

# Kuo-Han Hung

☎ +886-952663269 | ✉ khhung906@gmail.com | 🏠 khhung906.github.io | 🎓 Kuo-Han-Hung

## Education

### National Taiwan University (NTU)

Taipei, Taiwan

B.S. IN COMPUTER SCIENCE AND INFORMATION ENGINEERING

Sept. 2020 - Dec. 2024 (expected)

- Overall GPA: 4.24/4.3; Major GPA: 4.27/4.3; Ranking: 5/187 (top 2.6%)
- Awards: Bachelor Thesis Dean's Award, 3 times Dean's List Awards, 2 times Presidential Awards
- Relevant Courses: Foundations of AI, Machine Learning\*, Applied Deep Learning\*, Computer Vision\*, Reinforcement Learning\*, Security and Privacy of ML\*, Calculus, Linear Algebra, Probability (\*graduate-level)

### University of California, Berkeley

California, USA

SUMMER SESSION PROGRAM IN THE DEPT. OF EECS

June 2022 - Aug. 2022

- Course: CS161 Computer Security (GPA: 4.0/4.0)

## Research Experience

### IBM Thomas J. Watson Research Center

New York, USA

RESEARCH INTERN, ADVISOR: DR. PIN-YU CHEN

June 2024 - Sept. 2024

- Researched on *prompt injection detection for LLM agent using attention*. [\[under review\]](#)
  - Explored prompt injection attacks using explainable AI by analyzing changes in attention patterns.
  - Developed an attention-based detector with a 31% AUROC improvement over training-free methods.

### Communication and Multimedia Lab

Taipei, Taiwan

UNDERGRADUATE RESEARCHER, ADVISOR: PROF. WINSTON H. HSU

Feb. 2022 - June 2024

- Researched on *vision-instruction correlation rewards for long-horizon manipulation*. [\[ICMLW' 24\]](#)
  - Developed a hierarchical reward model using large language models and vision language models.
  - Outperformed the previous best method in long-horizon manipulation with a 43% higher success rate.
- Researched on *adaptable error detection for few-shot imitation policy*. [\[NeurIPS' 24\]](#)
  - Designed novel learning objectives that enabled the extraction of policy features to detect errors.
  - Achieved top performance in 17 out of 21 testing cases compared to strong baselines.

### Machine Intelligence and Understanding Lab

Taipei, Taiwan

UNDERGRADUATE RESEARCHER, ADVISOR: PROF. YUN-NUNG CHEN

Feb. 2022 - June 2024

- Benchmarked *multi-source retrieval and reasoning in visual question answering*. [\[under review\]](#)
- Researched on *open-domain conversational questions, answering using historical answers*. [\[AAACL' 22\]](#)
  - Employed knowledge distillation to enhance the efficacy of retrieving passages using historical replies.
  - Achieved state-of-the-art performance in open-domain conversational retrieval with 77.9 (R@5).

## Selected Publications (\*co-authorship)

- [1] **K. H. Hung\***, P. C. Lo\*, J. F. Yeh\*, H. Y. Hsu, Y. T. Chen, and W. H. Hsu, "VICtoR: Learning Hierarchical Vision-Instruction Correlation Rewards for Long-horizon Manipulation." The 41st International Conference on Machine Learning (ICML) ARLET Workshop, 2024.
- [2] J. F. Yeh, **K. H. Hung\***, P. C. Lo\*, C. M. Chung, T. H. Wu, H. T. Su, Y. T. Chen, and W. H. Hsu, "AED: Adaptable Error Detection for Few-shot Imitation Policy." The 38th Conference on Neural Information Processing Systems (NeurIPS), 2024.
- [3] **K. H. Hung**, C. Y. Ko, A. Rawat, I. H. Chung, W. H. Hsu, and P. Y. Chen, "Attention Tracker: Detecting Prompt Injection Attacks in LLMs." Preprint.
- [4] **K. H. Hung**, C. Zhang, and D. Yankov, "Customizable Routing with Learning from Past Recommendations." The 32nd International Conference on Advances in Geographic Information Systems (SIGSPATIAL), 2024.
- [5] H. C. Fang\*, **K. H. Hung\***, C. W. Huang, and Y. N. Chen, "Open-Domain Conversational Question Answering with Historical Answers." The 2nd Asian Chapter of the Association for Computational Linguistics (AAACL), 2022.

