Cheng-Han Lee

Personal Website

Email : chenghan.lee@utexas.edu
Phone : (+1) 512-921-3036

Education

- **The University of Texas at Austin** — Austin, USA
  Ph.D. Candidate of ECE (DICE track); GPA: 3.9 / 4.0; supervised by Prof. Alan Bovik — *Aug. 2022 - Now*
- **National Taiwan University** — Taipei, Taiwan
  Master of CS; GPA: 4.0 / 4.3 — *Sep. 2016 – Aug. 2018*
- **National Yang Ming Chiao Tung University** — Hsinchu, Taiwan
  Bachelor of ECE — *Sep. 2012 – June. 2016*

Work Experience

- **The University of Texas at Austin** — Austin, USA
  Graduate Research Assistant, Lab of Image/Video Engineering and in collaboration with YouTube — *Aug 2022 - Present*
  - **Deep Video Cropper**: Develop a new database to address the challenge. Design new ROI detection algorithms.
  - **Dynamic Video Outpainting**: Leverage our video cropper to enable video outpainting beyond center crops.
- **Amazon Inc.** — Bellevue, USA
  Applied Scientist Intern, Astro Home Robot Team — *May 2025 - Aug 2025*
  - **Advanced Frame Sampling For Video Understanding**: Conduct research on video question answering and grounding via multi-modal large language model.
- **Amazon Inc.** — Bellevue, USA
  Applied Scientist Intern, Astro Home Robot Team — *May 2024 - Aug 2024*
  - **Text-guided Video Reframing**: Conduct research on video reframing and video language modeling.
- **Mediatek Inc.** — Hsinchu, Taiwan
  AI Software Engineer, Computational Artificial Intelligent Department, Algorithm Team — *Mar 2021 - July 2022*
  - **Machine Learning Toolkits [link]**: Develop and test components of efficient DL toolkit which utilizes special IR to support several AI frameworks. Provide example codes and optimization report for AutoML on semantic segmentation. Enable the entire framework to support Tensorflow V2. Enable the AutoML toolkit to support device-aware searching. Enable the toolkit to fetch latency and power information from AI platforms.
  - **Neural Network Accelerator [link]**: Develop and test resizing operators in Mediatek AI Compiler which enables deep models to run parallel on multi-cores AI Processing Unit (Dimensity 9000).
- **Perfect Corp.** — Taipei, Taiwan
  Computer Vision Software Engineer, Magic Technology Team — *Jun 2020 - Feb 2021*
  - **Develop SkinCare AR algorithms [link]**: Implement high-resolution, realistic, light reflection considered skin texture/pore removal algorithms on mobile devices. Combine digital image processing methodology and deep learning-based skin/texture/pore segmentation models.
  - **Optimize model pruning and face alignment algorithms**: Implement state-of-the-art pruning methods on Perfect's deep face alignment models which is for real-time makeup rendering.
  - **Develop tagging tools**: Develop face alignment data tagging tools to reduce labeling time and enhance quality.
- **SenseTime Ltd.** — Shenzhen, China
  Trainee Computer Vision Researcher, SenseTime and CUHK Joint Lab — *Sep 2018 - April 2019*
  - **CelebAMask-HQ [2k+ stars, 300+ forks] [link]**: Present MaskGAN (CVPR'20) for diverse and interactive face manipulation and contribute a large-scale high-resolution face segmentation dataset called CelebAMask-HQ.
  - **Switchable Normalization on Generative Adversarial Network**: Discover the probability of learning to select different normalizers for different layers of a discriminator by meta-learning.
- **National Taiwan University** — Taipei, Taiwan
  Graduate Research Assistant, Communications and Multimedia Lab — *Sep 2016 - Aug 2018*
  - **Attribute Augmented Face Hallucination**: Present the first face hallucination model which exploit attribute information to assist ultra low resolution image recovering (CVPRW'18).

- **Pose Invariant Virtual Try-on Shoes**: Present the first shoes try-on framework which is based on key point detection and conditional image completion (ACCV'18).

- **QNAP Systems** Taipei, Taiwan
  Software Engineering Intern, Cloud Innovation Group *July 2016 - Aug 2016*
  - **Implement an universal cloud client of Slack**: Integrate the application of QNAP NAS (Network-Attached Storage), Connect To Cloud Drive, with Slack. Enable users to synchronize files between Slack and QNAP NAS.

Journal Publications

- **Text-guided Video Reframing**
  Lee, C.**, Liu, B., Su, C., Bovik, A. C., Sun, M., Kuo, C. **TIP (Under Review)
- **Subjective Portrait Region Cropping On Landscape Video with Temporal Smoothing**
  Lee, C.**, Mandal, M., Birkbeck, N., Wang, Y., Adsumilli, B., Bovik, A. C. **TIP (Under Review)
- **Is HDR Always Better Than SDR?**
  *Chen, B., **Lee, C.**, Chen, Y., Shang, Z., Wei, H., Bovik, A. C. **TIP (Under Review)***

Conference Publications

- **Memory Tree Guided Key Frame Querying for Efficient 3D Question Answering**
  *Huang, S., Chen, F., Tsao, L., **Lee, C.**, et al. **ICLR'26 (Under Review)***
- **HDRSDR-VQA: A Subjective Video Quality Dataset for HDR and SDR Evaluation** [link]
  *Chen, B., **Lee, C.**, Chen, Y., Shang, Z., Wei, H., Bovik, A. C. **PCS'25***
- **Subjective Portrait Region Cropping On Landscape Video Study** [link]
  Lee, C.**, Mandal, M., Birkbeck, N., Wang, Y., Adsumilli, B., Bovik, A. C. **ICIP'24
- **MaskGAN: Towards Diverse and Interactive Facial Image Manipulation** [link]
  Lee, C.**, Liu, Z., Wu, L., Luo, P. **CVPR'20, Top 1% most influential papers
- **PIVTONS : Pose Invariant Virtual Try-on Shoes with Conditional Image Completion** [link]
  *Chou, C., **Lee, C.**, Zhang, K., Lee, H., Hsu, W. H. **ACCV'18(Oral)***
- **Attribute Augmented Convolutional Neural Network for Face Hallucination** [link]
  Lee, C.**, Zhang, K., Lee, H., Cheng, C., Hsu, W. H. **CVPRW'18
- **Drone-View Building Identification by Cross-View Visual Learning** [link]
  *Chen, C., Kuo, Y., Lee, T., **Lee, C.**, Hsu, W. H. **CVPRW'18***

Course Projects

- **Expainable Image Quality Assessment Using LLM** [link]: Our explainable image quality assessment model offers in-depth reasoning of the quality scores it estimates. (UT ECE, generative model course)
- **AutoPortrait: Create Portrait Videos** [link]: A framework designed to convert a video in landscape mode to a portrait mode video. (UT ECE, digital video course)
- **HouseMonitor: A Real-Time Surveillance Android App** [link]: Users can control the angle of a remote camera and see real-time video of house on a mobile phone. (NYCU CS, embedding system course)
- **Intelligent Speech Emotion Recognition Android App**: Users can record a piece of voice and identify its emotion on a mobile phone. (NYCU ECE, spacial project course)

Programming Skills

- **Languages**: Python (4 years industry experience), C/C++ (2.5 years industry experience), JAVA (1 year industry experience), Objective-C (1 year industry experience)
- **Toolkits**: Pytorch, Tensorflow, FFmpeg, OpenCV, CUDA, Docker, Latex, Android, iOS
- **Platforms**: Linux, MacOS, Windows, Vim, Visual Studio, Android Studio, Xcode, Jira, Git