

# Jingdong Wang

Fellow of CAE, IEEE & IAPR · ACM Distinguished Member

wangjingdong@outlook.com · +86 13661386588 · jingdongwang2017.github.io · Google Scholar · DBLP

## Summary

---

Researcher in multimodal AI, computer vision, deep learning, and multimedia information retrieval.

Fellow of the Canadian Academy of Engineering (CAE), IEEE, IAPR, and AAIA; ACM Distinguished Member. Chief Scientist for Computer Vision (2021 – present); Senior Principal Researcher at Microsoft Research Asia (2007 – 2021). Best known for foundational contributions including HRNet, OCRNet, SPTAG, CAE, and SRA. Program Chair of ICCV 2025. Associate Editor for IEEE TPAMI, IJCV, and ACM TOMM.

## Research Interests

---

**Multimodal Understanding and Generation · Diffusion Models · World Models · Embodied Intelligence** · Vision Foundation Models · Self-Supervised Learning · Neural Architecture Design · Semantic Segmentation · Object Detection · Large-Scale Vector Search

## Education

---

**Ph.D. in Computer Science** 2004.09 – 2007.08

*The Hong Kong University of Science and Technology*

*Advisor: Long Quan*

**M.Eng. in Automation** 2001.09 – 2004.07

*Tsinghua University*

*Advisor: Changshui Zhang · Excellent Master Degree Thesis Award*

**B.Eng. in Automation** 1997.09 – 2001.07

*Tsinghua University*

*Rank: 1/150*

## Work Experience

---

**Chief Scientist for Computer Vision** 2021.09 – present

*Baidu*

**Senior Principal Research Manager** 2019.09 – 2021.08

*Microsoft Research Asia*

**Ph.D. Supervisor** 2017.09 – 2021.08

*University of Science and Technology, China*

**Principal Researcher** 2017.09 – 2019.08

*Microsoft Research Asia*

|                                                                                                                                  |                   |
|----------------------------------------------------------------------------------------------------------------------------------|-------------------|
| <b>Principal Researcher</b><br><i>Microsoft Research Asia</i>                                                                    | 2015.09 – 2017.08 |
| <b>Senior Researcher</b><br><i>Microsoft Research Asia</i>                                                                       | 2013.09 – 2015.08 |
| <b>Senior Researcher</b><br><i>Microsoft Research Asia</i>                                                                       | 2012.03 – 2013.08 |
| <b>Researcher</b><br><i>Microsoft Research Asia</i>                                                                              | 2007.09 – 2012.02 |
| <b>Research and Teaching Assistant</b><br><i>The Hong Kong University of Science and Technology</i><br><i>Advisor: Long Quan</i> | 2004.09 – 2007.08 |
| <b>Research Intern</b><br><i>Microsoft Research Asia</i><br><i>Mentor: Heung-Yeung Shum</i>                                      | 2005.07 – 2005.12 |
| <b>Research Intern</b><br><i>Microsoft Research Asia</i><br><i>Mentor: Heung-Yeung Shum</i>                                      | 2003.09 – 2004.05 |
| <b>Research Assistant</b><br><i>Tsinghua University</i><br><i>Advisor: Changshui Zhang</i>                                       | 2001.09 – 2004.07 |

## Professional Activities

---

### Program Chair

- ICCV 2025

### Conference Organizer

- ACMMM Asia 2024 (Sponsor Chair)
- ACMMM 2021 (Open Source Software Chair)
- ACMMM Asia 2020 (General Chair)
- ICME 2019 (Workshop Co-Chair)
- ACM ICMR 2014 (Special Session Co-Chair)

### Senior Area Chair

- CVPR 2024, ACMMM 2024/25/26

### Area/Track Chair

- ECCV 2026, ICML 2026, CVPR 2026, ICLR 2026, NeurIPS 2024, ECCV 2024, NeurIPS 2023, ICCV 2023, CVPR 2023, AAAI 2023, ICPR 2022, AAAI 2022, CVPR 2021, AAAI 2021, IJCAI 2021, ECCV 2020, CVPR 2020, IJCAI 2020, ICCV 2019, IJCAI 2019, ACMMM 2018, ICCV 2017, CVPR 2017, ECCV 2016, ACMMM 2015, ICPR 2018, ICIP 2017, ICME 2015, ICME 2012

## Senior Programme Committee Member

- AAAI 2019, IJCAI 2018, AAAI 2018

## Associate Editor / Editorial Board

- ACM Transactions on Multimedia Computing, Communications, and Applications (July 2023 – present)
- International Journal of Computer Vision (Aug. 2021 – present)
- IEEE Transactions on Pattern Analysis and Machine Intelligence (Sept. 2018 – present)
- IEEE Transactions on Circuits and Systems for Video Technology (Jan. 2018 – Jan. 2022)
- IEEE Transactions on Multimedia (May 2016 – May 2020)
- Neurocomputing (April 2015 – May 2017)
- Multimedia Tools and Applications (Oct. 2011 – present)

## Workshop Organizer

- Fifth Workshop on Neural Architecture Search (CVPR 2024)
- Workshop on Foundation Models: 2nd Workshop (CVPR 2024)
- Workshop on Foundation Models: 1st Foundation Model Challenge (CVPR 2023)
- Fourth Workshop on Neural Architecture Search (CVPR 2023)
- Billion-Scale Approximate Nearest Neighbor Search Challenge (NeurIPS 2021)
- 2nd International Workshop on Deep Learning for Pattern Recognition (ICPR 2020)
- 2nd Workshop on Target Re-Identification (CVPR 2019)
- 2nd International Workshop on Deep Learning for Pattern Recognition (ICPR 2018)
- 1st Workshop on Target Re-Identification (CVPR 2017)
- 1st International Workshop on Deep Learning for Pattern Recognition (ICPR 2016)
- Big Media Data: Understanding, Search, and Mining (ICDM 2015)
- Advances in Large-Scale Multimedia Data Collection, Mining and Retrieval (ICME 2012)

## Special Issue Guest Editor

- Human Pose Estimation and Its Applications (Machine Vision and Applications, 2023)
- Fine-grained Image Categorization (IEEE TPAMI, 2018)
- Intelligent Urban Computing with Big Data (Machine Vision and Applications, 2017)
- Big Media Data: Understanding, Search, and Mining (IEEE Trans. on Big Data, 2016)
- Ad Hoc Web Multimedia Analysis with Limited Supervision (Multimedia Tools and Applications, 2015)

## Honors and Awards

---

|                                                                      |      |
|----------------------------------------------------------------------|------|
| Outstanding Area Chair Award, CVPR                                   | 2026 |
| Editorial Contribution Award, Springer Nature/IJCV                   | 2025 |
| Fellow, Canadian Academy of Engineering (CAE)                        | 2024 |
| Fellow, IEEE                                                         | 2022 |
| Second Prize, Natural Science of China Society of Image and Graphics | 2022 |
| Fellow, AAIA                                                         | 2021 |
| Featured ACM Member                                                  | 2021 |
| Outstanding Service Award, ICME                                      | 2019 |
| ACM Distinguished Member                                             | 2018 |
| Fellow, IAPR                                                         | 2018 |

|                                                                |             |
|----------------------------------------------------------------|-------------|
| Best Paper Runner-Up, Workshop on Big Data Analytics           | 2018        |
| First Prize, Natural Science of Ministry of Education of China | 2017        |
| Best Paper Runner-Up, ACMMM                                    | 2015        |
| Outstanding Reviewer, CVPR                                     | 2014        |
| Gold Star Award, Microsoft                                     | 2009        |
| Graduate Fellowship, HKUST                                     | 2004 – 2007 |
| Excellent Master Degree Thesis Award, Tsinghua University      | 2004        |
| Excellent Master Scholarship, Tsinghua University              | 2003        |
| Graduate with Honors (Top 5%), Tsinghua University             | 2001        |
| Rockwell Scholarship, Tsinghua University                      | 2000        |
| Friend-Eastern Communications Scholarship, Tsinghua University | 1999        |
| Outstanding Student Scholarship, Tsinghua University           | 1998        |

### Competition Rankings

|                                                        |      |
|--------------------------------------------------------|------|
| Rank #1, nuScenes Tracking Task (Camera) Leaderboard   | 2022 |
| Rank #1, nuScenes Tracking Task (Lidar) Leaderboard    | 2022 |
| Rank #1, COCO Object Detection Leaderboard             | 2022 |
| Rank #1, Standard English and Chinese Benchmarks       | 2022 |
| Rank #1, Cityscapes Segmentation Leaderboard           | 2020 |
| Second Place, COCO Keypoints Detection Challenge, ECCV | 2018 |

### Publications

---

Google Scholar citations: 85,555, h-index: 115

<https://scholar.google.com/citations?user=z5SPCmgAAAAJ&hl=en>

#### Book

- [1] Jingdong Wang. Graph Based Image Segmentation: A Modern Approach. *ISBN-10: 3639110749, ISBN-13: 978-3639110746*. VDM Verlag Dr. Muller Aktiengesellschaft & Co. KG. 2008.

#### Book chapters

- [1] Jingdong Wang, **Jing Wang**, Gang Zeng, Rui Gan, Shipeng Li, and Baining Guo. Fast Neighborhood Graph Search using Cartesian Concatenation. *Multimedia Data Mining and Analytics: Disruptive Innovation. Edited by Aaron K. Baughman, Jiang Gao, Jia-Yu Pan, and Valery Petrushin*, 397-417, Springer, 2015.
- [2] Jingdong Wang, **Jing Wang**, Qifa Ke, Gang Zeng, and Shipeng Li. Fast Approximate K-Means via Cluster Closures. *Multimedia Data Mining and Analytics: Disruptive Innovation. Edited by Aaron K. Baughman, Jiang Gao, Jia-Yu Pan, and Valery Petrushin*, 373-395, Springer, 2015.
- [3] Jingdong Wang, **Yinghai Zhao**, and Xian-Sheng Hua. Image Search by Color Map. *Internet Multimedia Search and Mining. Edited by Xian-Sheng Hua, Marcel Worring, and Tat-Seng Chua*. Bentham Science Publishers, 2010.

#### Technical reports

- [1] Kai Chen, Jiaqi Wang, Jiangmiao Pang, Yuhang Cao, Yu Xiong, Xiaoxiao Li, Shuyang Sun, Wansen Feng, Ziwei Liu, Jiarui Xu, Zheng Zhang, Dazhi Cheng, Chenchen Zhu, Tianheng Cheng, Qijie Zhao, Buyu Li, Xin Lu, Rui Zhu, Yue Wu, Jifeng Dai, Jingdong Wang, Jianping Shi, Wanli Ouyang, Chen Change Loy, and Dahua Lin. MMDetection: Open MMLab Detection Toolbox and Benchmark. *CoRR*, abs/1906.07155, 2019.

- [2] **Liming Zhao**, Jingdong Wang, Xi Li, Zhuowen Tu, and Wenjun Zeng. On the connection of deep fusion to ensembling. *CoRR*, abs/1611.07718, 2016.
- [3] Jingdong Wang, **Zhen Wei**, **Ting Zhang**, and Wenjun Zeng. Deeply-fused nets. *CoRR*, abs/1605.07716, 2016.

**Top referred journals:** TPAMI (18), IJCV (18), TOG (3), TIP (10), TMM (7), TOMM (2), TKDE (1), ACM CS (1)

- [1] Ming Dai, Wenxuan Cheng, Jiang-Jiang Liu, Lingfeng Yang, Zhenhua Feng, Wankou Yang, Jingdong Wang. Improving Generalized Visual Grounding With Instance-Aware Joint Learning. *IEEE Trans. Pattern Anal. Mach. Intell.*, 2026.
- [2] Jiacheng Zhang, Jiaming Li, Xiangru Lin, Wei Zhang, Xiao Tan, Hongbo Gao, Jingdong Wang, Guanbin Li. Toward Efficient Semi-Supervised Object Detection With Detection Transformer. *IEEE Trans. Pattern Anal. Mach. Intell.*, 2026.
- [3] Yasheng Sun, Zhiliang Xu, Hang Zhou, Jiazhi Guan, Quanwei Yang, Kaisiyuan Wang, Borong Liang, Yingying Li, Haocheng Feng, Jingdong Wang, Ziwei Liu, Hideki Koike. Cosh-DiT: Co-Speech Gesture Video Synthesis via Hybrid Audio-Visual Diffusion Transformers. *International Journal of Computer Vision*, 2025.
- [4] Jiankai Sun, Chuanyang Zheng, Enze Xie, Zhengying Liu, Ruihang Chu, Jianing Qiu, Jiaqi Xu, Mingyu Ding, Hongyang Li, Mengzhe Geng, Yue Wu, Wenhai Wang, Junsong Chen, Zhangyue Yin, Xiaozhe Ren, Jie Fu, Junxian He, Wu Yuan, Qi Liu, Xihui Liu, Yu Li, Hao Dong, Yu Cheng, Ming Zhang, Pheng-Ann Heng, Jifeng Dai, Ping Luo, Jingdong Wang, Ji-Rong Wen, Xipeng Qiu, Yike Guo, Hui Xiong, Qun Liu, Zhenguo Li. A Survey of Reasoning with Foundation Models: Concepts, Methodologies, and Outlook. *ACM Comput. Surv.*, 2025.
- [5] Yiqun Zhao, Chenming Wu, Binbin Huang, Yihao Zhi, Chen Zhao, Jingdong Wang, Shenghua Gao. Surfel-Based Gaussian Inverse Rendering for Fast and Relightable Dynamic Human Reconstruction From Monocular Videos. *IEEE Trans. Pattern Anal. Mach. Intell.*, 2025.
- [6] Dingwen Zhang, Hao Li, Diqi He, Nian Liu, Lechao Cheng, Jingdong Wang, Junwei Han. Unsupervised Pre-Training With Language-Vision Prompts for Low-Data Instance Segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.*, 2025.
- [7] Wenhao Wu, Xiaohan Wang, Haipeng Luo, Jingdong Wang, Yi Yang, Wanli Ouyang. Cap4Video++: Enhancing Video Understanding With Auxiliary Captions. *IEEE Trans. Pattern Anal. Mach. Intell.*, 2025.
- [8] Zhuohang Dang, Minnan Luo, Jihong Wang, Chengyou Jia, Haochen Han, Herun Wan, Guang Dai, Xiaojun Chang, Jingdong Wang. Disentangled Noisy Correspondence Learning. *IEEE Trans. Image Processing*, 2025.
- [9] Yanpeng Sun, Qiang Chen, Jian Wang, Jingdong Wang, Zechao Li. Exploring Effective Factors for Improving Visual In-Context Learning. *IEEE Trans. Image Processing*, 2025.
- [10] Jiaxiang Tang, Kaisiyuan Wang, Hang Zhou, Xiaokang Chen, Dongliang He, Tianshu Hu, Jingtuo Liu, Ziwei Liu, Gang Zeng, Jingdong Wang. Real-Time Neural Radiance Talking Portrait Synthesis via Audio-Spatial Decomposition. *International Journal of Computer Vision*, 2025.
- [11] Zhengqi Zhao, Xiaohu Huang, Hao Zhou, Kun Yao, Errui Ding, Jingdong Wang, Xinggang Wang, Wenyu Liu, Bin Feng. Skim then Focus: Integrating Contextual and Fine-grained Views for Repetitive Action Counting. *International Journal of Computer Vision*, 2025.
- [12] Zhiyuan Yang, Xuekuan Wang, Wei Zhang, Xiao Tan, Jincheng Lu, Jingdong Wang, Errui Ding, Cairong Zhao. Fusion4DAL: Offline Multi-modal 3D Object Detection for 4D Auto-labeling. *International Journal of Computer Vision*, 2025.
- [13] Jiazheng Xing, Chao Xu, Yijie Qian, Yang Liu, Guang Dai, Baigui Sun, Yong Liu, Jingdong Wang. TryOn-Adapter: Efficient Fine-Grained Clothing Identity Adaptation for High-Fidelity Virtual Try-On. *International Journal of Computer Vision*, 2025.

