Ziyu Zhao

Mobile: (1)352-871-1917 Email: ziyuz@email.sc.edu Personal website: https://ziyuz-vision.github.io/

EDUCATION

Ph. D in Computer Science	University of South Carolina	*Sep 2021 - present*
MS in Computer Engineering (GPA: 3.72 / 4.0)	University of Florida	*Aug 2019 - May 2021*
BS in Mechanical Engineering (GPA: 91 / 100)	Xidian University	*Sep 2014-May 2018*

RESEARCH INTERESTED

My research focuses on computer vision, particularly multimodal learning, image inpainting, segmentation (2D/3D), and 3D reconstruction.

ACADEMIC EXPERIENCE

DPSeg: Dual-Prompt Cost Volume Learning for Open-Vocabulary Semantic Segmentation *Aug 2024 - Nov 2024*

- Developed a **dual-prompt framework** that fuses **visual** and **textual prompts** to mitigate the **domain gap** in open-vocabulary segmentation.
- Designed a **multi-scale cost volume-guided decoder** and **semantic-guided refinement** to enhance fine-grained spatial accuracy.
- Outperformed prior methods on *ADE20K-847*, *Pascal Context-459*, *COCO-Stuff-171*, and other OVSS benchmarks.

Cross-modal Few-shot 3D Point Cloud Semantic Segmentation via View Synthesis *Jan 2024 – May 2024*

- Proposed a cross-modal few-shot 3D segmentation framework to address **geometry incompleteness** caused by single-view support occlusion.
- Introduced a **multi-view synthesis pipeline** with depth and image inpainting, and a co-embedding network with adaptive prototype weighting.
- Achieved state-of-the-art performance on *S3DIS* and *ScanNet* under 1-shot and 5-shot settings

Few-shot point cloud semantic segmentation based on class-specific Transformer network *May 2023 – Nov 2023*

- Proposed a **stratified class-specific transformer framework** to reduce dependency on labeled 3D data and improve prototype representation.
- Designed a multi-scale transformer architecture that hierarchically aggregates query features conditioned on class-specific support representations.
- Achieved superior accuracy and 15% lower inference time on *S3DIS* and *ScanNet* compared to previous graph-based few-shot segmentation methods.

Cross-modal point cloud semantic segmentation *Aug 2022 – Dec 2022*

- Proposed the **first few-shot framework** that uses labeled 2D images as support to guide 3D point cloud segmentation, reducing 3D annotation cost.
- Converted 2D images into pseudo-3D support via depth estimation, and aligned features through a co-embedding network with masked prototypes.
- Demonstrated competitive results on *S3DIS* and *ScanNet* with minimal supervision.

PROFESSIONAL EXPERIENCE

Reviewer of CVPR, ICCV, AAAI, ECCV, ACM MM, TMM, TPAMI

Research Scientist Intern in PAII Inc. *Feb 2025 – present*

- Research and Development of computer vision technologies with focus on video processing and understanding, specifically **audio-driven image-to-video generation**—animating static images to synchronize with spoken audio, resulting in realistic talking-head videos.
- Machine Learning R & D in the fields of image generation based on *stable diffusion 3.5*.
- Research and development, deep learning modeling, data mining.

Graduate Instructional Assistant, University of South Carolina *Sep 2021 - present*

- Algorithms (Java, Python); General Applications Programming (HTML/CSS and JavaScript).
- Big Data Analytics (graduate level): My responsibilities included conducting hands-on lab sessions, guiding. students through complex concepts, and conducting regular code reviews. For example, I facilitated sessions on decision trees and introduced ensemble methods like Random Forest, emphasizing their relevance in handling large-scale datasets efficiently.
- Computer architecture; Embedded system; Operating system.

Greater University Tutoring Service, University of South Carolina *Sep 2022 – Dec 2022*

Undergraduate Program

- Provided academic tutoring in Calculus, Algebra and Chemistry.
- Provided academic advice to freshmen who had undecided major.

PUBLICATIONS

- **Zhao, Z.**, Li, X., Shi, L., Imanpour, N., Wang, S.
"DPSeg: Dual-Prompt Cost Volume Learning for Open-Vocabulary Semantic Segmentation."
Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2025. (**accepted**)
- **Zhao, Z.**, Li, X., Zhang, C., Cai, P., Wang, S.
"Crossmodal Few-shot 3D Point Cloud Semantic Segmentation via View Synthesis."
Proceedings of the 32nd ACM International Conference on Multimedia (ACM MM), 2024, pp. 2345–2353.
- **Zhao, Z.**, Wu, Z., Wu, X., Zhang, C., Wang, S.
"Crossmodal Few-shot 3D Point Cloud Semantic Segmentation."
Proceedings of the 30th ACM International Conference on Multimedia (ACM MM), 2022, pp. 4760–4768.
- Zhang, C., Wu, Z., Wu, X., **Zhao, Z.**, Wang, S.
"Few-Shot 3D Point Cloud Semantic Segmentation via Stratified Class-Specific Attention Based Transformer Network."
Proceedings of the AAAI Conference on Artificial Intelligence (AAAI), vol. 37, no. 3, 2023, pp. 3410–3417.
- **Zhao, Z.**, Li, X., Zhang, C., Cai, P., Wang, S.
"Leveraging Adaptive Implicit Presentation Mapping for Ultra High-Resolution Image Segmentation."
arXiv preprint, 2025.
- Cai, P., **Zhao, Z.**, Wang, S.
"Efficient Point Cloud Denoising via Direction-Aware Projection."
arXiv preprint, 2025.
- Lu, X., Li, S., **Zhao, Z.**, Xin, B.
"Modeling and Control of WEDM Process of Silicon Single Crystal."
Journal of Mechanical Engineering, vol. 54, no. 17, 2018, pp. 149–156. (ISSN: 0577-6686)

AWARDS

- **Achievement Award Scholarship**, Gainesville, FL *20th Aug 2021*
- **Achievement Award Scholarship**, Gainesville, FL *6th Jan 2021*
- **Achievement Award Scholarship**, Gainesville, FL *20th Aug 2020*
- **Silver Award** in The 3rd China College Students' "Internet Plus" Innovation & Entrepreneurship Competition (July, 2017)

SKILLS

- **Programming**: Python, C, C#, C++, Java, JavaScript, SQL, MATLAB
- **Software:** PyCharm, LT-spice, MATLAB, Mathematica, Keil uVision5, Eclipse, Quartus, ModelSim, Maxterm
- **Deep Learning Framework:** Proficient at PyTorch platform, familiar with TensorFlow, Keras and Caffe
- **Other skills:** Oscilloscope, Digital Multi-meter, Solder gun and paste