

YANZHAO WU

1101 N Fair Oaks Ave, Apt 213, Sunnyvale, CA, 94089, USA

✉ yanzhaowumail@gmail.com ☎ +1 404-279-2853 🏠 yanzhaowu.me

Education

Georgia Institute of Technology

Ph.D. in Computer Science

Aug. 2017 – May 2022

Atlanta, GA, USA

University of Science and Technology of China (USTC)

B.E. in Computer Science and Technology

Sep. 2013 – Jul. 2017

Hefei, Anhui, China

Research Interests

- Systems for Machine Learning
- Machine Learning for Systems
- Big Data Systems & Analytics
- Edge AI Systems

Experience

Meta Platforms, Inc.

Research Scientist

May 2022 – Present

Menlo Park, CA 94025

- **Model and Feature Exploration:** Explore and advance machine learning techniques and applications to improve overall system efficiency and performance.

Georgia Institute of Technology

Graduate Research Assistant

Aug. 2017 – May 2022

Atlanta, GA 30332

- **High-performance Object Detection on Edge Devices:** Build an efficient framework for supporting various object detection/tracking models and achieving high performance on multiple edge devices.
- **High Accuracy and Robust Ensemble of Deep Neural Networks:** Design and implement an ensemble framework to improve deep neural network accuracy and optimize inference robustness on GPUs and edge devices.
- **Semi-automatic Hyperparameter Tuning for Deep Neural Networks:** Accelerate deep learning training and improve the training efficiency via semi-automatic hyper-parameter tuning.
- **Experimental Analysis and Optimization of Deep Learning Frameworks:** Analyze the hyper-parameters and core components of Deep Learning (DL) and optimize DL frameworks by tuning data and hardware related parameters.

Facebook, Inc.

Software Engineer Intern

Summer 2020, Summer 2021

Menlo Park, CA 94025

- **Data-efficient Learning with DNN Ensembles:** Study the data efficiency of DNN ensemble models and design effective subsampling strategies to improve data efficiency for training ML models. (Summer 2021)
- **Pipeline Parallelism for Deep Learning Recommendation Models:** Apply pipeline parallelism into Facebook deep learning recommendation models to accelerate distributed recommendation model training. (Summer 2020)

IBM Research

Research Intern

Summer 2018, Summer 2019

San Jose, CA 95120

- **A Performance Study of Deep Learning with the IBM High-performance Storage System:** Conduct a comprehensive performance analysis of the IBM Comanche storage system with different storage devices, such as persistent memory and SSD, on popular deep learning workloads. (Summer 2019)
- **Accelerating Deep Learning with Direct-to-GPU Storage:** Integrate the IBM Direct-to-GPU storage system into Caffe to obtain over 2× performance improvement by reducing the overhead of data transmission. (Summer 2018)

Publications

- [1] *Selecting and Composing Learning Rate Policies for Deep Neural Networks*
Yanzhao Wu and Ling Liu
ACM Transactions on Intelligent Systems and Technology (TIST)
- [2] *Boosting Deep Ensemble Performance with Hierarchical Pruning*
Yanzhao Wu and Ling Liu
21st IEEE International Conference on Data Mining (ICDM 2021)

