Chia-Hsuan Hsu

Email | LinkedIn | GitHub

Education

National Yang Ming Chiao Tung University (NYCU) Sep 2026 – Present
MS in Electrical Engineering

National Taiwan University of Science and Technology (NTUST) Sep 2022 – Jun 2026
BS in Computer Science and Information Engineering

Research Interests

AI reasoning, reinforcement learning, and multimodal learning. Research outcomes include publications in a **Q1 journal** and **top-tier conferences**, and co-developing a fast, modular multimodal agent platform with **385+ GitHub stars and 48+ forks**. More projects and full CV: tongyu0924.github.io

Experience

AI engineer (Part-Time), Toppan Security – Taipei, Taiwan Jun 2025 – Dec 2025
- Researched and designed unsupervised methods for comparing model performance in detection tasks without ground truth.
- Designed image augmentation and style transfer methods for data generation and expansion.

Student Researcher, Far Eastern Memorial Hospital – New Taipei, Taiwan Jun 2025 – Present
- First author of an IEEE Transactions on Artificial Intelligence paper and a NeurIPS 2025 Workshop paper on reinforcement learning for improving medical reasoning in large language models (LLMs), evaluated on real-world clinical data
- Developed a hybrid concept-bottleneck architecture that jointly models graph-structured concept representations and residual temporal features for interpretable longitudinal prediction; ongoing work targeting a major AI/ML conference.

Student Researcher, National Taiwan University of Science and Technology May 2024 – Dec 2025
- Contributed to Industrial Technology Research Institute (ITRI) project on bridge defect detection with YOLOv8.
- Researched camera-based visual localization systems (VO/SLAM); first-author paper on visual odometry accepted at ARIS 2025, accepted by the College Student Research Scholarship (NSTC, Taiwan).

Student Researcher, National Academia Sinica – Taipei, Taiwan Sep 2023 – Mar 2025
- Researched and developed diffusion-based generative models for conditional image generation and identity-controllable face synthesis.

Student Organization – Google Developer Groups (GDG) On Campus
National Taiwan University of Science and Technology – Taipei, Taiwan Sep 2023 – Jun 2024
- Designed and delivered Computer Vision courses for 20+ participants; co-designed ML content for an AI camp.

Publications

*Equal contribution

[1] **Chia-Hsuan Hsu**, Jun-En Ding, Hsin-Ling Hsu, Chih-Ho Hsu, Shihao Yang, Li-Hung Yao, Chun-Chieh Liao, Feng Liu, Fang-Ming Hung. RGPO: Ranking-Guided Preference Optimization for Reliable Clinical Reasoning. Accepted at IEEE Transactions on Artificial Intelligence 2026 (IEEE TAI, Q1 journal). [Paper] [Code]

[2] **Chia-Hsuan Hsu**, Jun-En Ding, Hsin-Ling Hsu, Chih-Ho Hsu, Li-Hung Yao, Chun-Chieh Liao, Feng Liu, Fang-Ming Hung. RPRO: Ranked Preference Reinforcement Optimization for Enhancing Medical QA and Diagnostic Reasoning. Accepted at Conference on Neural Information Processing Systems (NeurIPS) Workshop 2025. [Paper]

[3] **Chia-Hsuan Hsu**, Hsin-Chun Lin, Sin-Ye Jhong, Hui-Che Hsu, Ming-Xian Hong, Yung-Yao Chen. VSTFusion-VO: Monocular Visual Odometry with Video Swin Transformer Multimodal Fusion.

Accepted at the International Conference on Advanced Robotics and Intelligent Systems (ARIS), 2025. Best Conference Paper Award: Honorary Mention. [Paper] [Code]

[4] **Chia-Hsuan Hsu**, Jui-Ming Yao. Learning to Refine Hidden States for Reliable LLM Reasoning. arXiv preprint. [Paper] [Code]

[5] Hsin-Ling Hsu, Zizheng Wang*, Donghua Zhang*, Nai-Chia Chen, Jerry Wang, Jun-En Ding, **Chia-Hsuan Hsu**, Guoan Wang, Feng Liu, Fang-Ming Hung, Chenwei Wu, Liyue Shen. MedAction: Towards Active Multi-turn Clinical Diagnostic LLMs. Accepted at Conference on Language Modeling (COLM), 2026. [Paper]

[6] Yung Wei Shueh*, Zhi-Jie Chen*, **Chia-Hsuan Hsu**, Hsin-Ling Hsu, Donghua Zhang, Chenwei Wu, Jun-En Ding, Tongze Zhang, Shihao Yang, Pengfei Hu, Fang-Ming Hung, Feng Liu. DIASENTINEL: An Auditable Multi-Agent System for Guideline-Grounded Diabetes Risk Screening. Accepted at Conference on Empirical Methods in Natural Language Processing (EMNLP), 2026 Demo Track. [Paper]

[7] Pin-Yen Chiu*, Dai-Jie Wu*, Po-Hsun Chu, **Chia-Hsuan Hsu**, Hsiang-Chen Chiu, Chih-Yu Wang, Jun-Cheng Chen. StyleDiT: A Unified Framework for Diverse Child and Partner Faces Synthesis with Style Latent Diffusion Transformer. Accepted to the IEEE International Conference on Automatic Face and Gesture Recognition (FG), 2026. [Paper]

Under Review / Ongoing Work

[8] Jun-En Ding, Yung Wei Shueh, Li-Hung Yao, Zhi-Jie Chen, **Chia-Hsuan Hsu**, Shihao Yang, Chenwei Wu, Pengfei Hu, Feng Liu, Fang-Ming Hung, Liqin Wang. Hierarchical Adaptive Retrieval and Planning for Multi-Specialty Clinical Assessment Generation from Heterogeneous EHRs. (Under Review)

[9] Encoding Is a Data Plane: GPU-Resident State Preparation for Simulator-Based Quantum Machine Learning. (Co-author, Ongoing Work on Apache Mahout QDP, 2,300+ GitHub stars and 1,000+ forks)

Open Source

Open Source Project

Spy-Search: A Modular Multimodal Agent for Fast Research QA Jun 2025 – Present
Received 385+ GitHub Stars and 48+ forks (GitHub Top 0.1%)

- Co-developing and maintaining Spy-Search, a modular multimodal agent using LangChain and LLMs for fast retrieval, QA, and vision-text reasoning. **Achieved sub-1.5s web search with content extraction; project has received 385+ GitHub stars and 48+ forks.** [Code]

Open Source Contributions

- **Microsoft AutoGen** — PR **#6057**: Introduced a `metadata` field in `AssistantAgentConfig` for user/application identity tracking.
- **Hugging Face Diffusers** — PRs **#11426**, **#11427**: improved image interpolation; PRs **#11431**, **#11475**: added `torch.compile()` test support for transformer-based video models.
- **20+ merged PRs** to AI/ML open-source projects, covering bug fixes, feature additions, tests, and documentation; other contributions include **scikit-learn**, **Apache Mahout**, **TensorZero**, **Argilla Distilabel**, **The Turing Way**, **Ersilia**, and **Harmony**.

Technologies

Languages: Python, C++, C, Go, JavaScript, TypeScript, HTML, CSS
Deep Learning Frameworks: PyTorch, TensorFlow, Keras