video-to-4D Generation and Motion Transfer; *Zijie Wu*; Chaohui Yu; Yanqin Jiang; Chenjie Cao; Fan Wang; Xiang Bai**

286. CRM: Single Image to 3D Textured Mesh with Convolutional Reconstruction Model; *Zhengyi Wang*; Yikai Wang; Yifei Chen; Chendong Xiang; Shuo Chen; Dajiang Yu; Chongxuan Li; Hang Su; Jun Zhu*

287. AttnZero: Efficient Attention Discovery for Vision Transformers; *Lujun Li; Zimian Wei*; Peijie Dong; Wenhan Luo; Wei Xue; Qifeng Liu*; Yike Guo**

288. RoofDiffusion: Constructing Roofs from Severely Corrupted Point Data via Diffusion; *Kyle Shih-Huang Lo*; Jorg Peters; Eric Spellman*

289. Per-Gaussian Embedding-Based Deformation for Deformable 3D Gaussian Splatting; *Jeongmin Bae; Seoha Kim; Youngsik Yun; Hahyun Lee; Gun Bang; Youngjung Uh**

290. DreamScene360: Unconstrained Text-to-3D Scene Generation with Panoramic Gaussian Splatting; *Shijie Zhou*; Zhiwen Fan; Dejia Xu; Haoran Chang; Pradyumna Chari; Tejas K Bharadwaj; Suya You; Zhangyang Wang; Achuta Kadambi*

291. DynMF: Neural Motion Factorization for Real-time Dynamic View Synthesis with 3D Gaussian Splatting; *Angelos Kratimenos*; Jiahui Lei; Kostas Daniilidis*

292. City-on-Web: Real-time Neural Rendering of Large-scale Scenes on the Web; *Kaiwen Song; Xiaoyi Zeng; Chenqu Ren; Juyong Zhang**

293. GeoGaussian: Geometry-aware Gaussian Splatting for Scene Rendering; *Yanyan Li*; Chenyu Lyu; Yan Di; Guangyao Zhai; Gim Hee Lee; Federico Tombari*

294. Cascade-Zero123: One Image to Highly Consistent 3D with Self-Prompted Nearby Views; *Yabo Chen; Jiemin Fang; Yuyang Huang; Taoran Yi; Xiaopeng Zhang*; Lingxi Xie; Xinggang Wang; Wenrui Dai*; Hongkai Xiong; Qi Tian*

295. End-to-End Rate-Distortion Optimized 3D Gaussian Representation; *Henan Wang*; Hanxin Zhu; Tianyu He; Runsen Feng; Jiajun Deng; Jiang Bian; Zhibo Chen*

296. EAGLES: Efficient Accelerated 3D Gaussians with Lightweight EncodingS; *Sharath Girish*; Kamal Gupta; Abhinav Shrivastava*

297. HO-Gaussian: Hybrid Optimization of 3D Gaussian Splatting for Urban Scenes; *Zhuopeng Li*; Yilin Zhang; Chenming Wu; Jianke Zhu*; Liangjun Zhang*

298. Improving Neural Surface Reconstruction with Feature Priors from Multi-View Images; *Xinlin Ren*; Chenjie Cao; Yanwei Fu*; Xiangyang Xue*

299. SG-NeRF: Neural Surface Reconstruction with Scene Graph Optimization; *Yiyang Chen; Siyan Dong*; Xulong Wang; Lulu Cai; Youyi Zheng; Yanchao Yang**

300. WoVoGen: World Volume-aware Diffusion for Controllable Multi-camera Driving Scene Generation; *Jiachen Lu; Ze Huang; Zeyu Yang; Zhang Jiahui; Li Zhang**

301. Generalizable Human Gaussians for Sparse View Synthesis; *YoungJoong Kwon*; Baole Fang; Yixing Lu; Haoye Dong; Cheng Zhang; Francisco Vicente Carrasco; Albert Mosella-Montoro; Jianjin Xu; Shingo J Takagi; Daeil Kim; Aayush Prakash; Fernando de la Torre*

302. Few-shot NeRF by Adaptive Rendering Loss Regularization; *Qingshan Xu*; Xuanyu Yi; Jianyao Xu; Wenbing Tao; Yew Soon Ong; Hanwang Zhang*

303. Invertible Neural Warp for NeRF; *Shin-Fang Chng*; Ravi Garg; Hemanth Saratchandran; Simon Lucey*

304. UniVoxel: Fast Inverse Rendering by Unified Voxelization of Scene Representation; *Shuang Wu; Songlin Tang; Guangming Lu; Jianzhuang Liu; Wenjie Pei**

305. PISR: Polarimetric Neural Implicit Surface Reconstruction for Textureless and Specular Objects; *Guangcheng Chen*; Yicheng He; Li He; Hong Zhang*

306. Neural Poisson Solver: A Universal and Continuous Framework for Natural Signal Blending; *Delong Wu; Hao Zhu; Qi Zhang; You Li; Xun Cao*; Zhan Ma**

307. 3iGS: Factorised Tensorial Illumination for 3D Gaussian Splatting; *Zhe Jun Tang*; Tat-Jen Cham*

308. BAD-Gaussians: Bundle Adjusted Deblur Gaussian Splatting; *Lingzhe Zhao; Peng Wang; Peidong Liu**

309. Thermal3D-GS: Physics-induced 3D Gaussians for Thermal Infrared Novel-view Synthesis; *Qian Chen; Shihao Shu; Xiangzhi Bai**

310. Gaussian in the wild: 3D Gaussian Splatting for Unconstrained Image Collections; *Dongbin Zhang*; Chuming Wang; Weitao Wang; Peihao Li; Minghan Qin; Haoqian Wang**

311. Flash-Splat: 3D Reflection Removal with Flash Cues and Gaussian Splats; *Mingyang Xie*; Haoming Cai; Sachin Shah; Yiran Xu; Brandon Y. Feng; Jia-Bin Huang; Christopher A. Metzler*

312. Physical-Based Event Camera Simulator; *Haiqian Han; Jiacheng Lyu; Jianing Li*; Henglu Wei; Cheng Li; Yajing Wei; SHU CHEN; Xiangyang Ji**

313. Edge-Guided Fusion and Motion Augmentation for Event-Image Stereo; *Fengan Zhao*; Qianang Zhou; Junlin Xiong**

314. REDIR: Refocus-free Event-based De-occlusion Image Reconstruction; *Qi Guo; Hailong Shi*; Huan Li; Jinsheng Xiao; Xingyu Gao**

315. High-Fidelity and Transferable NeRF Editing by Frequency Decomposition; *Yisheng He*; Weihao Yuan*; Siyu Zhu; Zilong Dong; Liefeng Bo; Qixing Huang*

316. Depth-Aware Blind Image Decomposition for Real-World Adverse Weather Recovery; *Chao Wang*; Zhedong Zheng; Ruijie Quan; Yi Yang*

317. Self-Supervised Underwater Caustics Removal and Descattering via Deep Monocular SLAM; *Jonathan Sauder*; Devis Tuia*

318. Real-data-driven 2000 FPS Color Video from Mosaicked Chromatic Spikes; *Siqi Yang*; Zhaojun Huang; Yakun Chang; Bin Fan; Zhaofei Yu; Boxin Shi*

319. Raindrop Clarity: A Dual-Focused Dataset for Day and Night Raindrop Removal; *Yeying Jin*; Xin Li; Jiadong Wang; Yan Zhan; Malu Zhang**

320. L-DiffER: Single Image Reflection Removal with Language-based Diffusion Model; *Yuchen Hong*; Haofeng Zhong*; Shuchen Weng; Jinxiu S Liang; Boxin Shi*

321. Joint RGB-Spectral Decomposition Model Guided Image Enhancement in Mobile Photography; *Kailai Zhou*; Lijing Cai; Yibo Wang; Mengya Zhang; Bihan Wen; Qiu Shen*; Xun Cao*

322. Exploiting Dual-Correlation for Multi-frame Time-of-Flight Denoising; *Guanting Dong*; Yueyi Zhang*; Xiaoyan Sun; Zhiwei Xiong*

323. DualDn: Dual-domain Denoising via Differentiable ISP; *Ruikang Li; Yujin Wang*; Shiqi Chen; Fan Zhang; Jinwei Gu; Tianfan Xue*

324. Contribution-based Low-Rank Adaptation with Pre-training Model for Real Image Restoration; *Dongwon Park; Hayeon Kim; Se Young Chun**

325. Seeing the Unseen: A Frequency Prompt Guided Transformer for Image Restoration; *Shihao Zhou; Jinshan Pan; Jinglei Shi*; Duosheng Chen; Lishen Qu; Jufeng Yang*

326. Functional Transform-Based Low-Rank Tensor Factorization for Multi-Dimensional Data Recovery; *Jian-Li Wang; Xi-Le Zhao**

327. BurstM: Deep Burst Multi-scale SR using Fourier Space with Optical Flow; *EungGu Kang*; Byeonghun Lee; Sunghoon Im; Kyong Hwan Jin*

328. LMT-GP: Combined Latent Mean-Teacher and Gaussian Process for Semi-supervised Low-light Image Enhancement; *Ye Yu; Fengxin Chen; Jun Yu*; Zhen Kan*

329. Hierarchical Separable Video Transformer for Snapshot Compressive Imaging; *Ping Wang*; Yulun Zhang; Lishun Wang; Xin Yuan**

330. Rethinking Video Deblurring with Wavelet-Aware Dynamic Transformer and Diffusion Model; *Chen Rao; Guangyuan Li; Zehua Lan; Jiakai Sun; Junsheng Luan; Wei Xing*; Lei Zhao*; Huaizhong Lin*; Jianfeng Dong; Dalong*

Zhang

331. An Optimal Control View of LoRA and Binary Controller Design for Vision Transformers; *Chi Zhang*; Jingpu Cheng; Qianxiao Li*

332. AdaDiffSR: Adaptive Region-aware Dynamic acceleration Diffusion Model for Real-World Image Super-Resolution; *Yuanting Fan; Chengxu Liu; Nengzhong Yin; Changlong Gao; Xueming Qian**

334. XPSR: Cross-modal Priors for Diffusion-based Image Super-Resolution; *Qu Yunpeng*; Kun Yuan; Kai Zhao; Qizhi Xie; Jinhua Hao; Ming Sun; Chao Zhou*

335. DenseNets Reloaded: Paradigm Shift Beyond ResNets and ViTs; *DongHyun Kim; Byeongho Heo; Dongyoon Han**

336. Robustness Tokens: Towards Adversarial Robustness of Transformers; *Brian Pulfer*; Yury Belousov; Slava Voloshynovskiy*

337. Isomorphic Pruning for Vision Models; *Gongfan Fang*; Xinyin Ma; Michael Bi Mi; Xinchao Wang**

338. Adaptive Multi-head Contrastive Learning; *Lei Wang*; Piotr Koniusz; Tom Gedeon; Liang Zheng*

339. AdaDiff: Accelerating Diffusion Models through Step-Wise Adaptive Computation; *Shengkun Tang*; Yaqing Wang; Caiwen Ding; Yi Liang; Yao Li; Dongkuan Xu*

340. Energy-induced Explicit quantification for Multi-modality MRI fusion; *Xiaoming Qi*; Yuan Zhang; Tong Wang; Guanyu Yang*; Yueming Jin*; Shuo Li*

