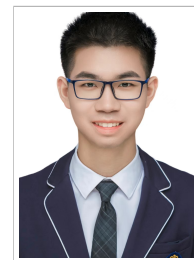


Basic Information

Name: Han Wu **Age:** 21
Phone: +86 19355693187 **Email:** wuhan9918@gmail.com
University: Shanghai Jiao Tong University **Major:** Automation (IEEE Honor Class)
Expected Graduation: Jul 2027 **GitHub:** github.com/HanWu9918



Education

Shanghai Jiao Tong University Aug 2023 – Jul 2027 (Expected)
*Full-time B.Eng. in Automation, **IEEE Honor Class**; Minor in Law (Intellectual Property)* *Shanghai, China*

- **School of Electronic Information and Electrical Engineering.**
- Relevant Coursework: Artificial Intelligence Principles and Applications, Machine Learning, Computer Vision and Robotics.
- English Proficiency: IELTS 7.0; CET-6 609; CET-4 630.
- Honors: Zhiyuan Honors Scholarship, Shanghai Jiao Tong University.

Core Competencies

- Programming: Proficient in Python; intermediate in C++; familiar with JavaScript, HTML/CSS, MATLAB and LaTeX.
- AI & Engineering: Experienced with PyTorch, OpenCV, FastAPI, LMDeploy and OpenCompass; proficient with Git and Linux; hands-on experience deploying LLM inference on multi-GPU clusters and building automated evaluation pipelines.
- Model Evaluation: Experienced in LLM-as-a-Judge evaluation (pointwise and pairwise), rubric design and human spot-check calibration for language and multimodal models.
- AI-Assisted Development: Skilled with Codex, GitHub Copilot and Trae; able to assess model limitations and design effective human-AI workflows.

Publications

Jize Wang, Han Wu, et al. *RouteMoA: Dynamic Routing without Pre-Inference Boosts Efficient Mixture-of-Agents*. Accepted as a long paper at the 64th Annual Meeting of the Association for Computational Linguistics (**ACL 2026**; CCF-A); Code: github.com/Jize-W/RouteMoA.

Teamwork at Alibaba Group. *FilmBench: A Film-Grade Benchmark for Cinematic Video Generation*. arXiv preprint; submitted to **AAAI 2027**; Code: github.com/Neo-yk/FilmOps.

Internship Experience

Alibaba Group Holding Limited Jun 2026 – Present
Foundation Model Evaluation Intern *Hangzhou, China*

Project 1 – FilmBench

- Contributed to FilmBench, a film-grade T2V/R2V benchmark submitted to **AAAI 2027**; evaluated HappyHorse, Seedance 2.0, Kling 3.0/Omni, Veo 3.1, Vidu and Hailuo on curated cinematic prompts.
- Assessed generated videos for cinematography, multi-shot continuity, instruction following and motion aesthetics, translating film-production concepts into consistent annotation guidelines.

Project 2 – QwenEditBench (Tongyi Lab Collaboration)

- Contributed to QwenEditBench, an open-source benchmark developed in collaboration with **Tongyi Lab**; evaluated Seedream 5.0 Lite, GPT Image 2 and other models, and refined the evaluation taxonomy for edit intent, instruction following, visual quality and failure cases.
- Helped build AI-agent-based evaluation sets and data-generation workflows through prompt design, case filtering and quality review, with plans to submit the project to **ICLR 2027**.

Project 3 – Qwen-Image Training Data Support

- Supported the Qwen-Image training team by organizing, filtering and quality-checking a subset of image-generation training data against shared delivery standards.

Research Experience

RouteMoA: Dynamic Routing for Efficient Mixture-of-Agents

May 2025 – Sep 2025

Core Member

- Proposed and validated a dynamic-routing MoA framework that selects high-potential candidate agents before inference and applies posterior correction through LLM self- and cross-evaluation.
- Created router supervision signals from ground-truth match scores and LLM-as-a-Judge evaluations; generated, cleaned and audited training data for an 86M-parameter mDeBERTaV3-base scorer.
- Built the evaluation stack with LMDeploy, FastAPI and OpenCompass, covering iFEval, GSM8K, AIME, GPQA, AlpacaEval, HumanEval, BBH and more than 20,000 QA examples.
- Ran model-pool comparisons and ablations against MoA and SMoA, showing stronger performance while reducing large-pool inference cost by 89.8% and latency by 63.6%.
- Contributed to experimental design, results analysis, artifact management and manuscript preparation for the **ACL 2026** long paper.

Aircraft Cabin Anomalous Sound Recognition and Classification

Jun 2024 – Sep 2024

Core Member, SJTU-COMAC Collaborative Project

Shanghai, China

- Collected and cleaned cabin audio from cruise, takeoff and landing, and extracted STFT and Mel-spectrogram features for classifying 12 types of anomalous sound.
- Expanded the dataset with SpecAugment and time-domain segment replacement, tailoring training and validation to realistic flight noise and onboard hardware constraints.
- Designed a lightweight MobileNetV3 classifier with efficient channel attention, depthwise separable convolutions, h-swish activation, EMA and early stopping.
- Validated the model with confusion matrices, per-class metrics and strong-noise tests; achieved 98.01% average accuracy, improved strong-noise accuracy by 10.66% and kept CPU inference under 20 ms for 0.4–1 s clips.

Selected Engineering Projects

Multi-Agent Production Planning System | *CoppeliaSim, LLM Agents*

May 2026 – Jun 2026

- Built a dual-line flexible-manufacturing simulator with natural-language order parsing, OpenAI-compatible local LLM interfaces and a rule-based baseline.
- Designed a three-agent planning architecture for vehicle and smartphone assembly, part transfer, collision-aware shuttle control, multi-batch production and quality inspection.
- Implemented AABB-based frame-by-frame collision checks and mapping logic for cloned part instances, enabling continuous transfer between production lines, coordinated assembly and end-to-end process simulation.
- Benchmarked multi-agent, single-agent and rule-based planning modes for execution stability and task completion.

AI Agents for Strategic Billiards Play | *MCTS, Bayesian Optimization, LLM Agents*

Feb 2026 – Jul 2026

- Designed a physics-grounded agent for eight-ball pool, translating table geometry, ghost-ball aiming, ray-circle collision prediction and shot difficulty into structured decision features.
- Built an iterative LLM replanning loop that revised shot plans from simulation feedback, and combined MCTS, Bayesian optimization and ensemble voting to improve stability.
- Compared geometric heuristics, MCTS, Bayesian optimization and LLM-agent approaches; the best specialized agent achieved an 86.75% win rate against baselines.

License Plate Recognition System | *OpenCV, PyTorch, LPRNet, PyQt5*

Nov 2024 – Jan 2025

- Built a complete recognition pipeline covering Gaussian blur, Canny edge detection, contour filtering, HSV color filtering, plate localization, cropping and character segmentation.
- Trained a lightweight LPRNet on CBLPRD330k with data augmentation, CTC loss and greedy decoding for end-to-end character recognition.
- Developed a PyQt5 GUI for local image upload and real-time camera input, and packaged the model, resources and application with PyInstaller.
- Delivered a deployable workflow from image preprocessing and inference to result visualization, achieving 86% character-level accuracy.