

Rui Zhao

Executive Director (L12, highest level) at SenseTime

Contact Information

40/F, Block A, Building 11,
Science and Technology Ecological Park,
Nanshan District, Shenzhen, Guangdong, China

☎ +86 188 2316 1367
✉ rzhao@link.cuhk.edu.hk
📧 zhaorui.xyz

Biography highlights

- Google Citation 15,000+, H-index 45, with three works cited over 1,000 times.
- Proposed the first deep learning method for person re-identification, two works on person re-id were selected as the Most Influential CVPR Papers by [PaperDigest](#).
- Ten-year experience of applying cutting-edge AI techniques to real-world industrial scenarios, and past works have been widely adopted in industrial products and applied to more than 100 cities with over 1 billion CNY revenue globally.
- Recent works cover a wide range of prevailing topics including Large Language Models, Vision-Language Models, LLM-based Agents, especially tool using and reasoning.
- Management experience over 200 people (mostly researchers and a few engineers).

Education

The Chinese University of Hong Kong (CUHK), Shatin, Hong Kong 2010-2015
Doctor of Philosophy, Electronic Engineering

University of Science and Technology of China (USTC), Hefei, China 2006-2010
Bachelor, Information Science and Electronic Engineering

Working Experience

2015 — now Executive Director (2019-now), Research Director (2018-2019), Senior Researcher (2015-2016), Smart City Business Group, SenseTime Group Ltd.

2015 — 2018 R&D Vice President (2016-2018), R&D Director (2015-2016), SenseNets Technology Ltd.

2024 — now Adjunct Professor, Tsinghua Shenzhen International Graduate School

2020 — 2022 Adjunct Professor, QingYuan Research Institute, Shanghai Jiao Tong University

2020 — 2022 Adjunct Professor, Shenzhen Institute of Adv. Tech., Chinese Academy of Sciences

Awards (Selected)

- Worldwide Top 0.5% Scholars ([ScholarGPS](#)) 2025
- Stanford's Top 2% Scientists ([Elsevier](#)) 2024
- Top 100 AI Chinese Young Scholars ([Baidu](#)) 2022
- AI 2000 Most Influential Scholar Award ([AMiner](#)) 2022
- Shenzhen May 1st Labor Medal ([SMFTU](#)) 2022
- Shenzhen Artificial Intelligence Scientific Progress Award (SAAI) 2022
- Guangdong Provincial Excellent Technology and Innovation Award (First Prize) 2020
- Wu Wen Jun AI Award for Scientific and Technology Progress 2019
- Black Sheep Award (SenseTime, 1‰) 2018
- Shenzhen Peacock Program Talent Award 2016
- ECCV Outstanding Reviewer Award 2016
- CVPR Doctoral Consortium Award (Mentor: Trevor Darrell) 2015
- Commendation for Outstanding Tutorial (CUHK) 2014
- ICCV Student Travel Grant 2013
- Second Prize in Chinese University Mathematical Contest in Modeling (USTC) 2009

PUBLICATIONS Google Scholar Profile: <https://scholar.google.com> (Citations: > 15,000; H-index: 45)

Selected Recent Publications

1. He, Weizhen, et al. "Instruct-ReID++: Towards Universal Purpose Instruction-Guided Person Re-identification." *IEEE Trans. on Pattern Analysis and Machine Intelligence (TPAMI)*. 2025.
2. Kong, Yilun, et al. "TPTU-v2: Boosting Task Planning and Tool Usage of Large Language Model-based Agents in Real-world Industry Systems." *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing (EMNLP): Industry Track*. 2024.
3. Zhang, Hengyuan, et al. "Balancing Speciality and Versatility: A Coarse to Fine Framework for Supervised Fine-tuning Large Language Model." *Findings of the Association for Computational Linguistics (ACL)*. 2024.
4. Liu, Xuejing, et al. "Deeply Coupled Cross-Modal Prompt Learning." *Findings of the Association for Computational Linguistics (ACL)*. 2023.