- [14] Huan Liu, Zichang Tan, Qiang Chen, Yunchao Wei, Yao Zhao, Jingdong Wang. Unified Frequency-Assisted Transformer Framework for Detecting and Grounding Multi-modal Manipulation. *International Journal of Computer Vision*, 2025.
- [15] Sifan Long, Zhen Zhao, Junkun Yuan, Zichang Tan, Jiangjiang Liu, Jingyuan Feng, Sheng-Sheng Wang, Jingdong Wang. Mutual Prompt Learning for Vision Language Models. *International Journal of Computer Vision*, 2025.
- [16] Yong Guo, Mingkui Tan, Zeshuai Deng, Jingdong Wang, Qi Chen, Jiezhong Cao, Yanwu Xu, Jian Chen: Towards Lightweight Super-Resolution With Dual Regression Learning. *IEEE Trans. Pattern Anal. Mach. Intell.*, 2024.
- [17] Keyao Wang, Guosheng Zhang, Haixiao Yue, Yanyan Liang, Mouxiao Huang, Gang Zhang, Junyu Han, Errui Ding, Jingdong Wang. CSDG-FAS: Closed-Space Domain Generalization for Face Anti-spoofing. *International Journal of Computer Vision*, 2024.
- [18] **Xiaokang Chen, Mingyu Ding**, Xiaodi Wang, Ying Xin, Shentong Mo, Yunhao Wang, Shumin Han, Ping Luo, Gang Zeng, and Jingdong Wang. Context Autoencoder for Self-Supervised Representation Learning. *International Journal of Computer Vision*, 2024.
- [19] Wenhao Wu, Zhun Sun, Yuxin Song, Jingdong Wang, and Wanli Ouyang. Transferring Vision-Language Models for Visual Recognition: A Classifier Perspective. *International Journal of Computer Vision*, 2024.
- [20] **Yifan Liu**, Changyong Shu, Jingdong Wang, and Chunhua Shen. Structured Knowledge Distillation for Dense Prediction. *IEEE Trans. Pattern Anal. Mach. Intell.*, 2023.
- [21] Yan Huang, Jingdong Wang, and Liang Wang. Few-Shot Image and Sentence Matching via Aligned Cross-Modal Memory. *IEEE Trans. Pattern Anal. Mach. Intell.*, 2022.
- [22] Yuhui Yuan, Lang Huang, Jianyuan Guo, Chao Zhang, Xilin Chen and Jingdong Wang. OCNet: Object Context for Semantic Segmentation. *International Journal of Computer Vision*, 2021.
- [23] Jingdong Wang, **Ke Sun, Tianheng Cheng, Borui Jiang, Chaorui Deng, Yang Zhao**, Dong Liu, Yadong Mu, Mingkui Tan, Xinggang Wang, Wenyu Liu, and Bin Xiao. Deep High-Resolution Representation Learning for Visual Recognition. *IEEE Trans. Pattern Anal. Mach. Intell.*, 2021.
- [24] **Jingchun Cheng**, Yuhui Yuan, Yali Li, Jingdong Wang, and Shengjin Wang. Learning to Segment Video Object With Accurate Boundaries. *IEEE Trans. Multimedia*, 2021.
- [25] **Dengpan Fu**, Bo Xin, Jingdong Wang, Dongdong Chen, Jianmin Bao, Gang Hua and Houqiang Li. Improving Person Re-Identification With Iterative Impression Aggregations. *IEEE Trans. Image Processing*, 2020.
- [26] **Zilong Huang**, Chunyu Wang, Xinggang Wang, Wenyu Liu, and Jingdong Wang. Semantic Image Segmentation by Scale-Adaptive Networks. *IEEE Trans. Image Processing*, 2019.
- [27] **Peng Tang**, Chunyu Wang, Xinggang Wang, Wenyu Liu, and Jingdong Wang. Object Detection in Videos by High Quality Object Linking. *IEEE Trans. Pattern Anal. Mach. Intell.*, 2020.
- [28] **Song Bai**, Zhichao Zhou, Jingdong Wang, Xiang Bai, Longin Jan Latecki and Qi Tian. Automatic Ensemble Diffusion for 3D Shape and Image Retrieval. *IEEE Trans. Image Processing*, 28(1):88-101, 2019.
- [29] Jingdong Wang and **Ting Zhang**. Composite Quantization. *IEEE Trans. Pattern Anal. Mach. Intell.*, 2019.
- [30] **Qiurui Wang**, Chun Yuan, Jingdong Wang and Wenjun Zeng. Learning Attentional Recurrent Neural Network for Visual Tracking. *IEEE Trans. Multimedia*, 2019.
- [31] **Hong Liu**, Rongrong Ji, Jingdong Wang, and Chunhua Shen. Ordinal Constraint Binary Coding for Approximate Nearest Neighbor Search. *IEEE Trans. Pattern Anal. Mach. Intell.*, 2019.

- [32] Jingdong Wang, **Ting Zhang**, **Jingkuan Song**, Nicu Sebe, and Heng Tao Shen. A Survey on Learning to Hash. *IEEE Trans. Pattern Anal. Mach. Intell.*, 40(4):769-790, 2018.
- [33] **Na Qi**, Yunhui Shi, Xiaoyan Sun, Jingdong Wang, Baocai Yin, and Junbin Gao. Multi-Dimensional Sparse Models. *IEEE Trans. Pattern Anal. Mach. Intell.*, 40(1):163-178, 2018.
- [34] Weiyao Lin, Yang Shen, Junchi Yan, Mingliang Xu, Jianxin Wu, Jingdong Wang, and Ke Lu. Learning Correspondence Structures for Person Re-identification. *IEEE Trans. Image Processing*, 26(5):2438-2453, 2017.
- [35] **Dapeng Chen**, Zejian Yuan, Jingdong Wang, Gang Hua, and Nanning Zheng. Exemplar-Guided Similarity Learning on Polynomial Feature Map for Person Re-Identification. *International Journal of Computer Vision*, 123(3):392-414, 2017.
- [36] Jingdong Wang, **Huaizu Jiang**, Zejian Yuan, Ming-Ming Cheng, Xiaowei Hu, and Nanning Zheng. Salient Object Detection: A Discriminative Regional Feature Integration Approach. *International Journal of Computer Vision*, 123(2):251-268, 2017.
- [37] **Lingxi Xie**, Jingdong Wang, Weiyao Lin, Bo Zhang, and Qi Tian. Towards Reversal-Invariant Image Representation. *International Journal of Computer Vision*, 123(2):226-250, 2017.
- [38] **Dapeng Chen**, Zejian Yuan, Gang Hua, Jingdong Wang, and Nanning Zheng. Multi-Timescale Collaborative Tracking. *IEEE Trans. Pattern Anal. Mach. Intell.*, 39(1):141-155, 2017.
- [39] **Liang Zheng**, Shengjin Wang, Jingdong Wang, and Qi Tian. Accurate image search with multi-scale contextual evidences. *International Journal of Computer Vision*, 120(1):1-13, 2016.
- [40] **Dingwen Zhang**, Junwei Han, Chao Li, Jingdong Wang, and Xuelong Li. Detection of co-salient objects by looking deep and wide. *International Journal of Computer Vision*, 120(2):215-232, 2016.
- [41] Xi Li, Xueyi Zhao, Zhongfei Zhang, Fei Wu, Yueting Zhuang, Jingdong Wang, and Xuelong Li. Joint multilabel classification with community-aware label graph learning. *IEEE Trans. Image Processing*, 25(1):484-493, 2016.
- [42] Weiyao Lin, **Yang Mi**, Weiyue Wang, Jianxin Wu, Jingdong Wang, and Tao Mei. A diffusion and clustering-based approach for finding coherent motions and understanding crowd scenes. *IEEE Trans. Image Processing*, 25(4):1674-1687, 2016.
- [43] Xi Li, Liming Zhao, Lina Wei, Ming-Hsuan Yang, Fei Wu, Yueting Zhuang, Haibin Ling, and Jingdong Wang. Deepsaliency: Multi-task deep neural network model for salient object detection. *IEEE Trans. Image Processing*, 25(8):3919-3930, 2016.
- [44] **Jingkuan Song**, Heng Tao Shen, Jianfeng Wang, Zi Huang, Nicu Sebe, and Jingdong Wang. A distance-computation-free search scheme for binary code databases. *IEEE Trans. Multimedia*, 18(3):484-495, 2016.
- [45] Jinhui Tang, **Xiangbo Shu**, Zechao Li, Guo-Jun Qi, and Jingdong Wang. Generalized deep transfer networks for knowledge propagation in heterogeneous domains. *ACM TOMCCAP*, 12(4s):68:1-68:22, 2016.
- [46] **Jianfeng Wang**, Jingdong Wang, Jingkuan Song, Xin-Shun Xu, Heng Tao Shen, and Shipeng Li. Optimized Cartesian k-means. *IEEE Trans. Knowl. Data Eng.*, 27(1):180-192, 2015.
- [47] **Lingxi Xie**, Jingdong Wang, Bo Zhang, and Qi Tian. Fine-grained image search. *IEEE Trans. Multimedia*, 17(5):636-647, 2015.
- [48] Jingdong Wang, **Naiyan Wang**, **You Jia**, Jian Li, Gang Zeng, Hongbin Zha, and Xian-Sheng Hua. Trinary-projection trees for approximate nearest neighbor search. *IEEE Trans. Pattern Anal. Mach. Intell.*, 36(2):388-403, 2014.
- [49] Jingdong Wang, **Huaizu Jiang**, **Yangqing Jia**, Xian-Sheng Hua, Changshui Zhang, and Long Quan. Regularized tree partitioning and its application to unsupervised image segmentation. *IEEE Trans. Image Processing*, 23(4):1909-1922, 2014.

- [50] Qinghua Huang, **Bisheng Chen**, Jingdong Wang, and Tao Mei. Personalized video recommendation through graph propagation. *ACM TOMCCAP*, 10(4):32:1–32:17, 2014.
- [51] **Peng Wang**, Gang Zeng, Rui Gan, Jingdong Wang, and Hongbin Zha. Structure-sensitive superpixels via geodesic distance. *International Journal of Computer Vision*, 103(1):1–21, 2013.
- [52] Houqiang Li, **Yang Wang**, Tao Mei, Jingdong Wang, and Shipeng Li. Interactive multimodal visual search on mobile device. *IEEE Trans. Multimedia*, 15(3):594–607, 2013.
- [53] **Tianjia Shao**, Weiwei Xu, Kun Zhou, Jingdong Wang, Dongping Li, and Baining Guo. An interactive approach to semantic modeling of indoor scenes with an RGBD camera. *ACM Trans. Graph*, 31(6):136:1–136:11, 2012.
- [54] Tie Liu, Zejian Yuan, Jian Sun, Jingdong Wang, Nanning Zheng, Xiaoou Tang, and Heung-Yeung Shum. Learning to detect a salient object. *IEEE Trans. Pattern Anal. Mach. Intell.*, 33(2):353–367, 2011.
- [55] **Xinmei Tian**, Linjun Yang, Jingdong Wang, Xiuqing Wu, and Xian-Sheng Hua. Bayesian visual reranking. *IEEE Trans. Multimedia*, 13(4):639–652, 2011.
- [56] Jingdong Wang, Fei Wang, Changshui Zhang, Helen C. Shen, and Long Quan. Linear neighborhood propagation and its applications. *IEEE Trans. Pattern Anal. Mach. Intell.*, 31(9):1600–1615, 2009.
- [57] Tie Liu, Jingdong Wang, Jian Sun, Nanning Zheng, Xiaoou Tang, and Heung-Yeung Shum. Picture collage. *IEEE Trans. Multimedia*, 11(7):1225–1239, 2009.
- [58] Long Quan, Jingdong Wang, Ping Tan, and Lu Yuan. Image-based modeling by joint segmentation. *International Journal of Computer Vision*, 75(1):135–150, 2007.
- [59] Ping Tan, Gang Zeng, Jingdong Wang, Sing Bing Kang, and Long Quan. Image-based tree modeling. *ACM Trans. Graph*, 26(3):87, 2007.
- [60] Long Quan, Ping Tan, Gang Zeng, Lu Yuan, Jingdong Wang, and Sing Bing Kang. Image-based plant modeling. *ACM Trans. Graph*, 25(3):599–604, 2006.

#### Referred journals

- [1] **Guotian Xie**, Kuiyuan Yang, Ting Zhang, Jingdong Wang, and Jianhuang Lai. Balanced Decoupled Spatial Convolution for CNNs. *IEEE Trans. Neural Netw. Learning Syst.*, 2019.
- [2] Shengkang Yu, Xi Li, Xueyi Zhao, Zhongfei Mark Zhang, Fei Wu, Jingdong Wang, Yue Ting Zhuang, and Xuelong Li. A bilinear ranking SVM for knowledge based relation prediction and classification. *IEEE Trans. Big Data*, 2018.
- [3] Ruoyu Liu, Shikui Wei, Yao Zhao, Zhenfeng Zhu, and Jingdong Wang. Multi-View Cross-Media Hashing with Semantic Consistency. *IEEE MultiMedia*, 2018.
- [4] **Baoguang Shi**, Xiang Bai, Wenyu Liu, and Jingdong Wang. Face Alignment with Deep Regression. *IEEE Trans. Neural Netw. Learning Syst.*, 29(1):183–194, 2018.
- [5] **Lingxi Xie**, Jingdong Wang, Bo Zhang, and Qi Tian. Incorporating visual adjectives for image classification. *Neurocomputing*, 182:48–55, 2016.
- [6] **Min Tan**, Baoyuan Wang, Zhaohui Wu, Jingdong Wang, and Gang Pan. Weakly supervised metric learning for traffic sign recognition in a lidar-equipped vehicle. *IEEE Trans. Intelligent Transportation Systems*, 17(5):1415–1427, 2016.
- [7] Congyan Lang, Jiashi Feng, Songhe Feng, Jingdong Wang, and Shuicheng Yan. Dual low-rank pursuit: Learning salient features for saliency detection. *IEEE Trans. Neural Netw. Learning Syst.*, 27(6):1190–1200, 2016.
- [8] **Shiyang Lu**, Tao Mei, Jingdong Wang, Jian Zhang, Zhiyong Wang, and Shipeng Li. Exploratory product image search with circle-to-search interaction. *IEEE Trans. Circuits Syst. Video Techn.*, 25(7):1190–1202, 2015.