341. Imaging with Confidence: Uncertainty Quantification for High-dimensional Undersampled MR Images; *Frederik Hoppe*; Claudio Mayrink Verdun; Hannah Sophie Laus; Sebastian Endt; Marion Irene Menzel; Felix Krahmer; Holger Rauhut*

342. Style-Extracting Diffusion Models for Semi-Supervised Histopathology Segmentation; *Mathias Öttl*; Frauke Wilm; Jana Steenpass; Jingna Qiu; Matthias Rübner; Prof Arndt Hartmann; Matthias W. Beckmann; Peter Fasching; Andreas K Maier; Ramona Erber; Bernhard Kainz; Katharina Breininger*

343. GenerateCT: Text-Conditional Generation of 3D Chest CT Volumes; *Ibrahim Ethem Hamamci*; Sezgin Er; Anjany Sekuboyina; Enis Simsar; Alperen Tezcan; Ayse Gulnihan Simsek; Sevval Nil Esirgun; Furkan Almas; Irem Dogan; Muhammed Furkan Dasdelen; Chinmay Prabhakar; Hadrien Reynaud; Sarthak Pati; Christian Bluethgen; Mehmet Kemal Ozdemir; Bjoern Menze*

344. I-MedSAM: Implicit Medical Image Segmentation with Segment Anything; *Xiaobao Wei; Jiajun Cao; Yizhu Jin; Ming Lu; Guangyu Wang; Shanghang Zhang**

19:00 – 00:00
Dinner Party - Hall 4

FRIDAY, 4TH OCTOBER

08:00 – 12:30
Registration - Badge Pickup

09:00 – 12:30
Exhibition - Level 0

08:30 – 10:30
Oral session 7A: Learning architectures, transfer, continual and long-tail - Gold Room
Chair: Tatiana Tommasi; Kai Han

1. On the Topology Awareness and Generalization Performance of Graph Neural Networks; *Junwei Su*; Chuan Wu* BEST PAPER CANDIDATE

2. Improving Knowledge Distillation via Regularizing Feature Direction and Norm; *Yuzhu Wang; Lechao Cheng*; Manni Duan; Yongheng Wang; Zunlei Feng; Shu Kong*

3. Spline-based Transformers; *Prashanth Chandran*; Agon Serifi*; Markus Gross; Moritz Bächer*

4. Anytime Continual Learning for Open Vocabulary Classification; *Zhen Zhu*; Yiming Gong; Derek Hoiem**

5. Weighted Ensemble Models Are Strong Continual Learners; *Imad Eddine MAROUF*; Subhankar Roy; Enzo Tartaglione; Stéphane Lathuilière*

6. COD: Learning Conditional Invariant Representation for Domain Adaptation Regression; *Hao-Ran Yang; Chuan-Xian Ren*; You-Wei Luo*

7. Echoes of the Past: Boosting Long-tail Recognition via Reflective Learning; Qihao Zhao; Yalun Dai; Shen Lin; Wei Hu; Fan Zhang*; Jun Liu

8. Chameleon: A Data-Efficient Generalist for Dense Visual Prediction in the Wild; *Donggyun Kim; Seongwoong Cho; Semin Kim; Chong Luo; Seunghoon Hong**

9. Mamba-ND: Selective State Space Modeling for Multi-Dimensional Data; *Shufan Li*; Aditya Grover; Harkanwar Singh*

10. HiT-SR: Hierarchical Transformer for Efficient Image Super-Resolution; *Xiang Zhang*; Yulun Zhang; Fisher Yu*

08:30 – 10:30
Oral session 7B: Adversarial learning and privacy - Auditorium
Chair: Andrés Bruhn; Venkatesh Babu Radhakrishnan

1. Prompt-Driven Contrastive Learning for Transferable Adversarial Attacks; *Hunmin Yang; Jongoh Jeong; Kuk-Jin Yoon**

2. Adversarial Robustification via Text-to-Image Diffusion Models; *Daewon Choi; Jongheon Jeong; Huiwon Jang; Jinwoo Shin**

3. Flatness-aware Sequential Learning Generates Resilient Backdoors; *Hoang Pham*; The-Anh Ta; Anh T Tran; Khoa D Doan*

4. A Closer Look at GAN Priors: Exploiting Intermediate Features for Enhanced Model Inversion Attacks; *Yixiang Qiu*; Hao Fang; Hongyao Yu; Bin Chen*; Meikang Qiu; Shu-Tao Xia*

5. Learning a Dynamic Privacy-preserving Camera Robust to Inversion Attacks; *Jiacheng Cheng*; Xiang Dai; Jia Wan; Nick Antipa; Nuno Vasconcelos*

6. R.A.C.E.: Robust Adversarial Concept Erasure for Secure Text-to-Image Diffusion Model; *Changhoon Kim*; Kyle Min*; Yezhou Yang*

7. Privacy-Preserving Adaptive Re-Identification without Image Transfer; *Hamza Rami*; Jhony H. Giraldo; Nicolas Winckler; Stéphane Lathuilière*

8. Images are Achilles' Heel of Alignment: Exploiting Visual Vulnerabilities for Jailbreaking Multimodal Large Language Models; *Yifan Li*; Hangyu Guo; Kun Zhou; Wayne Xin Zhao; Ji-Rong Wen*

9. Concept Arithmetics for Circumventing Concept Inhibition in Diffusion Models; *Vitali Petsiuk*; Kate Saenko*
BEST PAPER CANDIDATE

08:30 – 10:30
Oral session 7C: Optimization and theory - Silver Room
Chair: Qixing Huang; Vladislav Golyanik
1. A Direct Approach to Viewing Graph Solvability; *Federica Arrigoni*; Andrea Fusiello; Tomas Pajdla*

2. Convex Relaxations for Manifold-Valued Markov Random Fields with Approximation Guarantees; *Robin Kenis*; Emanuel Laude; Panagiotis Patrinos*

3. Flash Cache: Reducing Bias in Radiance Cache Based Inverse Rendering; *Benjamin Attal*; Dor Verbin; Ben Mildenhall; Peter Hedman; Jonathan T Barron; Matthew O'Toole; Pratul Srinivasan*

4. A Riemannian Approach for Spatiotemporal Analysis and Generation of 4D Tree-shaped Structures; *Tahmina Khanam; Mohammed Bennamoun; Guan Wang; Guanjin Wang; Ferdous Sohel; Farid Boussaid; Anuj Srivastava; Hamid Laga**

5. Physics-Based Interaction with 3D Objects via Video Generation; *Tianyuan Zhang*; Hong-Xing Yu; Rundi Wu; Brandon Y Feng; Changxi Zheng; Noah Snavely; Jiajun Wu; William T. Freeman*

6. Shape from Heat Conduction; *Sriram Narayanan*; Mani Ramanagopal; Mark Sheinin; Aswin C. Sankaranarayanan; Srinivasa G. Narasimhan*

7. Rasterized Edge Gradients: Handling Discontinuities Differentially; *Stanislav Pidhorskyi*; Tomas Simon; Gabriel Schwartz; He Wen; Yaser Sheikh; Jason Saragih* BEST PAPER CANDIDATE

8. ControlNet-XS: Rethinking the Control of Text-to-Image Diffusion Models as Feedback-Control Systems; *Denis Zavadski*; Johann-Friedrich Feiden; Carsten Rother*

9. Parrot: Pareto-optimal Multi-Reward Reinforcement Learning Framework for Text-to-Image Generation; *Seung Hyun Lee*; Yinxiao Li; Junjie Ke; Innfarn Yoo; Han Zhang; Jiahui Yu; Qifei Wang; Fei Deng; Glenn Entis; Junfeng He; Gang Li; Sangpil Kim; Irfan Essa; Feng Yang**

10. Model Stock: All we need is just a few fine-tuned models; *Dong-Hwan Jang; Sangdoo Yun; Dongyoon Han**

10:30 – 11:00
Coffee Break - Exhibition Area (Level 0)

10:30 – 12:30
Poster session 7
1. LLaVA-Plus: Learning to Use Tools for Creating Multimodal Agents; *Shilong Liu*; Hao Cheng; Haotian Liu; Hao Zhang; Feng Li; Tianhe Ren; Xueyan Zou; Jianwei Yang; Hang Su; Jun Zhu; Lei Zhang; Jianfeng Gao; Chunyuan Li**

2. Adapt2Reward: Adapting Video-Language Models to Generalizable Robotic Rewards via Failure Prompts; *Yanting Yang; Minghao Chen*; Qibo Qiu; Jiahao WU; Wenxiao Wang; Binbin Lin; Ziyu Guan; Xiaofei He*

3. Pre-trained Visual Dynamics Representations for Efficient Policy Learning; *Hao Luo*; Bohan Zhou; Zongqing Lu**

4. R^2-Bench: Benchmarking the Robustness of Referring Perception Models under Perturbations; *Xiang Li*; Kai Qiu; Jinglu Wang; Xiaohao Xu; Kashu Yamazaki; Hao Chen; Rita Singh; Xiaonan Huang; Bhiksha Raj*

5. Paying More Attention to Images: A Training-Free Method for Alleviating Hallucination in LVLMs; *Shi Liu*; Kecheng Zheng*; Wei Chen**

6. An Explainable Vision Question Answer Model via Diffusion Chain-of-Thought; *Chunhao LU; Qiang Lu*; Jake Luo*

7. SQ-LLaVA: Self-Questioning for Large Vision-Language Assistant; *Guohao Sun*; Can Qin; JIAMINAN WANG; Zeyuan Chen; Ran Xu; Zhiqiang Tao*

8. Fully Authentic Visual Question Answering Dataset from Online Communities; *Chongyan Chen*; Mengchen Liu; Noel C Codella; Yunsheng Li; Lu Yuan; Danna Gurari*

9. TrojVLM: Backdoor Attack Against Vision Language Models; *Weimin Lyu*; Lu Pang; Tengfei Ma; Haibin Ling; Chao Chen*

10. BEAF: Observing BEfore-AFter Changes to Evaluate Hallucination in Vision-language Models; *Moon Ye-Bin; Nam Hyeon-Woo; Wonseok Choi; Tae-Hyun Oh**

11. Prompt-Driven Contrastive Learning for Transferable Adversarial Attacks; *Hunmin Yang; Jongoh Jeong; Kuk-Jin Yoon**

12. Attention Prompting on Image for Large Vision-Language Models; *Runpeng Yu*; Weihao Yu*; Xinchao Wang**

13. Images are Achilles' Heel of Alignment: Exploiting Visual Vulnerabilities for Jailbreaking Multimodal Large Language Models; *Yifan Li*; hangyu guo; Kun Zhou; Wayne Xin Zhao; Ji-Rong Wen*

14. Agent3D-Zero: An Agent for Zero-shot 3D Understanding; *sha zhang; Di Huang; Jiajun Deng*; Shixiang Tang; Wanli Ouyang; Tong He*; Yanyong Zhang**

15. Diffusion-Refined VQA Annotations for Semi-Supervised Gaze Following; *Qiaomu Miao*; Alexandros Graikos; Jingwei Zhang; Sounak Mondal; Minh Hoai; Dimitris Samaras*

16. Meerkat: Audio-Visual Large Language Model for Grounding in Space and Time; *Sanjoy Chowdhury*; Sayan Nag; Subhrajyoti Dasgupta; Jun Chen; Mohamed Elhoseiny; Ruohan Gao; Dinesh Manocha*

17. LHRS-Bot: Empowering Remote Sensing with VGI-Enhanced Large Multimodal Language Model; *Dilxat Muhtar; Zhenshi Li; Feng Gu; Xueliang Zhang*; Pengfeng Xiao*

