

CHONG HOU CHOI

+852 6822-8258  Email  Personal Webpage  LinkedIn  GitHub

Work/Research Experiences

AI Algorithm Engineer, Mentor: Rui Liu

Dec 2024 - Jul 2026

Celia Large Model Application Lab, Huawei HKRC

Hong Kong

- Built a state-of-the-art video shot boundary detection pipeline based on DINOv3, outperforming traditional deep learning methods by over 30% and enabling automated cleansing of 5M+ video clips.
- Operated distributed AI infrastructure spanning 512+ NVIDIA Hopper GPUs, supporting large-scale multimodal data processing, model training, and evaluation.
- Built a batch data construction pipeline in Python for image restoration, using configurable old-photo degradation algorithms to synthesize real-world artifacts such as fading, scratches, stains, blur, and noise.
- Trained a Flux.1 Dev-based color tuning model on paired clean/degraded images and constructed the full experimentation workflow, including data ingestion, metadata validation, training job orchestration, checkpoint evaluation, and lightweight benchmarking; outperformed open-source SOTA image restoration baselines at the time.

Machine Learning Engineer (Intern), Mentor: Yi Wang

Jan 2022 - Nov 2022

Shanghai Artificial Intelligence Laboratory, General Vision Group

Beijing, China

- Debugged and optimized the **Temporal Action Localization** module for **INTERN 2.0**, a deep learning video recognition system, ensuring high-quality testing and analysis.
- Improved performance benchmarks on key datasets (THUMOS14, ActivityNet v1.3, and HACS) by 6.82% through optimized machine learning algorithms and system-level data validation techniques.

Publications

Boogu-Image-0.1: Boosting Open Agentic Multimodal Generation via Understanding under a Minimal Budget

Jul 2026

arXiv preprint, cs.CV and cs.AI

G. Chen et al., including C. H. Choi.

- Developed and deployed content-safety safeguards across the Boogu-Image-0.1 model family, supporting responsible text-to-image generation and instruction-based editing.

InsEdit: Towards Instruction-based Visual Editing via Data-Efficient Video Diffusion Models Adaptation

Apr 2026

arXiv preprint, cs.CV

Z. Rao, B. Zou, H. Che, X. He, C. H. Choi, Y. Li, R. Liu, and Q. Chen.

- Built a video-pair construction pipeline for instruction-based editing, generating aligned source-target videos with rewritten captions and edit instructions.
- Designed multi-stage quality controls—including response filtering, metadata validation, multi-round VQA alignment, and augmentation—to curate approximately 100K training pairs.

Capybara: A Unified Visual Creation Model

Feb 2026

Open-source model release, GitHub repository

Z. Rao, H. Che, Z. Hu, B. Zou, Y. Liu, X. He, C. H. Choi, et al.

- Extended the shared data pipeline across T2V, T2I, TV2V, and TI2I, adding task-specific labeling, sample construction, and modality-aware processing.
- Supported large-scale Capybara training on 256+ NVIDIA Hopper GPUs, coordinating data ingestion, distributed training runs, and checkpoint evaluation.
- Shipped FP8 weight-only quantization and custom ComfyUI nodes for model loading, prompt rewriting, generation, and image/video editing.

Education

University of Waterloo

Sep 2023 - Aug 2024

Master of Engineering in Electrical and Computer Engineering

Waterloo, ON, Canada

Tsinghua University

Sep 2019 - Jun 2023

Bachelor of Engineering in Automation Engineering

Beijing, China

Skills Summary

Programming: Python, C++, SQL, Bash

Machine Learning: PyTorch, CUDA, OpenCV, Diffusion Models, Multimodal Learning, Model Quantization

AI Development: Codex, Claude Code, LLM Agents

Data: Pandas, NumPy, Data Curation, ETL/Data Pipelines, Metadata Validation

Languages: English, Cantonese, Mandarin, Japanese