Journals

1. Wang, Yizhou, et al. "Hulk: A Universal Knowledge Translator for Human-Centric Tasks." *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 2025.
2. He, Weizhen, et al. "Instruct-ReID++: Towards Universal Purpose Instruction-Guided Person Re-identification." *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 2025.
3. Xie, Chi, et al. "RelationLMM: Large Multimodal Model as Open and Versatile Visual Relationship Generalist." *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 2025.
4. Chen, Jianqiu, et al. "ZeroPose: CAD-Prompted Zero-Shot Object 6D Pose Estimation in Cluttered Scenes." *IEEE Transactions on Circuits and Systems for Video Technology (TCSVT)*, 2024.
5. Li, Zhaowen, et al. "Efficient Masked Autoencoders with Self-Consistency." *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 2024.
6. Jiang, Haoyuan, et al. "A General Scenario-agnostic Reinforcement Learning for Traffic Signal Control." *IEEE Transactions on Intelligent Transportation Systems (TITS)*, 2024.
7. Jiang, Haoyuan, et al. "GESA: A General Scenario-Agnostic Reinforcement Learning for Traffic Signal Control." *IEEE Transactions on Intelligent Transportation Systems (TITS)*, 2024.
8. Liu, Kaijian, et al. "Relation-aware Distribution Representation Network for Person Clustering with Multiple Modalities." *IEEE Transactions on Multimedia (TMM)*, 2023.
9. Li, Shihua, et al. "Cocas+: Large-scale Clothes-Changing Person Re-identification with Clothes Templates." *IEEE Transactions on Circuits and Systems for Video Technology (TCSVT)*, 2022.
10. Ge, Yixiao, et al. "Structured Domain Adaptation with Online Relation Regularization for Unsupervised Person Re-ID." *IEEE Transactions on Neural Networks and Learning Systems (TNNLS)*, 2022.
11. Yuan, Haojie, et al. "AutoMA: Towards Automatic Model Augmentation for Transferable Adversarial Attacks." *IEEE Transactions on Multimedia (TMM)*, 2021.
12. Ge, Yixiao, et al. "Self-Paced Contrastive Learning with Hybrid Memory for Domain Adaptive Object Re-ID." *Advances in Neural Information Processing Systems (NeurIPS)*, 2020.
13. **R. Zhao**, W. Ouyang and X. Wang, "Person Re-identification by Saliency Learning." *in IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, vol.39, pp.356-370, 2017.
14. **R. Zhao** and X. Wang, "Counting Vehicles from Semantic Regions." *in IEEE Transactions on Intelligent Transportation Systems (TITS)*, vol.14, no.2, pp.1016-1022, March 2013.

