

Zhihan Zhu

1088 Xueyuan Avenue, Shenzhen, 518055, P.R. China | 12312326@mail.sustech.edu.cn | +86 150 0404 5391

Google Scholar | github.com/airtnzhuzh | my website

Education

Southern University of Science and Technology, B.S. in Electronic and Electrical Engineering (Expected June 2027) Sept 2023 – present

- Weighted Average Score: 91.63
- GPA: 3.85/4.0
- Student Ranking: 10/46
- Coursework: Linear Algebra, Probability and Statistics, Engineering Mathematics, AI and Machine Learning

Papers

Runge-Kutta Approximation and Decoupled Attention for Rectified Flow Inversion and Semantic Editing Sept 2025

Weiming Chen, **Zhihan Zhu**, Yijia Wang, Zhihai He

arXiv:2509.12888 [cs.CV], 2025. <https://doi.org/10.48550/arXiv.2509.12888>

Summary: Proposed a high-order inversion method for rectified flow models using a Runge–Kutta solver, and introduced Decoupled Diffusion Transformer Attention (DDTA) to disentangle multimodal attention for more precise semantic control. Achieved state-of-the-art fidelity and editability in image inversion and editing tasks.

Generative Semantic Coding for Ultra-Low Bitrate Visual Communication and Analysis Oct 2025

Weiming Chen, Yijia Wang, **Zhihan Zhu**, Zhihai He

arXiv:2510.27324 [cs.CV], 2025. <https://doi.org/10.48550/arXiv.2510.27324>

Summary: Developed a generative semantic coding framework integrating image generation with deep compression via rectified flow models. Enables ultra-low bitrate visual communication while maintaining high reconstruction quality and analysis accuracy for remote vision and interaction systems.

Patents

[1] Zhihai He, Weiming Chen, **Zhihan Zhu**, Yijia Wang, Jian Ouyang, *Image Editing Method Based on Attention Decoupling, Device, Terminal and Storage Medium*. App. No.: CN 202511281645.1, accepted by CNIPA.

[2] Zhihai He, Weiming Chen, Yijia Wang, **Zhihan Zhu**, *Template-Replacement Image Compression & Reconstruction Method, System and Storage Medium*. App. No.: CN 202511281423.X, accepted by CNIPA.

Research Experience

Diffusion Inversion & Image Editing Sept 2024 – May 2025

- Gained an in-depth understanding of inversion techniques for UNet-based and flow-based diffusion models, and proposed a higher-order inversion approach inspired by the Runge–Kutta method.
- Reimplemented and benchmarked 10+ diffusion inversion and image editing methods on benchmarks(PIE bench, DCI datasets), analyzing current challenges in diffusion-based image editing tasks.

Diffusion Transformer (DiT) & Flow-based Model Oct 2024 – present

- Reproduced the DiT architecture from scratch and studied its internal structure and design principles in depth.
- Built precise attention-level editing controls on top of FLUX by implementing hook functions to manipulate FLUX attention maps for fine-grained image editing.

One-step Diffusion for Super-Resolution May 2025 – present

- Reimplemented *MeanFlow* and developed an auxiliary **plug-in SR module** for image super-resolution; independently wrote and trained the full pipeline on HPC with 8×A100 GPUs, iterating through multiple rounds

of optimization and debugging.

- Designed several plug-in SR modules (based on MeanFlow and SDXL-Turbo) to enable one-step super-resolution conditioned on reference images, leveraging hook-based integration within the SDXL-Turbo pipeline.
- Constructed a custom super-resolution dataset based on DIV2K and Flickr2K to serve as both the training set and a reference dictionary for model development.

Projects

*Note: Several research-oriented projects with published results are presented in the **papers** section.*

One-step Diffusion for Image Super-Resolution

May 2025 – present

- Developed several **plug-in SR modules** based on MeanFlow and SDXL-Turbo to achieve one-step super-resolution.
- Implemented full training and evaluation pipelines; trained models on HPC clusters with 8×A100 GPUs.
- Created a custom SR dataset derived from DIV2K and Flickr2K for model training and reference retrieval.
- **Tools Used:** PyTorch, CUDA, HPC (LSF)

2048 Game Project — [CS109] Introduction to Computer Programming

Mar 2024 – Jun 2024 (2024 Spring Semester)

- Developed an intelligent **AI agent** for the 2048 game capable of autonomously playing the 2048 game using traditional algorithmic methods.
- Selected as one of the **Top 5% Excellent Projects** in the course.
- Open-sourced implementation on GitHub: github.com/airtnzhuzh/CS-109-Project
- **Tools Used:** Java, JavaFX, Algorithm Design, Git

Deep Learning-based Modulation Recognition

Sept 2024 – Dec 2024 (2024 Fall Semester)

- Built a **CNN-based system** to classify QPSK, 8-PSK, and 64-QAM modulation types with **100% accuracy**.
- Implemented full data generation and training pipelines using MATLAB and parallel computing.
- Visualized results with confusion matrices and signal spectrograms.
- **Tools Used:** MATLAB, Deep Learning Toolbox, Signal processing

“Yunjing Zhiyu” — Intelligent Aquaculture System

2025 – Guangdong

China International College Students’ Innovation Competition 2025

Provincial Gold Award

- Designed an intelligent aquaculture device capable of automatic fish feeding, algae removal, and real-time water temperature monitoring.

Awards & Honors

- **Gold Award (Guangdong Provincial Level, TOP 0.02%), “Yunjing Zhiyu” — Intelligent Aquaculture System, China International College Students’ Innovation Competition 2025**
- **6th Place**, Men’s 400m, 9th Southern University of Science and Technology Athletics Games
- **Outstanding Student**, Southern University of Science and Technology (SUSTech), 2025
- **Second-Class Scholarship for Academic Excellence**, Southern University of Science and Technology (SUSTech), Academic Year 2024–2025
- **Outstanding Student**, Southern University of Science and Technology (SUSTech), 2024
- **Third-Class Scholarship for Academic Excellence**, Southern University of Science and Technology (SUSTech), Academic Year 2023–2024

Technologies

Programming: Python, Java, C

Technologies: PyTorch, Hugging Face Diffusers, NumPy, OpenCV, Matplotlib, Stable Diffusion, FLUX, DiT (Diffusion Transformer), HPC (LSF, Multi-GPU Training), Git, Linux, LaTeX, STM32, Embedded Systems, Microcontroller, Android development

Language: Chinese (native), English (IELTS 6.5, taken Aug 16, 2023; certificate expired)