Yicong (Bryce) Chen

949-558-4534 | ychen2229@wisc.edu | bryce-chen.github.io

Education

University of Wisconsin - Madison Sept. 2020 - expect graduation: May 2024

B.S. in Computer Engineering: Machine Learning & Data Science option

B.S. in Computer Science (2nd Major)

- **GPA: 4.0 / 4.0**

Awards

- **Wisconsin Hilldale Undergraduate Research Fellowship** (PI: Prof. Kangwook Lee) May 2023
- UW-Madison Dean's Honors List Fall 2020 - Spring 2023

Publications

FedGP: Buffer-based Gradient Projection for Continual Federated Learning

*Shenghong Dai, **Yicong Chen**, Jy-yong Sohn, S M Iftekharul Alam, Ravikumar Balakrishnan, Suman Banerjee, Nageen Himayat, Kangwook Lee*

Federated Learning Systems (FLSys) Workshop @ MLSys 2023 • Oral Presentation • **Best Paper Award**

Under review at International Conference on Learning Representations (ICLR) 2024

Coded Prompts for Large Language Models

*Ziqian Lin, **Yicong Chen**, Yuchen Zeng, Kangwook Lee*

Robustness of Few-shot and Zero-shot Learning in Foundation Models (R0-FoMo) Workshop @ NeurIPS 2023

Zero-shot Improvement of Object Counting with CLIP

Ruisu Zhang, **Yicong Chen***, Kangwook Lee*

Robustness of Few-shot and Zero-shot Learning in Foundation Models (R0-FoMo) Workshop @ NeurIPS 2023

Research Experience

Undergraduate Researcher, advised by Prof. Kangwook Lee May 2022 - Present

Benchmarking Visual In-Context Learning for Multimodal Large Language Models (in progress)

- Explored the foundation model's in-context learning capabilities for the task pattern: x=text, y=image. This involves mapping from low-dimension input (text) to high-dimension output (image), a concept unexplored in current in-context learning literature.
- Aimed to establish a benchmark and assess the in-context capabilities of all current multimodal LLMs on it, paving the way for future work to further enhance the in-context abilities of multimodal LLMs.

Coded Prompts for Large Language Models

- Introduced coded prompts, inspired by coding theory, to process multiple inputs simultaneously in Large Language Models (LLMs), enhancing task performance. When viewing LLM inference as a noisy communication channel, coded prompt has the potential to protect against information lost.
- Validated this approach through experiments on two specially designed tasks for LLMs: maximum prime number and sentence toxicity, demonstrating the potential of coded prompts.

FedGP: Buffer-based Gradient Projection for Continual Federated Learning

- Designed a novel algorithm which mitigates forgetting by leveraging aggregated buffer gradients, ensuring the retention of prior knowledge across clients in Continual Federated Learning.
- Conducted rigorous evaluations on various datasets, including CV benchmark datasets, NLP classification datasets, and streaming decentralized automotive object detection dataset, achieving state-of-the-art results.
- Enhanced existing continual learning baselines using our method, further improving performance.

Zero-shot Improvement of Object Counting with CLIP

- Conducted analyses of CLIP's counting performance across diverse objects using a specially collected dataset, unveiling its non-uniform proficiency in counting different objects.
- Pioneered a zero-shot technique to extract a counting-specific vector from CLIP's text embedding space, enhancing the counting accuracy of CLIP by integrating this vector with the original prompt embedding.
- Demonstrated the method's versatility by showcasing its potential in guiding Text-to-Image generation models, enabling them to generate images aligned with counting prompts.

Mixed Sample Data Augmentation with Self-Distillation in Small Data Regimes

- Investigated the constraints of data availability in image classification and explored the crucial role of Mixed Sample Data Augmentation (MSDA) in scenarios with limited data for enhancing model generalization.
- Enhanced the efficacy of MSDA by introducing self-distillation (SD) for relabeling, thereby addressing the issue of inaccurate labeling common in most MSDA methods, such as mixup and cutmix.

Super-Resolution Emulation of Large Cosmological Fields with Diffusion Model

- Enhanced low-resolution cosmic data into high-resolution images using diffusion to aid dark matter research.

Undergraduate Researcher, advised by Prof. Dane Morgan Jan. 2022 - May 2022

Acceleration of molecular machine learning

- Integrated Nystroem into the kernel training process with FCHL and Kernel Ridge Regression.
- Reduced computational time to approximately one-third of the original model using Nystroem.

Projects

WISC-SP23 architecture microprocessor design Spring 2023

- Designed and implemented a 16-bit, 5-stage pipelined processor (WISC-SP23) using Verilog.
- Developed a two-way set associative instruction cache, a multi-cycle main memory, and other optimizations.

Online Twitter Bot Detection with Nature Language Processing Fall 2022

- Preprocessed and extracted features from 10k Twitter datasets using Bag of Words, TF-IDF, and Doc2Vec.
- Evaluated classifiers including Support Vector Machine, Logistic Regression, Naive Bayes, k-Nearest Neighbors, and Random Forest for their capacity to differentiate between bots' tweets and those of humans.

Skills

Language:	Python, Java, C++, C, MATLAB, Verilog, HTML/CSS, JavaScript
Tools:	ChatGPT, Latex, Wandb, Amazon Web Services, Jupyter Notebook, VS Code, Google Colab