Conferences

1. Liu, Chenxi, et al. "TimeCMA: Towards LLM-Empowered Multivariate Time Series Forecasting via Cross-Modality Alignment." Proceedings of the AAAI Conference on Artificial Intelligence. Vol. 39. No. 18. 2025.
2. Xu, Qianxiong, et al. "KITS: Inductive Spatio-Temporal Kriging with Increment Training Strategy." Proceedings of the AAAI Conference on Artificial Intelligence. Vol. 39. No. 12. 2025.
3. Zhong, Cheng, et al. "Graph Pooling via Dropping Task-Irrelevant Nodes." ICASSP 2025-2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2025.
4. Xu, Qianxiong, et al. "Hybrid Mamba for Few-Shot Segmentation." Advances in Neural Information Processing Systems 37 (2024): 73858-73883.
5. Kong, Yilun, et al. "TPTU-v2: Boosting Task Planning and Tool Usage of Large Language Model-based Agents in Real-world Industry Systems." Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing: Industry Track. 2024.
6. Ding, Chuanghao, et al. "SynthDoc: Bilingual Documents Synthesis for Visual Document Understanding." Proceedings of the 2nd Workshop on Large Generative Models Meet Multimodal Applications. 2024.
7. Xu, Qianxiong, et al. "Eliminating Feature Ambiguity for Few-shot Segmentation." European Conference on Computer Vision (ECCV). 2024.
8. Ruan, Jingqing, et al. "Coslight: Co-optimizing Collaborator Selection and Decision-Making to Enhance Traffic Signal Control." Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. 2024.
9. Yang, Sun, et al. "SQL-to-Schema Enhances Schema Linking in Text-to-SQL." International Conference on Database and Expert Systems Applications. 2024.
10. Zhang, Hengyuan, et al. "Balancing Speciality and Versatility: a Coarse to Fine Framework for Supervised Fine-Tuning Large Language Model." Findings of the Association for Computational Linguistics ACL 2024. 2024.
11. Zhang, Bin, et al. "Sequential Asynchronous Action Coordination in Multi-Agent Systems: A Stackelberg Decision Transformer Approach." Forty-first International Conference on Machine Learning. 2024.
12. Liu, Chenxi, et al. "Spatial-Temporal Large Language Model for Traffic Prediction." 2024 25th IEEE International Conference on Mobile Data Management (MDM). IEEE, 2024.
13. Li, Zhaowen, et al. "Self-Supervised Representation Learning from Arbitrary Scenarios." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2024.
14. He, Weizhen, et al. "Instruct-ReID: A Multi-Purpose Person Re-identification Task with Instructions." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2024.
15. Tai, Yan, et al. "Link-Context Learning for Multimodal LLMs." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2024.
16. Xie, Chi, et al. "Described Object Detection: Liberating Object Detection with Flexible Expressions." Advances in Neural Information Processing Systems 36 (2023): 79095-79107.
17. Li, Dedong, et al. "A critical perceptual pre-trained model for complex trajectory recovery." Proceedings of the 2nd ACM SIGSPATIAL International Workshop on Searching and Mining Large Collections of Geospatial Data. 2023.
18. Lin, Junpeng, et al. "Dynamic Causal Graph Convolutional Network for Traffic Prediction." 2023 IEEE 19th International Conference on Automation Science and Engineering (CASE). IEEE, 2023.
19. Wang, Luxuan, et al. "Correlated Time Series Self-supervised Representation Learning via Spatiotemporal Bootstrapping." 2023 IEEE 19th International Conference on Automation Science and Engineering (CASE). IEEE, 2023.