18. VISA: Reasoning Video Object Segmentation via Large Language Model; *Cilin Yan; Haochen Wang; Shilin Yan; Xiaolong Jiang; Yao Hu; Guoliang Kang*; Weidi Xie; Efstratios Gavves*

19. PALM: Predicting Actions through Language Models; *Sanghwan Kim*; Daoji Huang; Yongqin Xian; Otmar Hilliges; Luc Van Gool; Xi Wang*

20. LLaMA-VID: An Image is Worth 2 Tokens in Large Language Models; *Yanwei Li*; Chengyao Wang; Jiaya Jia*

21. BAM-DETR: Boundary-Aligned Moment Detection Transformer for Temporal Sentence Grounding in Videos; *Pilhyeon Lee*; Hyeran Byun*

22. Learning Chain of Counterfactual Thought for Bias-Robust Vision-Language Reasoning; *Yifeng Zhang; Ming Jiang; Qi Zhao**

23. APL: Anchor-based Prompt Learning for One-stage Weakly Supervised Referring Expression Comprehension; *Yaxin Luo; Jiayi Ji; Xiaofu Chen; Yuxin Zhang; Tianhe Ren; Gen Luo**

24. Exploring Pre-trained Text-to-Video Diffusion Models for Referring Video Object Segmentation; *Xuelu Feng; Dongdong Chen; Junsong Yuan; Chunming Qiao; Gang Hua; Zixin Zhu**

25. Generalizing to Unseen Domains via Text-guided Augmentation; *Daiqing Qi*; Handong Zhao; Aidong Zhang; Sheng Li*

26. Mamba-ND: Selective State Space Modeling for Multi-Dimensional Data; *Shufan Li*; Aditya Grover; Harkanwar Singh*

27. Free-ATM: Harnessing Free Attention Masks for Representation Learning on Diffusion-Generated Images; *David Junhao Zhang*; Mutian Xu; Jay Zhangjie Wu; Chuhui Xue; Wenqing Zhang; Xiaoguang Han; Song Bai; Mike Zheng Shou**

28. TOD3Cap: Towards 3D Dense Captioning in Outdoor Scenes; *Bu Jin; Yupeng Zheng*; Pengfei Li; Weize Li; Yuhang Zheng; Sujie Hu; Xinyu Liu; Jinwei Zhu; Zhijie Yan; Haiyang Sun; Kun Zhan; Peng Jia; Xiaoxiao Long; Yilun Chen; Hao Zhao*

29. FALIP: Visual Prompt as Foveal Attention Boosts CLIP Zero-Shot Performance; *Jiedong Zhuang; Jiaqi Hu; Lianrui Mu; Rui Hu; Xiaoyu Liang; Jiangnan Ye; Haoji Hu**

30. Audio-visual Generalized Zero-shot Learning the Easy Way; *Shentong Mo*; Pedro Morgado*

31. COM Kitchens: An Unedited Overhead-view Procedural Videos Dataset a Vision-Language Benchmark; *Atsushi Hashimoto*; Koki Maeda; Tosho Hirasawa; Jun Harashima; Leszek Rybicki; Yusuke Fukasawa; Yoshitaka Ushiku*

32. TrajPrompt: Aligning Color Trajectory with Vision-Language Representations; *Li-Wu Tsao*; Hao-Tang Tsui; Yu-Rou Tuan; Pei-Chi Chen; Kuan-Lin Wang; Jhih-Ciang Wu; Hong-Han Shuai*; Wen-Huang Cheng*

33. Soft Prompt Generation for Domain Generalization; *Shuanghao Bai*; Yuedi Zhang; Wanqi Zhou; Zhirong Luan; Badong Chen**

34. GTMS: A Gradient-driven Tree-guided Mask-free Referring Image Segmentation Method; *Haoxin Lv; Tianxiong Zhong; Sanyuan Zhao**

35. Reason2Drive: Towards Interpretable and Chain-based Reasoning for Autonomous Driving; *Ming Nie; Renyuan Peng; Chunwei Wang; Xinyue Cai; Jianhua Han; Hang Xu*; Li Zhang**

36. Towards Multi-modal Transformers in Federated Learning; *Guangyu Sun*; Matias Mendieta; Aritra Dutta; Xin Li; Chen Chen*

37. Latent-INR: A Flexible Framework for Implicit Representations of Videos with Discriminative Semantics; *Shishira R Maiya*; Anubhav Gupta; Matthew A Gwilliam; Max Ehrlich; Abhinav Shrivastava*

38. OmniSat: Self-Supervised Modality Fusion for Earth Observation; *Guillaume Astruc*; Nicolas Gonthier; Clement Mallet; Loic Landrieu*

39. Prompting Language-Informed Distribution for Compositional Zero-Shot Learning; *Wentao Bao*; Lichang Chen; Heng Huang; Yu Kong*

40. Plug and Play: A Representation Enhanced Domain Adapter for Collaborative Perception; *Tianyou Luo*; Quan Yuan*; Yuchen Xia; Guiyang Luo; Yujia Yang; Jinglin Li*

41. Lost in Translation: Modern Neural Networks Still Struggle With Small Realistic Image Transformations; *Ofir Shifman*; Yair Weiss*

42. TF-FAS: Twofold-Element Fine-Grained Semantic Guidance for Generalizable Face Anti-Spoofing; *Xudong Wang; Ke-Yue Zhang; Taiping Yao*; Qianyu Zhou; Shouhong Ding; Pingyang Dai*; Rongrong Ji*

43. Think before Placement: Common Sense Enhanced Transformer for Object Placement; *Yaxuan Qin; Jiayu Xu; Ruiping Wang*; Xilin Chen*

44. SPARO: Selective Attention for Robust and Compositional Transformer Encodings for Vision; *Ankit Vani*; Bac Nguyen; Samuel Lavoie; Ranjay Krishna; Aaron Courville*

45. Noise-assisted Prompt Learning for Image Forgery Detection and Localization; *Dong Li; Jiaying Zhu; Xueyang Fu*; Xun Guo; Yidi Liu; Gang Yang; Jiawei Liu; Zheng-Jun Zha*

46. MTA-CLIP: Language-Guided Semantic Segmentation with Mask-Text Alignment; *Anurag Das*; Xinting Hu; Li Jiang; Bernt Schiele*

47. Vision-Language Dual-Pattern Matching for Out-of-Distribution Detection; *Zihan Zhang; Zhuo Xu; Xiang Xiang**

48. LetsMap: Unsupervised Representation Learning for Label-Efficient Semantic BEV Mapping; *Nikhil Gosala*; Kürsat Petek; B Ravi Kiran; Senthil Yogamani; Paulo L. J. Drews-Jr; Wolfram Burgard; Abhinav Valada*

49. PromptIQA: Boosting the Performance and Generalization for No-Reference Image Quality Assessment via Prompts; *Zewen Chen; Haina Qin; Juan Wang; Chunfeng Yuan; Bing Li*; Weiming Hu; Leon Wang*

50. PPAD: Iterative Interactions of Prediction and Planning for End-to-end Autonomous Driving; *Zhili Chen; Maosheng Ye; Shuangjie Xu; Tongyi Cao; Qifeng Chen**

51. Hetecooper: Feature Collaboration Graph for Heterogeneous Collaborative Perception; *Congzhang Shao; Guiyang Luo*; Quan Yuan*; Yifu Chen; Yilin Liu; Gong Kexin; Jinglin Li*

52. Learning to Drive via Asymmetric Self-Play; *Chris Zhang*; Sourav Biswas; Kelvin Wong; Kion Fallah; Lunjun Zhang; Dian Chen; Sergio Casas; Raquel Urtasun*

53. Optimizing Diffusion Models for Joint Trajectory Prediction and Controllable Generation; *Yixiao Wang*; Chen Tang; Lingfeng Sun; Simone Rossi; Yichen Xie; Chensheng Peng; Thomas Hannagan; Stefano Sabatini; Nicola Poerio; Masayoshi TOMIZUKA; Wei Zhan*

54. Online Vectorized HD Map Construction using Geometry; *Zhixin Zhang; Yiyuan Zhang; Xiaohan Ding; Fusheng Jin*; Xiangyu Yue*

55. ProxyCLIP: Proxy Attention Improves CLIP for Open-Vocabulary Segmentation; *Mengcheng Lan; Chaofeng Chen; Yiping Ke; Xinjiang Wang; Litong Feng; Wayne Zhang**

56. Leveraging Enhanced Queries of Point Sets for Vectorized Map Construction; *Zihao Liu; Xiaoyu Zhang; Guangwei Liu; Ji Zhao*; Ningyi Xu**

57. CC-SAM: Enhancing SAM with Cross-feature Attention and Context for Ultrasound Image Segmentation; *Shreyank N Gowda*; David A Clifton*

58. O2V-Mapping: Online Open-Vocabulary Mapping with Neural Implicit Representation; *Muer Tie; Julong Wei; Zhengjun Wang; Ke Wu; Shanshuai Yuan; Kaizhao Zhang; Jie Jia; Jieru Zhao; Zhongxue Gan*; Wenchao Ding**

59. OccWorld: Learning a 3D Occupancy World Model for Autonomous Driving; *Wenzhao Zheng; Weiliang Chen; Yuanhui Huang; Borui Zhang; Yueqi Duan; Jiwen Lu**

60. T-Rex2: Towards Generic Object Detection via Text-Visual Prompt Synergy; *Qing Jiang*; Feng Li; Zhaoyang Zeng; Shilong Liu; Tianhe Ren; Lei Zhang**

61. OpenSight: A Simple Open-Vocabulary Framework for LiDAR-Based Object Detection; *Hu Zhang; xu jianhua; Tao Tang; Haiyang Sun; Xin Yu*; Zi Helen Huang*; Kaicheng Yu*

62. Better Call SAL: Towards Learning to Segment Anything in Lidar; *Aljosa Osep*; Tim Meinhardt; Francesco Ferroni; Neehar Peri; Deva Ramanan; Laura Leal-Taixé*

63. Cross-Domain Semantic Segmentation on Inconsistent Taxonomy using VLMs; *Jeongkee Lim; Yusung Kim**

64. Cross-Domain Few-Shot Object Detection via Enhanced Open-Set Object Detector; *Yuqian Fu*; Yu Wang; Yixuan Pan; Xingyu Qiu; Lian Huai; Zeyu Shangguan; Tong Liu; Yanwei Fu; Luc Van Gool; Xingqun Jiang*

65. Class-Agnostic Object Counting with Text-to-Image Diffusion Model; *Xiaofei Hui; Qian Wu; Hossein Rahmani; Jun Liu**

66. PDT Uav Target Detection Dataset for Pests and Diseases Tree; *Mingle Zhou; Rui Xing; Delong Han; Zhiyong Qi; Gang Li**

67. You Only Learn One Query: Learning Unified Human Query for Single-Stage Multi-Person Multi-Task Human-Centric Perception; *Sheng Jin; Shuhuai Li; Tong Li; Wentao Liu*; Chen Qian; Ping Luo**

68. Approaching Outside: Scaling Unsupervised 3D Object Detection from 2D Scene; *Ruiyang Zhang*; Hu Zhang; Hang Yu; Zhedong Zheng**

69. MTMamba: Enhancing Multi-Task Dense Scene Understanding by Mamba-Based Decoders; *Baijiong Lin*; Weisen Jiang; Pengguang Chen; Yu Zhang; Shu Liu; Yingcong Chen*