- [3] *Transparent Network Memory Storage for Efficient Container Execution in Big Data Clouds*
Juhyun Bae, Ling Liu, Ka-Ho Chow, **Yanzhao Wu**, Gong Su, and Arun Iyengar
2021 IEEE International Conference on Big Data (IEEE BigData 2021)
- [4] *Learning Text-Image Joint Embedding for Efficient Cross-Modal Retrieval with Deep Feature Engineering*
Zhongwei Xie, Ling Liu, **Yanzhao Wu**, Lin Li, and Luo Zhong
ACM Transactions on Information Systems (TOIS)
- [5] *Parallel Detection for Efficient Video Analytics at the Edge*
Yanzhao Wu, Ling Liu, and Ramana Kompella
2021 IEEE International Conference on Cognitive Machine Intelligence (CogMI 2021)
- [6] *RDMAbox : Optimizing RDMA for Memory Intensive Workload*
Juhyun Bae, Ling Liu, **Yanzhao Wu**, Gong Su, and Arun Iyengar
2021 International Conference on Collaborative Computing: Networking, Applications and Worksharing (CIC 2021)
- [7] *Gradient-Leakage Resilient Federated Learning*
Wenqi Wei, Ling Liu, **Yanzhao Wu**, Gong Su, and Arun Iyengar
41st IEEE International Conference on Distributed Computing Systems (ICDCS 2021)
- [8] *Boosting Ensemble Accuracy by Revisiting Ensemble Diversity Metrics*
Yanzhao Wu, Ling Liu, Zhongwei Xie, Ka-Ho Chow, and Wenqi Wei
2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2021)
- [9] *Learning TFIDF Enhanced Joint Embedding for Recipe-Image Cross-Modal Retrieval Service*
Zhongwei Xie, Ling Liu, **Yanzhao Wu**, Lin Li, and Luo Zhong
IEEE Transactions on Services Computing (TSC)
- [10] *Promoting High Diversity Ensemble Learning with EnsembleBench*
Yanzhao Wu, Ling Liu, Zhongwei Xie, Juhyun Bae, Ka-Ho Chow, and Wenqi Wei
2020 IEEE International Conference on Cognitive Machine Intelligence (CogMI 2020)
- [11] *Cross-Modal Joint Embedding with Diverse Semantics*
Zhongwei Xie, Ling Liu, **Yanzhao Wu**, Lin Li, and Luo Zhong
2020 IEEE International Conference on Cognitive Machine Intelligence (CogMI 2020)
- [12] *Memory Abstraction and Optimization for Distributed Executors*
Semih Sahin, Ling Liu, Wenqi Cao, Qi Zhang, Juhyun Bae, and **Yanzhao Wu**
2020 International Conference on Collaborative Computing: Networking, Applications and Worksharing (CIC 2020)
- [13] *Adversarial Deception in Deep Learning: Analysis and Mitigation*
Wenqi Wei, Ling Liu, Margaret Loper, Ka-Ho Chow, Mehmet Emre Gursoy, Stacey Truex, and **Yanzhao Wu**
2020 IEEE International Conference on Trust, Privacy and Security in Intelligent Systems and Applications (TPS 2020)
- [14] *Adversarial Objectness Gradient Attacks in Real-time Object Detection Systems*
Ka-Ho Chow, Ling Liu, Margaret Loper, Juhyun Bae, Mehmet Emre Gursoy, Stacey Truex, Wenqi Wei, and **Yanzhao Wu**
2020 IEEE International Conference on Trust, Privacy and Security in Intelligent Systems and Applications (TPS 2020)
- [15] *Efficient Orchestration of Host and Remote Shared Memory for Memory Intensive Workloads*
Juhyun Bae, Gong Su, Arun Iyengar, **Yanzhao Wu**, and Ling Liu
2020 International Symposium on Memory Systems (MEMSYS 2020)
- [16] *Understanding Object Detection Through An Adversarial Lens*
Ka-Ho Chow, Ling Liu, Mehmet Emre Gursoy, Stacey Truex, Wenqi Wei, and **Yanzhao Wu**
2020 European Symposium on Research in Computer Security (ESORICS 2020)
- [17] *A Framework for Evaluating Client Privacy Leakages in Federated Learning*
Wenqi Wei, Ling Liu, Margaret Loper, Ka-Ho Chow, Mehmet Emre Gursoy, Stacey Truex, and **Yanzhao Wu**
2020 European Symposium on Research in Computer Security (ESORICS 2020)
- [18] *Cross-layer Strategic Ensemble Defense against Adversarial Examples*
Wenqi Wei, Ling Liu, Margaret Loper, Ka Ho Chow, Emre Gursoy, Stacey Truex, and **Yanzhao Wu**
2020 International Conference on Computing, Networking and Communications (ICNC 2020)

