Ziyu Zhao

(1)352-871-1917 ziyuz@email.sc.edu ziyuz-vision.github.io

SUMMARY

PhD Computer Vision Engineer specializing in **computational photography** and **3D reconstruction**. Experienced in developing **image processing** (detection, segmentation, restoration) pipelines and collaborating with hardware teams to optimize vision systems for spatial applications. Passionate about advancing the frontier of **vision perception** and **generative AI**.

TECHINICAL SKILLS

- **Languages**: Python (Expert), Java , C++
- **Computer Vision**: Image Processing, Vision Transformers, Multimodal Learning
- **3D Technologies**: Gaussian Splatting, NeRF, COLMAP, Point Cloud Processing, Mesh Optimization, SMPL Human Modeling
- **Frameworks**: PyTorch (Expert), TensorFlow (Proficient), JAX, OpenCV
- **Software & Infrastructure**: AWS, Git, Docker, Linux, CI/CD, REST APIs, Unit Testing

PROFESSIONAL EXPERIENCE

Reviewer : CVPR, ICCV, AAAI, ECCV, ACM MM, TMM, TPAMI, etc.

Research Scientist Intern in PAII Inc. *Jan 2025 – Aug 2025*

◊ **Image Expansion System (Engineering)**

- Implemented **image preprocessing pipeline** (OpenCV) for input normalization and edge-aware padding, ensuring consistent quality for expansion algorithm;
- Collaborated with product engineering team to define image quality thresholds (artifacts limits, resolution standards) and integrated automated validation checks;
- Optimized memory handling for large-resolution inputs, reducing processing failures by ~25% in production deployment.

◊ **Image Asset Tagging Pipeline (Engineering)**

- Built batch **image preprocessing workflow** (color space standardization, resolution normalization) to improve downstream tagging accuracy;
- Developed **error-handling module** to filter low-quality inputs (blurry/dark images), reducing tagging errors by ~20%;
- Documented processing specifications for engineering team handoff during *YouTu* platform deployment.

◊ **Pose-Controlled Generation System (Research)**

- Evaluated diffusion-based frameworks (**Stable Diffusion 3.5, ControlNet**) for pose-conditioned generation;
- Optimized character-specific fine-tuning using **LoRA and DreamBooth**, reducing identity drift by ~25% in perceptual metrics while maintaining brand-color fidelity (**$\Delta E < 6$**);
- Integrated pose-aware alignment module with **OpenCV-based geometric warping**, ensuring visual consistency between generated characters and background scenes across multi-pose inputs.

RESEARCH EXPERIENCE

Dynamic Scene Reconstruction with 4D Gaussian Splatting *Mar 2025 – Aug 2025*

- Engineered **memory-optimized temporal coherence module** for 4D reconstruction pipeline, reducing VRAM usage by ~25% through hierarchical feature caching and enabling robust handling of moving elements (people, curtains, lighting changes) in spatial scans

- Designed **background-foreground decoupling strategy** using motion-aware segmentation, reducing reconstruction artifacts by 35% in scenes with dynamic objects
- Implemented feature alignment between geometry and appearance streams, improving visual consistency across novel viewpoints (PSNR +2.1dB)

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

- Proposed **cross-view consistency constraints** to resolve occlusions in sparse reconstruction scenarios (e.g., furniture behind pillars, partially visible fixtures)
- Integrated dual-prompt cost volume enabling zero-shot semantic understanding of novel objects without category-specific training
- **Achieved +12.3 mIoU over SOTA on various benchmarks.**

Efficient Semantic Enrichment for Sparse 3D Scans

- Developed **view synthesis pipeline** to generate multi-view inputs from sparse point clouds, improving segmentation robustness in occluded regions by 37%
- Designed **photometric consistency loss** to align synthesized views with real captures, reducing annotation noise for downstream reconstruction tasks
- Validated on **Habitat - Matterport 3D dataset**: enabled accurate room-type classification (kitchen/bathroom) with minimal supervision

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 Computer Vision and Pattern Recognition Conference (CVPR), 2025, pp. 25346-25356

- **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.

- 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.

- 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.**, 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.

- **Zhao, Z.**, Li, X., Zhang, C., Cai, P., Wang, S.
"Leveraging Adaptive Implicit Presentation Mapping for Ultra High-Resolution Image Segmentation." arXiv

- Cai, P., **Zhao, Z.**, Wang, S.
"Efficient Point Cloud Denoising via Direction-Aware Projection." arXiv

- 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)

EDUCATION

- **Ph.D. in Computer Science**, University of South Carolina *(Sep 2021 – **Expected** May 2026)*
- **M.S. in Computer Engineering (GPA: 3.72 / 4.0)**, University of Florida *(Aug 2019 – May 2021)*
- **B.S. in Mechanical Engineering (GPA: 91 / 100)**, Xidian University *(Sep 2014 – May 2018)*