Zhenyu Li

+966566152875 zhenyu.li.1@kaust.edu.sa https://zhyever.github.io/ Zhenyu Li Zhyever

Highlights

Highly Cited: **2K+** citations in total of my publications as a **third-year** PhD student
Widely Impactful: **~2K** stars on GitHub
Experienced: **2.5+** years of industrial intern experience. 5+ years of research experience.
Prestigious affiliations: Meta, TikTok, DiDi, SenseTime Research
Keywords: computer vision; deep learning; depth estimation; 3D reconstruction, diffusion model; 3D perception and understanding; multi-modality learning; semi-supervised learning; LLM; benchmarking

Experience

Research Intern *Zurich, Switzerland*
Meta *Aug. 2025 - Feb. 2026*

- Researched scalable egocentric body tracking algorithms for AR/VR (from both architecture design and data utilization aspects). It led to a ***CVPR 2026 Highlight*** paper.

Research Intern (*TopSeed* candidate) *Beijing, China*
TikTok *Jan. 2025 - July 2025*

- Researched benchmarking for depth estimation foundation models. The result was presented in the paper.
- Collaborated the development for DepthAnything 3 (***ICLR 2026 Oral***), the SoTA feed-forward 3D reconstruction method.

Elite Research Intern (Elite Camp 2022) *Beijing, China*
Didi Cargo (ADAS team) *Sep. 2022 - Mar. 2023*

- Researched semi-supervised algorithms for monocular 3D object detection. The solution improves the model performance with limited labeled data and was presented in the ***ICRA 2025*** paper.

Research Intern *Shanghai, China*
SenseTime Research (ADAS team) *Jan. 2022 - Jul. 2022*

- Researched unsupervised domain adaptation algorithms for monocular 3D object detection. It led to an ***ECCV 2024*** paper.
- Deployed the aforementioned unsupervised domain adaptation algorithm in an industrial project with GAC Group. Achieved excepted goals.
- Researched domain generalization algorithms for monocular 3D object detection (a follow-up work of the aforementioned paper). It was presented in the paper.
- Collaborated with other team members and researched 3D object detection. It led to an ***ICLR 2023*** paper.

Perception Algorithm Development Intern *Shanghai, China*
SenseTime Research (ADAS team) *Mar. 2021 - Sep. 2021*

- Built up a ReID dataset based on the ground-truth system, trained a ReID model, and developed the ReID model for the ADAS system.
- Built and deployed a multi-object tracking algorithm for the ADAS system (C++), including importing appearance representation from the ReID model and adopting the cascade association strategy.
- Researched multi-modal contrastive learning algorithms for spatial-aware visual representations to benefit 3D-related downstream tasks. It led to an ***AAAI 2022*** paper.
- Collaborated with other team members and researched 3D object detection. It led to an ***IJCAI 2022*** paper and an ***ECCV 2022*** paper.

Selected First Author Publications

- **EgoPoseFormer V2: Accurate Egocentric Human Motion Estimation for AR/VR**. ***CVPR 2026 (Highlight)***. **Zhenyu Li**, Sai Kumar Dwivedi, Filip Maric, Carlos Chacón, Nadine Bertsch, Filippo Arcadu, Tomas Hodan, Michael Ramamonjisoa, Peter Wonka, Amy Zhao, Robin Kips, Cem Keskin, Anastasia Tkach, Chenhongyi Yang. Project link.
- **PatchRefiner V2: Fast and Lightweight Real-Domain High-Resolution Metric Depth Estimation**. ***ICLR 2026***. **Zhenyu Li**, Wenqing Cui, Shariq Farooq Bhat, Peter Wonka. Project link.
- **Amodal Depth Anything: Amodal Depth Estimation in the Wild**. ***ICCV 2025***. **Zhenyu Li**, Mykola Lavreniuk, Jian Shi, Shariq Farooq Bhat, Peter Wonka. Project link.
- **PatchRefiner: Leveraging Synthetic Data for Real-Domain High-Resolution Monocular Metric Depth Estimation**. ***ECCV 2024***. **Zhenyu Li**, Shariq Farooq Bhat, Peter Wonka. Project link.
- **PatchFusion: An End-to-End Tile-Based Framework for High-Resolution Monocular Metric Depth Estimation**. ***CVPR 2024***. **Zhenyu Li**, Shariq Farooq Bhat, Peter Wonka. Project link (1k+ stars on GitHub).
- **Unsupervised Domain Adaptation for Monocular 3D Object Detection via Self-Training.** ***ECCV 2022***. **Zhenyu Li**, Zehui Chen, Ang Li, Liangji Fang, Qinhong Jiang, Xianming Liu, Junjun Jiang
- **SimIPU: Simple 2D Image and 3D Point Cloud Unsupervised Pre-Training for Spatial-Aware Visual Representations**. ***AAAI 2022***. **Zhenyu Li**, Zehui Chen, Ang Li, Liangji Fang, Qinhong Jiang, Xianming Liu, Junjun Jiang, Bolei Zhou, Hang Zhao. Project link.
- **BinsFormer: Revisiting Adaptive Bins for Monocular Depth Estimation.** ***Transactions on Image Processing***. **Zhenyu Li**, Xuyang Wang, Xianming Liu, Junjun Jiang. Project link.
- **DepthFormer: Exploiting Long-Range Correlation and Local Information for Accurate Monocular Depth Estimation.** ***Machine Intelligence Research***. **Zhenyu Li**, Zehui Chen, Xianming Liu, Junjun Jiang. Project link.
- **BenchDepth: Are We on the Right Way to Evaluate Depth Foundation Models?**. **Zhenyu Li**, Haotong Lin, Jiashi Feng, Peter Wonka, Bingyi Kang.
- **CriterAlign: Criterion-Centric Rationale Alignment for Code Preference Judging**. ***Under Review***. **Zhenyu Li**, Aleksandar Cvejic, Zehui Chen, Peter Wonka. Project link.

