

Hung-Ting Su

hungtingsu.tw@gmail.com • [Personal Website](#) • [Google Scholar](#) • [DBLP](#) • [LinkedIn](#)

[Reliable Multimodal AI](#) • [Robot Learning](#) • [LLM / MLLM Reasoning](#) • [Evaluation](#)

Summary

- Researcher-engineer building reliable multimodal AI systems for embodied and long-horizon decision-making across robotics, language, and vision.
- Current work spans robot-learning infrastructure, tactile and visual policy learning, VLA-style manipulation, emerging world-model directions, rollout evaluation, and failure-driven improvement for real-world manipulation.
- Academic research spans trustworthy LLM/VLM reasoning, long-form video understanding, narrative reasoning, embodied AI, and multimodal evaluation.

Experience

Senior Associate Researcher

2025–Present

Delta Robotics Innovation Center, Delta Electronics

Taiwan

- Build robot-learning infrastructure and embodied-AI prototypes that connect multimodal data, VLA-style policies, emerging world-model directions, evaluation loops, and real-world deployment.
- Develop tactile and visual policy-learning workflows for contact-rich manipulation, including failure analysis and targeted iteration from rollout outcomes.
- Collaborate across robotics, sensing, and AI components to move research systems toward robust end-to-end operation.

Postdoctoral Researcher

2023–2025

National Taiwan University (NTU)

Taiwan

- Led and mentored projects across embodied AI, trustworthy LLM/MLLM reasoning, long-form video understanding, and multimodal evaluation.
- Built research directions that connect navigation, planning, movie reasoning, automatic verification, and long-horizon evidence-seeking behavior.
- Led projects resulting in CoRL 2024, EMNLP 2024, and ACL 2026, and mentored student-led work accepted at NeurIPS 2024, ICCV 2025, EMNLP 2025, AAAI 2026, and ACL 2026.

Visiting Researcher

2022–2023

Columbia University, hosted by Prof. Shih-Fu Chang

New York, NY

- Developed early zero-shot causal video reasoning frameworks and helped establish a cross-institution collaboration between NTU and Columbia.
- This collaboration led to an EMNLP 2024 publication and strengthened later work on trustworthy multimodal reasoning.

Selected Publications

Full publication history: [Google Scholar](#).

- **VLN-NF: Feasibility-Aware Vision-and-Language Navigation with False-Premise Instructions** (*ACL 2026*) [paper] [project page] — Evidence-seeking embodied navigation with explicit abstention under false-premise or infeasible instructions. **Role: Lead**
- **Unveiling Narrative Reasoning Limits of Large Language Models with Trope in Movie Synopses** (*EMNLP 2024*) [paper] [code] — Analyzes narrative reasoning failure modes in LLMs and reduces unsupported chain-of-thought behavior. **Role: Lead**
- **Context-Aware Replanning with Pre-Explored Semantic Map for Object Navigation** (*CoRL 2024*) [paper] [code] — Zero-shot visual map refinement and semantic replanning for more robust embodied navigation. **Role: Lead**
- **AED: Adaptable Error Detection for Few-shot Imitation Policy** (*NeurIPS 2024*) [paper] — Policy monitoring for few-shot imitation learning with deployment-oriented error detection. **Role: Mentor**
- **MovieCORE: COgnitive REasoning in Movies** (*EMNLP 2025*) [paper] [project page] — Long-form movie benchmark for cognitive and social reasoning with multimodal models. **Role: Mentor**
- **HERMES: temporal-coHERent long-forM understanding with Episodes and Semantics** (*ICCV 2025*) [paper] [project page] — Improves long-form multimodal understanding through temporal coherence and episodic structure. **Role: Mentor**
- **Affordance-Guided Coarse-to-Fine Exploration for Base Placement in Open-Vocabulary Mobile Manipulation** (*AAAI 2026*) [paper] — Training-free VLM-guided planning for mobile-manipulation base placement in open-vocabulary settings. **Role: Mentor**
- **ADAPT: Benchmarking Commonsense Planning under Unspecified Affordance Constraints** (*ACL 2026*) [paper] — Benchmarking and evaluating planning under unspecified affordance and feasibility constraints. **Role: Mentor**

Education

PhD in Computer Science 2023
National Taiwan University (NTU), Advisor: Prof. Winston H. Hsu Taiwan

- Dissertation: *Multi-modal Video Comprehension and Beyond*
- Visiting researcher at Columbia University, hosted by Prof. Shih-Fu Chang.

MS in Computer Science 2017
National Taiwan University Taiwan

BS in Computer Science 2013
National Tsing Hua University Taiwan

Skills

- **Research Areas:** robot learning, embodied AI, multimodal reasoning, long-form video understanding, and evaluation under uncertainty.
- **Core Methods:** imitation learning, VLA-style policies, world-model-oriented planning, failure analysis, benchmark construction, data generation, and automatic verification.
- **Technical Stack:** Python, PyTorch, HuggingFace Transformers, ROS2, Linux, Docker, Git, RGB-D/3D perception, and tactile/contact sensing.

- **Collaboration:** student mentoring, cross-institution research coordination, technical writing, and publication-driven project leadership.

Academic Services

- Area Chair: *EACL 2026, ACL 2026*
- Reviewer: *ICML, NeurIPS, CVPR, ICCV, NAACL, ACL, EMNLP, CoRL, ICRA, IROS, AAAI, ACM Multimedia, TCSVT, and related venues*

References

References are available upon request.