Dr. Yanzhao Wu

Assistant Professor at Florida International University
11200 SW 8th St, CASE 212D, Miami, FL 33199, USA
yawu@fiu.edu +1 404-279-2853 yanzhaowu.me

Education

Georgia Institute of Technology **Aug. 2017 – May 2022**
Ph.D. in Computer Science *Atlanta, GA, USA*

University of Science and Technology of China (USTC) **Sep. 2013 – Jul. 2017**
B.E. in Computer Science and Technology *Hefei, Anhui, China*

Research Interests

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

Experience

Florida International University **Dec. 2022 – Present**
Assistant Professor in the Knight Foundation School of Computing and Information Sciences *Miami, FL*

Meta Platforms, Inc. **May 2022 – Dec. 2022**
Research Scientist in Ads Core ML *Menlo Park, CA*

- **Model and Feature Exploration:** Explore and advance machine learning techniques and applications to improve the overall efficiency and performance of large-scale Ads recommendation systems.

Georgia Institute of Technology **Aug. 2017 – May 2022**
Graduate Research/Teaching Assistant *Atlanta, GA*

- **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. **Summer 2020, Summer 2021**
Software Engineer Intern *Menlo Park, CA*

- **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 **Summer 2018, Summer 2019**
Research Intern *San Jose, CA*

- **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] *STDLens: Securing Federated Learning Against Model Hijacking Attacks*
Ka-Ho Chow, Ling Liu, Wenqi Wei, Fatih Ilhan, and **Yanzhao Wu**
2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2023)

[2] *Selecting and Composing Learning Rate Policies for Deep Neural Networks*
Yanzhao Wu and Ling Liu
ACM Transactions on Intelligent Systems and Technology (TIST)

[3] *Boosting Deep Ensemble Performance with Hierarchical Pruning*
Yanzhao Wu and Ling Liu
21st IEEE International Conference on Data Mining (ICDM 2021)

[4] *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)

[5] *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)

[6] *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)

[7] *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)

[8] *Gradient-Leakage Resilient Federated Learning*
Wenqi Wei, Ling Liu, **Yanzhao Wu**, Gong Su, and Arun Iyenger
41st IEEE International Conference on Distributed Computing Systems (ICDCS 2021)

[9] *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)

[10] *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)

[11] *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)

[12] *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)

[13] *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)

[14] *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)

[15] *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)

[16] *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)

[17] *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)

[18] *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)

[19] *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)

[20] *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)

[21] *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)

[22] *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)

[23] *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)

[24] *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)

[25] *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)

[26] *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)

[27] *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

Florida International University

Assistant Professor

- CAP4630/CAP5602: Artificial Intelligence (Intro to AI) (Spring 2023)

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

Professional Activities

- Program Committee: AAAI 2023, ICDCS 2023, IJCAI 2023
- Conference Reviwer: ICDE 2018, UCC 2018, BDCAT 2018, ICDCS 2019, WWW 2021, CVPR 2022, ECCV 2022, CVPR 2023, ICCV 2023, KDD 2023
- Journal Reviwer: IEEE TKDE, ACM TOIT, JISA, DCN, Computers & Security, Information Sciences, Knowledge-Based Systems

Awards

- IEEE CIC Best Paper Award, December 2021
- ICDM 2021 Student Attendance Award, December 2021
- College of Computing Student Travel Award, December 2020
- Qualified for Men's Singles in 2020 NCTTA South Regional Championships
- Outstanding Graduate Award (USTC), April 2017
- Fourth Place for 2016 ISC Student Cluster Competition, June 2016
- Excellent Student Scholarship (Top 3%, USTC), 2015-2016
- Leadership Scholarship, 2014-2015
- The Third Prize for Electromagnetism Paper Competition, June 2014

Open-source Projects

- EVA: Exploiting multi-model multi-device detection parallelism for fast video analytics at the edge.
- DP-Ensemble: Leveraging FQ-diversity metrics to identify high diversity ensemble teams with high performance.
- PipeDLRM: Using pipeline parallelism for training deep learning recommendation models.
- EnsembleBench: A set of tools for building high-quality ensembles for machine learning and deep learning models.
- LRBench: A semi-automatic learning rate tuning tool to improve the DNN training efficiency and accuracy.
- GTDLBench: A performance benchmark to measure and optimize mainstream deep learning frameworks.
- Comanche: Accelerating deep learning with Direct-to-GPU storage with a modified Caffe and DeepBench.
- CCAligner: A token based code clone detector for detecting large-gap copy-and-paste source codes.
- PRISM: Building the LTS and Game model checkers for PRISM, a widely applied model checker for system analysis.