CHENG CHIEN

🏠 Taipei City, Taiwan (R.O.C.)
📞 (+886) 02-29587999
📱 (+886) 0989819022
✉ chengchien1999@gmail.com
🌐 https://github.com/JamesQian1999
ⓔ https://www.linkedin.com/in/cchien1999/

Education

September 2022- 2024
Master of Science,
National Yang Ming Chiao Tung University, Taiwan, ROC

Computer Science
4.19 / 4.3 GPA

September 2018-June 2022
Bachelor of Science,
National Sun Yat-sen University, Taiwan, ROC

Computer Science and Engineering
4.01 / 4.3 GPA

Projects

Machine Learning
- Implement a system combine learned image compression codec and MLLMs
- Implement Conditional VAE for Video Prediction

Software Security
- The Study and Implementation of Secure Device Pairing and Access Control for Smart IoT Devices

Network
- Implementation of TCP Protocol using UDP
- Huffman Compression Algorithm

School Activities

September 2020 - June 2021
North Alumni Association,
National Sun Yat-sen University, Taiwan, ROC

Director
- Hosted events for over 100 people
- Organized numerical activities of our association

Febuary 2019 - February 2023
Taiwan Happy Coding Promotion Association, Taiwan, ROC

Outreach Instructor
- Promoting programming skills and inspiring Taiwanese children through engaging tech education.

Publications

- C.H. Kao, **C. Chien,** Y.J. Tseng, Y.H. Chen, A. Gnutti, S.Y. Lo, W.H. Peng, R. Leonardi, "ComNeck: Bridging Compressed Image Latents and Multimodal LLMs via Universal Transform-Neck," *arXiv preprint arXiv:2407.19651*, 2024.
- Y.H. Chen, Y.C. Weng, C.H. Kao, **C. Chien**, W.C. Chiu, W.H. Peng, "TransTIC: Transferring Transformer-based Image Compression from Human Visualization to Machine Perception," *IEEE International Conference on Computer Vision (ICCV)*, Oct. 2023.
- M.J. Chen, H.S. Xie, **C. Chien**, W.H. Peng, H.M. Hang, "Learned Hierarchical B-frame Coding with Adaptive Feature Modulation for YUV 4:2:0 Content," *IEEE International Symposium on Circuits and Systems (ISCAS)*, May 2023.
- C.H. Kao*, Y.H. Chen*, **C. Chien**, W.C. Chiu, and W.H. Peng, "Transformer-based Image Compression with Variable Image Quality Objectives," *Asia-Pacific Signal and Information Processing Association (APSIPA)*, Oct. 2023.

Honors and Awards

- Top Creativity AwardTop Creativity Award, IEEE ISCAS
- Taiwan Imaging-Tek Corporation Scholarship
- Excellent Student Awards, NSYSU
- Calculus Contest Awards, NSYSU
- Cathay Financial Holdings Customer-Children Scholarship
- New Taipei City Apartment Building Management Services Occupational Union Scholarship

Skills

Python, C++, Machine Learning, Multimodal Large Language Model (MLLM), Image Compression

Certifications

2021
Happy Coding Promotion Association

Outreach Instructor
Inspiring Taiwanese kids through fun tech education

2020
TOEIC

TOEIC Blue
Score 845

2019
CPE

Advanced Level
Collegiate Programming Examination

About

My name is Chien Cheng, and I hold a Master's degree in Computer Science from National Yang Ming Chiao Tung University. My master's research focuses on machine learning, computer vision, image compression, and multimodal large language models. My master's thesis, titled "ComNeck: Bridging Compressed Image Latents and Multimodal LLMs via Universal Transform-Neck," combines learning-base image compression with a multimodal language model. This work has been submitted to the international conference ICLR and is currently under review.

During my graduate studies, I participated in several National Science and Technology Council (NSTC) projects and collaborated with my lab members to publish research in top-tier international conferences such as ICCV, APSIPA ASC, and ISCAS. Through these experiences, I developed strong technical skills, cooperation, and communication abilities. I am proficient in Python and experienced in using PyTorch to build the machine learning model, with expertise in handling the entire workflow from data preprocessing to model training. Additionally, I studied in the Philippines for three months after graduation, where I demonstrated my adaptability to new environments and my capability to perform independently across different cultures and languages. This further strengthened my ability to communicate, coordinate, and resolve issues in English.

During my undergraduate studies, I maintained excellent academic performance, receiving multiple academic awards. At the same time, I served as the president of a student club, where I successfully organized large-scale events with over 100 participants. This role shows my strong ability on leadership, time management, and multitasking skills.

I earned my M.S. in Computer Science from National Yang Ming Chiao Tung University in 2024, under the guidance of Prof. Wen-Hsiao Peng. Prior to this, I completed my B.S. in Computer Science and Engineering at National Sun Yat-sen University in 2022.

Master's Thesis

This work first adapts compressed image latents to vision tasks using Multimodal Large Language Models (MLLMs). MLLMs have successfully extended their capabilities from text to other modalities, such as vision. However, their deployment on end devices is challenging due to their huge parameter scale. Although we can deploy MLLMs on cloud devices, it becomes necessary to use a high-effect compression system to transfer the image to the cloud. To address this issue, we propose a novel neural image compression system featuring a lightweight transform-neck and surrogate loss designed to adapt compressed image latents for various MLLM-based vision tasks. Our system is versatile and applicable to multiple scenarios. The neural image codec can be configured to (1) focus on human perception without updates, (2) address both human and machine perception by updating the network, or (3) fully update for machine perception. Our proposed surrogate loss that can be simply back-propagated through the partial visual encoder instead of the entire huge MLLM. Given that most MLLMs share a common visual encoder, the proposed transform-neck trained with the surrogate loss is universally applicable and suitable for various MLLMs. Experimental results demonstrate that our method achieves excellent rate-accuracy performance with less complexity.