70. Plain-Det: A Plain Multi-Dataset Object Detector; *Cheng Shi; Yuchen Zhu; Sibei Yang**

71. Background Adaptation with Residual Modeling for Exemplar-Free Class-Incremental Semantic Segmentation; *Anqi Zhang; Guangyu Gao**

72. Towards Reliable Evaluation and Fast Training of Robust Semantic Segmentation Models; *Francesco Croce*; Naman D. Singh; Matthias Hein**

73. MOD-UV: Learning Mobile Object Detectors from Unlabeled Videos; *Yihong Sun*; Bharath Hariharan*

74. DHR: Dual Features-Driven Hierarchical Rebalancing in Inter- and Intra-Class Regions for Weakly-Supervised Semantic Segmentation; *Sanghyun Jo; Fei Pan; In-Jae Yu; Kyungsu Kim**

75. Image-to-Lidar Relational Distillation for Autonomous Driving Data; *Anas Mahmoud*; Ali Harakeh; Steven Waslander*

76. Diffusion-Guided Weakly Supervised Semantic Segmentation; *Sung-Hoon Yoon; Hoyong Kwon; Jaeseok Jeong; Daehee Park; Kuk-Jin Yoon**

77. Eliminating Feature Ambiguity for Few-Shot Segmentation; *Qianxiong Xu*; Guosheng Lin; Chen Change Loy; Cheng Long; Ziyue Li; Rui Zhao*

78. Two-Stage Active Learning for Efficient Temporal Action Segmentation; *Yuhao Su; Ehsan Elhamifar**

79. Multi-scale Cross Distillation for Object Detection in Aerial Images; *Kun Wang; Zi Wang; Zhang Li*; Xichao Teng; Yang Li*

80. Semantic Diversity-aware Prototype-based Learning for Unbiased Scene Graph Generation; *Jaehyeong Jeon*; Kibum Kim; Kanghoon Yoon; Chanyoung Park*

81. Towards Adaptive Pseudo-label Learning for Semi-Supervised Temporal Action Localization; *Feixiang Zhou; Bryan Williams; Hossein Rahmani**

82. Co-Student: Collaborating Strong and Weak Students for Sparsely Annotated Object Detection; *Lianjun Wu; Jiangxiao Han; Zengqiang Zheng; Xinggang Wang**

83. GenView: Enhancing View Quality with Pretrained Generative Model for Self-Supervised Learning; *Xiaojie Li; Yibo Yang*; Xiangtai Li; Jianlong Wu*; Yue Yu; Bernard Ghanem; Min Zhang*

84. MICDrop: Masking Image and Depth Features via Complementary Dropout for Domain-Adaptive Semantic Segmentation; *Linyan Yang*; Lukas Hoyer*; Mark Weber; Tobias Fischer; Dengxin Dai; Laura Leal-Taixé; Daniel Cremers; Marc Pollefeys; Luc Van Gool*

85. On the Topology Awareness and Generalization Performance of Graph Neural Networks; *Junwei Su*; Chuan Wu*

BEST PAPER CANDIDATE

86. Learn from the Learnt: Source-Free Active Domain Adaptation via Contrastive Sampling and Visual Persistence; *Mengyao Lyu; Tianxiang Hao; Xinhao Xu; Hui Chen*; Zijia Lin; Jungong Han; Guiguang Ding**

87. ExMatch: Self-guided Exploitation for Semi-Supervised Learning with Scarce Labeled Samples; *Noo-ri Kim; Jin-Seop Lee; Jee-Hyong Lee**

88. Causal Subgraphs and Information Bottlenecks: Redefining OOD Robustness in Graph Neural Networks; *Weizhi An; Wenliang Zhong; Feng Jiang; Hehuan Ma; Junzhou Huang**

89. SelEx: Self-Expertise in Fine-Grained Generalized Category Discovery; *Sarah Rastegar*; Mohammadreza Salehi; Yuki M Asano; Hazel Doughty; Cees Snoek*

90. Integrating Markov Blanket Discovery into Causal Representation Learning for Domain Generalization; *Naiyu Yin*; Hanjing Wang; Yue Yu; Tian Gao; Amit Dhurandhar; Qiang Ji*

91. Dynamic Retraining-Updating Mean Teacher for Source-Free Object Detection; *Trinh Le Ba Khanh*; Huy-Hung Nguyen; Long Hoang Pham; Duong Nguyen-Ngoc Tran; Jae Wook Jeon**

92. Discover-then-Name: Task-Agnostic Concept Bottlenecks via Automated Concept Discovery; *Sukrut Rao*; Sweta Mahajan*; Moritz Böhle; Bernt Schiele*

93. Dynamic Data Selection for Efficient SSL via Coarse-to-Fine Refinement; *Aditay Tripathi*; Pradeep Shenoy; Anirban Chakraborty*

94. On the Approximation Risk of Few-Shot Class-Incremental Learning; *Xuan Wang; Zhong Ji*; Xiyao Liu; Yanwei Pang; Jungong Han*

95. Learning Representation for Multitask Learning through Self-Supervised Auxiliary Learning; *Seokwon Shin; Hyungrok Do; Youngdoo Son**

96. COD: Learning Conditional Invariant Representation for Domain Adaptation Regression; *Hao-Ran Yang; Chuan-Xian Ren*; You-Wei Luo*

97. Federated Learning with Local Openset Noisy Labels; *Zonglin Di*; Zhaowei Zhu; Xiaoxiao Li; Yang Liu**

98. DGR-MIL: Exploring Diverse Global Representation in Multiple Instance Learning for Whole Slide Image Classification; *Wenhui Zhu*; Xiwen Chen; Peijie Qiu; Aristeidis Sotiras; Abolfazl Razi; Yalin Wang*

99. Flexible Distribution Alignment: Towards Long-tailed Semi-supervised Learning with Proper Calibration; *Emanuel Sanchez Aimar*; Nathaniel D Helgesen; Yonghao Xu; Marco Kuhlmann; Michael Felsberg*

100. CLEO: Continual Learning of Evolving Ontologies; *Shishir Muralidhara*; Saqib Bukhari; Georg Dr. Schneider; Didier Stricker; René Schuster*

101. COIN-Matting: Confounder Intervention for Image Matting; *Zhaohe Liao; Jiangtong Li; Jun Lan; Huijia Zhu; Weiqiang Wang; Li Niu*; Liqing Zhang**

102. STAMP: Outlier-Aware Test-Time Adaptation with Stable Memory Replay; *Yu Yongcan; Lijun Sheng; Ran He; Jian Liang**

103. FairDomain: Achieving Fairness in Cross-Domain Medical Image Segmentation and Classification; *Yu Tian*; Congcong Wen; Min Shi; Muhammad Muneeb Afzal; Hao Huang; Muhammad Osama Khan; Yan Luo; Yi Fang; Mengyu Wang*

104. Echoes of the Past: Boosting Long-tail Recognition via Reflective Learning; *Qihao Zhao; Yalun Dai; Shen Lin; Wei Hu; Fan Zhang*; Jun Liu*

105. Weighted Ensemble Models Are Strong Continual Learners; *Imad Eddine MAROUF*; Subhankar Roy; Enzo Tartaglione; Stéphane Lathuilière*

106. FedHARM: Harmonizing Model Architectural Diversity in Federated Learning; *Anestis Kastellos*; Athanasios Psaltis; Charalampos Z Patrikakis; Petros Daras*

107. An accurate detection is not all you need to combat label noise in web-noisy datasets; *Paul Albert*; Kevin McGuinness; Eric Arazo; Tarun Krishna; Noel O Connor; Jack Valmadre*

108. Improving Knowledge Distillation via Regularizing Feature Direction and Norm; *Yuzhu Wang; Lechao Cheng*; Manni Duan; Yongheng Wang; Zunlei Feng; Shu Kong*

109. Unlocking the Potential of Federated Learning: The Symphony of Dataset Distillation via Deep Generative Latents; *Yuqi Jia; Saeed Vahidian*; Jingwei Sun; Jianyi Zhang; Vyacheslav Kungurtsev; Neil Zhenqiang Gong; Yiran Chen*

110. Model Stock: All we need is just a few fine-tuned models; *Dong-Hwan Jang; Sangdoo Yun; Dongyoon Han**

111. Anytime Continual Learning for Open Vocabulary Classification; *Zhen Zhu**; *Yiming Gong; Derek Hoiem**

112. Shedding More Light on Robust Classifiers under the lens of Energy-based Models; *Mujtaba Hussain Mirza**; *Maria Rosaria Briglia**; *Senad Beadini**; *Iacopo Masi**

113. Mew: Multiplexed Immunofluorescence Image Analysis through an Efficient Multiplex Network; *Sukwon Yun; Jie Peng; Alexandro E Trevino; Chanyoung Park; Tianlong Chen**

114. Deep Online Probability Aggregation Clustering; *Yuxuan Yan; Na Lu**; *Ruofan Yan*

115. Scissorhands: Scrub Data Influence via Connection Sensitivity in Networks; *Jing Wu**; *Mehrtash Harandi*

116. Dissolving Is Amplifying: Towards Fine-Grained Anomaly Detection; *Jian Shi**; *Pengyi Zhang; Ni Zhang; Hakim Ghazzai; Peter Wonka*

117. Group Testing for Accurate and Efficient Range-Based Near Neighbor Search for Plagiarism Detection; *Harsh Shah**; *Kashish Mittal; Ajit Rajwade**

118. MoEAD: A Parameter-efficient Model for Multi-class Anomaly Detection; *Shiyuan Meng; Wenchao Meng**; *Qihang Zhou; Shizhong Li; Weiye Hou; Shibo He*

119. FedHide: Federated Learning by Hiding in the Neighbors; *Hyunsin Park**; *Sungrack Yun*

120. SIMBA: Split Inference - Mechanisms, Benchmarks and Attacks; *Abhishek Singh**; *Vivek Sharma; Rohan Sukumaran; John J Mose; Jeffrey K Chiu; Justin Yu; Ramesh Raskar*

121. AdvDiff: Generating Unrestricted Adversarial Examples using Diffusion Models; *Xuelong Dai**; *Kaisheng Liang; Bin Xiao*

122. I Can't Believe It's Not Scene Flow!; *Ishan Khatri**; *Kyle Vedder**; *Neehar Peri; Deva Ramanan; James Hays*

123. GeneralAD: Anomaly Detection Across Domains by Attending to Distorted Features; *Luc P.J. Sträter**; *Mohammadreza Salehi; Efstratios Gavves; Cees G.M. Snoek; Yuki M. Asano*

124. SPAMming Labels: Efficient Annotations for the Trackers of Tomorrow; *Orcun Cetintas**; *Tim Meinhardt; Guillem Brasó; Laura Leal-Taixé*

125. Adaptive High-Frequency Transformer for Diverse Wildlife Re-Identification; *Chenyue Li; Shuoyi Chen; Mang Ye**

126. PaPr: Training-Free One-Step Patch Pruning with Lightweight ConvNets for Faster Inference; *Tanvir Mahmud**; *Burhaneddin Yaman; Chun-Hao Liu; Diana Marculescu*

127. Inter-Class Topology Alignment for Efficient Black-Box Substitute Attacks; *Lingzhuang Meng; Mingwen Shao**; *Yuanjian Qiao; Wenjie Liu*