## Work Experience

---

### Microsoft, Bing Team

Remote (team in California, USA)

APPLIED SCIENTIST INTERN

July 2023 – June 2024

- Integrated Bing Maps with Microsoft Copilot to create an enriched experience for conversational map queries.
- Developed a novel pathfinding algorithm that learns from history, achieving a 10x speedup. [\[SIGSPATIAL' 24\]](#)
- Invented and evaluated a context-aware map search system using LLMs. [\[SIGSPATIAL' 24\]](#)

### Cinnamon AI

Taipei, Taiwan

DEEP LEARNING INTERN

June 2022 - Aug. 2022

- Developed and implemented the first image-driven recipe retrieval model with adaptable ingredients, along with a cross-modal retrieval system featuring a vision-ingredients seq2seq architecture.

### Courtero

Remote (team in USA)

FULL-STACK DEVELOPER & FOUNDING MEMBER

Jan. 2022 - June 2022

- Developed map and rating systems for the Courtero app using React Native and Firebase.
- Partnered with students from Harvard and UIUC, gaining acceptance into the Harvard Innovation Labs.

## Honors and Awards

---

<b>Dean's Award</b> (for top two theses in EECS department), NTU Bachelor's Thesis Award	June 2024
<b>Presidential Award/Dean's List</b> (for top 2%/5% students), NTU	Fall'22, '23; Spring'21, '22, '23
<b>Lin Hsiung Chen Memorial Scholarship</b> (top scholarship in Taiwan)	Nov. 2023
<b>Second prize</b> , NTU CSIE Undergraduate Thesis Exhibition	June 2023
<b>Creativity and Entrepreneurship Excellence Award</b> , NTU D-school	June 2023
<b>Charity Trust Jason International Fund Scholarship</b> , issued by Acer Taiwan	Sept. 2022
<b>Honorable Award</b> , NTU CSIE Undergraduate Thesis Exhibition	June 2022
<b>Special Award</b> (top 5/300+ teams), LINE Fresh Hackathon	Dec. 2021

## Leadership and Activities

---

### Creativity and Entrepreneurship Program, NTU D-School

Taipei, Taiwan

TEAM LEADER

Sept. 2022 - June 2023

- Developed and presented detailed business plans for a startup specializing in AI-driven recruitment solutions, incorporating market analysis, financial projections, and go-to-market strategies.
- Awarded the Excellence Award (top 3 out of 10+ teams) in the 2023 class of D-School.

### NTU Google Developer Student Club

Taipei, Taiwan

CO-FOUNDER & ACADEMIC INSTRUCTOR

Aug. 2022 - June 2023

- Expanded the club from 3 to over 70 members by planning, structuring operations, and designing coursework
- Initiated and led a 20-person study group to develop AI kidney disease management, achieving first place in Taiwan in the Google Solution Challenge.

## Selected Projects

---

### VULNERABILITIES IN VLM-POWERED POLICIES [\[Slide\]](#)

- Developed a text and image-based adversarial attack method to manipulate embodied AI behavior, achieving a 40% increase in targeted attack success on VLM-powered policy networks.

### ZERO-SHOT TEXT BEHAVIOR RETRIEVAL [\[Report\]](#)

- Proposed a zero-shot behavior retrieval system using text-guided object detection and CLIP to retrieve task-relevant data from offline datasets, enabling policy training without expert data.

## Teaching Experience

---

TEACHER ASSISTANT, CSIE1000 INTRODUCTION TO COMPUTER SCIENCE

Sept. 2024 - Dec. 2024

- Designed eight homework across major fields in computer science to help students reinforce weak concepts.
- Led a weekly TA sessions to address student questions on lectures and assignments.

TEACHER ASSISTANT, CSIE5043 MACHINE LEARNING (GRADUATE-LEVEL)

Jan. 2023 - June 2023

- Designed and implemented advanced machine learning assignments on VC dimension and SVM.
- Organized a final project competition focused on music recommendation systems using tabular data.