20. Ruan, Jingqing, et al. "TPTU: Large Language Model-based AI Agents for Task Planning and Tool Usage." arXiv preprint arXiv:2308.03427 (2023).
21. Lan, Tian, et al. "MM-DAG: Multi-task DAG Learning for Multi-Modal Data-with Application for Traffic Congestion Analysis." Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. 2023.
22. Li, Dedong, Rui Zhao, and Fei Tan. "CWSeg: An Efficient and General Approach to Chinese Word Segmentation." Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 5: Industry Track). 2023.
23. Qi, Yu, et al. "Exploring Stochastic Autoregressive Image Modeling for Visual Representation." Proceedings of the AAAI Conference on Artificial Intelligence. Vol. 37. No. 2. 2023.
24. Lu, Jinghui, et al. "Punifiedner: A Prompting-based Unified NER System for Diverse Datasets." Proceedings of the AAAI Conference on Artificial Intelligence. Vol. 37. No. 11. 2023.
25. Zhang, Shaofeng, et al. "Patch-level Contrasting without Patch Correspondence for Accurate and Dense Contrastive Representation Learning." arXiv preprint arXiv:2306.13337 (2023).
26. Wu, Xiaoshi, et al. "Human Preference Score v2: A Solid Benchmark for Evaluating Human Preferences of Text-to-Image Synthesis." arXiv preprint arXiv:2306.09341 (2023).
27. Xiang, Yuhan, et al. "Trust Your Partner's Friends: Hierarchical Cross-Modal Contrastive Pre-Training for Video-Text Retrieval." IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). 2023.
28. Liu, Xuejing, et al. "Deeply Coupled Cross-Modal Prompt Learning." arXiv preprint arXiv:2305.17903 (2023).
29. Zhang, Bin, et al. "Stackelberg Decision Transformer for Asynchronous Action Coordination in Multi-agent Systems." arXiv preprint arXiv:2305.07856 (2023).
30. Chen, Jianqiu, et al. "ZeroPose: CAD-model-based Zero-Shot Pose Estimation." arXiv e-prints (2023): arXiv-2305.
31. Sun, Mingshan, et al. "Uni6Dv2: Noise Elimination for 6D Pose Estimation." International Conference on Artificial Intelligence and Statistics (ICAIS). 2023.
32. Jin, Guoqiang, et al. "SeqCo-DETR: Sequence Consistency Training for Self-Supervised Object Detection with Transformers." arXiv preprint arXiv:2303.08481 (2023).
33. Chen, Meilin, et al. "Saliency Guided Contrastive Learning on Scene Images." arXiv preprint arXiv:2302.11461 (2023).
34. Lin, Shaobo, Xingyu Zeng, and Rui Zhao. "Explore the Power of Dropout on Few-shot Learning." arXiv preprint arXiv:2301.11015 (2023).
35. Wu, Xiaoshi, et al. "Human Preference Score: Better Aligning Text-to-Image Models with Human Preference." Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). 2023.
36. Ruan, Jingqing, et al. "TPTU: Task Planning and Tool Usage of Large Language Model-based AI Agents." NeurIPS 2023 Foundation Models for Decision Making Workshop. 2023.
37. Xie, Chi, et al. "Exposing the Troublemakers in Described Object Detection." CoRR (2023).
38. Wu, Yixuan, et al. "Advancing Referring Expression Segmentation beyond Single Image." Proceedings of the IEEE/CVF International Conference on Computer Vision. 2023.
39. He, Weizhen, et al. "Retrieve Anyone: A General-Purpose Person Re-Identification Task with Instructions." arXiv preprint arXiv:2306.07520 (2023).
40. Chen, Jianqiu, et al. "3D Model-Based Zero-Shot Pose Estimation Pipeline." CoRR (2023).
41. Wang, Yuchao, et al. "Balancing Logit Variation for Long-tailed Semantic Segmentation." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2023.

42. Wu, Xiaoshi, et al. "CORA: Adapting Clip for Open-Vocabulary Detection with Region Prompting and Anchor Pre-matching." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2023.
43. Lin, Shaobo, et al. "Explore the Power of Synthetic Data on Few-shot Object Detection." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2023.
44. Tang, Shixiang, et al. "Humanbench: Towards General Human-Centric Perception with Projector Assisted Pretraining." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2023.
45. Ci, Yuanzheng, et al. "UniHCP: A Unified Model for Human-Centric Perceptions." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2023.
46. Lin, Shaobo, et al. "An Effective Crop-Paste Pipeline for Few-shot Object Detection." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2023.
47. Bao, Tianpeng, et al. "MIAD: A Maintenance Inspection Dataset for Unsupervised Anomaly Detection." Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). 2023.
48. Mao, Hangyu, et al. "Transformer in Transformer as Backbone for Deep Reinforcement Learning." arXiv preprint arXiv:2212.14538 (2022).
49. Wang, Yizhou, et al. "Unsupervised Object Detection Pretraining with Joint Object Priors Generation and Detector Learning." Advances in Neural Information Processing Systems (NeurIPS) 35 (2022): 12435-12448.
50. Chen, Zhiyang, et al. "Obj2seq: Formatting Objects as Sequences with Class Prompt for Visual Tasks." Advances in Neural Information Processing Systems (NeurIPS) 35 (2022): 2494-2506.
51. Tang, Shixiang, et al. "Unifying Visual Contrastive Learning for Object Recognition from a Graph Perspective." European Conference on Computer Vision. (ECCV). 2022.
52. Tang, Shixiang, et al. "Relative Contrastive Loss for Unsupervised Representation Learning." European Conference on Computer Vision. (ECCV). 2022.
53. Li, Xulin, et al. "Counterfactual Intervention Feature Transfer for Visible-infrared Person Re-Identification." European Conference on Computer Vision (ECCV). 2022.
54. Yang, Haiyang, et al. "Domain Invariant Masked Autoencoders for Self-Supervised Learning from Multi-Domains." European Conference on Computer Vision (ECCV). 2022.
55. Chen, Jianqiu, et al. "Geo6D: Geometric Constraints Learning for 6D Pose Estimation." arXiv preprint arXiv:2210.10959 (2022).
56. Mao, Zhenyu, et al. "Jointly Contrastive Representation Learning on Road Network and Trajectory." Proceedings of the 31st ACM International Conference on Information & Knowledge Management (CIKM). 2022.
57. Lin, Shaobo, Xingyu Zeng, and Rui Zhao. "A Unified Framework with Meta-dropout for Few-shot Learning." arXiv preprint arXiv:2210.06409 (2022).
58. Zhu, Dongsheng, et al. "SDA: Simple Discrete Augmentation for Contrastive Sentence Representation Learning." arXiv preprint arXiv:2210.03963 (2022).
59. Lu, Jinghui, et al. "What Makes Pre-trained Language Models Better Zero-Shot Learners?." arXiv preprint arXiv:2209.15206 (2022).
60. Zheng, Ye, et al. "Focus Your Distribution: Coarse-to-fine Non-contrastive Learning for Anomaly Detection and Localization." 2022 IEEE International Conference on Multimedia and Expo (ICME). IEEE, 2022.
61. Yang, Haiyang, et al. "Cycle-consistent Masked Autoencoder for Unsupervised Domain Generalization." The Eleventh International Conference on Learning Representations (ICLR). 2022.