- [9] Weiwei Xu, Zhouxu Shi, Mingliang Xu, Kun Zhou, Jingdong Wang, Bin Zhou, Jinrong Wang, and Zhenming Yuan. Transductive 3d shape segmentation using sparse reconstruction. *Comput. Graph. Forum*, 33(5):107–115, 2014.
- [10] Jingdong Wang, **Jiazhen Zhou**, **Hao Xu**, Tao Mei, Xian-Sheng Hua, and Shipeng Li. Image tag refinement by regularized latent Dirichlet allocation. *Computer Vision and Image Understanding*, 124:61–70, 2014.
- [11] **Shiyang Lu**, Tao Mei, Jingdong Wang, Jian Zhang, Zhiyong Wang, and Shipeng Li. Browse-to-search: Interactive exploratory search with visual entities. *ACM Trans. Inf. Syst.*, 32(4):18:1–18:27, 2014.
- [12] Jingdong Wang, **Zhe Zhao**, **Jiazhen Zhou**, Hao Wang, Bin Cui, and Guojun Qi. Recommending Flickr groups with social topic model. *Inf. Retr.*, 15(3-4):278–295, 2012.
- [13] **Tianjia Shao**, Weiwei Xu, KangKang Yin, Jingdong Wang, Kun Zhou, and Baining Guo. Discriminative sketch-based 3d model retrieval via robust shape matching. *Comput. Graph. Forum*, 30(7):2011–2020, 2011.
- [14] Jingdong Wang, **Liyan Jia**, and Xian-Sheng Hua. Interactive browsing via diversified visual summarization for image search results. *Multimedia Syst.*, 17(5):379–391, 2011.
- [15] Jingdong Wang, **Yinghai Zhao**, Xiuqing Wu, and Xian-Sheng Hua. A transductive multi-label learning approach for video concept detection. *Pattern Recognition*, 44(10-11):2274–2286, 2011.
- [16] Jingdong Wang and Xian-Sheng Hua. Interactive image search by color map. *ACM TIST*, 3(1):12:1–12:23, 2011.
- [17] **Zheng-Jun Zha**, Tao Mei, Jingdong Wang, Zengfu Wang, and Xian-Sheng Hua. Graph-based semi-supervised learning with multiple labels. *J. Visual Communication and Image Representation*, 20(2):97–103, 2009.
- [18] Fei Wang, Jingdong Wang, Changshui Zhang, and James T. Kwok. Face recognition using spectral features. *Pattern Recognition*, 40(10):2786–2797, 2007.
- [19] Jianguo Lee, Jingdong Wang, Changshui Zhang, and Zhaoqi Bian. Visual object recognition using probabilistic kernel subspace similarity. *Pattern Recognition*, 38(7):997–1008, 2005.

**Top refereed conferences:** CVPR (77), ICCV (31), ECCV (32), ACMMM (18), SIGGRAPH (2), SIGGRAPH ASIA (2), NeurIPS (20), ICLR (10), ICML (11), SIGIR (3), KDD (1), IJCAI (6), AAAI (10)

- [1] Weihao Bo, Shan Zhang, Yanpeng Sun, Jingjing Wu, Qunyi Xie, Xiao Tan, Kunbin Chen, Wei He, Xiaofan Li, Na Zhao, Jingdong Wang, Zechao Li. Agentic Learner with Grow-and-Refine Multimodal Semantic Memory. *CVPR 2026*.
- [2] Xuan Huang, Mochu Xiang, Zhelun Shen, Jinbo Wu, Chenming Wu, Chen Zhao, Kaisiyuan Wang, Hang Zhou, Shanshan Liu, Haocheng Feng, Wei He, Jingdong Wang. GenHOI: Towards Object-Consistent Hand–Object Interaction with Temporally Balanced and Spatially Selective Object Injection. *CVPR 2026*.
- [3] Zhiyuan Yan, Kaiqing Lin, Zongjian Li, Junyan Ye, Hui Han, Haochen Wang, Zhendong Wang, Bin Lin, Li Hao, Xinyan Xiao, Jingdong Wang, Haifeng Wang, Li Yuan. Unified Multimodal Models as Auto-Encoders. *CVPR 2026*.
- [4] Hui Li, Jiayue Lyu, Fu-Yun Wang, Kaihui Cheng, Siyu Zhu, Jingdong Wang. MixFlow Training: Alleviating Exposure Bias with Slowed Interpolation Mixture. *CVPR 2026*.
- [5] Yanpeng Sun, Jing Hao, Ke Zhu, Jiang-Jiang Liu, Xiaofan Li, Na Zhao, Zechao Li, Jingdong Wang. Enhancing Descriptive Captions with Visual Attributes for Multimodal Perception. *CVPR 2026*.
- [6] Jing Yang, Sen Yang, Boqiang Duan, Ming Dai, Wei Zhang, Xiao Tan, Kunbin Chen, Wei He, Jingdong Wang, Hanli Wang. Hugging Visual Prompt and Segmentation Tokens: Consistency Learning for Fine-Grained Visual Understanding in MLLMs. *CVPR 2026*.

- [7] Haoyuan Zhang, Keyao Wang, Guosheng Zhang, Haixiao Yue, Zhiwen Tan, Siran Peng, Tianshuo Zhang, Xiao Tan, Kunbin Chen, Wei He, Jingdong Wang, Ajian Liu, Xiangyu Zhu, Zhen Lei. From Intuition to Investigation: A Tool-Augmented Reasoning MLLM Framework for Generalizable Face Anti-Spoofing. *CVPR 2026*.
- [8] Mochu Xiang, Zhelun Shen, Xuesong Li, Jiahui Ren, Jing Zhang, Chen Zhao, Shanshan Liu, Haocheng Feng, Jingdong Wang, Yuchao Dai. RnG: A Unified Transformer for Complete 3D Modeling from Partial Observations. *CVPR 2026*.
- [9] Jiaye Li, Baoyou Chen, Hui Li, Zilong Dong, Jingdong Wang, Siyu Zhu. Head-wise Adaptive Rotary Positional Encoding for Fine-Grained Image Generation. *CVPR 2026*.
- [10] YuXin Song, Yu Lu, Haoyuan Sun, Huanjin Yao, Fanglong Liu, Yifan Sun, Haocheng Feng, Hang Zhou, Jingdong Wang. CoLoGen: Progressive Learning of Concept–Localization Duality for Unified Image Generation. *CVPR 2026*.
- [11] Mengmeng Wang, Dengyang Jiang, Liuzhuozheng Li, Yucheng Lin, Guojiang Shen, Xiangjie Kong, Yong Liu, Guang Dai, Jingdong Wang. VAE-REPA: Variational Autoencoder Representation Alignment for Efficient Diffusion Training. *CVPR 2026*.
- [12] Zedong Dan, Zijie Wang, Wei Zhang, Xiangru Lin, Weiming Zhang, Xiao Tan, Jingdong Wang, Liang Lin, Guanbin Li. OptiMVMap: Offline Vectorized Map Construction via Optimal Multi-vehicle Perspectives. *CVPR 2026*.
- [13] Ming Dai, Sen Yang, Boqiang Duan, Boyuan Tong, Jiedong Zhuang, Wankou Yang, Jingdong Wang. VideoSEG-O3: A Multi-turn Reinforcement Learning Framework for Reasoning Video Object Segmentation. *ICML 2026*.
- [14] Yuxuan Yao, Yuxuan Chen, Hui Li, Kaihui Cheng, Qipeng Guo, Yuwei Sun, Zilong Dong, Jingdong Wang, Siyu Zhu. Prompt Reinjection: Alleviating Prompt Forgetting in Multimodal Diffusion Transformers. *ICML 2026*.
- [15] Junbo Zhao, Ting Zhang, Can Li, Wei He, Jingdong Wang, Hua Huang. Non-Parametric Structural Priors for Geometry Theorem Prediction. *ICML 2026*.
- [16] Xiaojing Wei, Ting Zhang, Wei He, Jingdong Wang, Hua Huang. GeoLoom: High-quality Geometric Diagram Generation from Textual Input. *ICML 2026*.
- [17] Ze Feng, Sen Yang, Boqiang Duan, Wankou Yang, Jingdong Wang. EM-KD: Distilling Efficient Multimodal Large Language Model with Unbalanced Vision Tokens. *AAAI 2026*.
- [18] Chunyu Li, Jiaye Li, Ruiqiao Mei, Haoyuan Xia, Hao Zhu, Jingdong Wang, Siyu Zhu. Hallo-Live: Real-Time Streaming Audio-video Avatar Generation with Asynchronous Dual-stream and Human-Centric Preference Distillation. *ACM MM 2026*.
- [19] Weihao Bo, Shan Zhang, Yanpeng Sun, Jie Liu, Yongke Yao, Jinhao Du, Wei He, Kai Zou, Zechao Li, Jingdong Wang. Diagram-MMU: A Multi-Modal Benchmark for Scientific Diagrams. *ECCV 2026*.
- [20] Zizhao Chen, Ping Wei, Guang Dai, Jingdong Wang, Mengmeng Wang. From Draft to Draft-Free: One-Step Video Object Removal via Privileged Distillation and Fast Planting. *ECCV 2026*.
- [21] YuXin Song, Wenkai Dong, Shizun Wang, Qi Zhang, Song Xue, Tao Yuan, Hu Yang, Haocheng Feng, Hang Zhou, Xinyan Xiao, Jingdong Wang. Query-Kontext: An Unified Multimodal Model for Image Generation and Editing. *ECCV 2026*.
- [22] Lei Wang, YuXin Song, Ge Wu, Haocheng Feng, Hang Zhou, Jingdong Wang, Yaxing Wang, Jian Yang. RefAlign: Representation Alignment for Reference-to-Video Generation. *ECCV 2026*.
- [23] Xinyao Zhang, Wenkai Dong, YuXin Song, Bo Fang, Qi Zhang, Jing Wang, Fan Chen, Hui Zhang, Haocheng Feng, Yu Lu, Hang Zhou, Chun Yuan, Jingdong Wang. SAMA: Factorized Semantic Anchoring and Motion Alignment for Instruction-Guided Video. *ECCV 2026*.

- [24] Ming Dai, Sen Yang, Boqiang Duan, Wankou Yang, Jingdong Wang. MomentSeg: Moment-Centric Sampling for Enhanced Referring Video Object Segmentation. *ECCV 2026*.
- [25] Dengyang Jiang, Mengmeng Wang, Liuzhuozheng Li, Lei Zhang, Haoyu Wang, Wei Wei, Guang Dai, Yanning Zhang, Jingdong Wang. Representation Alignment for Diffusion Transformers without External Components. *ICLR 2026*.
- [26] Zanyi Wang, Dengyang Jiang, Liuzhuozheng Li, Sizhe Dang, Chengzu Li, Harry Yang, Guang Dai, Mengmeng Wang, Jingdong Wang. Deforming Videos to Masks: Flow Matching for Referring Video Segmentation. *ICLR 2026*.
- [27] Hui Li, Baoyou Chen, Li jiaye, Jingdong Wang, Siyu Zhu. Pyramid Patchification Flow for Visual Generation. *ICLR 2026*.
- [28] Sizhe Dang, Huan Liu, Mengmeng Wang, Xin Lai, Guang Dai, Jingdong Wang. VidEvo: Evolving Video Editing through Exhaustive Temporal Modeling. *IJCAI 2025*.
- [29] Guosheng Zhang, Keyao Wang, Haixiao Yue, Ajian Liu, Gang Zhang, Kun Yao, Errui Ding, Jingdong Wang. Interpretable Face Anti-Spoofing: Enhancing Generalization with Multimodal Large Language Models. *AAAI 2025*.
- [30] Jiahao Wang, Caixia Yan, Weizhan Zhang, Haonan Lin, Mengmeng Wang, Guang Dai, Tieliang Gong, Hao Sun, Jingdong Wang. SpotActor: Training-Free Layout-Controlled Consistent Image Generation. *AAAI 2025*.
- [31] Zirui Wang, Mengmeng Wang, Jade Dai, Teli Ma, Guo-Jun Qi, Yong Liu, Guang Dai, Jingdong Wang. DynaMind: Reasoning over Abstract Video Dynamics for Embodied Decision-Making. *ICML 2025*.
- [32] Yuzhe Yao, Jun Chen, Zeyi Huang, Haonan Lin, Mengmeng Wang, Guang Dai, Jingdong Wang. Manifold Constraint Reduces Exposure Bias in Accelerated Diffusion Sampling. *ICLR 2025*.
- [33] Jing Yang, Minyue Jiang, Sen Yang, Xiao Tan, Yingying Li, Errui Ding, Jingdong Wang, Hanli Wang. MGMapNet: Multi-Granularity Representation Learning for End-to-End Vectorized HD Map Construction. *ICLR 2025*.
- [34] Jiahao Cui, Hui Li, Yao Yao, Hao Zhu, Hanlin Shang, Kaihui Cheng, Hang Zhou, Siyu Zhu, Jingdong Wang. Hallo2: Long-Duration and High-Resolution Audio-Driven Portrait Image Animation. *ICLR 2025*.
- [35] Zebin You, Xinyu Zhang, Hanzhong Guo, Jingdong Wang, Chongxuan Li. Are Images Indistinguishable to Humans Also Indistinguishable to Classifiers? *CVPR 2025*.
- [36] Jialun Liu, Jinbo Wu, Xiaobo Gao, Jiakui Hu, Bojun Xiong, Xing Liu, Chen Zhao, Hongbin Pei, Haocheng Feng, Yingying Li, Errui Ding, Jingdong Wang. TexGarment: Consistent Garment UV Texture Generation via Efficient 3D Structure-Guided Diffusion Transformer. *CVPR 2025*.
- [37] Jiahao Cui, Hui Li, Yun Zhan, Hanlin Shang, Kaihui Cheng, Yuqi Ma, Shan Mu, Hang Zhou, Jingdong Wang, Siyu Zhu. Hallo3: Highly Dynamic and Realistic Portrait Image Animation with Video Diffusion Transformer. *CVPR 2025*.
- [38] Mengmeng Wang, Zeyi Huang, Xiangjie Kong, Guojian Shen, Guang Dai, Jingdong Wang, Yong Liu: Action Detail Matters. Refining Video Recognition with Local Action Queries. *CVPR 2025*.
- [39] Yingying Fan, Quanwei Yang, Kaisiyuan Wang, Hang Zhou, Yingying Li, Haocheng Feng, Errui Ding, Yu Wu, Jingdong Wang. Re-HOLD: Video Hand Object Interaction Reenactment via adaptive Layout-instructed Diffusion Model. *CVPR 2025*.
- [40] Ke Zhu, Yu Wang, Yanpeng Sun, Qiang Chen, Jiangjiang Liu, Gang Zhang, Jingdong Wang. Continual SFT Matches Multimodal RLHF with Negative Supervision. *CVPR 2025*.

