Zichen Zhao

Ph.D. Applicant

✉ jeremy0228@nuaa.edu.cn   📞 18742311211

📍 Nanjing University of Aeronautics and Astronautics, CEIE, Nanjing, China

Education

- **B.S. in [Electronic Information Science and Technology]**   Expected Jun 2026
  - Nanjing University of Aeronautics and Astronautics, College of Electronic Information and Engineering   • GPA: 3.5/5.0 (Ranked top 28% in major)

Research Experience

- **Research Assistant**   Mar. 2025 - Present

  - Project: HandsOnVLM-based Hand-Object Interaction, supervised by Prof. Wentong Li. Focused on enhancing the model's trajectory prediction ability in embodied interaction tasks.
  - Contribution: Improved the Trajectory Decoder by introducing a diffusion-based mechanism to increase prediction diversity and accuracy, and explored its applicability to VLA tasks.
  - Progress:Conducted literature review and reproduction of diffusion-based trajectory models; proposed applying diffusion processes to enhance feature representation, in order to strengthen model expressiveness and performance. Through this work, I gained a deeper understanding of generative modeling principles, developed stronger implementation and problem-solving skills, and learned how to bridge theoretical ideas with practical innovations. currently replacing the module within HandsOnVLM and performing experimental evaluation.
  - Expected Outcome: The work is planned to be submitted to CVPR 2025 in November.

- **Independent Project: Improving YOLO11 Detection Head with Progressive Multi-Scale Feature Fusion for Small Object Detection**   Mar. 2025 – Present

  - Project Overview: Focused on enhancing YOLO11's performance on small object detection, particularly in aerial imagery scenarios where recognizing tiny objects is challenging.
  - Contribution: Modified the original YOLO11 architecture in which the backbone generates P1–P5 feature maps but the detection head only uses P3–P5. Designed an improved detection head supporting P2–P5 inputs, inspired by AFPN, and introduced a progressive fusion strategy to better integrate multi-level features and mitigate semantic gaps.
  - Results: On the VisDrone2019 dataset, achieved performance improvements with mAP50 rising from 34.8% to 38.1% and mAP50–95 increasing from 20.4% to 22.9%.

Technical Skills

- **Programming Languages:** Python, MATLAB, C++, Golang,

- **Software & Tools:** PyTorch, TensorFlow, LaTeX, Git, Docker,

- **Languages:** English (Professional working proficiency, CET-6: 500)

Awards & Honors

- **Lanqiao Cup Software Competition (C/C++ A Group)** 15th Edition, Jiangsu Province Third Prize, 2024

- **Nanjing University of Aeronautics and Astronautics Math Modeling Competition** Second Prize, 2024

- **Huashu Cup National College Student Mathematical Modeling Competition** Excellence Award, 2024

- **Academic Excellence Scholarship** 2022–2025 (consecutive three academic years)

- **Outstanding Student Scholarship** 2022–2025 (consecutive three academic years)

Internship Experience

- **Zhejiang University CS Summer Camp (Research Program)** Jul. 2025 – Aug. 2025 Hangzhou, China

  - Developed features for text-to-image, multi-image generation, short video, and long video generation, along with corresponding front-end pages.
  - Built WeChat MiniProgram pages to address issues with non-navigable links on the mini program and unsupported online playback in Coze, by calling APIs from workflows deployed on Coze.
  - Integrated Volcano Engine Object Storage to temporarily store media resources and enable asynchronous processing of long videos, and developed custom plugins in Coze IDE to utilize these services.