62. Zhang, Shaofeng, et al. "Contextual Image Masking Modeling via Synergized Contrasting without View Augmentation for Faster and Better Visual Pretraining." The Eleventh International Conference on Learning Representations (ICLR). 2022.
63. Lin, Shaobo, et al. "Three-stage Training Pipeline with Patch Random Drop for Few-shot Object Detection." Proceedings of the Asian Conference on Computer Vision (ACCV). 2022.
64. Zhang, Shaofeng, et al. "Align Representations with Base: A New Approach to Self-supervised Learning." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2022.
65. Jiang, Xiaoke, et al. "Uni6D: A Unified CNN Framework without Projection Breakdown for 6D Pose Estimation." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2022.
66. Li, Zhaowen, et al. "UniVIP: A Unified Framework for Self-Supervised Visual Pre-training." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2022.
67. Wang, Zhikang, et al. "Feature Erasing and Diffusion Network for Occluded Person Re-Identification." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2022.
68. Wang, Yuchao, et al. "Semi-Supervised Semantic Segmentation using Unreliable Pseudo-labels." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2022.
69. Zhiyang Chen, Yousong Zhu, Fan Yang, Wei Li, Chaoyang Zhao, Zhaowen Li, Liwei Wu, **Rui Zhao**, Jinqiao Wang, Ming Tang, "Decouple Spatial and Channel Dimension in MLP" submitted to International Conference on Computer Vision (**AAAI**) 2022.
70. Dedong Li, Zhenyu Mao, Ziyue Li, **Rui Zhao**, "RANG-AT&T: RoAd Network Graph Augmented Transformer for Trajectory Representation Learning" submitted to International Conference on Computer Vision (**AAAI**) 2022.
71. Zhenyu Mao, Dedong Li, Ziyue Li, **Rui Zhao**, "Trajectory-guided Contrastive Representation Learning on Road Network Graph" submitted to International Conference on Computer Vision (**AAAI**) 2022.
72. Kaijian Liu, Hong Zhang, Shixiang Tang, Lihong Zheng, Ziyue Li, **Rui Zhao**, "MAGNET: Modality-Agnostic distribution Graph NETWORK for Person Clustering" submitted to International Conference on Computer Vision (**AAAI**) 2022.
73. Shijie Yu, Feng Zhu, Dapeng Chen, **Rui Zhao**, Haobin Chen, Shixiang Tang, Jingguo Zhu, Yu Qiao, "Multiple Domain Experts Collaborative Learning: Multi-Source Domain Generalization For Person Re-Identification" submitted to International Conference on Computer Vision (**AAAI**) 2022.
74. Shaobo Lin, Xingyu Zeng, Dapeng Chen, **Rui Zhao**, "MDFL: A Unified Framework with Meta-dropout for Few-shot Learning" submitted to International Conference on Computer Vision (**AAAI**) 2022.
75. Shijie Yu, Shihua Li, Haobin Chen, Feng Zhu, **Rui Zhao**, Yu Qiao, "SRViP: A Video Benchmark for Cloth Changing Person Re-Identification" submitted to Neural Information Processing Systems (**NeurIPS**) 2022.

