Junhao Zhu

jz5770@nyu.edu ◇ Github: JunhaoZhu0220 ◇ Personal Page

EDUCATION

New York University Shanghai — Shanghai, China
B.S. in Computer Science, Minor in Interactive Media Arts — 2022 – 2026
Cumulative GPA: 3.728 Major GPA: 3.845
Relevant Coursework: Machine Learning (A), Discrete Mathematics (A), Probability and Statistics (A)

New York University — New York City, NY
Sep 2024 – Dec 2024
Relevant Coursework: Computer Graphics (Graduate Level) (A), Linear Algebra (A)

New York University Abu Dhabi — Abu Dhabi, United Arab Emirates
Jan 2025 – May 2025
Relevant Coursework: Computer Vision and Pattern Recognition (A), Artificial Intelligence and Human Decision (A-)

RESEARCH EXPERIENCE

Syntax-Guided Cascade Discrete Diffusion for Text Generation — Jun 2025 – Present
Advisor: Prof. Shengjie Wang & Prof. Hongyi Wen — New York University Shanghai, China

- Devised a prioritization scheme that parses whole sentences with constituency trees, assigning a 3-tier importance level to each token.
- Designed a 3-stage cascade diffusion architecture where each stage exclusively generates tokens of a specific priority level, effectively building the sentence from a core grammatical structure to full contextual detail.

Gaussian Quantized Variational Bayesian for Continuous and Discrete Generation — Oct 2024 – Present
Advisor: Prof. Shengjie Wang & Prof. Hongyi Wen — New York University Shanghai, China

- Introduced a **Gaussian Quantization (GQ) method** to resolve the critical issue of codebook collapse in variational autoencoders, enabling higher codebook engagement and richer latent representations.
- Integrated a collapse-robust autoencoder into autoregressive decoders (Transformer, Mamba), delivering 21% and 76% improvements in face-image generation quality across two comparable codebook configurations

Continuous Diffusion with VQ-VAE for Symbolic Music Generation — Jun 2024 – Oct 2024
Advisor: Prof. Shengjie Wang & Prof. Hongyi Wen — New York University Shanghai, China

- Developed a novel tokenization scheme that directly maps piano audio to a bar-aligned symbolic representation, bypassing the need for complex, multi-stage processing pipelines.
- Integrated this tokenizer with a standard Diffusion Transformer (DiT), creating a streamlined, non-hierarchical framework for symbolic music generation.
- This approach outperformed a baseline hierarchical model by achieving a **3x faster inference speed** while preserving generative quality, demonstrating a more efficient paradigm for high-fidelity music synthesis.

PROJECTS

Audio Classification with Residual Network — Nov 2023
CSCI-SHU 360 Machine Learning

- Developed a custom Residual Network (ResNet) in **PyTorch** to classify gender from audio signals, achieving **80% accuracy** on a dataset of 5,000 song clips.
- Secured a **top 10% ranking** among all participants, demonstrating high model performance and effective implementation.

HONORS

Dean's List for Academic Year, **NYU Shanghai** — AY 2024 & 2025

SKILLS

Programming Languages: Python, R, Javascript, C, LaTeX, Markdown, HTML
Skills: Git, Docker, High Performance Computing
Machine Learning: PyTorch, Numpy, Pandas, NLTK, Transformers