Zhenyu Li

+966566152875 Homepage Email Google Scholar Linkedin GitHub

Highlights

Publication Record: First-author papers at CVPR, ICCV, ECCV, and ICLR, including a *CVPR 2026 Highlight*, with research spanning depth estimation, 3D perception, reconstruction, and egocentric vision. Recent work further explores LLM-as-a-judge, complemented by collaborative research on generative video, 3D generative modeling, and reward-based generation.
Research Impact: 2K+ citations and 2K+ GitHub stars across open-source research projects.
Industry Experience: 2.5+ years of research internships at **Meta, ByteDance, DiDi, and SenseTime**, with work spanning AR/VR, 3D foundation models, and autonomous driving.
Research Trajectory:

- *2021–2023*: monocular depth estimation, **autonomous-driving** perception across monocular and multi-sensor 3D object detection, and multi-modal 3D representation learning
- *2023–2025*: advanced geometric vision through high-resolution and amodal depth estimation
- *2025–2026*: **3D foundation models**, **articulated-object reconstruction**, post-training for generative models (reinforcement learning, GRPO), and **egocentric** vision for **AR/VR**

Experience

Research Intern *Zurich, Switzerland*
Meta (Meta Reality Lab) *Aug. 2025 - Feb. 2026*

- Designed a scalable egocentric full-body pose estimation system for AR/VR, with a focus on model architecture and data efficiency. The resulting model outperformed the previous SoTA by 22.4% and led to EgoPoseFormer V2, a *CVPR 2026 Highlight* paper.
- Developed a semi-supervised learning pipeline that leveraged 70 million unlabeled egocentric frames, improving accuracy by 12% while maintaining low-latency inference for deployment on AR/VR devices.

Research Intern (*TopSeed* candidate) *Beijing, China*
ByteDance Seed (Foundation Model) *Jan. 2025 - Jul. 2025*

- Led the development of BenchDepth, a downstream-task benchmark for depth foundation models, deriving insights into data scaling, synthetic supervision, and fine-tuning strategies.
- Contributed to Depth Anything 3, a feed-forward 3D reconstruction system and an *ICLR 2026 Oral* paper, through benchmark design, large-scale evaluation and result analysis, and network architecture refinement.

Elite Research Intern (Elite Camp 2022) *Beijing, China*
DiDi Cargo (Autonomous Driving) *Sep. 2022 - Mar. 2023*

- Developed an Augment-and-Criticize framework for semi-supervised monocular 3D object detection, combining robust pseudo-label aggregation with learned dynamic sample selection; improved MonoDLE by 3.32 AP_{3D} points on KITTI and led to 3DSeMo, an *ICRA 2025* paper.

Research Intern *Shanghai, China*
SenseTime Research (Autonomous Driving) *Jan. 2022 - Jul. 2022*

- Developed self-training methods for cross-domain monocular 3D object detection, resulting in STMono3D, an *ECCV 2024* paper, and its domain-generalization extension DGMono3D.
- Deployed the domain adaptation pipeline in an industrial collaboration with GAC Group, meeting project performance targets, and contributed to BEVDistill, an *ICLR 2023* paper.

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

- Built and deployed an ADAS multi-object tracking pipeline, including ReID dataset construction, model training, and a C++ tracker with appearance-based cascade association.
- Researched multimodal representation learning and 3D perception, contributing to SimIPU, AutoAlign, and AutoAlignV2, published at *AAAI 2022*, *IJCAI 2022*, and *ECCV 2022*.

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.
- **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. Project link.
- **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.
- **DivRL: Disentangled Self-Similarity Rewards for Diverse Subject-Driven Generation**. ***ECCV 2026***. Qian Wang, **Zhenyu Li**, Abdelrahman Eldesokey, Peter Wonka.
- **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 2026***. Jian Shi, Qian Wang, **Zhenyu Li**, Wenqing Cui, Ramzi Idoughi, Peter Wonka. Project link.
- **ImmersePro: End-to-End Stereo Video Synthesis via Implicit Disparity Learning**. ***ICML 2026***. Jian Shi, **Zhenyu Li**, Peter Wonka. Project link.
- **Lari: Layered Ray Intersections for Single-View 3D Geometric Reasoning**. ***ICML 2026***. Rui Li, Biao Zhang, **Zhenyu Li**, Federico Tombari, Peter Wonka. Project link.
- **Artic-O: End-to-End Articulated Object Reconstruction via Latent Geometry Learning**. ***Siggraph Asia 2026***. Xuyang Wang, **Zhenyu Li**, Jian Ding, Habib Slim, Peter Wonka, Hongdong Li, *etc.*
- **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 multiple monocular depth estimation methods with reproduced results for fair comparisons
- **900+** stars and 100+ forks 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, KAUST CEMSE, awarded to top students for academic achievement 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