- [41] Jiazhi Guan, Kaisiyuan Wang, Zhiliang Xu, Quanwei Yang, Yasheng Sun, Shengyi He, Borong Liang, Yukang Cao, Yingying Li, Haocheng Feng, Errui Ding, Jingdong Wang, Youjian Zhao, Hang Zhou, Ziwei Liu. AudCast: Audio-Driven Human Video Generation by Cascaded Diffusion Transformers. *CVPR 2025*.
- [42] Hui Li, Mingwang Xu, Yun Zhan, Shan Mu, Jiaye Li, Kaihui Cheng, Yuxuan Chen, Tan Chen, Mao Ye, Jingdong Wang, Siyu Zhu. OpenHumanVid: A Large-Scale High-Quality Dataset for Enhancing Human-Centric Video Generation. *CVPR 2025*.
- [43] Ziyue Zhu, Shenlong Wang, Jin Xie, Jiang-jiang Liu, Jingdong Wang, Jian Yang. VoxelSplat: Dynamic Gaussian Splatting as an Effective Loss for Occupancy and Flow Prediction. *CVPR 2025*.
- [44] Jialong Zuo, Jiahao Hong, Feng Zhang, Changqian Yu, Hanyu Zhou, Changxin Gao, Nong Sang, Jingdong Wang. PLIP: Language-Image Pre-training for Person Representation Learning. *NeurIPS 2024*.
- [45] Jie Zhu, Yixiong Chen, Mingyu Ding, Ping Luo, Leye Wang, Jingdong Wang. MoLE: Enhancing Human-centric Text-to-image Diffusion via Mixture of Low-rank Experts. *NeurIPS 2024*.
- [46] Chuyang Zhao, Yuxin Song, Junru Chen, Kang Rong, Haocheng Feng, Gang Zhang, Shufan Ji, Jingdong Wang, Errui Ding, Yifan Sun. Octopus: A Multi-modal LLM with Parallel Recognition and Sequential Understanding. *NeurIPS 2024*.
- [47] Mengxi Zhang, Wenhao Wu, Yu Lu, Yuxin Song, Kang Rong, Huanjin Yao, Jianbo Zhao, Fanglong Liu, Haocheng Feng, Jingdong Wang, Yifan Sun. Automated Multi-level Preference for MLLMs. *NeurIPS 2024*.
- [48] Huanjin Yao, Wenhao Wu, Taojiannan Yang, Yuxin Song, Mengxi Zhang, Haocheng Feng, Yifan Sun, Zhiheng Li, Wanli Ouyang, Jingdong Wang. Dense Connector for MLLMs. *NeurIPS 2024*.
- [49] Quanwei Yang, Jiazhi Guan, Kaisiyuan Wang, Lingyun Yu, Wenqing Chu, Hang Zhou, ZhiQiang Feng, Haocheng Feng, Errui Ding, Jingdong Wang, Hongtao Xie. ShowMaker: Creating High-Fidelity 2D Human Video via Fine-Grained Diffusion Modeling. *NeurIPS 2024*.
- [50] Yanmin Wu, Jiarui Meng, Haijie Li, Chenming Wu, Yahao Shi, Xinhua Cheng, Chen Zhao, Haocheng Feng, Errui Ding, Jingdong Wang, Jian Zhang. OpenGaussian: Towards Point-Level 3D Gaussian-based Open Vocabulary Understanding. *NeurIPS 2024*.
- [51] Haonan Lin, Wenbin An, Jiahao Wang, Yan Chen, Feng Tian, Mengmeng Wang, Qianying Wang, Guang Dai, Jingdong Wang. Flipped Classroom: Aligning Teacher Attention with Student in Generalized Category Discovery. *NeurIPS 2024*.
- [52] Haonan Lin, Yan Chen, Jiahao Wang, Wenbin An, Mengmeng Wang, Feng Tian, Yong Liu, Guang Dai, Jingdong Wang, Qianying Wang. Schedule Your Edit: A Simple yet Effective Diffusion Noise Schedule for Image Editing. *NeurIPS 2024*.
- [53] Mingxiang Liao, Hannan Lu, Qixiang Ye, Wangmeng Zuo, Fang Wan, Tianyu Wang, Yuzhong Zhao, Jingdong Wang, Xinyu Zhang. Evaluation of Text-to-Video Generation Models: A Dynamics Perspective. *NeurIPS 2024*.
- [54] Zhe Liu, Jinghua Hou, Xinyu Wang, Xiaoqing Ye, Jingdong Wang, Hengshuang Zhao, Xiang Bai. LION: Linear Group RNN for 3D Object Detection in Point Clouds. *NeurIPS 2024*.
- [55] Jiazhi Guan, Quanwei Yang, Kaisiyuan Wang, Hang Zhou, Shengyi He, Zhiliang Xu, Haocheng Feng, Errui Ding, Jingdong Wang, Hongtao Xie, Youjian Zhao, Ziwei Liu. TALK-Act: Enhance Textural-Awareness for 2D Speaking Avatar Reenactment with Diffusion Model. *SIGGRAPH Asia 2024*.
- [56] Chengyou Jia, Minnan Luo, Xiaojun Chang, Zhuohang Dang, Mingfei Han, Mengmeng Wang, Guang Dai, Sizhe Dang, Jingdong Wang. Generating Action-conditioned Prompts for Open-vocabulary Video Action Recognition. *ACM Multimedia 2024*.
- [57] Xiuzhe Wu, Yang-Tian Sun, Handi Chen, Hang Zhou, Jingdong Wang, Zhengzhe Liu, Xiaojuan Qi. 3D-Aware Text-Driven Talking Avatar Generation. *ECCV 2024*.

- [58] 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. ReSyncer: Rewiring Style-Based Generator for Unified Audio-Visually Synced Facial Performer. *ECCV 2024*.
- [59] Hao Li, Yuanyuan Gao, Chenming Wu, Dingwen Zhang, Yalun Dai, Chen Zhao, Haocheng Feng, Errui Ding, Jingdong Wang, Junwei Han. GGRt: Towards Pose-Free Generalizable 3D Gaussian Splatting in Real-Time. *ECCV 2024*.
- [60] Penghui Du, Yu Wang, Yifan Sun, Luting Wang, Yue Liao, Gang Zhang, Errui Ding, Yan Wang, Jingdong Wang, Si Liu. LaMI-DETR: Open-Vocabulary Detection with Language Model Instruction. *ECCV 2024*.
- [61] Yuzhe Yao, Feng Tian, Jun Chen, Haonan Lin, Guang Dai, Yong Liu, Jingdong Wang. Timestep-Aware Correction for Quantized Diffusion Models. *ECCV 2024*.
- [62] Jinghua Hou, Tong Wang, Xiaoqing Ye, Zhe Liu, Shi Gong, Xiao Tan, Errui Ding, Jingdong Wang, Xiang Bai. OPEN: Object-Wise Position Embedding for Multi-view 3D Object Detection. *ECCV 2024*.
- [63] Rui Zhang, Xiangru Lin, Wei Zhang, Jincheng Lu, Xuekuan Wang, Xiao Tan, Yingying Li, Errui Ding, Jingdong Wang, Guanbin Li. Interactive 3D Object Detection with Prompts. *ECCV 2024*.
- [64] Zhe Liu, Jinghua Hou, Xiaoqing Ye, Tong Wang, Jingdong Wang, Xiang Bai. SEED: A Simple and Effective 3D DETR in Point Clouds. *ECCV 2024*.
- [65] Dingyuan Zhang, Dingkan Liang, Zichang Tan, Xiaoqing Ye, Cheng Zhang, Jingdong Wang, Xiang Bai. Make Your ViT-Based Multi-view 3D Detectors Faster via Token Compression. *ECCV 2024*.
- [66] 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. IRGen: Generative Modeling for Image Retrieval. *ECCV 2024*.
- [67] Xingyu Wan, Chengquan Zhang, Pengyuan Lyu, Sen Fan, Zihan Ni, Kun Yao, Errui Ding, and Jingdong Wang. Towards Unified Multi-granularity Text Detection with Interactive Attention. *ICML 2024*.
- [68] Zhiyu Yao, Jian Wang, Haixu Wu, Jingdong Wang, and Mingsheng Long. Mobile Attention: Mobile-Friendly Linear-Attention for Vision Transformers. *ICML 2024*.
- [69] hexi Peng, Tianjia Shao, Yong Liu, Jingke Zhou, Yin Yang, Jingdong Wang, and Kun Zhou. RTG-SLAM: Real-time 3D Reconstruction at Scale using Gaussian Splatting. *SIGGRAPH 2024*.
- [70] Chuyang Zhao, Yifan Sun, Wenhao Wang, Qiang Chen, Errui Ding, Yi Yang, and Jingdong Wang. MS-DETR: Efficient DETR Training with Mixed Supervision. *CVPR 2024*.
- [71] Huan Liu, Zichang Tan, Chuangchuang Tan, Yunchao Forgery-aware Adaptive Transformer for Generalizable Synthetic Image Detection. *CVPR 2024*.
- [72] Hao Li, Dingwen Zhang, Yalun Dai, Nian Liu, Lechao Cheng, Li Jingfeng, Jingdong Wang, and Junwei Han. GP-NeRF: Generalized Perception NeRF for Context-Aware 3D Scene Understanding . *CVPR 2024*.
- [73] Yanpeng Sun, Jiahui Chen, Shan Zhang, Xinyu Zhang, Qiang Chen, Gang Zhang, Errui Ding, Jingdong Wang, and Zechao Li. VRP-SAM: SAM with Visual Reference Prompt. *CVPR 2024*.
- [74] Haochen Han, Qinghua Zheng, Guang Dai, Minnan Luo, and Jingdong Wang. Learning to Rematch Mismatched Pairs for Robust Cross-Modal Retrieval. *CVPR 2024*.
- [75] Wenjie Wang, Yehao Lu, Guangcong Zheng, Shuigenzhan, Xiaoqing Ye, Zichang Tan, Jingdong Wang, Gaoang Wang, and Xi Li. BEVspread: Spread Voxel Pooling for Bird’s-Eye-View Representation in Vision-based Roadside 3D Object Detection. *CVPR 2024*.
- [76] Jiacheng Zhang, Jiaming Li, Xiangru Lin, Wei Zhang, Xiao Tan, Junyu Han, Errui Ding, Jingdong Wang, and Guanbin Li. Decoupled Pseudo-labeling for Semi-Supervised Monocular 3D Object Detection. *CVPR 2024*.

- [77] Zhuohang Dang, Minnan Luo, Chengyou Jia, Guang Dai, Xiaojun Chang, and Jingdong Wang. Noisy Correspondence Learning with Self-Reinforcing Errors Mitigation. *AAAI 2024*.
- [78] Chengyou Jia, Minnan Luo, Zhuohang Dang, Guang Dai, Xiaojun Chang, Mengmeng Wang, and Jingdong Wang. SSMG: Spatial-Semantic Map Guided Diffusion Model for Free-Form Layout-to-Image Generation. *AAAI 2024*.
- [79] Keyao Wang, Guosheng Zhang, Haixiao Yue, Ajian Liu, Gang Zhang, Haocheng Feng, Junyu Han, Errui Ding, and Jingdong Wang. Multi-Domain Incremental Learning for Face Presentation Attack Detection. *AAAI 2024*.
- [80] Mengmeng Wang, Jiazheng Xing, Boyuan Jiang, Jun Chen, Jianbiao Mei, Xingxing Zuo, Guang Dai, Jingdong Wang, and Yong Liu. A Multimodal, Multi-Task Adapting Framework for Video Action Recognition. *AAAI 2024*.
- [81] Junkun Yuan, Xinyu Zhang, Hao Zhou, Jian Wang, Zhongwei Qiu, Zhiyin Shao, Shaofeng Zhang, Sifan Long, Kun Kuang, Kun Yao, Junyu Han, Errui Ding, Lanfen Lin, Fei Wu, and Jingdong Wang. HAP: Structure-Aware Masked Image Modeling for Human-Centric Perception. *NeurIPS 2023*.
- [82] Linyan Huang, Zhiqi Li, Chonghao Sima, Wenhai Wang, Jingdong Wang, Yu Qiao, and Hongyang Li. Leveraging Vision-Centric Multi-Modal Expertise for 3D Object Detection. *NeurIPS 2023*.
- [83] Zhengdong Hu, Yifan Sun, Jingdong Wang, and Yi Yang. DAC-DETR: Divide the Attention Layers and Conquer. *NeurIPS 2023*.
- [84] Kaisiyuan Wang, Hang Zhou, Qianyi Wu, Jiaxiang Tang, Zhiliang Xu, Borong Liang, Tianshu Hu, Errui Ding, Jingtuo Liu, Ziwei Liu, and Jingdong Wang. Efficient Video Portrait Reenactment via Grid-based Codebook. *SIGGRAPH 2023*.
- [85] **Bo Fang**, Wenhao Wu, Chang Liu, Yu Zhou, Yuxin Song, Weiping Wang, Xiangbo Shu, Xiangyang Ji, and Jingdong Wang. UATVR: Uncertainty-Adaptive Text-Video Retrieval. *ICCV 2023*.
- [86] **Lizhao Liu**, Zhuangwei Zhuang, Shangxin Huang, Xunlong Xiao, Tianhang Xiang, Cen Chen, Jingdong Wang, and Mingkui Tan. CPCM: Contextual Point Cloud Modeling for Weakly-supervised Point Cloud Semantic Segmentation. *ICCV 2023*.
- [87] **Jiixiang Tang**, Hang Zhou, Xiaokang Chen, Tianshu Hu, Errui Ding, Jingdong Wang, and Gang Zeng. Delicate Textured Mesh Recovery from NeRF via Adaptive Surface Refinement. *ICCV 2023*.
- [88] **Huan Liu**, Qiang Chen, Zichang Tan, Jiang-Jiang Liu, Jian Wang, Xiangbo Su, Xiaolong Li, Kun Yao, Junyu Han, Errui Ding, Yao Zhao, and Jingdong Wang. Group Pose: A Simple Baseline for End-to-End Multi-person Pose Estimation. *ICCV 2023*.
- [89] Qiang Chen, Xiaokang Chen, Jian Wang, Shan Zhang, Kun Yao, Haocheng Feng, Junyu Han, Errui Ding, Gang Zeng, and Jingdong Wang. Group DETR: Fast DETR Training with Group-Wise One-to-Many Assignment. *ICCV 2023*.
- [90] **Shan Zhang**, **Jinhao Du**, Qiang Chen, Haifeng Le, Yanpeng Sun, Yao Ni, Jian Wang, Bin He, and Jingdong Wang.  $\sigma$ -Adaptive Decoupled Prototype for Few-Shot Object Detection. *ICCV 2023*.
- [91] Wenhao Wu, Yuxin Song, Zhun Sun, Jingdong Wang, Chang Xu, and Wanli Ouyang. What Can Simple Arithmetic Operations Do for Temporal Modeling? *ICCV 2023*.
- [92] **Zhiyin Shao**, Xinyu Zhang, Changxing Ding, Jian Wang, and Jingdong Wang Unified Pre-training with Pseudo Texts for Text-To-Image Person Re-identification. *ICCV 2023*.
- [93] **Jiaming Li**, Xiangru Lin, Wei Zhang, Xiao Tan, Yingying Li, Junyu Han, Errui Ding, Jingdong Wang, and Guanbin Li. Gradient-based Sampling for Class Imbalanced Semi-supervised Object Detection. *ICCV 2023*.

