ZIHAN ZHOU

PhD Applicant for Fall 2026
zhzhjycs@bu.edu | andyjobs.zhou@gmail.com | (315) 878-1254
linkedin.com/in/zihan-zhou-andy

EDUCATION

Boston University, Boston, Massachusetts — Sep 2023 - May 2025
Master of Science in Computer Science

Hobart and William Smith Colleges, Geneva, New York — Aug 2019 - May 2023
Bachelor of Science in Computer Science

RESEARCH EXPERIENCE

Graduate Research Assistant - Computer Vision & Deep Learning — Aug 2024 - Present
Boston University | Supervisor: Prof. Reza Rawassizadeh

- Leading two major research initiatives in computer vision and machine learning optimization
- **Pose Estimation Pipeline:** Building end-to-end 'Camera to 3D Human Model' system for motion capture
- Explored high-fidelity models including Meta's Sapiens (308 keypoints) and DensePose
- Implemented pose matching algorithms using MediaPipe, Dynamic Time Warping, and k-NN
- Currently optimizing models through Knowledge Distillation and LoRA-C fine-tuning
- **GradES Project:** Developed gradient-based early stopping algorithm for transformers
- Achieved 1.57-7.22x training speedup with improved accuracy on language and multimodal tasks

Graduate Research Assistant - Federated Learning — Feb 2025 - Present
Boston University | Supervisor: Prof. Avi Mohan

- Validated federated learning convergence theory through dual-track research methodology
- Built UltraFlwr framework integrating Flower and YOLOv8 for distributed object detection
- Discovered FedAvg achieves 12.3% better performance than FedSGD with 80% fewer rounds
- **Key Discovery:** Batch Normalization is essential for federated learning (18.6% performance swing)
- Proved BatchNorm prevents gradient diversity explosion in heterogeneous settings

PUBLICATIONS & PAPERS IN PROGRESS

Published/Submitted:

• **Wen, Q., Zeng, X., Zhou, Z., Liu, S., Hosseinzadeh, M., Su, N., Rawassizadeh, R.** (2025). GradES: Significantly Faster Training in Transformers with Gradient-Based Early Stopping. *Submitted to MLSys 2026*. arXiv:2509.01842

In Progress:

• Pose Estimation Pipeline: End-to-end Camera to 3D Human Model System (Target: CVPR 2026)
• Federated Learning: Critical Role of Batch Normalization in Heterogeneous Settings (Target: ICLR 2026)

PROFESSIONAL EXPERIENCE

Software Engineering Intern - GenAI Jun 2024 - Oct 2024
Bragr

- Developed RAG planner utilizing recursive retrieval for model prompting and fine-tuning
- Architected data pipeline using Azure Cosmos DB, LangChain, and Azure Functions
- Implemented semantic chunking for enhanced embedding retrieval efficiency

Large Language Processing Research Intern Jan 2023 - May 2023
Hobart and William Smith Colleges

- Conducted NLP research using Word2Vec embeddings and PyTorch models
- Developed theoretical framework for semantic construction in language models

SELECTED PROJECTS

Pose Matching - Real-Time Fitness Analysis via Edge Computing Sep - Dec 2024

- Developed edge computing application for real-time fitness pose matching with 60% accuracy
- Integrated OpenPose, MinHashing, and Dynamic Time Warping with sliding window analysis

Resume Advisor - AI-Powered Career Matching Platform Sep - Dec 2024

- Designed scalable platform using vector embeddings for job matching
- Implemented knowledge graph with chain-of-thought reasoning

Full Stack E-commerce Platform Jan - May 2024

- Built React frontend with Material-UI and Node.js/Express REST API backend
- Implemented state management and integrated with DummyJSON API

TECHNICAL SKILLS

Languages: Python, Java, C/C++, JavaScript, HTML/CSS, SQL, CUDA, PHP
ML/AI: PyTorch, TensorFlow, Transformers, Computer Vision (OpenPose, MediaPipe), NLP, Federated Learning
Frameworks: React.js, Node.js, Express.js, Spring/SpringBoot, Django, LangChain
Cloud & Tools: Azure, AWS, GCP, Docker, Git, Linux, MongoDB, MySQL, PostgreSQL, Redis