

Zhipeng Zhang

Portfolio: <http://zhipengzhang.cn/>
Click to Google Scholar

Email: zhipeng.zhang.cv@outlook.com
Mobile: 132-6950-1016

RESEARCH INTERESTS

2D vision and multi-modal understanding like detection, tracking, segmentation, grounding; 3D understanding like monocular/multi-view BEV object detection, multi-sensor fusion, and end-to-end autonomous driving; Embodied AI

EDUCATION

- **NLPR, Institute of Automation, Chinese Academy of Sciences (CASIA)** Beijing, China
Ph.D in Pattern recognition and intelligent system Sep. 2017 - July 2022
- **University of Electronic Science and Technology of China (UESTC)** Chengdu, China
Bachelor in Opto-Electronics Information Science and Engineering Sep. 2013 - July 2017

EMPLOYMENT HISTORY

- **Research Intern** Microsoft Research Asia (MSRA)
Major in video object tracking Sep. 2019 - June 2020
- **Senior Researcher in Autonomous Driving** DiDi & KargoBot
In charge of vision-only, multi-sensor BEV perception and end-to-end autonomous driving July 2022 - March 2025
- **Assistant Professor** School of AI, Shanghai Jiaotong University, China
2D and 3D perception, autonomous driving, embodied intelligence April 2025 -

WORK EXPERIENCE

Dr. Zhipeng Zhang is an Assistant Professor at the School of AI, Shanghai Jiao Tong University, specializing in 2D/3D perception, autonomous driving, embodied intelligence, and multimodal large models. Prior to his academic role, he was a Senior Researcher and Engineer at DiDi KargoBot Autonomous Driving, where he directed projects on camera-only Bird's-Eye View (BEV) detection, multi-sensor fusion object detection, and end-to-end autonomous driving.

HONORS AND AWARDS

- President Award of Chinese Academy of Sciences
- Huawei Talent Youth Program
- National Scholarship for Ph.D
- Runner-up of VOT2019RGBT challenge
- Champion of VOT2020RGBT
- Runner-ups of Short-term and Real-time Tracks in VOT2020 challenge
- Top5 (2,000+ teams) of Panda MOT challenge in Tianchi2021 (supervisor)

PUBLICATION SUMMARY

Zhipeng has published more than 30 high-quality papers in top conferences and journals such as CVPR, ICCV, ECCV, NeurIPS and TPAMI. in which he has served as the first author, co-first author, or corresponding author in more than 20 papers. His representative works (as first author), SiamDW and Ocean, published in CVPR 2019 and ECCV 2020 respectively, have garnered over **1300** and **900** citations. In total, his work has been cited more than **7000** times.

PAPER LIST

Representative Work (as First Author, co-First Author, corresponding author and supervisor)

- **[First Author] Zhipeng Zhang**, Houwen Peng, “Deeper and Wider Siamese Networks for Real-Time Visual Tracking” in IEEE Conference on Computer Vision and Pattern Recognition (**CVPR2019**), **Oral, Cite 1100+**
- **[First Author] Zhipeng Zhang**, Houwen Peng, Jianlong Fu, Bing Li, Weiming Hu, “Ocean: Object-aware Anchor-free Tracking” in European Conference on Computer Vision (**ECCV2020**), **Cite 730+**
- **[First Author] Zhipeng Zhang**, Yihao Liu, Xiao Wang, Bing Li, Weiming Hu, “Learn to Match: Automatic Matching Networks Design for Visual Tracking” in IEEE International Conference on Computer Vision (**ICCV2021**)
- **[First Author] Zhipeng Zhang**, Yufan Liu, Bing Li, Weiming Hu, Houwen Peng, “Towards Accurate Pixelwise Object Tracking via Attention Retrieval” in IEEE Transactions on Image Processing (**TIP2021**)
- **[Corresponding and Co-First Author] Mingzhe Guo***, **Zhipeng Zhang*†**, Heng Fan, “Divert more attention to vision-language object tracking” in IEEE Transactions on Pattern Analysis and Machine Intelligence (**TPAMI2024**)