- [94] **Jiazheng Xing**, Mengmeng Wang, Yudi Ruan, Bofan Chen, Yaowei Guo, Boyu Mu, Guang Dai, Jingdong Wang, and Yong Liu. Boosting Few-shot Action Recognition with Graph-guided Hybrid Matching. *ICCV 2023*.
- [95] **Sifan Long, Zhen Zhao, Junkun Yuan**, Zichang Tan, Jiang-Jiang Liu, Shengsheng Wang, Luping Zhou, and Jingdong Wang. Task-Oriented Multi-Modal Mutual Learning for Vision-Language Models. *ICCV 2023*.
- [96] **Xiang Guo**, Jiadai Sun, Yuchao Dai, Guanying Chen, Xiaoqing Ye, Xiao Tan, Errui Ding, Yumeng Zhang, and Jingdong Wang. Forward Flow for Novel View Synthesis of Dynamic Scenes. *ICCV 2023*.
- [97] **Shuo Li**, Yue He, Weiming Zhang, Wei Zhang, Xiao Tan, Junyu Han, Errui Ding, and Jingdong Wang. CFCG: Semi-Supervised Semantic Segmentation via Cross-Fusion and Contour Guidance Supervision. *ICCV 2023*.
- [98] **Zhen Zhao**, Lihe Yang, Sifan Long, Jimin Pi, Luping Zhou, and Jingdong Wang. Augmentation Matters: A Simple-Yet-Effective Approach to Semi-Supervised Semantic Segmentation. *CVPR 2023*.
- [99] **Zhen Zhao**, Sifan Long, Jimin Pi, Jingdong Wang, and Luping Zhou. Instance-Specific and Model-Adaptive Supervision for Semi-Supervised Semantic Segmentation. *CVPR 2023*.
- [100] **Chang Liu**, Weiming Zhang, Xiangru Lin, Wei Zhang, Xiao Tan, Junyu Han, Xiaomao Li, Errui Ding, and Jingdong Wang. Ambiguity-Resistant Semi-Supervised Learning for Dense Object Detection. *CVPR 2023*.
- [101] Wenhao Wu, Xiaohan Wang, Haipeng Luo, Jingdong Wang, Yi Yang, and Wanli Ouyang. Bidirectional Cross-Modal Knowledge Exploration for Video Recognition With Pre-Trained Vision-Language Models. *CVPR 2023*.
- [102] **Jiazhi Guan**, Zhanwang Zhang, Hang Zhou, Tianshu Hu, Kaisiyuan Wang, Dongliang He, Haocheng Feng, Jingtuo Liu, Errui Ding, Ziwei Liu, and Jingdong Wang. StyleSync: High-Fidelity Generalized and Personalized Lip Sync in Style-Based Generator. *CVPR 2023*.
- [103] **Kaixin Xiong**, Shi Gong, Xiaoqing Ye, Xiao Tan, Ji Wan, Errui Ding, Jingdong Wang, and Xiang Bai. CAPE: Camera View Position Embedding for Multi-View 3D Object Detection. *CVPR 2023*.
- [104] **Zhongwei Qiu**, Qiansheng Yang, Jian Wang, Haocheng Feng, Junyu Han, Errui Ding, Chang Xu, Dongmei Fu, and Jingdong Wang. PSVT: End-to-End Multi-Person 3D Pose and Shape Estimation With Progressive Video Transformers. *CVPR 2023*.
- [105] **Sifan Long, Zhen Zhao**, Jimin Pi, Shengsheng Wang, and Jingdong Wang. Beyond Attentive Tokens: Incorporating Token Importance and Diversity for Efficient Vision Transformers. *CVPR 2023*.
- [106] **Jiacheng Zhang**, Xiangru Lin, Wei Zhang, Kuo Wang, Xiao Tan, Junyu Han, Errui Ding, Jingdong Wang, and Guanbin Li. Semi-DETR: Semi-Supervised Object Detection With Detection Transformers. *CVPR 2023*.
- [107] Wenhao Wu, Haipeng Luo, Bo Fang, Jingdong Wang, and Wanli Ouyang. Cap4Video: What Can Auxiliary Captions Do for Text-Video Retrieval? *CVPR 2023*.
- [108] Yuechen Yu, Yulin Li, Chengquan Zhang, Xiaoqiang Zhang, Zengyuan Guo, Xiameng Qin, Kun Yao, Junyu Han, Errui Ding, and Jingdong Wang. StrucTexTv2: Masked Visual-Textual Prediction for Document Image Pre-training. *ICLR 2023*.
- [109] **Xiaohu Huang**, Hao Zhou, Jian Wang, Haocheng Feng, Junyu Han, Errui Ding, Jingdong Wang, Xinggang Wang, Wenyu Liu, and Bin Feng. Graph Contrastive Learning for Skeleton-based Action Recognition. *ICLR 2023*.
- [110] Kaisiyuan Wang, Changcheng Liang, Hang Zhou, Jiayang Tang, Qianyi Wu, Dongliang He, Zhibin Hong, Jingtuo Liu, Errui Ding, Ziwei Liu, and Jingdong Wang. Robust Video Portrait Reenactment via Personalized Representation Quantization. *AAAI 2023*.
- [111] Haixiao Yue, Keyao Wang, Guosheng Zhang, Haocheng Feng, Junyu Han, Errui Ding, and Jingdong Wang. Cyclically Disentangled Feature Translation for Face Anti-spoofing. *AAAI 2023*.

- [112] Pengyuan Lyu, Weihong Ma, Hongyi Wang, Yuechen Yu, Chengquan Zhang, Kun Yao, Yang Xue, and Jingdong Wang. GridFormer: Towards Accurate Table Structure Recognition via Grid Prediction. In *Proceedings of the 2023 ACM Conference on Multimedia Conference, MM 2023*.
- [113] Jian Wang, Chenhui Gou, Qiman Wu, Haocheng Feng, Junyu Han, Errui Ding, and Jingdong Wang. RTFormer: Efficient Design for Real-Time Semantic Segmentation with Transformer. In *Thirty-sixth Conference on Neural Information Processing Systems, NeurIPS 2022*.
- [114] Jiazhi Guan, Hang Zhou, Zhibin Hong, Errui Ding, Jingdong Wang, Chengbin Quan, Youjian Zhao. Delving into Sequential Patches for Deepfake Detection. In *Thirty-sixth Conference on Neural Information Processing Systems, NeurIPS 2022*.
- [115] **Yanpeng Sun**, Qiang Chen, Xiangyu He, Jian Wang, Haocheng Feng, Junyu Han, Errui Ding, Jian Cheng, Zechao Li, Jingdong Wang. Singular Value Fine-tuning: Few-shot Segmentation requires Few-parameters Fine-tuning. In *Thirty-sixth Conference on Neural Information Processing Systems, NeurIPS 2022*.
- [116] **Mingyu Ding**, Bin Xiao, Noel Codella, Ping Luo, Jingdong Wang, Lu Yuan. DaViT: Dual Attention Vision Transformers. *ECCV 2022*.
- [117] **Shi Gong**, Xiaoqing Ye, Xiao Tan, Jingdong Wang, Errui Ding, Yu Zhou, Xiang Bai. GitNet: Geometric Prior-Based Transformation for Birds-Eye-View Segmentation. *ECCV 2022*.
- [118] **Yang Bai**, Desen Zhou, Songyang Zhang, Jian Wang, Errui Ding, Yu Guan, Yang Long, Jingdong Wang. Action Quality Assessment with Temporal Parsing Transformer. *ECCV 2022*.
- [119] Teng Xi, Yifan Sun, Deli Yu, Bi Li, Nan Peng, Gang Zhang, Xinyu Zhang, Zhigang Wang, Jinwen Chen, Jian Wang, Lufei Liu, Haocheng Feng, Junyu Han, Jingtuo Liu, Errui Ding, Jingdong Wang. UFO: Unified Feature Optimization. *ECCV 2022*.
- [120] Linfeng Li, Minyue Jiang, Yue Yu, Wei Zhang, Xiangru Lin, Yingying Li, Xiao Tan, Jingdong Wang, Errui Ding. Diverse Learner: Exploring Diverse Supervision for Semi-supervised Object Detection. *ECCV 2022*.
- [121] Zhiliang Xu, Hang Zhou, Zhibin Hong, Ziwei Liu, Jiaming Liu, Zhizhi Guo, Junyu Han, Jingtuo Liu, Errui Ding, Jingdong Wang. StyleSwap: Style-Based Generator Empowers Robust Face Swapping. *ECCV 2022*.
- [122] **Haoran Wang**, Dongliang He, Wenhao Wu, Boyang Xia, Min Yang, Fu Li, Yunlong Yu, Zhong Ji, Errui Ding, Jingdong Wang. CODER: Coupled Diversity-Sensitive Momentum Contrastive Learning for Image-Text Retrieval. *ECCV 2022*.
- [123] Yasheng Sun, Hang Zhou, Kaisiyuan Wang, Qianyi Wu, Zhibin Hong, Jingtuo Liu, Errui Ding, Jingdong Wang, Ziwei Liu, and Hideki Koike. Masked Lip-Sync Prediction by Audio-Visual Contextual Exploitation in Transformers. *SIGGRAPH ASIA 2022*.
- [124] **Borong Liang**, Yan Pan, Zhizhi Guo, Hang Zhou, Zhibin Hong, Xiaoguang Han, Junyu Han, Jingtuo Liu, Errui Ding, Jingdong Wang. Expressive Talking Head Generation with Granular Audio-Visual Control. *CVPR 2022*.
- [125] **Mengjun Cheng**, Yipeng Sun, Longchao Wang, Xiongwei Zhu, Kun Yao, Jie Chen, Guoli Song, Junyu Han, Jingtuo Liu, Errui Ding, Jingdong Wang. ViSTA: Vision and Scene Text Aggregation for Cross-Modal Retrieval. *CVPR 2022*.
- [126] Qiang Chen, Qiman Wu, Jian Wang, Qinghao Hu, Tao Hu, Errui Ding, Jian Cheng, Jingdong Wang. MixFormer: Mixing Features across Windows and Dimensions. *CVPR 2022*.
- [127] Xinyu Zhang, Dongdong Li, Zhigang Wang, Jian Wang, Errui Ding, Javen Qinfeng Shi, Zhaoxiang Zhang, Jingdong Wang. Implicit Sample Extension for Unsupervised Person Re-Identification. *CVPR 2022*.
- [128] Licheng Tang, Yiyang Cai, Jiaming Liu, Zhibin Hong, Mingming Gong, Minhu Fan, Junyu Han, Jingtuo Liu, Errui Ding, Jingdong Wang. Few-Shot Font Generation by Learning Fine-Grained Local Styles. *CVPR 2022*.