Selected Co-Author Publications

- **Depth Anything 3: Recovering the Visual Space from Any Views**. ***ICLR 2026 Oral***. Haotong Lin, Sili Chen, Junhao Liew, Donny Y. Chen, **Zhenyu Li**, Guang Shi, Jiashi Feng, Bingyi Kang. Project link.
- **Any Resolution Any Geometry: From Multi-View To Multi-Patch**. ***CVPR 2026***. Wenqing Cui, **Zhenyu Li**, Mykola Lavreniuk, Jian Shi, Ramzi Idoughi, Xiangjun Tang, Peter Wonka. Project link.
- **DissolveStereo: Coarse Depth Injection for Zero-Shot Stereo Video Generation**. ***SIGGRAPH, TOG 2026***. Jian Shi, Qian Wang, **Zhenyu Li**, Wenqing Cui, mboxRamzi Idoughi, Peter Wonka. Code.
- **Lari: Layered Ray Intersections for Single-View 3D Geometric Reasoning**. ***ICML 2026***. Rui Li, Biao Zhang, **Zhenyu Li**, Federico Tombari, Peter Wonka. Project link.
- **Learning from Noisy Data for Semi-Supervised 3D Object Detection**. ***ICCV 2023***. Zehui Chen, **Zhenyu Li**, Shuo Wang, Dengpan Fu, Feng Zhao.
- **BEVDistill: Cross-Modal BEV Distillation for Multi-View 3D Object Detection**. ***ICLR 2023***. Zehui Chen, **Zhenyu Li**, Shiquan Zhang, Liangji Fang, Qinhong Jiang, Feng Zhao. Code.
- **Deformable Feature Aggregation for Dynamic Multi-Modal 3D Object Detection**. ***ECCV 2022***. Zehui Chen, **Zhenyu Li**, Shiquan Zhang, Liangji Fang, Qinhong Jiang, Feng Zhao. Code.

Projects

Monocular Depth Estimation Toolbox *Link*

- Developed several monocular depth estimation methods with reproduced results for fair comparisons
- **900+** stars on GitHub

Education

King Abdullah University of Science and Technology (KAUST) — Thuwal, Saudi Arabia
PhD in Computer Science, Advisor: Prof. Peter Wonka — *Sep. 2023 - Present*

Harbin Institute of Technology (HIT) — Harbin, China
MS in Computer Science, Advisor: Prof. Junjun Jiang. Program rank - 2/80+ — *Sep. 2021 - Jul. 2023*

Harbin Institute of Technology (HIT) — Harbin, China
BS in Computer Science, Advisor: Prof. Junjun Jiang. GPA - 89.58/100 — *Sep. 2017 - Jul. 2021*

Awards

Dean's List Award, award to top students by KAUST CEMSE for their academic achievements 2024
ICCV VCL 2023 Multitask Learning for Robustness Challenge, 1st place 2023
ECCV SSLAD 2022 3D Object Detection Challenge, 3rd Place 2022
ECCV Mobile AI & AIM 2022 Monocular Depth Estimation Challenge, 2nd Place 2022
China National Scholarship 2022

Skills

Tools & Frameworks: Pytorch (profecient), Tensorflow (familiar), OpenMMLab Toolbox, *etc.*
Languages: Chinese (Mother Tongue), English (Fluent)
Teamwork: Can collaborate with team members with diverse backgrounds, resulting in enhanced efficiency.
Problem-solving: Proficient in identifying issues and implementing solutions that exceed project objectives.
Adaptability for Projects: Can take advantages of previous experiences to adapt to various new projects.
Time Management: Expert in prioritizing tasks and managing schedules to maximize productivity.