Xu Zheng

PhD Candidate @ HKUST(GZ) | Doctoral Researcher @ INSAIT
Tel: +86 15942357285 | Email: zhengxu128@gmail.com | Website | Google Scholar | LinkedIn

Education

- Institute for Computer Science, Artificial Intelligence and Technology (INSAIT) — Sofia, Bulgaria
  Doctoral Researcher – Advisors: Prof. Luc Van Gool& Dr. Danda Pani Paudel — Feb. 2025 – Present
- The Hong Kong University of Science and Technology (HKUST) — Guangzhou, China
  PhD Candidate – Advisors: Prof. Xuming Hu& Prof. Raymond Chi-Wing Wong — Sept. 2023 – Present
- Northeastern University, School of Computer Science and Engineering — Shenyang, China
  M.Eng. in Information and Communication Engineering – Advisor: Prof. Chong Fu — Sept. 2020 – Jun. 2023
- Northeastern University, School of Computer Science and Engineering — Shenyang, China
  B.Eng. in Communication Engineering – Advisor: Prof. Chong Fu — Sept. 2015 – Jun. 2019

Research Interests

- Multimodal Artificial Intelligence:
  Vision-Language Models: Any2Seg (ECCV 2024) / ExACT (CVPR 2024), etc.
  Foundation Models & Representation Learning: UniBind (CVPR 2024) / EventBind (ECCV 2024).
  Embodied Understanding & Reasoning: OSR-Bench (Arxiv 2025) / 360-R1 (Arxiv 2025).
  Unimodal & Multimodal Bias: DebiasedSeg (ICCV 2025) / MAGIC (ECCV 2024).
- Computer Vision:
  Generative AI: RealRAG (ICML 2025) / TransDiff (Arxiv 2025).
  Scene Understanding: OmniSAM (ICCV 2025) / 360SFUDA++ (TPAMI 2024).
  Knowledge Distillation: CIARD (ICCV 2025) / C2VKD (PR 2024).
  Novel Sensors: DATR (ICCV 2023) / EventDance (CVPR 2024).

Teaching Experience

- Trends in Computer Vision (Graduate Course, English), Teaching Assistant, HKUST (Guangzhou), Spring 2024.
- Applied Statistics (Undergraduate Course, English), Teaching Assistant, HKUST (Guangzhou), Spring 2025.

Publications (Google Scholar Citations: 1K+)

First Author (* denotes co-first author):

1. ICCV 2025. Reducing Unimodal Bias in Multimodal Semantic Segmentation with Multi-Scale Functional Entropy Regularization.

2. ICCV 2025* Highlight Paper. OmniSAM: Omnidirectional Segment Anything Model for UDA in Panoramic Semantic Segmentation.

3. TPAMI 2024. 360SFUDA++: Towards Source-Free UDA for Panoramic Segmentation by Learning Reliable Category Prototypes.

4. ECCV 2024 Oral Presentation. Learning Modality-Agnostic Representation for Semantic Segmentation from Any Modalities.

5. ECCV 2024. Centering the Value of Every Modality: Towards Efficient and Resilient Modality-Agnostic Segmentation.

6. CVPR 2024. EventDance: Unsupervised Cross-Modal Source-Free Adaptation for Event-Based Object Recognition.

7. CVPR 2024. Semantics, Distortion, and Style Matter: Source-Free UDA for Panoramic Segmentation.

8. CVPR 2024*. UniBind: LLM-Augmented Unified and Balanced Representation Space to Bind Them All.

9. ICCV 2023. Look at the Neighbor: Distortion-Aware Unsupervised Domain Adaptation for Panoramic Segmentation.

10. CVPR 2023. Both Style and Distortion Matter: Dual-Path Unsupervised Domain Adaptation for Panoramic Segmentation.

11. ICRA 2024. Transformer–CNN Cohort: Semi-Supervised Semantic Segmentation by the Best of Both Students.

12. PR 2024. Distilling Efficient Vision Transformers from CNNs for Semantic Segmentation.

13. CBM 2022. Uncertainty-Aware Deep Co-Training for Semi-Supervised Medical Image Segmentation.

Corresponding Author:

1. AAAI 2026 Multimodal Robust Prompt Distillation for 3D Point Cloud Models.

2. AAAI 2026 SatireDecoder: Visual Cascaded Decoupling for Enhancing Satirical Image Comprehension.

3. ICCV 2025. CIARD: Cyclic Iterative Adversarial Robustness Distillation.

4. IROS 2025. Unveiling the Potential of SAM2 for RGB–Thermal Semantic Segmentation with Language Guidance.

5. ACM MM Asia 2025. Omnidirectional Spatial Modeling from Correlated Panoramas.

Co-Author:

1. TCSVT 2025. CLIP-to-Seg Distillation for Zero-Shot Semantic Segmentation.

2. NeurIPS 2025. Domain-RAG: Retrieval-Guided Compositional Image Generation for Cross-Domain Few-Shot Detection.

3. NeurIPS 2025. Don't Just Chase "Highlighted Tokens" in MLLMs: Revisiting Visual Holistic Context Retention.

4. ICCV 2025. Unlocking Constraints: Source-Free Occlusion-Aware Seamless Segmentation.

5. ACL 2025 Findings. MMUnLearner: Reformulating Multimodal Machine Unlearning in the Era of MLLMs.

6. ICML 2025. RealRAG: Retrieval-Augmented Realistic Image Generation via Self-Reflective Contrastive Learning.

7. CVPR 2025 @ TMM-OpenWorld Best Paper Award. Benchmarking Multimodal Semantic Segmentation under Sensor Failures.

8. ECCV 2024. EventBind: Learning a Unified Representation for Event-Based Open-World Understanding.

Preprints (Selected)

First Author:

1. Arxiv 2025. Marrying Autoregressive Transformer and Diffusion with Multi-Reference Autoregression.

2. Arxiv 2025. Multimodal Spatial Reasoning in the Large Model Era: A Survey and Benchmarks.

3. Arxiv 2025. MLLMs Are Deeply Affected by Modality Bias.

4. Arxiv 2025. Are Multimodal LLMs Ready for Omnidirectional Spatial Reasoning?

5. Arxiv 2025. Retrieval-Augmented Generation and Understanding in Vision: A Survey.

Honors & Awards

- Liaoning Province Outstanding Master's Thesis – 2023 (Top 2%)
- Liaoning Province Outstanding Graduate – 2023, 2019 (Top 2%)
- China National Scholarship – 2022 (Top 2%)
- Northeastern University First-Class Scholarship – 2020–2022 (Top 20%)

Professional Service

- Reviewer for journals including:
  International Journal of Computer Vision (IJCV), IEEE TIP, IEEE TNNLS, IEEE TMM, IEEE TCI, Neurocomputing, IVC, CBM, MVA, etc.
- Program Committee / Reviewer for:
  CVPR, ICCV, ECCV, NeurIPS, ICML, ICLR, AAAI, ACM MM, etc.