- [129] Changyong Shu, Hemao Wu, Hang Zhou, Jiaming Liu, Zhibin Hong, Changxing Ding, Junyu Han, Jingtuo Liu, Errui Ding, Jingdong Wang. Few-Shot Head Swapping in the Wild. *CVPR 2022*.
- [130] Desen Zhou, Zhichao Liu, Jian Wang, Leshan Wang, Tao Hu, Errui Ding, Jingdong Wang. Human-Object Interaction Detection via Disentangled Transformer. *CVPR 2022*.
- [131] **Mingyu Ding**, Yuqi Huo, Haoyu Lu, Linjie Yang, Zhe Wang, Zhiwu Lu, Jingdong Wang, and Ping Luo. Learning Versatile Neural Architectures by Propagating Network Codes. *ICLR 2022*.
- [132] **Qi Han, Zejia Fan**, Qi Dai, Lei Sun, Ming-Ming Cheng, Jiaying Liu, and Jingdong Wang. On the Connection between Local Attention and Dynamic Depth-wise Convolution. *ICLR 2022*.
- [133] Bo Ju, Zhikang Zou, Xiaoqing Ye, Minyue Jiang, Xiao Tan, Errui Ding, Jingdong Wang. Paint and Distill: Boosting 3D Object Detection with Semantic Passing Network. In *Proceedings of the 2022 ACM Conference on Multimedia Conference, MM 2022*.
- [134] Dongdong Li, Zhigang Wang, Jian Wang, Xinyu Zhang, Errui Ding, Jingdong Wang, and Zhaoxiang Zhang. Self-Guided Hard Negative Generation for Unsupervised Person Re-Identification. *IJCAI 2022*.
- [135] Qi Chen, Bing Zhao, Haidong Wang, Mingqin Li, Chuanjie Liu, Zengzhong Li, Mao Yang, and Jingdong Wang. SPANN: Highly-efficient Billion-scale Approximate Nearest Neighbor Search. In *Thirty-fifth Conference on Neural Information Processing Systems, NeurIPS 2021*.
- [136] Yuhui Yuan, Rao Fu, Lang Huang, Weihong Lin, Chao Zhang, Xilin Chen, and Jingdong Wang. HRFormer: High-Resolution Transformer for Dense Prediction. In *Thirty-fifth Conference on Neural Information Processing Systems, NeurIPS 2021*.
- [137] Jianming Ye, Shiliang Zhang, and Jingdong Wang. Hybrid Network Compression via Meta-Learning. In *Proceedings of the 2021 ACM Conference on Multimedia Conference, MM 2021*.
- [138] **Depu Meng, Xiaokang Chen, Zejia Fan**, Gang Zeng, Houqiang Li, Yuhui Yuan, Lei Sun, and Jingdong Wang. Conditional DETR for Fast Training Convergence. In *IEEE/CVF International Conference on Computer Vision, ICCV 2021*.
- [139] Xiaosen Wang, Xuanran He, Jingdong Wang, and Kun He. Admix: Enhancing the Transferability of Adversarial Attacks. In *IEEE/CVF International Conference on Computer Vision, ICCV 2021*.
- [140] **Changqian Yu**, Bin Xiao, Changxin Gao, Lu Yuan, Lei Zhang, Nong Sang, and Jingdong Wang. Lite-HRNet: A Lightweight High-Resolution Network. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2021*.
- [141] **Zigang Geng**, Ke Sun, Bin Xiao, Zhaoxiang Zhang, and Jingdong Wang. Bottom-Up Human Pose Estimation Via Disentangled Keypoint Regression. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2021*.
- [142] **Xiaokang Chen**, Yuhui Yuan, Gang Zeng, and Jingdong Wang. Semi-Supervised Semantic Segmentation with Cross Pseudo Supervision. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2021*.
- [143] **Baifeng Shi**, Dinghuai Zhang, Qi Dai, Zhanxing Zhu, Yadong Mu, and Jingdong Wang. Informative Dropout for Robust Representation Learning: A Shape-bias Perspective. In *Proceedings of the 37th International Conference on Machine Learning, ICML 2020*.
- [144] Yuhui Yuan, Xilin Chen, and Jingdong Wang. Object-Contextual Representations for Semantic Segmentation. In *European Conference Computer Vision, ECCV 2020*.
- [145] Yuhui Yuan, Jingyi Xie, Xilin Chen, and Jingdong Wang. SegFix: Model-Agnostic Boundary Refinement for Segmentation. In *European Conference Computer Vision, ECCV 2020*.

- [146] **Yifan liu**, Chunhua Shen, Changqian Yu, and Jingdong Wang Efficient Semantic Video Segmentation with Per-frame Inference. In *European Conference Computer Vision*, ECCV 2020.
- [147] Fangyun Wei, Xiao Sun, Hongyang Li, Jingdong Wang, and Stephen Lin. Point-Set Anchors for Object Detection, Instance Segmentation and Pose Estimation. In *European Conference Computer Vision*, ECCV 2020.
- [148] **Bowen Cheng**, Bin Xiao, Jingdong Wang, Honghui Shi, Thomas S. Huang, and Lei Zhang. HigherHRNet: Scale-Aware Representation Learning for Bottom-Up Human Pose Estimation. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, CVPR 2020.
- [149] **Baifeng Shi**, Qi Dai, Yadong Mu, and Jingdong Wang. Weakly-Supervised Action Localization by Generative Attention Modeling. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, CVPR 2020.
- [150] Yong Guo, Jian Chen, Jingdong Wang, Qi Chen, Jiezhong Cao, Zeshuai Deng, Yanwu Xu, and Minghui Tan. Closed-Loop Matters: Dual Regression Networks for Single Image Super-Resolution. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, CVPR 2020.
- [151] **Jianing Li**, Jingdong Wang, Qi Tian, Wen Gao, and Shiliang Zhang Global-Local Temporal Representations For Video Person Re-Identification. In *IEEE/CVF International Conference on Computer Vision*, ICCV 2019.
- [152] **Haibo Qiu**, Chunyu Wang, Jingdong Wang, Naiyan Wang, and Wenjun Zeng Cross View Fusion for 3D Human Pose Estimation. In *IEEE/CVF International Conference on Computer Vision*, ICCV 2019.
- [153] **Ke Sun**, Bin Xiao, Dong Liu, and Jingdong Wang. Deep High-Resolution Representation Learning for Human Pose Estimation. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, CVPR 2019.
- [154] **Yifan Liu**, Ke Chen, Chris Liu, Zengchang Qin, Zhenbo Luo, and Jingdong Wang. Structured Knowledge Distillation for Semantic Segmentation. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, CVPR 2019.
- [155] Shengze Yu, Xin Wang, Wenwu Zhu, Peng Cui, and Jingdong Wang. Disparity-preserved Deep Cross-platform Association for Cross-platform Video Recommendation. In *IJCAI* 2019.
- [156] Ruochen Fan, Ming-Ming Cheng, Qibin Hou, Tai-Jiang Mu, Jingdong Wang, Shi-Min Hu. S4Net: Single Stage Salient-Instance Segmentation. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, CVPR 2019.
- [157] **Xuguang Duan**, Wenbing Huang, Chuang Gan, Jingdong Wang, Wenwu Zhu, and Junzhou Huang Weakly Supervised Dense Event Captioning in Videos. In *Proceedings of Advances in Neural Information Processing Systems*, NIPS 2018.
- [158] Bin Liu, Yue Cao, Mingsheng Long, Jianmin Wang, and Jingdong Wang. Deep Triplet Quantization. In *Proceedings of the 2018 ACM Conference on Multimedia Conference*, MM 2018.
- [159] Hao Xiao, Weiyao Lin, Bin Sheng, Ke Lu, Junchi Yan, Jingdong Wang, Errui Ding, Yihao Zhang, Hongkai Xiong. Group Re-Identification: Leveraging and Integrating Multi-Grain Information. In *Proceedings of the 2018 ACM Conference on Multimedia Conference*, MM 2018.
- [160] **Yumin Suh**, Jingdong Wang, Siyu Tang, Tao Mei, and Kyoung Mu Lee. Part-Aligned Bilinear Representations for Person Re-identification. In *European Conference Computer Vision*, ECCV 2018.
- [161] **Liming Zhao**, **Mingjie Li**, **Depu Meng**, Xi Li, Zhuowen Tu, Zhaoxiang Zhang, Yueting Zhuang, and Jingdong Wang. Deep Convolutional Neural Networks with Merge-and-Run Mappings. In *Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence*, IJCAI 2018.
- [162] **Guotian Xie**, Jingdong Wang, Ting Zhang, Jianhuang Lai, Richang Hong, and Guo-Jun Qi. IGCv2: Interleaved Structured Sparse Convolutional Neural Networks. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, CVPR 2018.

- [163] **Zilong Huang**, Xinggang Wang, Jiasi Wang, Wenyu Liu, and Jingdong Wang. Weakly-Supervised Semantic Segmentation Network with Deep Seeded Region Growing. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2018*.
- [164] Guo-Jun Qi, Liheng Zhang, Hao Hu, Marzieh Edraki, Jingdong Wang, and Xian-Sheng Hua. Global versus Localized Generative Adversarial Nets. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2018*.
- [165] **Guotian Xie**, Ting Zhang, Kuiyuan Yang, Jianhuang Lai, and Jingdong Wang. Coupled Convolutions for CNNs. In *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence, New Orleans, Louisiana, USA, February 2-7, 2018*, 2018.
- [166] **Ting Zhang**, Guo-Jun Qi, Bin Xiao, and Jingdong Wang. Interleaved Group Convolutions. In *IEEE International Conference on Computer Vision, ICCV 2017, Venice, Italy, October 22-29, 2017*, pages 4383-4392, 2017.
- [167] **Liming Zhao**, Xi Li, Yueting Zhuang, and Jingdong Wang. Deeply-Learned Part-Aligned Representations for Person Re-Identification. In *IEEE International Conference on Computer Vision, ICCV 2017, Venice, Italy, October 22-29, 2017*, pages 3239-3248, 2017.
- [168] **Ke Sun**, Cuiling Lan, Junliang Xing, Jingdong Wang, Wenjun Zeng, Dong Liu. Human Pose Estimation using Global and Local Normalization. In *IEEE International Conference on Computer Vision, ICCV 2017, Venice, Italy, October 22-29, 2017*, pages 5600-5608, 2017.
- [169] **Song Bai**, Zhichao Zhou, Jingdong Wang, Xiang Bai, Longin Jan Latecki, Qi Tian. Ensemble Diffusion for Retrieval. In *IEEE International Conference on Computer Vision, ICCV 2017, Venice, Italy, October 22-29, 2017*, pages 774-783, 2017.
- [170] **Lixin Liao**, Yao Zhao, Shikui Wei, and Jingdong Wang. Finding the Secret of CNN Parameter Layout under Strict Size Constraint. In *Proceedings of the 2017 ACM on Multimedia Conference, MM 2017, Mountain View, CA, USA, October 23-27, 2017*, pages 997-1005, 2017.
- [171] Guo-Jun Qi, Jiliang Tang, Jingdong Wang, and Jiebo Luo. Mixture Factorized Ornstein-Uhlenbeck Processes for Time-Series Forecasting. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Halifax, NS, Canada, August 13 - 17, 2017*, pages 987-995, 2017.
- [172] **Gangming Zhao**, Jingdong Wang, and Zhaoxiang Zhang. Random Shifting for CNN: a Solution to Reduce Information Loss in Down-Sampling Layers. In *Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence, IJCAI 2017, Melbourne, Australia, August 19-25, 2017*, pages 3476-3482, 2017.
- [173] **Lingxi Xie**, **Liang Zheng**, Jingdong Wang, Alan L. Yuille, and Qi Tian. Interactive: Inter-layer activeness propagation. In *2016 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2016, Las Vegas, NV, USA, June 27-30, 2016*, pages 270-279, 2016.
- [174] **Xiaojuan Wang**, **Ting Zhang**, Guo-Jun Qi, Jinhui Tang, and Jingdong Wang. Supervised quantization for similarity search. In *2016 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2016, Las Vegas, NV, USA, June 27-30, 2016*, pages 2018-2026, 2016.
- [175] **Ting Zhang** and Jingdong Wang. Collaborative quantization for cross-modal similarity search. In *2016 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2016, Las Vegas, NV, USA, June 27-30, 2016*, pages 2036-2045, 2016.
- [176] **Lingxi Xie**, Jingdong Wang, **Zhen Wei**, Meng Wang, and Qi Tian. Disturblabel: Regularizing CNN on the loss layer. In *2016 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2016, Las Vegas, NV, USA, June 27-30, 2016*, pages 4753-4762, 2016.
- [177] Lingxi Xie, Qi Tian, John Flynn, Jingdong Wang, and Alan L. Yuille. Geometric neural phrase pooling: Modeling the spatial co-occurrence of neurons. In *Computer Vision - ECCV 2016 - 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part I*, pages 645-661, 2016.

- [178] **Liang Zheng**, Zhi Bie, Yifan Sun, Jingdong Wang, Chi Su, Shengjin Wang, and Qi Tian. MARS: A video benchmark for large-scale person re-identification. In *Computer Vision - ECCV 2016 - 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part VI*, pages 868–884, 2016.
- [179] **Qi Dai**, Jianguo Li, Jingdong Wang, and Yu-Gang Jiang. Binary optimized hashing. In *Proceedings of the 2016 ACM Conference on Multimedia Conference, MM 2016, Amsterdam, The Netherlands, October 15-19, 2016*, pages 1247–1256, 2016.
- [180] Jian Liang, Zhihang Li, Dong Cao, Ran He, and Jingdong Wang. Self-paced cross-modal subspace matching. In *Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval, SIGIR 2016, Pisa, Italy, July 17-21, 2016*, pages 569–578, 2016.
- [181] **Dapeng Chen**, Zejian Yuan, Gang Hua, Nanning Zheng, and Jingdong Wang. Similarity learning on an explicit polynomial kernel feature map for person re-identification. In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2015, Boston, MA, USA, June 7-12, 2015*, pages 1565–1573, 2015.
- [182] **Dingwen Zhang**, Junwei Han, Chao Li, and Jingdong Wang. Co-saliency detection via looking deep and wide. In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2015, Boston, MA, USA, June 7-12, 2015*, pages 2994–3002, 2015.
- [183] **Ting Zhang**, Guo-Jun Qi, Jinhui Tang, and Jingdong Wang. Sparse composite quantization. In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2015, Boston, MA, USA, June 7-12, 2015*, pages 4548–4556, 2015.
- [184] **Lingxi Xie**, Jingdong Wang, Weiyao Lin, Bo Zhang, and Qi Tian. RIDE: reversal invariant descriptor enhancement. In *2015 IEEE International Conference on Computer Vision, ICCV 2015, Santiago, Chile, December 7-13, 2015*, pages 100–108, 2015.
- [185] **Liang Zheng**, Liyue Shen, Lu Tian, Shengjin Wang, Jingdong Wang, and Qi Tian. Scalable person re-identification: A benchmark. In *2015 IEEE International Conference on Computer Vision, ICCV 2015, Santiago, Chile, December 7-13, 2015*, pages 1116–1124, 2015.
- [186] **Yang Shen**, Weiyao Lin, Junchi Yan, Mingliang Xu, Jianxin Wu, and Jingdong Wang. Person re-identification with correspondence structure learning. In *2015 IEEE International Conference on Computer Vision, ICCV 2015, Santiago, Chile, December 7-13, 2015*, pages 3200–3208, 2015.
- [187] **Botong Wu**, Qiang Yang, Wei-Shi Zheng, Yizhou Wang, and Jingdong Wang. Quantized correlation hashing for fast cross-modal search. In *Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence, IJCAI 2015, Buenos Aires, Argentina, July 25-31, 2015*, pages 3946–3952, 2015.
- [188] **Xiangbo Shu**, Guo-Jun Qi, Jinhui Tang, and Jingdong Wang. Weakly-shared deep transfer networks for heterogeneous-domain knowledge propagation. In *Proceedings of the 23rd Annual ACM Conference on Multimedia Conference, MM '15, Brisbane, Australia, October 26 - 30, 2015*, pages 35–44, 2015.
- [189] **Lingxi Xie**, Jingdong Wang, Baining Guo, Bo Zhang, and Qi Tian. Orientational pyramid matching for recognizing indoor scenes. In *2014 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2014, Columbus, OH, USA, June 23-28, 2014*, pages 3734–3741, 2014.
- [190] **Weiyue Wang**, Weiyao Lin, Yuanzhe Chen, Jianxin Wu, Jingdong Wang, and Bin Sheng. Finding coherent motions and semantic regions in crowd scenes: A diffusion and clustering approach. In *Computer Vision - ECCV 2014 - 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part I*, pages 756–771, 2014.
- [191] **Ting Zhang**, **Chao Du**, and Jingdong Wang. Composite quantization for approximate nearest neighbor search. In *Proceedings of the 31th International Conference on Machine Learning, ICML 2014, Beijing, China, 21-26 June 2014*, pages 838–846, 2014.

