Brian Nlong Zhao

Los Angeles, Shanghai

briannlz@usc.edu | +1(213)610-4699 | github.com/briannlongzhao
briannlongzhao@gmail.com | (+86)14721002407 | briannlongzhao.github.io

Education

Aug 2018 – May 2024

University of Southern California, Los Angeles, CA

- M.S. in Computer Science, Progressive Degree Program (3.9/4)
- B.S. in Computer Engineering and Computer Science, *Summa Cum Laude* (3.94/4)
- B.A. in Applied and Computational Mathematics, *Summa Cum Laude* (3.95/4)
- Minor in Astronomy (3.87/4) and Music Recording (4/4)

Major Courses:
Machine Learning, Computer Vision, Natural Language Processing, Distributed Computing, Operating System, System-on-Chip, Computer Architecture, Computer Network, Software Development, Algorithms, Data Structures, Statistics, Optimization Theory, ML Theory, Probability Theory, Combinatorics, Numerical Methods, Modeling and Computer Simulation, Linear Algebra, Multivariable Calculus

Research Interests

I am interested in Multimodal Generative AI and Computer Vision in general, and their applications towards human-centric and data-centric AI. As a newcommer to the field of AI research, I have broad interests and open to other topics as well. Some of the specific interests include:

- Controllable and creative image/text/3D generation
- Multimodal pretraining, learning, large language models and their applications.
- Controllable text and image synthetic data generation.
- Other general computer vision tasks such as detection, tracking, matching, super-resolution, calibration.

Selected Publications

DreamDistribution: Prompt Distribution Learning for Text-to-Image Diffusion Models
Brian Nlong Zhao, Yuhang Xiao, Jiashu Xu, Xinyang Jiang, Yifan Yang, Dongsheng Li, Laurent Itti, Vibhav Vineet, Yunhao Ge
Submitted to ECCV 2024. [project] [paper] [code]

Large Multimodal Model for Real-World Radiology Report Generation
Brian Nlong Zhao, Xinyang Jiang, Xufang Luo, Yifan Yang, Bo Li, Zilong Wang, Javier Alvarez-Valle, Matthew P. Lungren, Dongsheng Li, Lili Qiu
Submitted to ECCV 2024. [paper]

Beyond Generation: Harnessing Text to Image Models for Object Detection and Segmentation
Yunhao Ge, Jiashu Xu, Brian Nlong Zhao, Neel Joshi, Laurent Itti, Vibhav Vineet.
arXiv preprint, 2023. [paper] [code]

EM-Paste: EM-guided Cut-Paste for Image-level Weakly Supervised Instance Segmentation
Yunhao Ge, Jiashu Xu, Brian Nlong Zhao, Laurent Itti, Vibhav Vineet.
arXiv preprint, 2022. [paper] [code]

Progressive Motion Coherence for Remote Sensing Image Matching.
Yizhang Liu, Brian Nlong Zhao, Shengjie Zhao, Lin Zhang.
IEEE Transactions on Geoscience and Remote Sensing, 2022. [paper]

Scene Text Image Super-Resolution via Parallelly Contextual Attention Network
Cairong Zhao, Shuyang Feng, Brian Nlong Zhao, Zhijun Ding, Jun Wu, Fuming Shen, and Hengtao Shen.
ACM Multimedia 2021. [paper] [code]

Research and Work Experience

May 2023 – Present
Microsoft Research Asia, Shanghai, China
Research Intern. Advisor: Dr. Xinyang Jiang, Dr. Dongsheng Li
- Working on adapting Multimodal LLM for medical image and report generation task.
- Build an instruction-tuning medical report generation dataset leveraging GPT for real-world radiology report generation tasks with context input.
- Submitted a paper as first author to ECCV 2024.

Apr 2022 – Present
iLab, University of Southern California, Los Angeles, CA
Research Assistant. Advisor: Dr. Yunhao Ge, Prof. Laurent Itti
- Lead a project on personalizing text-to-image model using prompt distribution learning.
- Submitted a paper as first author to ECCV 2024.
- Co-authored two papers on synthetic dataset for weekly supervised object detection and image segmentation.
- Designed and implemented algorithms to extract foreground segments from image and algorithms for language-driven image foreground and background augmentation.