128. Data Poisoning Quantization Backdoor Attack; *Tran Huynh**; *Anh Tran; Khoa Doan; Tung Pham*

130. Event Trojan: Asynchronous Event-based Backdoor Attacks; *Ruofei Wang**; *Qing Guo; Haoliang Li; Renjie Wan**

132. BKDSNN: Enhancing the Performance of Learning-based Spiking Neural Networks Training with Blurred Knowledge Distillation; *Zekai Xu; Kang You; Qinghai Guo; Xiang Wang; Zhezhi He**

133. An Incremental Unified Framework for Small Defect Inspection; *Jiaqi Tang; Hao Lu; Xiaogang Xu; Ruizheng Wu; Sixing Hu; Tong Zhang; Tsz Wa Cheng; Ming Ge; Ying-Cong Chen**; *Fugee Tsung*

134. CLAMP-ViT: Contrastive Data-Free Learning for Adaptive Post-Training Quantization of ViTs; *Akshat Ramachandran**; *Souvik Kundu**; *Tushar Krishna**

135. PQ-SAM: Post-training Quantization for Segment Anything Model; *Xiaoyu Liu**; *Xin Ding; Lei Yu; Yuanyuan Xi; Wei Li; Zhijun Tu; jie hu; Hanting Chen; Baoqun YIN; Zhiwei Xiong**

136. ELSE: Efficient Deep Neural Network Inference through Line-based Sparsity Exploration; *Zeqi Zhu**; *Alberto Garcia-Ortiz; Luc Waeijen; Egor Bondarev; Arash Pourtaherian; Orlando Moreira*

137. A Framework for Efficient Model Evaluation through Stratification, Sampling, and Estimation; *Riccardo Fogliato**; *Pratik Patil; Mathew Monfort; Pietro Perona*

138. LPViT: Low-Power Semi-structured Pruning for Vision Transformers; *Kaixin Xu**; *Zhe Wang**; *Chunyun Chen; Xue Geng; Jie Lin; Xulei Yang; Min Wu**; *Xiaoli Li; Weisi Lin**

139. Statewide Visual Geolocalization in the Wild; *Florian Fervers*; Sebastian Bullinger; Christoph Bodensteiner; Michael Arens; Rainer Stiefelhagen*

140. iNeMo: Incremental Neural Mesh Models for Robust Class-Incremental Learning; *Tom Fischer*; Yaoyao Liu; Artur Jesslen; Noor Ahmed; Prakhar Kaushik; Angtian Wang; Alan Yuille; Adam Kortylewski; Eddy Ilg*

141. Parameter-Efficient and Memory-Efficient Tuning for Vision Transformer: A Disentangled Approach; *Taolin Zhang; Jiawang Bai; Zhihe Lu; Dongze Lian; genping wang*; Xinchao Wang*; Shu-Tao Xia*

142. Dropout Mixture Low-Rank Adaptation for Visual Parameters-Efficient Fine-Tuning; *Zhengyi Fang; Yue Wang; Ran Yi*; Lizhuang Ma*

143. FreeAugment: Data Augmentation Search Across All Degrees of Freedom; *Tom Bekor*; Niv Nayman; Lihi Zelnik-Manor*

144. Characterizing Model Robustness via Natural Input Gradients; *Adrian Rodriguez-Munoz*; Tongzhou Wang; Antonio Torralba*

145. Chameleon: A Data-Efficient Generalist for Dense Visual Prediction in the Wild; *Donggyun Kim; Seongwoong Cho; Semin Kim; Chong Luo; Seunghoon Hong**

146. A Closer Look at GAN Priors: Exploiting Intermediate Features for Enhanced Model Inversion Attacks; *Yixiang Qiu*; Hao Fang; Hongyao Yu; Bin Chen*; Meikang Qiu; Shu-Tao Xia*

147. Flatness-aware Sequential Learning Generates Resilient Backdoors; *Hoang Pham*; The-Anh Ta; Anh T Tran; Khoa D Doan*

148. Convex Relaxations for Manifold-Valued Markov Random Fields with Approximation Guarantees; *Robin Kenis*; Emanuel Laude; Panagiotis Patrinos*

149. A Riemannian Approach for Spatiotemporal Analysis and Generation of 4D Tree-shaped Structures; *Tahmina Khanam; Mohammed Bennamoun; Guan Wang; Guanjin Wang; Ferdous Sohel; Farid Boussaid; Anuj Srivastava; Hamid Laga**

150. Physics-Based Interaction with 3D Objects via Video Generation; *Tianyuan Zhang*; Hong-Xing Yu; Rundi Wu; Brandon Y Feng; Changxi Zheng; Noah Snavely; Jiajun Wu; William T. Freeman*

151. Spline-based Transformers; *Prashanth Chandran*; Agon Serifi*; Markus Gross; Moritz Bächer*

152. ControlNet-XS: Rethinking the Control of Text-to-Image Diffusion Models as Feedback-Control Systems; *Denis Zavadski*; Johann-Friedrich Feiden; Carsten Rother*

153. Parrot: Pareto-optimal Multi-Reward Reinforcement Learning Framework for Text-to-Image Generation; *Seung Hyun Lee*; Yinxiao Li; Junjie Ke; Innfarn Yoo; Han Zhang; Jiahui Yu; Qifei Wang; Fei Deng; Glenn Entis; Junfeng He; Gang Li; Sangpil Kim; Irfan Essa; Feng Yang**

154. Concept Arithmetics for Circumventing Concept Inhibition in Diffusion Models; *Vitali Petsiuk*; Kate Saenko*
BEST PAPER CANDIDATE

155. R.A.C.E.: Robust Adversarial Concept Erasure for Secure Text-to-Image Diffusion Model; *Changhoon Kim*; Kyle Min*; Yezhou Yang*

156. Adversarial Robustification via Text-to-Image Diffusion Models; *Daewon Choi; Jongheon Jeong; Huiwon Jang; Jinwoo Shin**

157. HiT-SR: Hierarchical Transformer for Efficient Image Super-Resolution; *XIANG ZHANG*; Yulun Zhang; Fisher Yu*

158. Improving Feature Stability during Upsampling -- Spectral Artifacts and the Importance of Spatial Context; *Shashank Agnihotri*; Julia Grabinski; Margret Keuper*

159. Region-Aware Sequence-to-Sequence Learning for Hyperspectral Denoising; *JiaHua Xiao; Yang Liu; Xing Wei**

160. Towards Certifiably Robust Face Recognition; *Seunghun Paik; Dongsoo Kim; Chanwoo Hwang; Sunpill Kim; Jae Hong Seo**

161. denoiSplit: a method for joint microscopy image splitting and unsupervised denoising; *Ashesh Ashesh*; Florian Jug**

162. Probabilistic Image-Driven Traffic Modeling via Remote Sensing; *Scott Workman*; Armin Hadzic*

163. VideoMamba: Spatio-Temporal Selective State Space Model; *Jinyoung Park*; Hee-Seon Kim; Kangwook Ko; Minbeom Kim; Changick Kim*

164. Bi-TTA: Bidirectional Test-Time Adapter for Remote Physiological Measurement; *Haodong LI*; Hao LU; Yingcong Chen**

165. Oulu Remote-photoplethysmography Physical Domain Attacks Database (ORPDAD); *Marko Savic; Guoying Zhao**

166. TAPTR: Tracking Any Point with Transformers as Detection; *Hongyang Li*; Hao Zhang; Shilong Liu; Zhaoyang Zeng; Tianhe Ren; Feng Li; Lei Zhang**

167. EcoMatcher: Efficient Clustering Oriented Matcher for Detector-free Image Matching; *Peiqi Chen*; Lei Yu; Yi Wan*; Yongjun Zhang*; Jian Wang; Liheng Zhong; Jingdong Chen; Ming Yang*

168. VP-SAM: Taming Segment Anything Model for Video Polyp Segmentation via Disentanglement and Spatio-temporal Side Network; *Zhixue Fang; Yuzhi Liu; Huisi Wu*; Jing Qin*

169. Privacy-Preserving Adaptive Re-Identification without Image Transfer; *Hamza Rami*; Jhony H. Giraldo; Nicolas Winckler; Stéphane Lathuilière*

170. PanoVOS: Bridging Non-panoramic and Panoramic Views with Transformer for Video Segmentation; *Shilin Yan*; Xiaohao Xu; Renrui Zhang; Lingyi Hong; wenchao chen; Wenqiang Zhang; Wei Zhang**

171. Exploring Reliable Matching with Phase Enhancement for Night-time Semantic Segmentation; *Yuwen Pan*; Rui Sun; Naisong Luo; Tianzhu Zhang; Yongdong Zhang*

172. GAReT: Cross-view Video Geolocalization with Adapters and Auto-Regressive Transformers; *Manu S Pillai*; Mamshad Nayeem Rizve; Mubarak Shah*

173. Semi-Supervised Video Desnowing Network via Temporal Decoupling Experts and Distribution-Driven Contrastive Regularization; *Hongtao Wu; Yijun Yang; Angelica I Aviles-Rivero; Jingjing Ren; Sixiang Chen; Haoyu Chen; Lei Zhu**

174. Make Your ViT-based Multi-view 3D Detectors Faster via Token Compression; *Dingyuan Zhang; Dingkang Liang*; Zichang Tan; Xiaoqing Ye; Cheng Zhang; Jingdong Wang; Xiang Bai**

175. Embracing Events and Frames with Hierarchical Feature Refinement Network for Object Detection; *Hu Cao; Zehua Zhang; Yan Xia; Xinyi Li; Jiahao Xia; Guang Chen*; Alois C. Knoll*

176. MAD-DR: Map Compression for Visual Localization with Matchness Aware Descriptor Dimension Reduction; *Qiang Wang**

177. ConGeo: Robust Cross-view Geo-localization across Ground View Variations; *Li Mi; Chang Xu*; Javiera Castillo Navarro; SYRIELLE MONTARIOL; Wen Yang; Antoine Bosselut; Devis Tuia*

178. Human Motion Forecasting in Dynamic Domain Shifts: A Homeostatic Continual Test-time Adaptation Framework; *Qiongjie Cui*; Huaijiang Sun; Bin Li; Jianfeng Lu; Weiqing Li*

179. Tensorial template matching for fast cross-correlation with rotations and its application for tomography; *Antonio Martinez-Sanchez*; Ulrike Homberg; J. M. Almira; Harold Phelippeau*

180. Motion and Structure from Event-based Normal Flow; *Zhongyang Ren; Bangyan Liao; Delei Kong; Jinghang Li; Peidong Liu; Laurent Kneip; Guillermo Gallego; Yi Zhou**

181. Towards Robust Event-based Networks for Nighttime via Unpaired Day-to-Night Event Translation; *Yuhwan Jeong; Hoonhee Cho; Kuk-Jin Yoon**

182. MetaWeather: Few-Shot Weather-Degraded Image Restoration; *Youngrae Kim*; Younggeol Cho; Thanh-Tung Nguyen; Seunghoon Hong; Dongman Lee**

183. Learning a Dynamic Privacy-preserving Camera Robust to Inversion Attacks; *Jiacheng Cheng*; Xiang Dai; Jia Wan; Nick Antipa; Nuno Vasconcelos*

184. Deep Patch Visual SLAM; *Lahav Lipson*; Zachary Teed; Jia Deng*

185. How Far Can a 1-Pixel Camera Go? Solving Vision Tasks using Photoreceptors and Computationally Designed Visual Morphology; *Andrei Atanov*; Rishubh Singh; Jiawei Fu; Isabella Yu; Andrew Spielberg; Amir Zamir*

186. SparseLIF: High-Performance Sparse LiDAR-Camera Fusion for 3D Object Detection; *Hongcheng Zhang; Liu Liang; Pengxin Zeng*; Xiao Song; Zhe Wang*