76. Shixiang Tang, Lei Bai, Feng Zhu, **Rui Zhao**, Wanli Ouyang, “Relative Contrastive Loss for Self-supervised Representation Learning” submitted to Neural Information Processing Systems (**NeurIPS**) 2022.
77. Zhaowen Li, Zhiyang Chen, Fan Yang, Wei Li, Yousong Zhu, Chaoyang Zhao, Rui Deng, Liwei Wu, **Rui Zhao**, Ming Tang, Jinqiao Wang, “MST: Masked Self-Supervised Transformer for Visual Representation” submitted to Neural Information Processing Systems (**NeurIPS**) 2022.
78. Ye Zheng, Xiang Wang, Rui Deng, Tianpeng Bao, **Rui Zhao**, Liwei Wu, “Focus Your Distribution: Coarse-to-Fine Non-Contrastive Learning for Anomaly Detection and Localization” submitted to Neural Information Processing Systems (**NeurIPS**) 2022.
79. Chen Zhao, Yixiao Ge, Feng Zhu, **Rui Zhao**, Hongsheng Li, Mathieu Salzmann, “Progressive Correspondence Pruning by Consensus Learning” in International Conference on Computer Vision (**ICCV**) 2021.
80. Bo Zhao, Shixiang Tang, Dapeng Chen, Hakan Bilen, **Rui Zhao**, “Continual Representation Learning for Biometric Identification” in *Winter Conference on Applications of Computer Vision* (**WACV**) 2021.
81. Yixiao Ge, Feng Zhu, **Rui Zhao**, Hongsheng Li, “Structured Domain Adaptation for Unsupervised Person Re-identification”, in *arXiv:2003.06650*, 2020.
82. Yixiao Ge, Haibo Wang, Feng Zhu, **Rui Zhao**, Hongsheng Li, “Self-supervising Fine-grained Region Similarities for Large-scale Image Localization” in *European Conference on Computer Vision* (**ECCV**) 2020.
83. Xiao Zhang, **Rui Zhao**, Yu Qiao, Hongsheng Li, “RBF-Softmax: Learning Deep Representative Prototypes with Radial Basis Function Softmax” in *European Conference on Computer Vision* (**ECCV**) 2020.
84. Shijie Yu, Shihua Li, Dapeng Chen, **Rui Zhao**, Junjie Yan, Yu Qiao, “COCAS: A Large-Scale Clothes Changing Person Dataset for Re-identification,” in *IEEE International Conference on Computer Vision and Pattern Recognition* (**CVPR**) 2020.
85. Senhui Guo, Jing Xu, Dapeng Chen, Chao Zhang, Xiaogang Wang, **Rui Zhao**, “Density-Aware Feature Embedding for Face Clustering,” in *IEEE International Conference on Computer Vision and Pattern Recognition* (**CVPR**) 2020.
86. Lei Yang, Dapeng Chen, Xiaohang Zhan, **Rui Zhao**, Chen Change Loy, Dahua Lin, “Learning to Cluster Faces via Confidence and Connectivity Estimation,” in *IEEE International Conference on Computer Vision and Pattern Recognition* (**CVPR**) 2020.
87. Suichan Li, Dapeng Chen, Bin Liu, Nenghai Yu, **Rui Zhao**, “Memory-Based Neighbourhood Embedding for Visual Recognition,” in *IEEE International Conference on Computer Vision* (**ICCV**) 2019.
88. X. Zhang, **R. Zhao**, Y. Qiao, X. Wang, and H. Li, “AdaScale: Adaptively Scaling Cosine Logits for Effectively Learning Deep Face Representations,” in *IEEE International Conference on Computer Vision and Pattern Recognition* (**CVPR**) 2019.