- **[Corresponding and Co-First Author]** Mingzhe Guo*, **Zhipeng Zhang***†, Heng Fan, Liping Jing, “Cyclic Refiner: Object-Aware Temporal Representation Learning for Multi-View 3D Detection and Tracking” in International Journal of Computer Vision (**IJCV2024**)
- **[Corresponding Author]** Hanshi Wang, **Zhipeng Zhang**†, Weming Hu, “A-Teacher: Asymmetric Network for 3D Semi-Supervised Object Detection” in IEEE Conference on Computer Vision and Pattern Recognition (**CVPR2024**)
- **[Co-First Author]** Chao Liang*, **Zhipeng Zhang***, Xue Zhou, Weiming Hu, “Rethinking the competition between detection and reid in multiobject tracking” in IEEE Transactions on Image Processing (**TIP2022**)
- **[Co-First Author]** Chao Liang*, **Zhipeng Zhang***, Xue Zhou, Weiming Hu, “One More Check: Making “Fake Background” Be Tracked Again” in Association for the Advancement of Artificial Intelligence (**AAAI2021**)
- **[Co-First Author]** Mingzhe Guo*, **Zhipeng Zhang***, Heng Fan, Liping Jing, “Divert more attention to vision-language tracking” in Advances in Neural Information Processing Systems (**NeurIPS2022**)
- **[Co-First Author]** Mingzhe Guo*, **Zhipeng Zhang***, Heng Fan, Liping Jing, Weiming Hu, “Learning target-aware representation for visual tracking via informative interactions” in International Joint Conference on Artificial Intelligence (**Long Oral, IJCAI2022**)
- **[Co-First Author]** Weiming Bai*, **Zhipeng Zhang***, Bing Li, Weiming Hu, “AUNet: Learning Relations Between Action Units for Face Forgery Detection” in IEEE Conference on Computer Vision and Pattern Recognition (**CVPR2022**)
- **[Co-First Author]** Weiming Bai*, **Zhipeng Zhang***, Bing Li, Weiming Hu, “Robust texture-aware computer-generated image forensic: Benchmark and algorithm” in IEEE Transactions on Image Processing (**TIP2021**)
- **[Corresponding and Co-First Author]** Mingzhe Guo*, **Zhipeng Zhang***†, Liping Jing, “End-to-End Autonomous Driving without Costly Modularization and 3D Manual Annotation” in Advances in Neural Information Processing Systems (**NeurIPS2024**) under review with score of 98533
- **[Corresponding Author]** Ziteng Xue, Mingzhe Guo, **Zhipeng Zhang**†, “Occlusion Inhibitor: Multi-View 3D Object Detection by Correlation Learning with Multi-modal Prototypes” in Advances in Neural Information Processing Systems (**NeurIPS2024**) under review with score of 8654
- **[Co-First Author]** Weiming Bai*, **Zhipeng Zhang***, Bing Li, Weiming Hu, “Face Forgery Detection via Exploring Degradation Inconsistency” in (**AAAI2025**) under review
- **[First Author]** **Zhipeng Zhang**, Zhenyu Li, Heng Fan, Yuan He, Ke Wang, Xianming Liu, Junjun Jiang, “Augment and Criticize: Exploring Informative Samples for Semi-Supervised Monocular 3D Object Detection” (Arxiv)

Some Other Publication Work (as co-author)

- Liting Lin, Heng Fan, **Zhipeng Zhang**, Haibin Ling, “Tracking Meets LoRA: Faster Training, Larger Model, Stronger Performance” in European Conference on Computer Vision (**ECCV2024**)
- Liting Lin, Heng Fan, **Zhipeng Zhang**, Haibin Ling, “Swintrack: A simple and strong baseline for transformer tracking” in Advances in Neural Information Processing Systems (**NeurIPS2022**)
- Xiao Wang, Xiujuan Shu, **Zhipeng Zhang**, Bo Jiang, “Towards more flexible and accurate object tracking with natural language: Algorithms and benchmark” in IEEE Conference on Computer Vision and Pattern Recognition (**CVPR2021**)
- Hubin Tan, Xiang Zhang, **Zhipeng Zhang**, “Nocal-Siam: Refining Visual Features and Response with Advanced Non-local Blocks for Real-time Siamese Tracking” in IEEE Transactions on Image Processing (**TIP2021**)
- Linfeng Zhang, Yukang Shi, Hung-Shuo Tai, **Zhipeng Zhang**, “Structured Knowledge Distillation Towards Efficient and Compact Multi-View 3D Detection” in British Machine Vision Conference (**BMVC2023 Oral**)
- Xiao Wang, Jianing Li, Lin Zhu, **Zhipeng Zhang**, Zhe Chen, “Visevent: Reliable object tracking via collaboration of frame and event flows” (**IEEE Transactions on Cybernetics**)
- Yihao Liu, Anran Liu, Jinjin Gu, **Zhipeng Zhang**, Wenhao Wu, Yu Qiao, Chao Dong, “Discovering Distinctive” Semantics” in Super-Resolution Networks” (Arxiv)
- L Peng, J Gao, X Liu, W Li, S Dong, **Zhipeng Zhang**, H Fan, L Zhang, “VastTrack: Vast Category Visual Object Tracking” in Advances in Neural Information Processing Systems (**NeurIPS2024**) under review

Conference Reviewer

- International Conference on Learning Representations (ICLR): 2025
- The IEEE / CVF Computer Vision and Pattern Recognition Conference (CVPR): 2021, 2022, 2023, 2024, 2025
- The Conference on Neural Information Processing Systems (Neurips): 2022, 2023, 2024
- The International Conference on Computer Vision (ICCV): 2023, 2025
- The European Conference on Computer Vision (ECCV): 2022, 2024
- The International Conference on Machine Learning (ICML): 2025
- The AAAI Conference on Artificial Intelligence (AAAI): 2022, 2023, 2024, 2025

Journal Reviewer

- IEEE Transactions on Pattern Analysis and Machine Intelligence (T-PAMI)
- IEEE Transactions on Image Processing (TIP)
- IEEE Transactions on Multimedia (TMM)
- IEEE Transactions on Circuits and Systems for Video Technology (TCSVT)