Kai-Po (Bryant) **Chang**

PH.D. STUDENT · GRADUATE INSTITUTE OF COMMUNICATION ENGINEERING · NATIONAL TAIWAN UNIVERSITY

f11942093@ntu.edu.tw | kaipoc0810-personal-page.netlify.app/ | kai-po-chang-14051a191 | Google Scholar

RESEARCH INTERESTS

- **Embodied VLM Agents:** RL for multi-turn reflection capability via imagination (ongoing), RL for spatialtemporal-based cognitive map reasoning (ongoing).
- **Truthworthy VLMs:** RL from natural language explanation feedback (RAPPER) *[ICLR'24]*, self-augmented hallucinative contrastive alignment (SANTA) *[WACV'26]*, contrastive decoding for temporal hallucination mitigation (SEASON) *[arXiv'25]*.

Education

National Taiwan University (NTU) *2022 - present*

PH.D. STUDENT IN GRADUATE INSTITUTE OF COMMUNICATION ENGINEERING (DATA SCIENCE)

- Research advisor: Prof. Yu-Chiang Frank Wang

National Tsing Hua University (NTHU) *2018 - 2022*

B.S. IN COMPUTER SCIENCE

Research Experiences

NVIDIA Research *Jul. 2024 - Jul. 2025*

RESEARCH SCIENTIST INTERN, MENTORED BY DR. FRED YANG & PROF. YU-CHIANG FRANK WANG

- Proposed SANTA *[WACV'26]*, a post-training scheme that exposes object and action hallucination tendencies via contrastive negative captions and alignment, improving **VLM faithfulness** by 2–3% in F1 while preserving captioning performance.

Vision and Learning Lab, NTU *June. 2022 - present*

PH.D. STUDENT, SUPERVISED BY PROF. YU-CHIANG FRANK WANG

- Led multiple projects on **trustworthy VLMs**: (1) RAPPER *[ICLR'24]*, a post-training alignment paradigm achieving around +7% / +2% SOTA gains on metrics measuring **plausible** and **faithful** response, (2) SANTA *[WACV'26]*; (3) SEASON *[arXiv'25]*, a contrastive decoding strategy reducing **temporal hallucination** by 3% while improving conventional video understanding tasks by 0.5%.
- Developed a diffusion model (DM) **unlearning** framework for concept erasure (Receler) *[ECCV'24]*, and proposed the first training framework for text-to-video DM to generate **customized multi-subject and motion video** (VideoMage) *[CVPR'25]*.
- Enhanced VLMs with **temporal reasoning** via temporal anchors as an additional modality (TA-Prompting) *[WACV'26]*, and designed a **domain-specialized MoE** framework for knowledge editing to mitigate catastrophic forgetting (ARM) *[ACL'25]*.
- Published **2 first author**, **2 co-first author**, and 4 co-author papers at top conferences.
- Received the **Top Reviewer Award** at NeurIPS 2025; reviewer for TAPMI, CVPR (2024–26), ICLR (2025–26), and NeurIPS 2025.

Publications (* indicates equal contribution.)

[1] **K.-P. Chang**, W.-Y. Cheng, C.-P. Huang, F.-E. Yang, Y.-C. F. Wang, "Mitigating Object and Action Hallucinations in Multimodal LLMs via Self-Augmented Contrastive Alignment", *In WACV 2026*.

[2] W.-Y. Cheng*, **K.-P. Chang***, C.-P. Huang, F.-E. Yang, Y.-C. F. Wang, "TA-Prompting: Enhancing Video Large Language Models for Dense Video Captioning via Temporal Anchors", *In WACV 2026*.

[3] H.-C. Lin, Y.-C. Yu, **K.-P. Chang**, Y.-C. F. Wang, "EMLoC: Emulator-based Memory-efficient Fine-tuning with LoRA Correction", *In NeurIPS 2025*.

[4] Y.-J. Cheng, Y.-C. Yu, **K.-P. Chang**, Y.-C. F. Wang, "Serial Lifelong Editing via Mixture of Knowledge Experts", *In ACL 2025*.

[5] C.-P. Huang, Y.-S. Wu, H.-K. Chung, **K.-P. Chang**, F.-E. Yang, Y.-C. F. Wang, "VideoMage: Multi-Subject and Motion Customization of Text-to-Video Diffusion Models", *In CVPR 2025*.

[6] **K.-P. Chang**, C.-P. Huang, W.-Y. Cheng, F.-E. Yang, C.-Y. Wang, Y.-H. Lai, Y.-C. F. Wang, "RAPPER: Reinforced Rationale-Prompted Paradigm for Natural Language Explanation in Vision Question Answering", *In ICLR 2024*.

[7] C.-P. Huang*, **K.-P. Chang***, C.-T. Tsai, Y.-H. Lai, F.-E. Yan, Y.-C. F. Wang, "Receler: Reliable Concept Erasing of Text-to-Image Diffusion Models via Lightweight Erasers", *In ECCV 2024*.

[8] Y.-C. Yu, C.-P. Huang, J.-J. Chen, **K.-P. Chang**, Y.-H. Lai, F.-E. Yang, Y.-C. F. Wang, "Dual-Teacher Memory Retrieval for Continual Learning on Vision-Language Models", *In ECCV 2024*.