89. X. Zhang, **R. Zhao**, J. Yan, M. Gao, Y. Qiao, X. Wang, and H. Li, "P2SGrad: Refined Gradients for Optimizing Deep Face Models," in *IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)* 2019.
90. J. Xu, **R. Zhao**, F. Zhu, and W. Ouyang, "Attention-Aware Compositional Network for Person Re-identification," in *IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)* 2018.
91. Z. Zhao, H. Li, **R. Zhao** and X. Wang, "Crossing-line Crowd Counting with Two-phase Deep Neural Networks," in *European Conference on Computer Vision (ECCV)* 2016.
92. **R. Zhao**, W. Ouyang, H. Li and X. Wang, "Saliency Detection by Multi-Context Deep Learning," in *IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)* 2015.
93. **R. Zhao**, W. Ouyang and X. Wang, "Learning Mid-level Filters for Person Re-Identification," in *IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)* 2014.
94. W. Li, **R. Zhao**, T. Xiao and X. Wang, "DeepReid: Deep Filter Pairing Neural Network for Person Re-Identification," in *IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)* 2014.
95. **R. Zhao**, W. Ouyang and X. Wang, "Person Re-Identification by Saliency Matching," in *IEEE International Conference on Computer Vision (ICCV)* 2013.
96. **R. Zhao**, W. Ouyang and X. Wang, "Unsupervised Saliency Learning for Person Re-Identification," in *IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)* 2013.
97. W. Li, **R. Zhao**, and X. Wang, "Human Reidentification with Transferred Metric Learning," in *Asian Conference on Computer Vision (ACCV)* 2012.
98. L. Zhuang, **R. Zhao**, N. Yu and B. Liu, "SVD Based Linear Filtering in DCT Domain," in *IEEE International Conference on Image Processing (ICIP)* 2010.

arXivs

1. Chen, Xiaxu, et al. "On the Suitability of Reinforcement Fine-Tuning to Visual Tasks." arXiv preprint arXiv:2504.05682 (2025).
2. Chen, Yuming, et al. "Re-Aligning Language to Visual Objects with an Agentic Workflow." arXiv preprint arXiv:2503.23508 (2025).
3. Fang, Rongyao, et al. "GOT: Unleashing Reasoning Capability of Multimodal Large Language Model for Visual Generation and Editing." arXiv preprint arXiv:2503.10639 (2025).
4. Wei, Anning, et al. "DTPPO: Dual-Transformer Encoder-Based Proximal Policy Optimization for Multi-UAV Navigation in Unseen Complex Environments." arXiv preprint arXiv:2410.15205 (2024).
5. Fang, Rongyao, et al. "Puma: Empowering unified mllm with multi-granular visual generation." arXiv preprint arXiv:2410.13861 (2024).
6. Kong, Yilun, et al. "QPO: Query-dependent Prompt Optimization via Multi-Loop Offline Reinforcement Learning." arXiv preprint arXiv:2408.10504 (2024).
7. Wang, Shiqi, et al. "Reward Difference Optimization For Sample Reweighting In Offline RLHF." arXiv preprint arXiv:2408.09385 (2024).
8. Wang, Shiqi, et al. "Offline RLHF Methods Need More Accurate Supervision Signals." arXiv e-prints (2024): arXiv-2408.

