Huancheng Chen

huanchengch@gmail.com | citychan.github.io | huanchengch | CityChan

Education

The University of Texas at Austin, Austin, Texas 2020 – 2025

Ph.D. in Electrical and Computer Engineering

South China University of Technology, Guangzhou, China 2015 – 2019

B.E. in Electrical Engineering and Automation

Research Overview

My research focuses on trustworthy machine learning, with an emphasis on developing scalable and efficient learning algorithms. In my Ph.D. dissertation, I worked on:

- **Continual learning** with large pretrained foundation models;
- **On-device learning** across data/system-heterogeneous networks;
- **Model compression** by pruning, quantization and knowledge distillation.

In AI industry, I am working on:

- **Data curation** for synthesizing tool-use trajectories to fine-tune LLM agents;
- **LLM post-training** to enhance LLM agent's tool-use capabilities and efficiency.

Industrial Experience

Accenture, Mountain View, California June. 2025 – present

Senior Research Scientist at Advanced AI Center

AI Refinery Platform Development:

- Core contributor to the Training-as-a-Service functionality (training/evaluation/data curation), an end-to-end pipeline that enables clients to customize LLMs with their own data.
- Main developer of the Kubernetes cluster for building, deploying, and maintaining docker images of post-training algorithms across different scenarios.

Agent Foundation Model Post-Training:

- Core contributor to Accenture's LLM agent foundation model for long-horizon tool use, search/research, and coding. Implement the SFT/RL pipeline and automated task trajectory synthesis in MCP-Bench.

Sony AI, Tokyo, Japan May. 2024 – Aug. 2024

Research Intern at PPML Team

- Designed a zero-shot method to enhance spatial semantics in layout-to-image using diffusion models.

Sony AI, Austin, Texas Fed. 2024 – May. 2024

Research Intern at PPML Team

- Proposed a continual learning approach using LoRA for sequential fine-tuning of vision foundation models.

Toyota InfoTech Lab, Mountain View, California May. 2022 – Aug. 2022

Research Intern at Infrastructure & Data Platform

- Applied knowledge distillation to mitigate model drift problem in data-heterogeneous distributed learning.

Bell Labs, Murray Hill, New Jersey Jan. 2022 – May. 2022

Research Intern at Mathematics & Algorithms Research Group

- Developed a U-2-Net-based framework for removing irrelevant background to improve flaw detection.

Publications

(⋆ indicates equal contribution)

[NeurIPS] **H. Chen**, H. Vikalo, "Heterogeneity-Guided Client Sampling: Towards Fast and Efficient Non-IID Federated Learning", *Advances in Neural Information Processing Systems*, 2024.

[ICML] **H. Chen**, H. Vikalo, "Recovering Labels from Local Updates in Federated Learning", *International Conference on Machine Learning*, 2024.

[CVPR] **H. Chen**, H. Vikalo, "Mixed-Precision Quantization for Federated Learning on Resource-Constrained Heterogeneous Devices", *Conference on Computer Vision and Pattern Recognition*, 2024.

[ICLR] **H. Chen**, H. Vikalo, J. Wang, "The Best of Both Worlds: Accurate Global and Personalized Models through Federated Learning with Data-Free Hyper-Knowledge Distillation", *International Conference on Learning Representations*, 2023.

[CVPRW] **H. Chen**, H. Vikalo, "Federated Learning in Non-IID Settings Aided by Differentially Private Synthetic Data", *Conference on Computer Vision and Pattern Recognition Workshop*, 2023. **Oral**

[ICCVW] A. Mohamed⋆, **H. Chen**⋆, "Skeleton-Graph: Long-Term 3D Motion Prediction From 2D Observations Using Deep Spatio-Temporal Graph CNNs", *International Conference on Computer Vision Workshop*, 2021.

Preprints

(⋆ indicates equal contribution)

H. Chen⋆, S. Cha⋆, H. Vikalo, "Task-Agnostic Federated Continual Learning via Replay-Free Gradient Projection", *Under Review*, 2025.

H. Chen, J. Li, W. Zhuang, H. Vikalo, L. Lyu, "Training-Free Layout-to-Image Generation with Marginal Attention Constraints", *Under Review*, 2024.

H. Chen, J. Li, W. Zhuang, C. Chen, L. Lyu, "Dual low-rank adaptation for continual learning with pre-trained models", *Under Review*, 2024.

Skills

- Platforms: Linux, Azure, AWS, GCP
- Programming: Python == LaTeX== Matlab > Java > C/C++
- Frameworks: Torch, Tensorflow, Git, Docker, Kubernetes, Skypilot

Community Services

- Reviewer for Conference:
  ICML(22,23,24,25), NeurIPS(22,23,24,25), ICLR(24,25), IJCAI(24,25), AAAI(25), CVPR(24,25).
- Reviewer for Journal:
  IEEE Transactions on Mobile Computing