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, Deep Learning, Data Science & AI for Business, Database Systems, The Global Economy

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

College of Engineering, Computer Science — *Aug 2023 - May 2024*

- **GPA: 3.83/4.0**
- **Courses:** Deep Neural Networks, Intro to Computer Vision and Computational Photography, Optimization Model in Engineering

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**

Research & Project Experience

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

Graduation Project, ShanghaiTech University — *May 2025*

- Proposed an MPMO diffusion framework using SMPL parameters and 6D rotation, achieving a **FID of 8.95** on selected dataset.
- Introduced kinematic constraints (velocity/acceleration) to enforce physical plausibility, **enhancing temporal coherence** by up to 24.6%.
- Integrated PointNet++ object encodings into an **AdaLN-Zero modulated DiT**, maintaining temporal feature consistency and reducing position errors.

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*

- Co-created a large-scale MPMO interaction dataset comprising **(181M+ frames)** captured via synchronized RGB and object-mounted IMUs.
- Engineered a SAM-based masking tool, generating pixel-accurate initial masks to bootstrap the **whole 42 views tracking pipeline** across 199 sequences.
- Constructed 3D human ground-truth for all **199 sequences**, optimizing SMPL parameters via ViTPose multi-view keypoint matching and tracking.
- **Accepted by CVPR 2024 (Highlight - Top 3% of all submissions)**. (arXiv:2404.00299v2 [cs.CV])

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

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

- Led the reproduction of a pipeline to generate **FGSM** adversarial perturbations against SlimSAM, benchmarking model robustness.
- Evaluated attack effectiveness on a 100-image subset of the SA-1B dataset, successfully degrading segmentation performance by **dropping mIoU to 79.32%**.
- Built a Gradio **interactive demo** enabling real-time parameter tuning and side-by-side visualization for qualitative inspection of attack failure modes.

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

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

- **Fine-tuned SD1.5 via LoRA** to enforce character identity consistency across generations, reduced FID by 15%
- Integrated ControlNet with **3D human pose solutions** to enforce spatial consistency, successfully mitigating multi-view geometric artifacts.
- Engineered the diffusion-based **Depth Control pipeline**, projecting depth priors via camera extrinsics to regulate body proportions for Gaussian Splatting.

Internship Experience

Hundsun Technologies Inc. — *Hangzhou, Zhejiang, China*

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

- Constructed a **Document Layout Analysis benchmarking pipeline** on the Chinese CDLA dataset, evaluating 7 open-source and commercial models.
- Fine-tuned state-of-the-art multimodal networks (LayoutLMv3, VGT) and object detection models (YOLO), achieving a peak **mAP@50:95 of 88.4** with VGT.
- Quantified the speed-accuracy trade-off, demonstrating YOLO achieved a **103× speedup** with a 10% mAP drop compared to the VGT model.

Skills

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

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