Jin Huang (Hugo) AI Research Engineer at Huawei R&D UK

Website: trenza1ore.github.io **LinkedIn:** HugoHuang123 **GitHub:** Trenza1ore **Email:** pondypondo@outlook.com
ORCID **Google Scholar** **OpenReview**

Publication

2023 **BEV@DC: Bird's-Eye View Assisted Training for Depth Completion**
Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2023)
Wending Zhou, Xu Yan, Yinghong Liao, Yuankai Lin, Jin Huang, Gangming Zhao, Shuguang Cui, Zhen Li
Computer Vision (Depth Completion)

2025 **LLM Shots: Best Fired at System or User Prompts?**
WWW '25: Companion Proceedings of the ACM on Web Conference 2025
Umut Halil, Jin Huang, Damien Graux, Jeff Z Pan
Natural Language Processing (In-Context Learning, Prompt Engineering)

Current Position

Research Engineer / System Engineer @ Huawei R&D UK (Engineer A, Mid-level) **Nov. 2024 – Current**

Engineering & Deployment:

- Received **Star of Edinburgh Award** for my contribution to Huawei's openJiuwen platform / SDK, led development for Graph-based Agent Memory and the internal demo for Context Engine – a key component of openJiuwen.
- Helped with code review & urgent bug fix for the initial release of openJiuwen, led development for multiple memory features to its close-sourced version, deployed to Huawei Cloud Product Line and one of the Top-4 Banks in China.

Research Projects:

- To aid smaller LLMs in tool-integrated reasoning for complex STEM questions, developed a technique to improve instruction following and designed an automated data creation pipeline for powering a theorem-based RAG system.
- Designed an interactive Program-of-Thoughts approach, enhancing LLM agents with feedbacks generated from code execution results, improved LLM agents' performance on complex STEM & Math reasoning tasks, reached 74% accuracy on TheoremQA dataset with 14 & 32B Qwen2.5 models without fine-tuning.
- Explored workflow generation with automated feedback (error messages & different kinds of natural language feedback), improved generation accuracy of ComfyUI workflows by 25.8%.

Education

Master's Degree	**MSc Artificial Intelligence**	University of Edinburgh	Sep. 2023 – Aug. 2024
Bachelor's Degree	**BSc Computer Science**	Cardiff University	Sep. 2020 – Jun. 2023

Projects

ViZDoom
(1.9k stars, 292k downloads)
RL & CV Environment
Top-5 Contributor

Cited in thousands of papers, ViZDoom allows developing AI that play Doom using only visual info. Intended for research in Computer Vision & Reinforcement Learning.
I built automated pipeline for Python binding's type-hinting support, in-engine object categorization (Semantic Segmentation) and improved / fixed various stuffs like seeding behavior, building process (GitHub Action & CMake), documentation & examples, etc.

Stable Baselines3
(12k stars, 15M downloads)
RL Framework
Contributor

Stable Baselines3 (SB3) is a set of reliable implementations of Reinforcement Learning algorithms in PyTorch. It is the next major version of Stable Baselines. One of the most popular Reinforcement Learning frameworks.
Discovered & merged enhancement for rollout buffers, reduced memory usage.

SB3 Extra Buffers
(22 stars, 6.3k downloads)
RAM-Saving Tool for RL
Author

A collection of extra Stable Baselines3 buffer classes. Reducing memory usage drastically with minimal overhead via vectorized implementation of lossless compression algorithms. Experiments conducted with Atari games show > 90% memory save with negligible latency.
Project featured in Stable Baselines3's documentation.

SegDoom
(Master's Thesis Project)
Play Doom with RL & CV
Author

Enhancing Reinforcement Learning in 3D Environments through Semantic Segmentation: A Case Study in ViZDoom explored the potential of applying Semantic Segmentation (SS) to 3D video game environment and analyzed issues in previous literature. Experiments conducted in ViZDoom showed significant improvement with SS mask as augmentation to RGB input, and decrease in memory usage was observed if SS is used as a replacement to RGB (see sb3-extra-buffers). Novel visualization methods were developed to evaluate RL agents' suitability for data collection. Custom Semantic Segmentation dataset was created.

Std Raytracer
(Personal Project)
Ray-tracer / Path-tracer
Author

A ray-traced / path-traced renderer with support for custom models & textures, built from scratch with only C++ 17 standard library and nlohmann's json parser. Multi-threaded with bounding volume hierarchy acceleration, alternate frame rendering option, detailed logs and progress bar for user-friendliness. Support loading simple animation from Blender. An early version was submitted as Computer Graphics Rendering coursework at University of Edinburgh and got an excellent grade of 97%.

Past Internships

Backend Developer *SenseTime*	Designed a tree-based RAG system for DevOps manuals, it preserves the structured nature of source materials and made use of keywords & synonyms to further provide alignment. Outperform official RAG implementation in SenseNova (SenseTime's LLM) at the time.	Jul. 2023 – Aug. 2023 Shenzhen, China
Research Assistant *Cardiff University*	Developed a dataset creation tool in Unity Engine for synthesizing image sequences with mirror segmentation.	Jun. 2022 – Aug. 2022 Cardiff, UK
Game Developer *GALA Sports*	Unpaid internship for study purpose, studied with programming and 3D art team. Developed a 3D game demo and learnt cloth simulation.	Jul. 2018 – Aug. 2018 Shenzhen, China

Past Voluntary Works

CS Tutor *SCIE (High School)*	Worked as Computer Science tutor for GCSE students in my high school, Shenzhen College of International Education. All students that entered with C/D grade improved to above B in IGCSE Exam.	Sep. 2018 – Jun. 2019 Shenzhen, China

Skills

General Workplace Skill	Communication, Teamwork, Translation (Between English and Chinese), Presentation
AI R&D Specific Skill	Bridging gaps between scientists' and engineers' mindsets during mettings & discussions
Field of Research	Natural Language Processing, Reinforcement Learning, Computer Vision
Programming Language	Python, C++, C#, Java, MATLAB, ACS, etc.
Scientific Computing	NumPy, Numba, PyTorch, Sci-kit Learn, Gensim, FAISS, etc.
Hosting Services/Demos	Docker, Podman, Flask, FastAPI, Gradio, ECharts, Vis.js
Vector/Graph Database	Milvus, Chroma, Qdrant, Nebula, Neo4j

A Beautiful Render by My Raytracer, I Hope You Have a Great Day :-)