187. A Direct Approach to Viewing Graph Solvability; *Federica Arrigoni*; Andrea Fusiello; Tomas Pajdla*

188. Learning Where to Look: Self-supervised Viewpoint Selection for Active Localization using Geometrical Information; *Luca Di Giammarino*; Boyang Sun; Giorgio Grisetti; Marc Pollefeys; Hermann Blum; Daniel Barath*

189. milliFlow: Scene Flow Estimation on mmWave Radar Point Cloud for Human Motion Sensing; *Fangqiang Ding*; Zhen Luo; Peijun Zhao; Chris Xiaoxuan Lu*

190. CliffPhys: Camera-based Respiratory Measurement using Clifford Neural Networks; *Omar Ghezzi*; Giuseppe Boccignone; Giuliano Grossi; Raffaella Lanzarotti; Alessandro D'Amelio*

191. PACE: Pose Annotations in Cluttered Environments; *Yang You*; kai xiong; Zhening Yang; Zhengxiang Huang; Junwei Zhou; Ruoxi Shi; Zhou FANG; Adam Harley; Leonidas Guibas; Cewu Lu**

192. ScatterFormer: Efficient Voxel Transformer with Scattered Linear Attention; *Chenhang He*; Ruihuang Li; Guowen Zhang; Lei Zhang*

193. Category-level Object Detection, Pose Estimation and Reconstruction from Stereo Images; *Chuanrui Zhang*; Yonggen Ling*; Minglei Lu; Minghan Qin; Haoqian Wang**

194. Zero-Shot Image Feature Consensus with Deep Functional Maps; *Xinle Cheng; Congyue Deng*; Adam Harley; Yixin Zhu*; Leonidas Guibas**

195. Occupancy as Set of Points; *Yiang Shi; Tianheng Cheng; Qian Zhang; Wenyu Liu; Xinggang Wang**

196. Domain-Adaptive 2D Human Pose Estimation via Dual Teachers in Extremely Low-Light Conditions; *Yihao Ai*; Yifei Qi; Bo Wang; Yu Cheng; Xinchao Wang; Robby T. Tan*

197. RSL-BA: Rolling Shutter Line Bundle Adjustment; *Yongcong Zhang; Bangyan Liao; Yifei Xue; Lu Chen; Peidong Liu; Yizhen Lao**

198. 3D Hand Pose Estimation in Everyday Egocentric Images; *Aditya Prakash*; Ruisen Tu; Matthew Chang; Saurabh Gupta*

199. Hyperion – A fast, versatile symbolic Gaussian Belief Propagation framework for Continuous-Time SLAM; *David Hug*; Ignacio Alzugaray; Margarita Chli*

200. Benchmarks and Challenges in Pose Estimation for Egocentric Hand Interactions with Objects; *Zicong Fan; Takehiko Ohkawa*; Linlin Yang; Nie Lin; Zhishan Zhou; Shihao Zhou; Jiajun Liang; Zhong Gao; Xuanyang Zhang; Xue Zhang; Fei Li; Liu Zheng; Feng Lu; Karim Abou Zeid; Bastian Leibe; Jeongwan On; Seungryul Baek; Aditya Prakash; Saurabh Gupta; Kun He; Yoichi Sato; Otmar Hilliges; Hyung Jin Chang; Angela Yao*

201. AddBiomechanics Dataset: Capturing the Physics of Human Motion at Scale; *Keenon Werling*; Janelle M Kaneda; Tian Tan; Rishi Agarwal; Six Skov; Tom Van Wouwe; Scott Uhlrich; Scott Delp; Karen Liu; Nicholas A Bianco; Carmichael Ong; Antoine Falisse; Shardul Sapkota; Aidan Jai Chandra; Joshua A Carter; Ezio Preatoni; Benjamin J Fregly; Jennifer Hicks*

202. SFPNet: Sparse Focal Point Network for Semantic Segmentation on General LiDAR Point Clouds; *Yanbo Wang*; Wentao Zhao; Cao Chuan; Tianchen Deng; Jingchuan Wang; Weidong Chen**

203. MAP-ADAPT: Real-Time Quality-Adaptive Semantic 3D Maps; *Jianhao Zheng*; Daniel Barath; Marc Pollefeys; Iro Armeni**

204. Segment, Lift and Fit: Automatic 3D Shape Labeling from 2D Prompts; *Jianhao Li; Tianyu Sun; Zhongdao Wang*; Enze Xie; Bailan Feng; Hongbo Zhang; Ze Yuan; Ke Xu; Jiaheng Liu*; Ping Luo*

205. DG-PIC: Domain Generalized Point-In-Context Learning for Point Cloud Understanding; *Jincen Jiang; Qianyu Zhou; Yuhang Li; Xuequan Lu*; Meili Wang*; Lizhuang Ma; Jian Chang; Jian Jun Zhang*

206. Self-supervised Shape Completion via Involution and Implicit Correspondences; *Mengya Liu*; Ajad Chhatkuli; Janis Postels; Luc Van Gool; Federico Tombari*

207. PARE-Net: Position-Aware Rotation-Equivariant Networks for Robust Point Cloud Registration; *Runzhao Yao; Shaoyi Du*; Wenting Cui; Canhui Tang; Chengwu Yang*

208. AEDNet: Adaptive Embedding and Multiview-Aware Disentanglement for Point Cloud Completion; *Zhiheng Fu; Longguang Wang; Lian Xu; Zhiyong Wang; Hamid Laga; Yulan Guo*; Farid Boussaid; Mohammed Bennamoun*

209. Flowed Time of Flight Radiance Fields; *Mikhail Okunev*; Marc Mapeke; Benjamin Attal; Christian Richardt; Matthew O'Toole; James Tompkin*

210. SGS-SLAM: Semantic Gaussian Splatting For Neural Dense SLAM; *Mingrui Li; Shuhong Liu; Heng Zhou; Guohao Zhu; Na Cheng; Tianchen Deng; Hongyu Wang**

211. GaussReg: Fast 3D Registration with Gaussian Splatting; *Jiahao Chang*; Yinglin Xu; Yihao Li; Yuantao Chen; Wensen Feng; Xiaoguang Han*

212. DiffusionDepth: Diffusion Denoising Approach for Monocular Depth Estimation; *Yiqun Duan*; Xianda Guo*; Zheng Zhu*

213. Spatially-Variant Degradation Model for Dataset-free Super-resolution; *SHAOJIE GUO; Haofei Song; Qingli Li; Yan Wang**

214. DiffCD: A Symmetric Differentiable Chamfer Distance for Neural Implicit Surface Fitting; *Linus Härenstam-Nielsen*; Lu Sang; Abhishek Saroha; Nikita Araslanov*; Daniel Cremers**

215. UniINR: Event-guided Unified Rolling Shutter Correction, Deblurring, and Interpolation; *Yunfan Lu*; Guoqiang Liang; Yusheng Wang; Lin Wang; Hui Xiong**

216. Global-to-Pixel Regression for Human Mesh Recovery; *Yabo Xiao; Mingshu HE*; Dongdong Yu*

217. Unrolled Decomposed Unpaired Learning for Controllable Low-Light Video Enhancement; *Lingyu Zhu; Wenhan Yang; Baoliang Chen; Hanwei Zhu; Zhangkai Ni; Qi Mao; Shiqi Wang**

218. SpaRP: Fast 3D Object Reconstruction and Pose Estimation from Sparse Views; *Chao Xu; Ang Li; Linghao Chen; Yulin Liu; Ruoxi Shi; Hao Su*; Minghua Liu**

219. Dual-Camera Smooth Zoom on Mobile Phones; *Renlong Wu; Zhilu Zhang*; Yu Yang; Wangmeng Zuo*

220. Image Demoireing in RAW and sRGB Domains; *Shuning Xu; Binbin Song; Xiangyu Chen; Xina Liu; Jiantao Zhou**

221. Leveraging Near-Field Lighting for Monocular Depth Estimation from Endoscopy Videos; *Akshay Paruchuri*; Samuel Ehrenstein; Shuxian Wang; Inbar Fried; Stephen Pizer; Marc Niethammer; Roni Sengupta*

222. CriSp: Leveraging Tread Depth Maps for Enhanced Crime-Scene Shoeprint Matching; *Samia Shafique*; Shu Kong; Charless Fowlkes*

223. Self-Training Room Layout via Geometry-aware Ray-casting; *Bolivar Solarte*; Chin-Hsuan Wu*; Jin-Cheng Jhang*; Jonathan Lee*; Yi-Hsuan Tsai*; Min Sun**

224. GenRC: Generative 3D Room Completion from Sparse Image Collections; *Ming-Feng Li*; Yueh-Feng Ku; Hong-Xuan Yen; Chi Liu; Yu-Lun Liu; Albert Y Chen; Cheng-Hao Kuo; Min Sun*

225. 6DGS: 6D Pose Estimation from a Single Image and a 3D Gaussian Splatting Model; *Matteo Bortolon*; Theodore Tsesmelis; Stuart James; Fabio Poiesi; Alessio Del Bue*

226. ClusteringSDF: Self-Organized Neural Implicit Surfaces for 3D Decomposition; *Tianhao Wu*; Chuanxia Zheng; Qianyi Wu; Tat-Jen Cham*

227. Forecasting Future Videos from Novel Views via Disentangled 3D Scene Representation; *Sudhir Yarram*; Junsong Yuan*

228. CoSIGN: Few-Step Guidance of ConSIstency Model to Solve General INverse Problems; *Jiankun Zhao; Bowen Song; Liyue Shen**

229. Image-adaptive 3D Lookup Tables for Real-time Image Enhancement with Bilateral Grids; *Wontae Kim*; Nam Ik Cho**

230. Surface-Centric Modeling for High-Fidelity Generalizable Neural Surface Reconstruction; *Rui Peng; Shihe Shen; Kaiqiang Xiong; Huachen Gao; Jianbo Jiao; Xiaodong Gu; Ronggang Wang**

231. Plug-and-Play Learned Proximal Trajectory for 3D Sparse-View X-Ray Computed Tomography; *Romain Vo*; Julie Escoda; Caroline Vienne; Etienne Decenciere*

232. Soft Shadow Diffusion (SSD): Physics-inspired Learning for 3D Computational Periscopy; *Fadlullah A Raji*; John Murray-Bruce**

233. Towards Architecture-Agnostic Untrained Networks Priors for Image Reconstruction with Frequency Regularization; *Yilin Liu; Yunkui Pang; Jiang Li; Yong Chen; Pew-Thian Yap**

234. Single-Mask Inpainting for Voxel-based Neural Radiance Fields; *Jiafu Chen*; Tianyi Chu; Jiakai Sun; Wei Xing; Lei Zhao*

235. MVPGS: Excavating Multi-view Priors for Gaussian Splatting from Sparse Input Views; *Wangze Xu; Huachen Gao; Shihe Shen; Rui Peng; Jianbo Jiao; Ronggang Wang**

236. AutoDIR: Automatic All-in-One Image Restoration with Latent Diffusion; *Yitong Jiang*; Zhaoyang Zhang; Tianfan Xue; Jinwei Gu**

237. Co-synthesis of Histopathology Nuclei Image-Label Pairs using a Context-Conditioned Joint Diffusion Model; *Seonghui Min; Hyun-Jic Oh; Won-Ki Jeong**