9. Zhao, Fufangchen, et al. "SimCT: A Simple Consistency Test Protocol in LLMs Development Lifecycle." arXiv preprint arXiv:2407.17150 (2024).
10. Gu, Jiawei, et al. "CMR Scaling Law: Predicting Critical Mixture Ratios for Continual Pre-training of Language Models." arXiv preprint arXiv:2407.17467 (2024).
11. Jiang, Haoyuan, et al. "GuideLight: Industrial Solution" Guidance for More Practical Traffic Signal Control Agents." arXiv preprint arXiv:2407.10811 (2024).
12. Chen, Sirui, et al. "CLEAR: Can Language Models Really Understand Causal Graphs?." arXiv preprint arXiv:2406.16605 (2024).
13. Guo, Zhaojun, et al. "What Makes Good Few-shot Examples for Vision-Language Models?." arXiv preprint arXiv:2405.13532 (2024).
14. Chen, Sirui, et al. "Causal Evaluation of Language Models." arXiv preprint arXiv:2405.00622 (2024).
15. Lan, Tian, et al. "Multifun-DAG: Multivariate Functional Directed Acyclic Graph." arXiv preprint arXiv:2404.13836 (2024).
16. Jiang, Haoyuan, et al. "X-light: Cross-city Traffic Signal Control using Transformer on Transformer as Meta Multi-Agent Reinforcement Learner." arXiv preprint arXiv:2404.12090 (2024).
17. Li, Zhishuai, et al. "PET-SQL: A Prompt-Enhanced Two-Round Refinement of Text-to-SQL with Cross-consistency." arXiv preprint arXiv:2403.09732 (2024).
18. Zhang, Bin, et al. "Benchmarking the Text-to-SQL Capability of Large Language Models: A Comprehensive Evaluation." arXiv preprint arXiv:2403.02951 (2024).
19. Hong, Zijin, et al. "Next-Generation Database Interfaces: A Survey of LLM-based Text-to-SQL." arXiv preprint arXiv:2406.08426 (2024).
20. Li, Zhishuai, et al. "Non-neighbors Also Matter to Kriging: A New Contrastive-Prototypical Learning." arXiv preprint arXiv:2401.12681 (2024).
21. Mao, Hangyu, et al. "PDIT: Interleaving Perception and Decision-Making Transformers for Deep Reinforcement Learning." arXiv preprint arXiv:2312.15863 (2023).
22. Lu, Jiaming, et al. "Dualight: Enhancing Traffic Signal Control by Leveraging Scenario-Specific and Scenario-Shared Knowledge." arXiv preprint arXiv:2312.14532 (2023).
23. Li, Zhishuai, et al. "Visiontraj: A Noise-Robust Trajectory Recovery Framework based on Large-Scale Camera Network." arXiv preprint arXiv:2312.06428 (2023).
24. Zhang, Bin, et al. "Controlling Large Language Model-based Agents for Large-Scale Decision-Making: An Actor-Critic Approach." arXiv preprint arXiv:2311.13884 (2023).
25. Kong, Yilun, et al. "Tptu-v2: Boosting task planning and tool usage of large language model-based agents in real-world systems." arXiv preprint arXiv:2311.11315 (2023).
26. Sun, Yuhan, et al. "To be or not to be? An Exploration of Continuously Controllable Prompt Engineering." arXiv preprint arXiv:2311.09773 (2023).
27. Liu, Xuejing, et al. "What Large Language Models Bring to Text-rich VQA?." arXiv preprint arXiv:2311.07306 (2023).
28. Sui, Guanghu, et al. "Reboost Large Language Model-based Text-to-SQL, Text-to-Python, and Text-to-Function—with Real Applications in Traffic Domain." arXiv preprint arXiv:2310.18752 (2023).
29. Liang, Zhixuan, et al. "Mean-AP Guided Reinforced Active Learning for Object Detection." arXiv preprint arXiv:2310.08387 (2023).
30. Dang, Ronghao, et al. "InstructDET: Diversifying Referring Object Detection with Generalized Instructions." arXiv preprint arXiv:2310.05136 (2023).
31. Wu, Xiaoshi, et al. "Better Aligning Text-to-Image Models with Human Preference." arXiv preprint arXiv:2303.14420 1.3 (2023).