Jun 2022 – Mar 2023
Visual Intelligence and Multimedia Analytics Laboratory, USC Information Sciences Institue, Marina del Rey, CA
Research Intern. Advisor: Joe Mathai, Prof. Peter Beerel,
- Participated in a project funded by DARPA, improved the architecture and implemented the training and testing pipeline for a transformer-based object tracking model on different datasets.
- Implemented and improved tracking algorithm and post-processing pipelines, improved mIDF1 score on BDD100K dataset from 27.2 to 49.8, on par with state-of-the-art.
- Helped migrate the model to FPGA for in-pixel processing and demonstration.

Jun 2020 – Dec 2022
School of Software Engineering, Tongji University, Shanghai, China
Visiting Research Assistant. Advisor: Prof. Lin Zhang
- Co-authored multiple papers on remote sensing, feature matching, camera calibration and image super-resolution.
- Implemented multiple algorithms for data and image processing, transformation, augmentation needed for experiments.

Jul 2020 – Oct 2020
ByteDance, Shanghai, China
Research and Development Intern.
- Joined a group developing ByteRTC, a real-time audio and video communication SDK used by both in-house software, such as TikTok and Lark, and business partners' software.
- Helped implement software architecture components including MessageBus for module decoupling and LogReport for event logging and debugging.
- Led a group of 5 interns on maintaining the software architecture and writing unit tests for the SDK project using GoogleTest; improved unit test code coverage from 10% to 60%.
- Composed a guide on writing unit tests specifically for ByteRTC and conducted a presentation to the group and leaders.

Jun 2018 – Aug 2018
PingAn Technology, Shanghai, China
Research and Development Intern.
- Joined a data science group working on a house price prediction competition on Kaggle.
- Built and tested several regression models decision tree models using scikit-learn.
- Implemented data cleaning and preprocessing programs for price prediction models.

Technical Skills

- Programming language: Python, C, C++, MATLAB, Verilog, Shell
- Machine Learning: PyTorch, Hugging Face, scikit-learn, Detectron2, Weights & Biases

- Other tools: LaTeX, Markdown, Git, Linux

Honors

Oct 2021	**USC Steven and Kathryn Sample Renaissance Scholar Distinction**
Sep 2021	**USC Associates Senior Scholars**
Jul 2020	**USC Academic Achievement Award**
Jan 2019–Jun 2022	**Dean's List, USC Viterbi School of Engineering** **Dean's List, USC Dornsife College of Letters, Arts and Sciences**

Academic Projects

Oct 2022–Nov 2022	**Text-to-Image Generative Model using Variational Autoencoder** Worked on an individual project of a simple text-to-image generative model by conditioning a Variational Autoencoder and integrating Contrastive Language-Image Pre-Training to it.
Aug 2022–Dec 2022	**PrevWORKS** Worked in a group of 4 to develop a website for a company that aims to provide a platform for companies and employees to manage work injuries and health issues. Mainly responsible for the database and the statistical learning module.
Mar 2022–May 2022	**A Parallel Implementation of Histogram of Oriented Gradient (HOG) Algorithm** Designed a parallel implementation of HOG algorithm using C and CUDA, examined its performance on different hardware environments and reached speedup of 50 compared to serial implementation.
Oct 2021–Dec 2021	**Greedy Scheduling for Cloud Computing Clusters** Designed and simulated a simple greedy scheduling algorithm to schedule tasks on computing clusters based on CPU and memory parameters of virtual machines.
Jul 2021–Aug 2021	**Image Colorization on CIFAR-10** Built and trained an image colorization deep CNN model that outputs colorized images from gray-scale images using TensorFlow.
Nov 2020–Dec 2020	**Bit Racing FPGA Game** Designed and implemented a simple racing game runs on Digilent Nexys 4 FPGA board with a VGA monitor.
Apr 2020–May 2020	**Remote CPU Temperature Monitoring System** An IoT system that can monitor and graph real-time CPU information on a remote machine and Raspberry Pi based on MQTT protocol. Implemented the code for obtaining and publishing data on host machine, message subscription and real-time processing and visualization on remote machine, and message subscription and LED monitoring on Raspberry Pi, produced 4-page documentation.
Mar 2020–May 2020	**353NET** Implemented a simulate internet system using self-designed protocols, consists of network, transport, and application layer.
Sep 2019–Dec 2019	**Rebound: A Laser Chess Game** An online multiplayer laser chess game. Led a team of 6, responsible for overall code organization, implementation of game logic, music and sound effects; produced 32-page documentation.