238. CaesarNeRF: Calibrated Semantic Representation for Few-Shot Generalizable Neural Rendering; *Haidong Zhu; Tianyu Ding*; Tianyi Chen; Ilya Zharkov; Ram Nevatia; Luming Liang*

239. High-Resolution and Few-shot View Synthesis from Asymmetric Dual-lens Inputs; *Ruikang Xu; Mingde Yao; Yue Li; Yueyi Zhang; Zhiwei Xiong**

240. IntrinsicAnything: Learning Diffusion Priors for Inverse Rendering Under Unknown Illumination; *Xi Chen*; Sida Peng; Dongchen Yang; Yuan Liu; Bowen Pan; Chengfei Lyu; Xiaowei Zhou**

241. Deep Polarization Cues for Single-shot Shape and Subsurface Scattering Estimation; *Chenhao Li*; Trung Thanh Ngo; Hajime Nagahara*

242. Diffusion-Generated Pseudo-Observations for High-Quality Sparse-View Reconstruction; *Xinhang Liu*; Jiaben Chen; Shiu-Hong Kao; Yu-Wing Tai; Chi-Keung Tang*

243. QueryCDR: Query-based Controllable Distortion Rectification Network for Fisheye Images; *Pengbo Guo; Chengxu Liu; Xingsong Hou*; Xueming Qian*

244. Mini-Splatting: Representing Scenes with a Constrained Number of Gaussians; *Guangchi Fang; Bing Wang**

245. 2S-ODIS: Two-Stage Omni-Directional Image Synthesis by Geometric Distortion Correction; *Atsuya Nakata*; Takao Yamanaka**

246. DPA-Net: Structured 3D Abstraction from Sparse Views via Differentiable Primitive Assembly; *Fenggen Yu*; Yiming Qian; Xu Zhang; Francisca Gil-Ureta; Brian Jackson; Eric Bennett; Hao Zhang*

247. Shape from Heat Conduction; *Sriram Narayanan*; Mani Ramanagopal; Mark Sheinin; Aswin C. Sankaranarayanan; Srinivasa G. Narasimhan*

248. Relightable 3D Gaussians: Realistic Point Cloud Relighting with BRDF Decomposition and Ray Tracing; *Jian Gao; chun gu; Youtian Lin; Zhihao Li; Hao Zhu; Xun Cao; Li Zhang*; Yao Yao**

249. Flash Cache: Reducing Bias in Radiance Cache Based Inverse Rendering; *Benjamin Attal*; Dor Verbin; Ben Mildenhall; Peter Hedman; Jonathan T Barron; Matthew O'Toole; Pratul Srinivasan*

250. NGP-RT: Fusing Multi-Level Hash Features with Lightweight Attention for Real-Time Novel View Synthesis; *Yubin Hu; Xiaoyang Guo; Yang Xiao; Jingwei Huang; Yong-Jin Liu**

251. Multiscale Sliced Wasserstein Distances as Perceptual Color Difference Measures; *Jiaqi He; Zhihua Wang; Leon Wang; Tsein-I Liu; Yuming Fang; Qilin Sun*; Kede Ma*

252. Rasterized Edge Gradients: Handling Discontinuities Differentially; *Stanislav Pidhorskyi*; Tomas Simon; Gabriel Schwartz; He Wen; Yaser Sheikh; Jason Saragih* BEST PAPER CANDIDATE

253. CompGS: Smaller and Faster Gaussian Splatting with Vector Quantization; *K L Navaneet*; Kossar Pourahmadi Meibodi; Soroush Abbasi Koohpayegani; Hamed Pirsiavash*

254. Freeview Sketching: View-Aware Fine-Grained Sketch-Based Image Retrieval; *Aneeshan Sain*; Pinaki Nath Chowdhury; Subhadeep Koley; Ayan Kumar Bhunia; Yi-Zhe Song*

255. Data Augmentation via Latent Diffusion for Saliency Prediction; *Bahar Aydemir*; Deblina Bhattacharjee; Tong Zhang; Mathieu Salzmann; Sabine Süsstrunk*

256. Segmentation-guided Layer-wise Image Vectorization with Gradient Fills; *Hengyu Zhou; Hui Zhang*; Bin Wang**

257. Taming CLIP for Fine-grained and Structured Visual Understanding of Museum Exhibits; *Ada-Astrid Balauca*; Danda Pani Paudel; Kristina Toutanova; Luc Van Gool*

258. EpipolarGAN: Omnidirectional Image Synthesis with Explicit Camera Control; *Christopher May*; Daniel Aliaga*

259. GVGEN: Text-to-3D Generation with Volumetric Representation; *Xianglong He; Junyi Chen; Sida Peng; Di Huang; Yangguang Li; Xiaoshui Huang; Chun Yuan*; Wanli Ouyang; Tong He**

260. GRM: Large Gaussian Reconstruction Model for Efficient 3D Reconstruction and Generation; *Yinghao Xu*; Zifan Shi; Wang Yifan; Hansheng Chen; Ceyuan Yang; Sida Peng; Yujun Shen; Gordon Wetzstein*

261. Thinking Outside the BBox: Unconstrained Generative Object Compositing; *Gemma Canet Tarrés*; Zhe Lin; Zhifei Zhang; Jianming Zhang; Yizhi Song; Dan Ruta; Andrew Gilbert; John Collomosse; Soo Ye Kim*

262. SemanticHuman-HD: High Resolution Semantic disentangled 3D Human Generation; *Peng Zheng; Tao Liu; Zili Yi; Rui Ma**

263. High-Fidelity Modeling of Generalizable Wrinkle Deformation; *Jingfan Guo; Jae Shin Yoon; Shunsuke Saito; Takaaki Shiratori; Hyun Soo Park**

264. ReLoo: Reconstructing Humans Dressed in Loose Garments from Monocular Video in the Wild; *Chen Guo*; Tianjian Jiang; Manuel Kaufmann; Chengwei Zheng; Julien Valentin; Jie Song*; Otmar Hilliges*

265. Enhancing Plausibility Evaluation for Generated Designs with Denoising Autoencoder; *Jiajie Fan*; Amal Trigui*; Thomas Bäck; Hao Wang*

266. StructLDM: Structured Latent Diffusion for 3D Human Generation; *Tao Hu; Fangzhou Hong; Ziwei Liu**

267. Skeleton-based Group Activity Recognition via Spatial-Temporal Panoramic Graph; *Zhengcen Li; Xinle Chang; Yueran Li; Jingyong Su**

268. Towards Physical World Backdoor Attacks against Skeleton Action Recognition; *Qichen Zheng; Yi Yu; SIYUAN YANG*; Jun Liu; Kwok-Yan Lam; Alex Kot*

269. MacDiff: Unified Skeleton Modeling with Masked Conditional Diffusion; *Lehong Wu*; Lilang Lin; Jiahang Zhang; Yiyang Ma; Jiaying Liu**

270. Tree-D Fusion: Simulation-Ready Tree Dataset from Single Images with Diffusion Priors; *Jae Joong Lee; Bosheng Li; Sara M Beery; Jonathan Huang; Songlin Fei; Raymond A. Yeh; Bedrich Benes**

271. Decomposed Vector-Quantized Variational Autoencoder for Human Grasp Generation; *zhao zhe*; Mengshi Qi; Huadong Ma*

272. DyFADet: Dynamic Feature Aggregation for Temporal Action Detection; *Le Yang*; Ziwei Zheng; Yizeng Han; Hao Cheng; Shiji Song; Gao Huang; Fan Li*

273. Loc3Diff: Local Diffusion for 3D Human Head Synthesis and Editing; *Yushi Lan*; Feitong Tan; Qiangeng Xu; Di Qiu; Kyle Genova; Zeng Huang; Rohit Pandey; Sean Fanello; Thomas Funkhouser; Chen Change Loy; Yinda Zhang**

274. PAV: Personalized Head Avatar from Unstructured Video Collection; *Akin Caliskan*; Berkay Kicanaoglu; Hyeongwoo Kim*

275. Expressive Whole-Body 3D Gaussian Avatar; *Gyeongsik Moon*; Takaaki Shiratori; Shunsuke Saito*

276. Language-Driven Physics-Based Scene Synthesis and Editing via Feature Splatting; *Ri-Zhao Qiu*; Ge Yang; Weijia Zeng; Xiaolong Wang*

277. High-Quality Mesh Blendshape Generation from Face Videos via Neural Inverse Rendering; *Xin Ming; Jiawei Li; Jingwang Ling; Libo Zhang; Feng Xu**

278. Unsupervised Multi-modal Medical Image Registration via Invertible Translation; *Mengjie Guo**

279. Hierarchically Structured Neural Bones for Reconstructing Animatable Objects from Casual Videos; *Subin Jeon; In Cho; Minsu Kim; Woong Oh Cho; Seon Joo Kim**

280. Region-Adaptive Transform with Segmentation Prior for Image Compression; *Yuxi Liu*; Wenhan Yang; Huihui Bai; Yunchao Wei; Yao Zhao*

281. Spherical World-Locking for Audio-Visual Localization in Egocentric Videos; *Heeseung Yun*; Ruohan Gao; Ishwarya Ananthabhotla; Anurag Kumar; Jacob Donley; Chao Li; Gunhee Kim; Vamsi Krishna Ithapu; Calvin Murdock**

282. ReALFRED: An Embodied Instruction Following Benchmark in Photo-Realistic Environments; *Taewoong Kim; Cheolhong Min; Byeonghwi Kim; Jinyeon Kim; Wonje Jeung; Jonghyun Choi**

283. DIM: Dyadic Interaction Modeling for Social Behavior Generation; *Minh Tran*; Di Chang; Maksim Siniukov; Mohammad Soleymani*

284. S^3D-NeRF: Single-Shot Speech-Driven Neural Radiance Field for High Fidelity Talking Head Synthesis; *Dongze Li*; Kang Zhao*; Wei Wang*; Yifeng Ma; Bo Peng; Yingya Zhang; Jing Dong*

285. Explorative Inbetweening of Time and Space; *Haiwen Feng*; Zheng Ding; Zhihao Xia; Simon Niklaus; Victoria Fernandez Abrevaya; Michael J. Black; Xuaner Zhang*

286. ZeroI2V: Zero-Cost Adaptation of Pre-Trained Transformers from Image to Video; *Xinhao Li; Yuhan Zhu; Limin Wang**

287. DGE: Direct Gaussian 3D Editing by Consistent Multi-view Editing; *Minghao Chen*; Iro Laina; Andrea Vedaldi*

288. Length-Aware Motion Synthesis via Latent Diffusion; *Alessio Sampieri*; Alessio Palma; Indro Spinelli; Fabio Galasso*

289. IAM-VFI : Interpolate Any Motion for Video Frame Interpolation with motion complexity map; *Kihwan Yoon*; Yong Han Kim; Sungjei Kim*; Jinwoo Jeong**

290. Text-Guided Video Masked Autoencoder; *David Fan*; Jue Wang; Shuai Liao; Zhikang Zhang; Vimal Bhat; Xinyu Li*

291. WildVidFit: Video Virtual Try-On in the Wild via Image-Based Controlled Diffusion Models; *Zijian He; Peixin Chen; Guangrun Wang; Guanbin Li*; Philip Torr; Liang Lin*

292. TCAN: Animating Human Images with Temporally Consistent Pose Guidance using Diffusion Models; *Jeongho Kim*; Min-Jung Kim*; Junsoo Lee; Jaegul Choo**

293. Motion-Guided Latent Diffusion for Temporally Consistent Real-world Video Super-resolution; *Xi Yang*; Chenhang He; Jianqi Ma; Lei Zhang*

