Zitong Huang

213-569-6645 | chuang95@usc.edu | linkedin.com/in/zitong-huang | zitonghuangcynthia.github.io

EDUCATION

University of Southern California   Sept.2022-May 2026(Expected)
Viterbi School of Engineering
Bachelor of Science Computer Engineering and Computer Science
  GPA: 3.95/4.0

  Honors: MHI ECE Undergraduate Scholar (top 6 selected from Electrical and Computer Engineering department), USC Women in Science and Engineering Summer Research Scholarship, Dean's List, Engineering Honor Program, Seventh place in the third 'Jittor AI Challenge' organized by Beijing National Research Center and the Joint Laboratory of Tsinghua University and Tencent

Standardization Examination: TOEFL:111(R30+L28+**S28**+W25)

PUBLICATIONS

Zitong Huang, Mansooreh Montazerin, and Ajitesh Srivastava. "Simultaneous Weight and Architecture Optimization for Neural Networks." NeurIPS 2024 Workshop on Fine-Tuning in Modern Machine Learning (FITML). Accepted. arXiv:2410.08339. https://arxiv.org/abs/2410.08339

Zitong Huang, Mansooreh Montazerin, and Ajitesh Srivastava. "SWAT-NN: Simultaneous Weights and Architecture Training for Neural Networks in a Latent Space." 2025 IEEE International Conference on Big Data. Accepted. arXiv:2506.08270. https://arxiv.org/abs/2506.08270

RESEARCH INTEREST

- Efficient Deep Learning, Automated Machine Learning, Neural Architecture Search, Model Compression and Pruning, Lightweight Post-Training Adaptation for Large Language Models

EXPERIENCE

Enhancing Factuality and Reasoning in Large Language Models with Explainable and Efficient Token-Level Supervision
Advisor: Robin Jia, Assistant Professor   Aug.2025-Present

- Reformulated the LLM next-token prediction task into a token-level supervision framework using a binary classification objective, enabling more expressive and cleaner training signals
- Trained and evaluated multiple LLM families on a 30k-instance token-level dataset derived from FAVA to analyze model behavior under token-level factual supervision
- Improved generalization and reduced hallucination in long-context summarization, achieving an average 3% factuality gain on the AggreFact OOD dataset compared with fine-tuned baselines
- Leveraged LoRA-based lightweight adaptation as an efficient post-training steering mechanism, enhancing both efficiency and explainability in model alignment

USC Viterbi School of Engineering and Network Science Institute, Northeastern University
Surrogate Graph Neural Modeling of Epidemic Dynamics for Efficient Trajectory Prediction
Advisors: Ajitesh Srivastava, Research Assistant Professor, Alessandro Vespignani, Director and Sternberg Family Distinguished Professor   Jul.2025-Sept.2025

- Proposed a surrogate modeling approach for metapopulation epidemic dynamics, enabling efficient prediction of system trajectories without directly solving nonlinear ODEs
- Constructed graph-based simulations from unlinearized ODE formulations and trained a GNN-encoded neural ODE model on five different epidemic models, achieving accurate time-series predictions
- Extended the framework to aggregate variables of interest by inserting auxiliary nodes in the graph representation, allowing flexible and efficient prediction of arbitrary trajectory functions

USC Viterbi School of Engineering

Project1: Sparse and Interpretable Symbolic Regression via Latent Neural Embedding Optimization
Advisor: Ajitesh Srivastava, Research Assistant Professor   Jul.2025-Present

- Replaced the traditional ADMM-based pruning process with a custom Neural Architecture Search (NAS) framework in an existing symbolic regression pipeline
- Encoded diverse symbolic expressions into a unified matrix-based representation to enable latent embedding and efficient search across varying expression structures
- Integrated a pretrained symbolic expression classifier that trained on 25,000 expressions sampled from GPT and randomly constructed formulas into the NAS-guided pruning process to guide the search toward human-plausible and interpretable expressions

Project2: Simultaneous Weights and Architecture Training for Neural Networks in a Latent Space
Advisor: Ajitesh Srivastava, Research Assistant Professor   Jan.2024-May.2025

- Studied techniques of AutoML and Neural Architecture Search through extensive literature study, including search space design, differentiable search methods and surrogate performance prediction
- Proposed a novel gradient-descent based neural network search method that can simultaneously search for sparse and compact neural network architectures and weights, reducing the number of non-zero weights by an average of 7% per task compared to DARTS, while maintaining comparable performance across 54 regression datasets
- Implemented and evaluated a multi-scale encoder-decoder framework, conducting large-scale training on a dataset of 1.5M samples to validate scalability and robustness
- Designed a custom loss function to promote computational efficiency in the searched neural networks
- Fine-tuned and adapted pretrained models, including large language models (e.g., GPT-based) and other deep learning architectures, and applied LoRA-based lightweight adaptation to improve computational efficiency and scalability

School of Information Science and Technology, Tsinghua University
Enhancing Unsupervised Semantic Segmentation with SAM Backbone
Advisor: Shimin Hu, Professor   Jun.2023-Dec.2023

- Investigated Large-Scale Unsupervised Semantic Segmentation by replacing ResNet backbone with SAM and designing a novel task workflow, enhancing Mean Intersection over Union from 29 to 35
- Solved the camera projection problem under both rotational and translational adjustments, which is currently being utilized in the generation of panoramic image projects

LEADERSHIP AND INVOLVEMENT

USC Center for AI in Society's Student Branch (CAIS++)
Cohort Member and Team Leader   Aug.2023-Present

- Directed a 7-member team in developing a benchmark for bimanual object manipulation using Robosuite and ManiSkill, establishing standardized evaluation protocols for robotic learning tasks
- Supervised a 3-member team in building a resume classification system, benchmarking architectures such as RNNs, LSTMs, and achieving notable improvements in classification accuracy and model robustness

ACADEMIC PROJECTS

Web Application Using Java and MySQL   Aug.2023-Dec.2023

- Led a team in a software engineering course to develop a web-based fitness application
- Implemented a back end with Java and MySQL, managing data for user profiles, workout goals, and privacy settings, while developing a responsive frontend with interactive visualizations for seamless user experience

SKILLS

Computer Programming: C++, Python, Pytorch, MATLAB, HTML, CSS, MeshLab, Java