Yunkai Ji

yj2927@nyu.edu | linkedin.com/in/yunkai-ji | github.com/gyunkai

Education

New York University Shanghai Aug 2023 – May 2027

Bachelor of Science in Computer Science

- GPA: 3.8/4.0
- Research interests: vision–language models, computer-use agent, diffusion-based generative models, and representation learning.

New York University Aug 2025 – May 2026

Study Away

- GPA: 3.8/4.0
- Coursework (Graduate): computer vision, natural language processing.

Research and Projects

LoRA Representation Learning June 2025 – Present

Research Assistant, supervised by Prof. Hongyi Wen, MAPS Lab, NYU Shanghai

- Proposed a novel black-box representation learning framework that encodes LoRA adapters into **Difference Vectors** by analyzing the semantic shift between base and adapted model outputs.
- Designed an **Iterative Probe algorithm** with dynamic stopping criteria (variance stability and cosine direction) to extract robust representations with minimal computational cost.
- Formulated the **LoRA Quality Index (LQI)**, an unsupervised metric combining cosine mean shifts and PCA eigenvalue concentration to quantify generation quality without human annotation.
- Developed a downstream classification pipeline to auto-tag LoRAs (Content vs. Style) and validated the metric's effectiveness by correlating LQI with Civitai popularity on the ImageGem dataset.

Flow-Based Diffusion Model Analysis June 2025 – Dec 2025

Research Assistant, supervised by Prof. Hongyi Wen and Prof. Shengjie Wang, NYU Shanghai

- Contributed to the empirical analysis of flow matching models, assisting in identifying the **two-stage training dynamics** (Navigation vs. Refinement).
- Explored **decoder-only training strategies** and noise injection techniques to enhance fine-grained image generation.
- Conducted extensive ablation studies on ImageNet generation, optimizing **FID scores** through customized training objectives and verifying the trade-off between memorization and generalization.

Text-Native Vision: Do LLMs Understand ASCII Images? | Project Paper Sep 2025 – Dec 2025

CSCI-GA 2271 Computer Vision Final Project, mentored by Prof. Saining Xie

- Constructed an MNIST/ImageNet-to-ASCII benchmark to evaluate text-only visual recognition, revealing that SOTA LLMs (e.g., Llama-3, Gemini) perform near random guessing despite rigorous tokenization controls.
- Proposed a **stride-aware 2D relative positional embedding** scheme that injects row-column inductive biases into attention mechanisms, effectively recovering spatial structure from flattened character sequences.
- Engineered a knowledge distillation pipeline transferring CLIP visual semantics into a lightweight **51.6M-parameter** student model, achieving **39% accuracy** on ASCIIEval and significantly outperforming 70B+ parameter baselines.

Publications and Preprints

From Navigation to Refinement: Revealing the Two-Stage Nature of Flow-based Diffusion Models through Oracle Velocity

Haoming Liu, Jinnuo Liu, Yanhao Li, Liuyang Bai, **Yunkai Ji**, Yuanhe Guo, Shenji Wan, Hongyi Wen

Under review (CVPR 2026)

Technical Skills

Languages: Python, C/C++
ML: PyTorch, Hugging Face (diffusers/transformers), Weight&Bias
Data: NumPy, pandas, Matplotlib
Tools: Git, LaTeX