Sihan Ji

jq2099@sjtu.edu.cn | 2199-rgb.github.io | github.com/2199-rgb

Open to research internships and collaborations

Education

Shanghai Jiao Tong University Sept. 2024 – June 2028 (Expected)

B.Eng. in Information Engineering, School of Information and Electronic Engineering *Shanghai, China*

- **GPA:** 3.79/4.3 | **Rank:** 34/144 overall, 23/144 in 2025–26 | **Comprehensive Evaluation:** 16/144 in 2025–26
- **Research Interests:** Image Restoration, Video Editing and Generation, World Models, Diffusion Transformers

Publications & Manuscripts

Sihan Ji, Wenxin Chen, Yue Shi, Jiezhang Cao. *Recovering What Matters, Transferring Only What Is Reliable: Toward Precise Reference-Based Super-Resolution.* arXiv:2611.33455, 2026. **Under review at CVPR 2027.** [arXiv]

Sihan Ji, Yue Shi, Wenjun Zhang, Jiezhang Cao. *TRACE: Training-Free Temporal Reference Attention for Correspondence-Aware Editing without Edited Anchor Frames.* **Under review at AAAI 2027**.

Research Experience

Undergraduate Research Assistant, Deep Pixel Group Apr. 2026 – Present

Shanghai Jiao Tong University, advised by Prof. Jiezhang Cao *Shanghai, China*

- Research on **image restoration, video editing and generation**, reusing pretrained generative priors for reference-guided generation.
- Work with **Diffusion Transformers (DiTs)**, rectified-flow models, attention-based correspondence, multimodal conditioning and reference feature transfer.
- Led two first-author projects: **TRACE** (AAAI 2027, under review) and **PreciseRef** (submitted to CVPR 2027).

Undergraduate Research Assistant, School of Microelectronics Sept. 2025 – Feb. 2026

Shanghai Jiao Tong University *Shanghai, China*

- Implemented hardware acceleration for **neural-network activation functions** in **Verilog**.

Research Projects

PreciseRef: Precise Reference-Based Super-Resolution | *First Author* Under review at CVPR 2027

- Proposed **PreciseRef**, a generative RefSR framework on a pretrained **FLUX DiT** that recovers reference-specific content (text, codes, identity cues) faithfully rather than merely plausibly.
- Designed **Asymmetric Twin-Attention (ATA)**: target and reference tokens share one joint attention space while the reference keeps its own shape, enabling fine-grained retrieval **without explicit alignment**.
- Introduced **Dual-Scale Reference Gating (DRG)**: Semantic Relevance Gating filters globally irrelevant references, and Bidirectional Consistency Gating suppresses unreliable local correspondences.
- Built **VeriRef**, a benchmark for exact reference-dependent recovery and robustness to mismatched references. Reached **85.6% character accuracy** and **76.3% exact code decoding** (vs. 17.4% and 0% for prior methods), with state-of-the-art perceptual quality on **CUFED5** and **WR-SR**.

TRACE: Temporal Reference Attention for Video Editing | *First Author* Under review at AAAI 2027

- Developed a **training-free** framework that transfers the identity and appearance of a reference object into a source video, preserving motion and structure **without an edited anchor frame**.
- Proposed **Temporal Reference Attention**, which treats the reference image as a **virtual temporal frame** so that pretrained video-diffusion attention builds implicit correspondence between the reference and every video frame.
- Designed a source-identity suppression strategy combining re-noising, conditional feature manipulation and **flow-based inversion** to stop the original object's identity from leaking into the edit.
- Improved reference consistency (**DINO-I 0.5503 → 0.6257**, **CLIP-I 0.7319 → 0.7725**) while keeping motion consistency above **0.98**.

Technical Skills

Programming: Python, C/C++, MATLAB, Verilog, Bash

Deep Learning: PyTorch, Diffusion Models, Rectified Flow, Diffusion Transformers (DiT), Multimodal Models

Tools: Linux, Codex, Claude Code, LaTeX, CUDA / GPU Computing