- [19] *Demystifying Learning Rate Policies for High Accuracy Training of Deep Neural Networks*
Yanzhao Wu, Ling Liu, Juhyun Bae, Ka-Ho Chow, Arun Iyengar, Calton Pu, Wenqi Wei, Lei Yu, and Qi Zhang
2019 IEEE International Conference on Big Data (IEEE BigData 2019)
- [20] *Denoising and Verification Cross-Layer Ensemble Against Black-box Adversarial Attacks*
Ka-Ho Chow, Wenqi Wei, **Yanzhao Wu**, and Ling Liu
2019 IEEE International Conference on Big Data (IEEE BigData 2019)
- [21] *Deep Neural Network Ensembles against Deception: Ensemble Diversity, Accuracy and Robustness*
Ling Liu, Wenqi Wei, Ka-Ho Chow, Margaret Loper, Emre Gursoy, Stacey Truex, and **Yanzhao Wu**
16th IEEE International Conference on Mobile Adhoc and Sensor Systems (MASS 2019)
- [22] *Memory Disaggregation: Research Problems and Opportunities*
Ling Liu, Wenqi Cao, Semih Sahin, Qi Zhang, Juhyun Bae, and **Yanzhao Wu**
39th IEEE International Conference on Distributed Computing Systems (ICDCS 2019)
- [23] *Experimental Characterizations and Analysis of Deep Learning Frameworks*
Yanzhao Wu, Wenqi Cao, Semih Sahin, and Ling Liu
2018 IEEE International Conference on Big Data (IEEE BigData 2018)
- [24] *Benchmarking Deep Learning Frameworks: Design Considerations, Metrics and Beyond*
Ling Liu, **Yanzhao Wu**, Wenqi Wei, Wenqi Cao, Semih Sahin, and Qi Zhang
38th IEEE International Conference on Distributed Computing Systems (ICDCS 2018)
- [25] *A Comparative Measurement Study of Deep Learning as a Service Framework*
Yanzhao Wu, Ling Liu, Calton Pu, Wenqi Cao, Semih Sahin, Wenqi Wei, and Qi Zhang
IEEE Transactions on Services Computing (TSC)
- [26] *CCAligner: A Token Based Large-Gap Clone Detector*
Pengcheng Wang, Jeffrey Svajlenko, **Yanzhao Wu**, Yun Xu, and Chanchal K. Roy
40th International Conference on Software Engineering (ICSE 2018)

Teaching

Georgia Institute of Technology

Graduate Teaching Assistant

- CS6220 Big Data Systems and Analytics (Fall 2021)
- CS6675/CS4675 Advanced Internet Computing Systems and Application Development (Spring 2018, Spring 2019, Spring 2020, Spring 2021, Spring 2022)
- CS6235/CS4220 Embedded Systems and Real-Time Systems (Fall 2018)

Guest Lectures

- Deep Learning Hyperparameter Optimization with GTDLBench and LRBench (CS6220 in Fall 2019)
- Introduction to Emulab, Hadoop and Spark (CS6675/4675 in Spring 2019 and Spring 2020, CS6220 in Fall 2020)
- Introduction to Amazon AWS and Google Colab (CS6675/4675 in Spring 2021, CS6220 in Fall 2021)

University of Science and Technology of China

Undergraduate Teaching Assistant

- CS1001A Computer Programming A (Fall 2015)

Talks

- 2021 IEEE International Conference on Cognitive Machine Intelligence (CogMI), Virtual, Dec. 13-15, 2021
- 2021 IEEE International Conference on Data Mining (ICDM), Auckland, New Zealand, Dec. 7-10, 2021
- 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Nashville, TN, June 19–25, 2021
- 2020 IEEE International Conference on Cognitive Machine Intelligence (CogMI), Atlanta, GA, Dec. 1-3, 2020
- 2019 IEEE International Conference on Big Data (IEEE BigData), Los Angeles, CA, Dec. 9-12, 2019
- 2018 IEEE International Conference on Big Data (IEEE BigData), Seattle, WA, Dec. 10-13, 2018

Reviewer

- Conference: ICDE 2018, UCC 2018, BDCAT 2018, ICDCS 2019, WWW 2021, CVPR 2022, ECCV 2022, AAAI 2023
- Journal: IEEE TKDE, ACM TOIT, JISA, DCN, Computers & Security, Information Sciences