Yufan Zhou

China | yfzhou@stu.hit.edu.cn | 19819825480 | wiserzhou.github.io | github.com/WiserZhou

Education

Harbin Institute of Technology, BE in Software Engineering   Sep. 2022 – Jun. 2026
- GPA: 3.9/4.0; Rank: 2/231
- **Technical Skills:** Python; Java; C/C++; PyTorch; Software Development; Linux

Research Experience

Internship, Westlake University, ENCODE Lab by Huan Wang   May. 2024 – Nov. 2024
- Conducted foundational research in Generative AI for image synthesis, contributing to advancements in personalization, dataset generation for knowledge distillation, scalable data expansion, and concept blending techniques to improve generative model performance and versatility.

Internship, VAST & HKU, MMLab by Yuan-Pei Cao & Xihui Liu   Mar. 2025 - Jun. 2025
- Investigated 3D part-aware generation, focusing on producing high-quality 3D assets with precise part segmentation, enabling more structured representation and controllable generation for downstream 3D understanding tasks.

Internship, NTU, MMLab by Xingang Pan   Jul. 2025 - Oct. 2025
- Explored 3D generation and reconstruction pipelines integrated with 6D pose estimation, developing methods to fuse generative modeling with geometric reconstruction and accurately estimate object poses from generated or reconstructed 3D scenes.

Publications

OmniPart: Part-Aware 3D Generation with Semantic Decoupling and Structural Cohesion   SIGGRAPH Asia 2025

Yunhan Yang*, *Yufan Zhou*, Yuan-Chen Guo, Zi-Xin Zou, Yukun Huang, Ying-Tian Liu, Hao Xu, Ding Liang, Yan-Pei Cao, Xihui Liu

- Proposed OmniPart, a two-stage part-aware 3D generation framework that achieves semantic decoupling of individual components and enforces structural cohesion across parts, enabling controllable and coherent multi-part 3D objects.
- Designed a structure planning module (autoregressive part bounding box predictor) and a spatially-conditioned latent generation module, ensuring parts are semantically distinct yet geometrically consistent within the overall object layout.

Masked Temporal Interpolation Diffusion for Procedure Planning in Instructional Videos   ICLR 2025

Yufan Zhou, Zhaobo Qi, Lingshuai Lin, Junqi Jing, Tingting Chai, Beichen Zhang, Shuhui Wang, Weigang Zhang

- Proposed a latent space temporal interpolation module to generate intermediate latent features, enhancing visual supervision with richer mid-state details for task-specific end-to-end training.
- Introduced a task-adaptive masked proximity loss to improve temporal coherence, prioritize accurate reasoning near start and end states, and filter out irrelevant action predictions for contextually aware sequences.

FreeBlend: Feedback-driven Latent Interpolation Diffusion for Concept Blending   NeurIPS Workshop 2025

Yufan Zhou, Haoyu Shen*, Huan Wang

- Proposed a novel framework for feedback-driven latent interpolation to enable effective concept blending, improving both semantic coherence and diversity in the generated outputs.
- Addressed the challenges of blending two concepts with distinct semantics and dissimilar shapes, ensuring meaningful and visually cohesive results.

Projects

LLM agent for Virtual Digital Person QA-question

- Developed an LLM-powered agent software that allows users to ask questions about travel experiences in specific cities. The agent provides professional and well-informed answers and presents them through a dramatic animated virtual digital person who teaches and explains the content in an engaging manner. The software offers both Vue.js and Android interfaces, utilizing the Baichuan-2 model along with Retrieval-Augmented Generation (RAG) technology to deliver dynamic context-aware responses.
- Tools Used: Python, PyTorch, Vue, JavaScript

AutoDroid for AccessibilityService AutoDroid

- AutoDroid is an intelligent smartphone automation assistant that uses large language models and accessibility services to automate tasks across any app, focusing on privacy protection, ease of use for disabled users, and optimizing system performance through innovative interface design, knowledge integration, and request efficiency.
- Tools Used: Java, XML, Android

Simple Paging Browser Framework Paging-Browser

- Focused on developing a simple, pagination-based browser using the Qt framework, including web browsing, bookmark management, and special effects handling.
- Tools Used: Qt, C++, MySQL

Wholesale and Retail Management System Wholesale System

- Designed for wholesalers and retailers, featuring a wide range of functionalities including user, department, role, and menu management, along with order, product, and supplier management, all supported by Vue, Spring Boot, Redis, and JWT for efficient multi-terminal authentication and system configuration.
- Tools Used: Java, XML, Spring Boot, Redis, Vue

Award

- National Scholarship
- Finalist, Mathematical Contest in Modeling
- National Third Prize in Computer Design Contest
- National Third Prize in Physics Experiment Competition
- First Prize, the Olympic Mathematics Competition
- First-Class Scholarship
- Huawei Smart Base Scholarship
- Second Prize in Software Innovation
- Bronze Medal, ACM Programming Contest for Freshmen
- Second Prize in the Blue Bridge Cup
- Third Prize in the China College Students Service Outsourcing Innovation and Entrepreneurship Competition
- Third Prize in the Artificial Intelligence Competition
- Third Prize in the Internet+ Competition
- Tencent AI Introductory Course Certificate
- Excellent League Member
- Outstanding Individual in Social Practice