Zhanhe Shi

New York University

✉ zhanhe.shi@outlook.com | ⌂ bio.zhanheshi.com | ⭘ github.com/saltfish-len

Education

New York University   *New York, NY, USA*

Master of Science in Information Systems (MSIS)   *Sep 2025 - Present*

- **GPA: 4.0/4.0**
- **Courses:** Fundamental Algorithms, Data Communications & Networks, Data Science & AI for Business
- **In Progress:** Deep Learning, Database Systems, The Global Economy

University of California, Berkeley   *Berkeley, CA, USA*

College of Engineering, Berkeley Extension Global Exchange Program, Computer Science   *Aug 2023 - May 2024*

- **GPA: 3.83/4.0**
- **Courses:** Designing, Visualizing and Understanding Deep Neural Networks, Intro to Computer Vision and Computational Photography

ShanghaiTech University   *Shanghai, China*

School of Information Science and Technology, Bachelor of Computer Science and Technology   *Sep 2021 - Jun 2025*

- **Courses:** Machine Learning, Econometric Analysis Methods and Modeling, Mathematical Modeling
- **Honors & Awards: Outstanding Student, 2023-2024**

Internship Experience

Hundsun Technologies Inc.   *Hangzhou, Zhejiang, China*

Intern, Junior Software Engineer   *May 2024 - Jul 2024*

- Conducted research on existing major datasets and related models for document layout analysis.
- Evaluated the accuracy and performance of several open-source models on the Chinese dataset CDLA.
- Retrained LayoutLMv3, YOLO, and VGT models, comparing their performance and accuracy.

Research & Project Experience

Multiple Human-Object Interaction Generation with Conditional Diffusion   *Shanghai, China*

Graduation Project   *May 2025*

- Proposed a generation framework for multiple human-object interaction, implemented a temporal Transformer with a conditional diffusion model.
- Incorporated PointNet++ object point-cloud encoding and a discrete quantity encoder to inject geometry and entity-count priors.
- Implement an AdaLN-Zero modulated temporal fusion Transformer to condition feature normalization on diffusion timesteps.

HOI-M3 : Capture Multiple Humans and Objects Interaction within Contextual Environment   *Shanghai, China*

Third Author, Frontier Science Research Base on Intelligent Human-Machine Collaboration and Interaction   *Aug 2023 - Mar 2024*

- Assisted in capturing a multi-human multi-object interaction dataset.
- Designed and developed a Segment Anything Model based annotation tool to track masks for people and objects.
- Employed ViTPose to detect single-view human body keypoints and performed multi-view matching to optimize the human body SMPL model.
- **Accepted by CVPR 2024 (Highlight)**. (arXiv:2404.00299v2 [cs.CV] 2 Apr 2024)

FGSM-Based Attack on SAM Model   *Berkeley, CA, USA*

Team Leader, University of California, Berkeley   *Apr 2024*

- Implemented FGSM (Fast Gradient Sign Method) to attack SlimSAM model and generate adversarial samples.
- Conducted quantitative evaluations of adversarial attacks on SlimSAM by calculating mIoU between original and adversarial masks.
- Developed a Gradio-based web interface to support image uploads and visualize attack masks based on selected weights.

3D Character Generation Using ControlNet and LoRA   *Berkeley, CA, USA*

Team Member, University of California, Berkeley   *Nov 2023 - Dec 2023*

- Used LoRA to fine-tune the existing text-to-image diffusion model, enhanced the consistency of characters features.
- Combined ControlNet with 3D human pose solution, controlled the spatial consistency in multi-view images through depth and key points.
- Generated 3D character models using Gaussian splatting.

Skills

Programming   Python (NumPy, Matplotlib, PyTorch, OpenCV), C/C++, MATLAB

Miscellaneous   Linux, LaTeX, Microsoft Office, Git, Tencent Cloud, Simulation of Urban Mobility, Wind, Blender