294. Learned Image Enhancement via Color Naming; *David Serrano-Lozano*; Luis Herranz; Michael S Brown; Javier Vazquez-Corral*

295. FreeMotion: A Unified Framework for Number-free Text-to-Motion Synthesis; *Ke Fan; Junshu Tang; Weijian Cao; Ran Yi*; Moran Li; Jingyu Gong; Jiangning Zhang; Yabiao Wang; Chengjie Wang; Lizhuang Ma**

296. Towards Open Domain Text-Driven Synthesis of Multi-Person Motions; *Mengyi Shan; Lu Dong; Yutao Han; Yuan Yao; Tao Liu; Ifeoma Nwogu; Guo-Jun Qi; Mitchell K Hill**

297. ObjectDrop: Bootstrapping Counterfactuals for Photorealistic Object Removal and Insertion; *Daniel Winter*; Matan Cohen; Shlomi Fruchter; Yael Pritch; Alex Rav-Acha; Yedid Hoshen**

298. ControlNet++: Improving Conditional Controls with Efficient Consistency Feedback; *Ming Li*; Taojiannan Yang; Huafeng Kuang; Jie Wu; Zhaoning Wang; Xuefeng Xiao; Chen Chen*

299. SignGen: End-to-End Sign Language Video Generation with Latent Diffusion; *Fan Qi*; Yu Duan; Changsheng Xu; Huaiwen Zhang**

300. Lossy Image Compression with Foundation Diffusion Models; *Lucas Relic*; Roberto Azevedo; Markus Gross; Christopher Schroers**

301. Disentangled Clothed Avatar Generation from Text Descriptions; *Jionghao Wang*; Yuan Liu; Zhiyang Dou; Zhengming Yu; Yongqing Liang; Cheng Lin; Rong Xie; Li Song*; Xin Li; Wenping Wang**

302. VividDreamer: Invariant Score Distillation for Hyper-Realistic Text-to-3D Generation; *Wenjie Zhuo*; Fan Ma; Hehe Fan; Yi Yang*

303. Pix2Gif: Motion-Guided Diffusion for GIF Generation; *Hitesh Kandala*; Jianfeng Gao; Jianwei Yang*

304. FreeDiff: Progressive Frequency Truncation for Image Editing with Diffusion Models; *Wei WU*; Qingnan Fan; Shuai Qin; Hong Gu; Ruoyu Zhao; Antoni Chan**

305. DATENeRF: Depth-Aware Text-based Editing of NeRFs; *Sara Rojas Martinez*; Julien Philip; Kai Zhang; Sai Bi; Fujun Luan; Bernard Ghanem; Kalyan Sunkavalli*

306. Score Distillation Sampling with Learned Manifold Corrective; *Thiemo Alldieck*; Nikos Kolotouros; Cristian Sminchisescu*

307. DNI: Dilutional Noise Initialization for Diffusion Video Editing; *Sunjae Yoon; Gwanhyeong Koo; Ji Woo Hong; Chang D. Yoo**

308. FRDiff : Feature Reuse for Universal Training-free Acceleration of Diffusion Models; *Junhyuk So; Jungwon Lee; Eunhyeok Park**

309. SmartControl: Enhancing ControlNet for Handling Rough Visual Conditions; *Xiaoyu Liu; Yuxiang Wei; Ming Liu*; Xianhui Lin; Peiran Ren; xuansong xie; Wangmeng Zuo*

310. Learning Quantized Adaptive Conditions for Diffusion Models; *Yuchen Liang*; Yuchuan Tian; Lei Yu; Huaao Tang; Jie Hu; Xiangzhong Fang; Hanting Chen**

311. Region-Native Visual Tokenization; *Mengyu Wang*; Yuyao Huang; Henghui Ding; Xinlong Wang; Tiejun Huang; Yao Zhao; Yunchao Wei; Shuicheng Yan*

312. ST-LDM: A Universal Framework for Text-Grounded Object Generation in Real Images; *Xiangtian Xue; Jiasong Wu*; Youyong Kong; Lotfi Senhadji; Huazhong Shu*

313. Concept Sliders: LoRA Adaptors for Precise Control in Diffusion Models; *Rohit Gandikota*; Joanna Materzynska; Tingrui Zhou; Antonio Torralba; David Bau*

314. Affine steerers for structured keypoint description; *Georg Bökman*; Johan Edstedt; Michael Felsberg; Fredrik Kahl*

315. PanGu-Draw: Advancing Resource-Efficient Text-to-Image Synthesis with Time-Decoupled Training and Reusable Coop-Diffusion; *Guansong Lu*; Yuanfan Guo; Jianhua Han; Minzhe Niu; Yihan Zeng; Songcen Xu; Wei Zhang; Hang Xu; Zhao Zhong; Zeyi Huang*

316. Factorizing Text-to-Video Generation by Explicit Image Conditioning; *Rohit Girdhar*; Mannat Singh; Andrew Brown; Quentin Duval; Samaneh Azadi; Sai Saketh Rambhatla; Mian Akbar Shah; Xi Yin; Devi Parikh; Ishan Misra*

317. MixDQ: Memory-Efficient Few-Step Text-to-Image Diffusion Models with Metric-Decoupled Mixed Precision Quantization; *Tianchen Zhao*; Xuefei Ning; Tongcheng Fang; Enshu Liu; Guyue Huang; Zinan Lin; Shengen Yan; Guohao Dai; Yu Wang*

318. LCM-Lookahead for Encoder-based Text-to-Image Personalization; *Rinon Gal*; Or Lichter; Elad Richardson; Or Patashnik; Amit Bermano; Gal Chechik; Danny Cohen-Or*

319. Improving image synthesis with diffusion-negative sampling; *Alakh Desai*; Nuno Vasconcelos*

320. MoMA: Multimodal LLM Adapter for Fast Personalized Image Generation; *Kunpeng Song*; Yizhe Zhu*; Bingchen Liu*; Qing Yan*; Ahmed Elgammal*; Xiao Yang**

321. Visual Text Generation in the Wild; *Yuanzhi Zhu; Jiawei Liu; Feiyu Gao; Wenyu Liu*; Xinggang Wang; Peng Wang; Fei Huang; Cong Yao; Zhibo Yang**

322. DreamReward: Aligning Human Preference in Text-to-3D Generation; *Junliang Ye; Fangfu Liu; Qixiu Li; Zhengyi Wang; Yikai Wang; Xinzhou Wang; Yueqi Duan*; Jun Zhu**

323. ReCON: Training-Free Acceleration for Text-to-Image Synthesis with Retrieval of Concept Prompt Trajectories; *Chen-Yi Lu*; Shubham Agarwal; Md Mehrab Tanjim; Kanak Mahadik; Anup Rao; Subrata Mitra; Shiv K Saini; Saurabh Bagchi; Somali Chaterji*

324. Idea2Img: Iterative Self-Refinement with GPT-4V for Automatic Image Design and Generation; *Zhengyuan Yang*; Jianfeng Wang; Linjie Li; Kevin Lin; Chung-Ching Lin; Zicheng Liu; Lijuan Wang*

325. Powerful and Flexible: Personalized Text-to-Image Generation via Reinforcement Learning; *Fanyue Wei; Wei Zeng; Zhenyang Li; Dawei Yin; Lixin Duan; Wen Li**

326. MTKD: Multi-Teacher Knowledge Distillation for Image Super-Resolution; *Yuxuan Jiang*; Chen Feng; Fan Zhang; David Bull*

327. Spherical Linear Interpolation and Text-Anchoring for Zero-shot Composed Image Retrieval; *Young Kyun Jang*; Dat B Huynh; Ashish Shah; Wen-Kai Chen; Ser-Nam Lim**

328. TIBET: Identifying and Evaluating Biases in Text-to-Image Generative Models; *Aditya Chinchure*; Pushkar Shukla*; Gaurav Bhatt; Kiri Salij; Kartik Hosanagar; Leonid Sigal; Matthew Turk*

329. Navigating Text-to-Image Generative Bias across Indic Languages; *Surbhi Mittal*; Arnav Sudan; Mayank Vatsa*; Richa Singh; Tamar Glaser; Tal Hassner*

330. Safeguard Text-to-Image Diffusion Models with Human Feedback Inversion; *Sanghyun Kim*; Seohyeon Jung; Balhae Kim; Moonseok Choi; Jinwoo Shin; Juho Lee**

335. Contourlet Residual for Prompt Learning Enhanced Infrared Image Super-Resolution; *Xingyuan Li; Jinyuan Liu*; ZHIXIN CHEN; Yang Zou; Long Ma; Xin Fan; Risheng Liu*

336. FairViT: Fair Vision Transformer via Adaptive Masking; *Bowei Tian; Ruijie Du; Yanning Shen**

337. Protecting NeRFs' Copyright via Plug-And-Play Watermarking Base Model; *Qi Song*; Ziyuan Luo; Ka Chun Cheung; Simon See; Renjie Wan*

338. Using My Artistic Style? You Must Obtain My Authorization; *Xiuli Bi; Haowei Liu; Weisheng Li; Bo Liu*; Bin Xiao*

339. Finding a needle in a haystack: A Black-Box Approach to Invisible Watermark Detection; *Minzhou Pan*; Zhenting Wang; Xin Dong; Vikash Sehwag; Lingjuan Lyu; Xue Lin*

340. Robust-Wide: Robust Watermarking against Instruction-driven Image Editing; *Runyi Hu; Jie Zhang*; Ting Xu; Jiwei Li; Tianwei Zhang*

341. ColorMNet: A Memory-based Deep Spatial-Temporal Feature Propagation Network for Video Colorization; *Yixin Yang; Jiangxin Dong; Jinhui Tang; Jinshan Pan**

342. RCS-Prompt: Learning Prompt to Rearrange Class Space for Prompt-based Continual Learning; *Longrong Yang; Hanbin Zhao; Yunlong Yu*; Xiaodong Zeng; Xi Li**

DIAMOND SPONSORS

BENDING SPOONS

Google Research

HPC-AI.COM
PREMIUM GPU, ANYTIME

PLATINUM SPONSORS

Lightning AI
Creators of PyTorch Lightning

Meta

GOLD SPONSORS

ANT RESEARCH

Apple

Baidu 百度

ByteDance

HUAWEI

iit ISTITUTO ITALIANO DI TECNOLOGIA

SILVER SPONSORS

TENCENT ROBOTICS X

Weights & Biases

BRONZE SPONSORS

CHEN TIANQIAO & CHRISSY INSTITUTE

INSAIT Institute for Computer Science, Artificial Intelligence and Technology

EXHIBITORS

3dMD

Advex

covision media

eNCORD

evs

IO INDUSTRIES

Labelbox

LatticeFlow

LIGHTLY

MDPI Academic Open Access Publishing since 1996

meshcapade

MOVE 4D

ONESOURCE CLOUD

parallel domain

SuperAnnotate

THREE LINES OF CODE

VOXEL51

STARTUP EXHIBITORS

abaka·ai

CL CINEMERSIVELABS

ITALAI

KEYLABS

Living Optics

TENYKS

IET The Institution of Engineering and Technology

人工智能与数据科学学院

Visual Layer

ORGANIZING SECRETARIAT

AIM GROUP INTERNATIONAL

AIM Group International - Milan Office
Viale E. Forlanini, 23 - 20134 Milan
T +39 02 56601.1 Mail: info@aimgroup.eu
aimgroupinternational.com