- [192] **Jianfeng Wang**, Heng Tao Shen, Shuicheng Yan, Nenghai Yu, Shipeng Li, and Jingdong Wang. Optimized distances for binary code ranking. In *Proceedings of the ACM International Conference on Multimedia, MM '14, Orlando, FL, USA, November 03 - 07, 2014*, pages 517–526, 2014.
- [193] **Huaizu Jiang**, Jingdong Wang, Zejian Yuan, Yang Wu, Nanning Zheng, and Shipeng Li. Salient object detection: A discriminative regional feature integration approach. In *2013 IEEE Conference on Computer Vision and Pattern Recognition, Portland, OR, USA, June 23-28, 2013*, pages 2083–2090, 2013.
- [194] **Peng Wang**, Jingdong Wang, Gang Zeng, Weiwei Xu, Hongbin Zha, and Shipeng Li. Supervised kernel descriptors for visual recognition. In *2013 IEEE Conference on Computer Vision and Pattern Recognition, Portland, OR, USA, June 23-28, 2013*, pages 2858–2865, 2013.
- [195] **Naiyan Wang**, Jingdong Wang, and Dit-Yan Yeung. Online robust non-negative dictionary learning for visual tracking. In *IEEE International Conference on Computer Vision, ICCV 2013, Sydney, Australia, December 1-8, 2013*, pages 657–664, 2013.
- [196] **Jing Wang**, Jingdong Wang, Gang Zeng, Rui Gan, Shipeng Li, and Baining Guo. Fast neighborhood graph search using Cartesian concatenation. In *IEEE International Conference on Computer Vision, ICCV 2013, Sydney, Australia, December 1-8, 2013*, pages 2128–2135, 2013.
- [197] **Honghui Zhang**, Jingdong Wang, Ping Tan, Jinglu Wang, and Long Quan. Learning crfs for image parsing with adaptive subgradient descent. In *IEEE International Conference on Computer Vision, ICCV 2013, Sydney, Australia, December 1-8, 2013*, pages 3080–3087, 2013.
- [198] **Quannan Li**, Jingdong Wang, David P. Wipf, and Zhuowen Tu. Fixed-point model for structured labeling. In *Proceedings of the 30th International Conference on Machine Learning, ICML 2013, Atlanta, GA, USA, 16-21 June 2013*, pages 214–221, 2013.
- [199] **Jianfeng Wang**, Jingdong Wang, Nenghai Yu, and Shipeng Li. Order preserving hashing for approximate nearest neighbor search. In *ACM Multimedia Conference, MM '13, Barcelona, Spain, October 21-25, 2013*, pages 133–142, 2013.
- [200] Xian-Sheng Hua, Linjun Yang, Jingdong Wang, **Jing Wang**, Ming Ye, Kuansan Wang, Yong Rui, and Jin Li. Clickage: towards bridging semantic and intent gaps via mining click logs of search engines. In *ACM Multimedia Conference, MM '13, Barcelona, Spain, October 21-25, 2013*, pages 243–252, 2013.
- [201] **Jing Wang**, Jingdong Wang, Gang Zeng, Zhuowen Tu, Rui Gan, and Shipeng Li. Scalable k-nn graph construction for visual descriptors. In *2012 IEEE Conference on Computer Vision and Pattern Recognition, Providence, RI, USA, June 16-21, 2012*, pages 1106–1113, 2012.
- [202] **Junyang Lu, Jiazhen Zhou**, Jingdong Wang, Tao Mei, Xian-Sheng Hua, and Shipeng Li. Image search results refinement via outlier detection using deep contexts. In *2012 IEEE Conference on Computer Vision and Pattern Recognition, Providence, RI, USA, June 16-21, 2012*, pages 3029–3036, 2012.
- [203] **Jing Wang**, Jingdong Wang, Qifa Ke, Gang Zeng, and Shipeng Li. Fast approximate k-means via cluster closures. In *2012 IEEE Conference on Computer Vision and Pattern Recognition, Providence, RI, USA, June 16-21, 2012*, pages 3037–3044, 2012.
- [204] **Peng Wang**, Jingdong Wang, Gang Zeng, Jie Feng, Hongbin Zha, and Shipeng Li. Salient object detection for searched web images via global saliency. In *2012 IEEE Conference on Computer Vision and Pattern Recognition, Providence, RI, USA, June 16-21, 2012*, pages 3194–3201, 2012.
- [205] **Naiyan Wang, Tiansheng Yao**, Jingdong Wang, and Dit-Yan Yeung. A probabilistic approach to robust matrix factorization. In *Computer Vision - ECCV 2012 - 12th European Conference on Computer Vision, Florence, Italy, October 7-13, 2012, Proceedings, Part VII*, pages 126–139, 2012.
- [206] Jingdong Wang and Shipeng Li. Query-driven iterated neighborhood graph search for large scale indexing. In *Proceedings of the 20th ACM Multimedia Conference, MM '12, Nara, Japan, October 29 - November 02, 2012*, pages 179–188, 2012.

- [207] **Jianping Shi, Xiang Ren**, Guang Dai, Jingdong Wang, and Zhihua Zhang. A non-convex relaxation approach to sparse dictionary learning. In *The 24th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2011, Colorado Springs, CO, USA, 20-25 June 2011*, pages 1809–1816, 2011.
- [208] Gang Zeng, **Peng Wang**, Jingdong Wang, Rui Gan, and Hongbin Zha. Structure-sensitive superpixels via geodesic distance. In *IEEE International Conference on Computer Vision, ICCV 2011, Barcelona, Spain, November 6-13, 2011*, pages 447–454, 2011.
- [209] **Hao Xu**, Jingdong Wang, Zhu Li, Gang Zeng, Shipeng Li, and Nenghai Yu. Complementary hashing for approximate nearest neighbor search. In *IEEE International Conference on Computer Vision, ICCV 2011, Barcelona, Spain, November 6-13, 2011*, pages 1631–1638, 2011.
- [210] Bin Cheng, Guangcan Liu, Jingdong Wang, ZhongYang Huang, and Shuicheng Yan. Multi-task low-rank affinity pursuit for image segmentation. In *IEEE International Conference on Computer Vision, ICCV 2011, Barcelona, Spain, November 6-13, 2011*, pages 2439–2446, 2011.
- [211] **Yang Wang**, Tao Mei, Jingdong Wang, Houqiang Li, and Shipeng Li. JIGSAW: interactive mobile visual search with multimodal queries. In *Proceedings of the 19th International Conference on Multimedia 2011, Scottsdale, AZ, USA, November 28 - December 1, 2011*, pages 73–82, 2011.
- [212] **Wenhao Lu**, Jingdong Wang, Xian-Sheng Hua, Shengjin Wang, and Shipeng Li. Contextual image search. In *Proceedings of the 19th International Conference on Multimedia 2011, Scottsdale, AZ, USA, November 28 - December 1, 2011*, pages 513–522, 2011.
- [213] **Nobuyuki Morioka** and Jingdong Wang. Robust visual reranking via sparsity and ranking constraints. In *Proceedings of the 19th International Conference on Multimedia 2011, Scottsdale, AZ, USA, November 28 - December 1, 2011*, pages 533–542, 2011.
- [214] **Dan Zhang**, Jingdong Wang, and Luo Si. Document clustering with universum. In *Proceeding of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR 2011, Beijing, China, July 25-29, 2011*, pages 873–882, 2011.
- [215] **You Jia**, Jingdong Wang, Gang Zeng, Hongbin Zha, and Xian-Sheng Hua. Optimizing kd-trees for scalable visual descriptor indexing. In *The Twenty-Third IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2010, San Francisco, CA, USA, 13-18 June 2010*, pages 3392–3399, 2010.
- [216] **Hao Xu**, Jingdong Wang, Xian-Sheng Hua, and Shipeng Li. Image search by concept map. In *Proceeding of the 33rd International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR 2010, Geneva, Switzerland, July 19-23, 2010*, pages 275–282, 2010.
- [217] Jingdong Wang, **Yangqing Jia**, Xian-Sheng Hua, Changshui Zhang, and Long Quan. Normalized tree partitioning for image segmentation. In *2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2008), 24-26 June 2008, Anchorage, Alaska, USA, 2008*.
- [218] **Zheng-Jun Zha**, Xian-Sheng Hua, Tao Mei, Jingdong Wang, Guo-Jun Qi, and Zengfu Wang. Joint multi-label multi-instance learning for image classification. In *2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2008), 24-26 June 2008, Anchorage, Alaska, USA, 2008*.
- [219] **Xinmei Tian**, Linjun Yang, Jingdong Wang, Yichen Yang, Xiuqing Wu, and Xian-Sheng Hua. Bayesian video search reranking. In *Proceedings of the 16th International Conference on Multimedia 2008, Vancouver, British Columbia, Canada, October 26-31, 2008*, pages 131–140, 2008.
- [220] Jianxiong Xiao, Jingdong Wang, Ping Tan, and Long Quan. Joint affinity propagation for multiple view segmentation. In *IEEE 11th International Conference on Computer Vision, ICCV 2007, Rio de Janeiro, Brazil, October 14-20, 2007*, pages 1–7, 2007.
- [221] Fei Wang, Jingdong Wang, Helen C. Shen, and Changshui Zhang. Semi-supervised classification using linear neighborhood propagation. In *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2006), 17-22 June 2006, New York, NY, USA, 2006*, pages 160–167, 2006.

- [222] Jingdong Wang, Long Quan, Jian Sun, Xiaoou Tang, and Heung-Yeung Shum. Picture collage. In *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2006)*, 17-22 June 2006, New York, NY, USA, pages 347–354, 2006.
- [223] Jianguo Lee, Jingdong Wang, Changshui Zhang, and Zhaoqi Bian. Probabilistic tangent subspace: a unified view. In *Machine Learning, Proceedings of the Twenty-first International Conference (ICML 2004)*, Banff, Alberta, Canada, July 4-8, 2004.

## Refereed conferences

- [1] Xiaosen Wang, Jiadong Lin, Han Hu, Jingdong Wang, and Kun He. Boosting Adversarial Transferability through Enhanced Momentum. In *British Machine Vision Conference, BMVC 2021*, 2021.
- [2] **Ke Sun, Mingjie Li**, Dong Liu, and Jingdong Wang. IGCv3: Interleaved Low-Rank Group Convolutions for Efficient Deep Neural Network. In *British Machine Vision Conference, BMVC 2018*, 2018.
- [3] **Liang Zheng**, Yali Zhao, Shengjin Wang, Jingdong Wang, and Qi Tian. On the Large-Scale Transferability of the Convolutional Neural Network. In Workshop on Big Data Analytics for Social Computing (BDASC’18), Melbourne, Australia, June 3-6, 2018.
- [4] Yuhui Yuan, Kuiyuan Yang, Jianyuan Guo, Chao Zhang, and Jingdong Wang. Feature Incay for Representation Regularization. In ICLR, 2018.
- [5] **Gangming Zhao**, Zhaoxiang Zhang, He Guan, Peng Tang, and Jingdong Wang. Rethinking ReLU to Train Better CNNs. In ICPR, 2018.
- [6] **Zhansheng Jiang, Lingxi Xie**, Xiaotie Deng, Weiwei Xu, and Jingdong Wang. Fast nearest neighbor search in the hamming space. In *MultiMedia Modeling - 22nd International Conference, MMM 2016, Miami, FL, USA, January 4-6, 2016, Proceedings, Part I*, pages 325–336, 2016.
- [7] Mengyin Wang, Zechao Li, Xiangbo Shu, Jingdong Wang, and Jinhui Tang. Deep kinship verification. In *17th IEEE International Workshop on Multimedia Signal Processing, MMSP 2015, Xiamen, China, October 19-21, 2015*, pages 1–6, 2015.
- [8] Shengnan Cai, Jingdong Wang, and Long Quan. How fashion talks: Clothing-region-based gender recognition. In *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications - 19th Iberoamerican Congress, CIARP 2014, Puerto Vallarta, Mexico, November 2-5, 2014. Proceedings*, pages 515–523, 2014.
- [9] **Na Qi**, Yunhui Shi, Xiaoyan Sun, Jingdong Wang, and Wenpeng Ding. Two dimensional analysis sparse model. In *IEEE International Conference on Image Processing, ICIP 2013, Melbourne, Australia, September 15-18, 2013*, pages 310–314, 2013.
- [10] **Na Qi**, Yunhui Shi, Xiaoyan Sun, Jingdong Wang, and Baocai Yin. Two dimensional synthesis sparse model. In *Proceedings of the 2013 IEEE International Conference on Multimedia and Expo, ICME 2013, San Jose, CA, USA, July 15-19, 2013*, pages 1–6, 2013.
- [11] Yingwei Pan, Ting Yao, Kuiyuan Yang, Houqiang Li, Chong-Wah Ngo, Jingdong Wang, and Tao Mei. Image search by graph-based label propagation with image representation from DNN. In *ACM Multimedia Conference, MM ’13, Barcelona, Spain, October 21-25, 2013*, pages 397–400, 2013.
- [12] **Peng Wang, Dongqing Zhang**, Gang Zeng, and Jingdong Wang. Contextual dominant color name extraction for web image search. In *2012 IEEE International Conference on Multimedia and Expo Workshops, Melbourne, Australia, July 9-13, 2012*, pages 319–324, 2012.
- [13] **Weiwen Tu**, Rong Pan, and Jingdong Wang. Similar image search with a tiny bag-of-delegates representation. In *Proceedings of the 20th ACM Multimedia Conference, MM ’12, Nara, Japan, October 29 - November 02, 2012*, pages 885–888, 2012.

- [14] **Bisheng Chen**, Jingdong Wang, Qinghua Huang, and Tao Mei. Personalized video recommendation through tripartite graph propagation. In *Proceedings of the 20th ACM Multimedia Conference, MM '12, Nara, Japan, October 29 - November 02, 2012*, pages 1133–1136, 2012.
- [15] **Shiyang Lu**, Tao Mei, Jingdong Wang, Jian Zhang, Zhiyong Wang, David Dagan Feng, Jian-Tao Sun, and Shipeng Li. Browse-to-search. In *Proceedings of the 20th ACM Multimedia Conference, MM '12, Nara, Japan, October 29 - November 02, 2012*, pages 1323–1324, 2012.
- [16] **Jing Wang**, Jingdong Wang, Xian-Sheng Hua, and Shipeng Li. Scalable similar image search by joint indices. In *Proceedings of the 20th ACM Multimedia Conference, MM '12, Nara, Japan, October 29 - November 02, 2012*, pages 1325–1326, 2012.
- [17] **Peng Wang, Dongqing Zhang**, Jingdong Wang, Zhong Wu, Xian-Sheng Hua, and Shipeng Li. Color filter for image search. In *Proceedings of the 20th ACM Multimedia Conference, MM '12, Nara, Japan, October 29 - November 02, 2012*, pages 1327–1328, 2012.
- [18] **Huaizu Jiang**, Jingdong Wang, Zejian Yuan, Tie Liu, and Nanning Zheng. Automatic salient object segmentation based on context and shape prior. In *British Machine Vision Conference, BMVC 2011, Dundee, UK, August 29 - September 2, 2011. Proceedings*, pages 1–12, 2011.
- [19] Jingdong Wang and Xian-Sheng Hua. Web-scale image search by color sketch. In *Proceedings of the 19th International Conference on Multimedia 2011, Scottsdale, AZ, USA, November 28 - December 1, 2011*, pages 751–752, 2011.
- [20] **Hao Xu**, Jingdong Wang, Xian-Sheng Hua, and Shipeng Li. Hybrid image summarization. In *Proceedings of the 19th International Conference on Multimedia 2011, Scottsdale, AZ, USA, November 28 - December 1, 2011*, pages 1217–1220, 2011.
- [21] **Xiangang Cheng**, Jingdong Wang, Liang-Tien Chia, and Xian-Sheng Hua. Learning to combine multi-resolution spatially-weighted co-occurrence matrices for image representation. In *Proceedings of the 2010 IEEE International Conference on Multimedia and Expo, ICME 2010, 19-23 July 2010, Singapore*, pages 631–636, 2010.
- [22] Yan Wang, Tao Mei, Jingdong Wang, and Xian-Sheng Hua. Dynamic video collage. In *Advances in Multimedia Modeling, 16th International Multimedia Modeling Conference, MMM 2010, Chongqing, China, January 6-8, 2010. Proceedings*, pages 793–795, 2010.
- [23] **Hao Xu**, Jingdong Wang, Xian-Sheng Hua, and Shipeng Li. Interactive image search by 2d semantic map. In *Proceedings of the 19th International Conference on World Wide Web, WWW 2010, Raleigh, North Carolina, USA, April 26-30, 2010*, pages 1321–1324, 2010.
- [24] **Hao Xu**, Jingdong Wang, Xian-Sheng Hua, and Shipeng Li. Summarizing tagged image collections by cross-media representativeness voting. In *Proceedings of the 2009 IEEE International Conference on Multimedia and Expo, ICME 2009, June 28 - July 2, 2009, New York City, NY, USA*, pages 922–925, 2009.
- [25] **Hao Xu**, Jingdong Wang, Xian-Sheng Hua, and Shipeng Li. Tag refinement by regularized LDA. In *Proceedings of the 17th International Conference on Multimedia 2009, Vancouver, British Columbia, Canada, October 19-24, 2009*, pages 573–576, 2009.
- [26] **Yang Hu**, Jingdong Wang, Nenghai Yu, and Xian-Sheng Hua. Maximum margin clustering with pairwise constraints. In *Proceedings of the 8th IEEE International Conference on Data Mining (ICDM 2008), December 15-19, 2008, Pisa, Italy*, pages 253–262, 2008.
- [27] **Yangqing Jia**, Jingdong Wang, Changshui Zhang, and Xian-Sheng Hua. Augmented tree partitioning for interactive image segmentation. In *Proceedings of the International Conference on Image Processing, ICIP 2008, October 12-15, 2008, San Diego, California, USA*, pages 2292–2295, 2008.

- [28] Jingdong Wang, **Xinmei Tian**, Linjun Yang, Zheng-Jun Zha, and Xian-Sheng Hua. Optimized video scene segmentation. In *Proceedings of the 2008 IEEE International Conference on Multimedia and Expo, ICME 2008, June 23-26 2008, Hannover, Germany*, pages 301–304, 2008.
- [29] **Zheng-Jun Zha**, Tao Mei, Jingdong Wang, Zengfu Wang, and Xian-Sheng Hua. Graph-based semi-supervised learning with multi-label. In *Proceedings of the 2008 IEEE International Conference on Multimedia and Expo, ICME 2008, June 23-26 2008, Hannover, Germany*, pages 1321–1324, 2008.
- [30] **Xinmei Tian**, Linjun Yang, Jingdong Wang, Xiuqing Wu, and Xian-Sheng Hua. Transductive video annotation via local learnable kernel classifier. In *Proceedings of the 2008 IEEE International Conference on Multimedia and Expo, ICME 2008, June 23-26 2008, Hannover, Germany*, pages 1509–1512, 2008.
- [31] Jingdong Wang, **Yinghai Zhao**, Xiuqing Wu, and Xian-Sheng Hua. Transductive multi-label learning for video concept detection. In *Proceedings of the 1st ACM SIGMM International Conference on Multimedia Information Retrieval, MIR 2008, Vancouver, British Columbia, Canada, October 30-31, 2008*, pages 298–304, 2008.
- [32] **Yangqing Jia**, Jingdong Wang, Changshui Zhang, and Xian-Sheng Hua. Finding image exemplars using fast sparse affinity propagation. In *Proceedings of the 16th International Conference on Multimedia 2008, Vancouver, British Columbia, Canada, October 26-31, 2008*, pages 639–642, 2008.
- [33] **Dan Zhang**, Jingdong Wang, Fei Wang, and Changshui Zhang. Semi-supervised classification with universum. In *Proceedings of the SIAM International Conference on Data Mining, SDM 2008, April 24-26, 2008, Atlanta, Georgia, USA*, pages 323–333, 2008.
- [34] Tao Mei, Zheng-Jun Zha, Yuan Liu, Meng Wang, Guo-Jun Qi, Xinmei Tian, Jingdong Wang, Linjun Yang, and Xian-Sheng Hua. MSRA att TRECVID 2008: High-level feature extraction and automatic search. In *TRECVID 2008 workshop participants notebook papers, Gaithersburg, MD, USA, November 2008*, 2008.
- [35] Xing Yi, Changshui Zhang, and Jingdong Wang. Multi-view EM algorithm and its application to color image segmentation. In *Proceedings of the 2004 IEEE International Conference on Multimedia and Expo, ICME 2004, 27-30 June 2004, Taipei, Taiwan*, pages 351–354, 2004.
- [36] Jingdong Wang, Jianguo Lee, and Changshui Zhang. Kernel trick embedded gaussian mixture model. In *Algorithmic Learning Theory, 14th International Conference, ALT 2003, Sapporo, Japan, October 17-19, 2003, Proceedings*, pages 159–174, 2003.
- [37] Jianguo Lee, Jingdong Wang, and Changshui Zhang. Color image segmentation: Kernel do the feature space. In *Machine Learning: ECML 2003, 14th European Conference on Machine Learning, Cavtat-Dubrovnik, Croatia, September 22-26, 2003, Proceedings*, pages 253–264, 2003.
- [38] Jingdong Wang, Jianguo Lee, and Changshui Zhang. Kernel GMM and its application to image binarization. In *Proceedings of the 2003 IEEE International Conference on Multimedia and Expo, ICME 2003, 6-9 July 2003, Baltimore, MD, USA*, pages 533–536, 2003.

## Tutorials and Seminars

---

- Context Autoencoder for Scalable Self-Supervised Representation Pretraining — Keynote, SSLAD@ECCV 2022
- Context Autoencoder for Scalable Self-Supervised Representation Pretraining — Keynote, NAS@CVPR 2022
- Dense Prediction with Transformers — Invited Talk, Neural Architectures@ICCV 2021
- Semantic Segmentation, Object Detection, and High-Resolution Backbone with Transformers — Keynote, Vis-Drone@ICCV 2021
- High-Resolution Networks — Keynote, BigMM 2020
- High-Resolution Networks — Keynote, CVWC@ICCV 2019
- Neural Architecture Construction and Search — Tutorial, ICME 2019
- High-Resolution Networks — Keynote, CEFRL@CVPR 2019

- Interleaved Group Convolutions for Efficient Deep Neural Networks — Keynote, ACM ICIMCS 2018
- Person Re-identification: Datasets and Representations — Keynote, PCM 2018
- Person Re-identification: Recent Advances and Challenges — Tutorial, FG 2018
- CNN Architecture Design: From Deeper to Wider — Keynote, ACM ICIMCS 2017

## Representative Product Transfer

---

- Context Auto-Encoder (CAE v2) — Baidu AI Cloud, 2023
- Group DETR — Baidu AI Cloud, 2023
- Context Auto-Encoder (CAE v1) — Baidu AI Cloud, 2022
- MaskOCR — Baidu OCR Products, 2022
- HRNet — Microsoft Form Recognizer, 2021
- SPTAG — Bing Platform, 2018
- Fast Human Detection — XiaoIce, 2017
- How-Old Fashion — XiaoIce, 2016
- How-Old 2.0 — XiaoIce, 2015
- Celebrity for Chat — XiaoIce, 2015, 2017
- Neighborhood Graph Construction and Search — Bing Image Search, 2015
- Neighborhood Graph Search — Ads Selection, 2014
- XiaoIce Book Recognition — Weibo, 2014
- Image Match — Bing Image Search, 2014
- Visual Attribute Enhanced Image Ranker — Bing Image Search, 2012
- Dominant Color Filter — Bing Image Search, 2012
- Image Search by Color Map — Bing Image Search, 2010

## Interns and Students

---

- |                              |                             |
|------------------------------|-----------------------------|
| • Huan Liu (BJTU'23)         | • Boqi Chen (BU'19)         |
| • Jiayang Tang (PKU'23)      | • Tianheng Cheng (HUST'18)  |
| • Jiyang Qi (HUST'22)        | • Chaorui Deng (SCUT)       |
| • Jie Zhu (PKU'22)           | • Manli Shu (USTC'18)       |
| • Mingyu Ding (HKU'22)       | • Borui Jiang (PKU'18)      |
| • Yan Liu (ZJU'21)           | • Jingchun Cheng (THU'18)   |
| • Xiaosong Zhao (UCAS'21)    | • Weijian Xu (UCSD'18)      |
| • Yifan Zhang (SCUT'21)      | • Yifan Liu (BUAA'18)       |
| • Juhong Min (POSTECH'20–21) | • Yao Qin (UCSD'18)         |
| • Zejia Fan (PKU'20–21)      | • Ke Sun (USTC'17–18)       |
| • Qi Han (NKU'20–21)         | • Shichen Liu (THU'17–18)   |
| • Zhuoyi Huang (Tongji'20)   | • Depu Meng (USTC'17–18)    |
| • Xiaokang Chen (PKU'20–21)  | • Mingji Li (USTC'17–18)    |
| • Changqian Yu (HUST'20)     | • Junran Peng (IA,CAS'17)   |
| • Shaobo Min (USTC'20)       | • Chong Shang (THU'17)      |
| • Baifeng Shi (PKU)          | • Guotian Xie (SYSU'17)     |
| • Chaoqun Wang (USTC'19)     | • Peng Tang (HUST'16–17)    |
| • Hung-Jen Chen (NCTU'19)    | • Song Bai (HUST'16–17)     |
| • Jingyi Xie (USTC'19)       | • Gangming Zhao (IA,CAS'17) |
| • Zigang Geng (USTC'19)      | • Kexin Zhang (PKU'17)      |

- Yumin Suh (SNU'16–17)
- Yajie Xing (PKU'16–17)
- Shuang Li (THU'16)
- Ryosuke Furuta (UTokyo'16)
- Yanzhao Wu (USTC'16)
- Jun He (HFUT'16)
- Liheng Zhang (UCF'16)
- Liming Zhao (ZJU'16)
- Zilong Huang (HUST'16)
- Xun Huang (BUAA'16)
- Xiaojuan Wang (SYSU'15)
- Zhen Wei (SJTU'15)
- Wei Tang (IA,CAS'15)
- Qiqi Hou (XJTU'15)
- Ning Xu (UIUC'15)
- Liang Zheng (THU'14–15)
- Jiaming Song (THU'14)
- Tianxiao Shen (THU'14)
- Jianqiu Ji (THU'14)
- Fan Liu (NJUST'14)
- Baoguang Shi (HUST'14)
- Dapeng Chen (XJTU'14)
- Falong Shen (PKU'13–14)
- Lingxi Xie (THU'13)
- Ting Zhang (USTC'13)
- Xianjie Chen (UCLA'13)
- Chongxuan Li (THU'13)
- Chao Du (THU'13–14)
- Jingkuan Song (UQ'12–13)
- Chuang Gan (BUAA'12–13)
- Zhansheng Jiang (SJTU'12–13)
- Yupeng Gu (USTC'12–13)
- Shiyu Chang (UIUC'12)
- Jianfeng Wang (USTC'12–13)
- Weiwen Tu (SYSU'11)
- Tiansheng Yao (ZJU'11)
- Dongqing Zhang (SJTU'11)
- Peng Wang (PKU'11–13)
- Xianghang Liu (UNSW'11)
- Jiazhen Zhou (PKU'11) → Columbia
- Junyang Lu (PKU'11) → Columbia
- Xiang Ren (ZJU'11) → UIUC
- Nobuyuki Morioka (UNSW'11) → Google
- Naiyan Wang (ZJU'10–11) → HKUST
- Huaizu Jiang (XJTU'10–11)
- Shijie Xu (BUAA'10)
- Jing Wang (PKU'10–13)
- Bin Cheng (NUS'10)
- Biao Wang (THU'10)
- Wenhao Lu (THU'10)
- Zhe Zhao (PKU'09–10) → Michigan
- Xiangang Chen (NTU'09)
- Jue Xie (Monash'09)
- Hao Xu (USTC'08–09)
- You Jia (PKU'09) → CMU/Facebook
- Liyan Jia (THU'09) → Cornell
- Heng Zhang (PKU'08–09)
- Yang Hu (USTC'08)
- Darui Li (USTC'08)
- Yinghai Zhao (USTC'08)
- Xuan Huang (USTC'07–08)
- Jianxiong Xiao (HKUST)
- Fei Wang (THU'05–06) → IBM Research
- Yangqing Jia (THU'07) → UC Berkeley
- Dan Zhang (THU'07,10) → Purdue
- Mingjie Qian (THU'09) → UIUC

## Contact

---

Email: [wangjingdong@outlook.com](mailto:wangjingdong@outlook.com)  
 Website: <https://jingdongwang2017.github.io>  
 Google Scholar: <https://scholar.google.com/citations?user=z5SPCmgAAAAJ>  
 DBLP: <https://dblp.uni-trier.de/pid/49/3441.html>  
 ORCID: 0000-0002-4888-4445  
 Phone: +86 13661386588