Max Zuo

max.zuo@gmail.com | maxzuo.ai | (978) 460-4107
github.com/maxzuo | /in/max-zuo/ | /users/7871685

Education

GEORGIA INSTITUTE OF TECHNOLOGY — **Atlanta, GA**

College of Computing

BS: Aug '18 – May '21
GPA: 4.00 / 4.00
MS: Aug '21 – Dec '22
GPA: 4.00 / 4.00

- Candidate for MS, Computer Science with a specialization in Machine Learning
- Relevant courses: *OOP, Data Structures & Algorithms, Artificial Intelligence, Machine Learning, Probability & Statistics, Combinatorics, Networking, Algorithms Honors, Computer Vision, NLP, Machine Learning Theory, Interactive Robot Learning, Human Machine Learning, Deep Learning, Cognitive Science*

Work Experience

GOOGLE — **MTV, CA**

Software Engineering Intern — *May '22 – Aug '22*

- Worked on the machine learning research teams **Tensorflow Model Garden** & **Tensorflow Vision** under CoreML to code, train, and improve open-vocabulary object detection models.
  - Worked on implementing the **ViLD** object detection framework.
- Presented papers to help keep the team updated on state-of-the-art works in different areas, including object detection, panoptic segmentation, and general vision techniques.
  - Papers presented to the team included: **CMT-Deeplab**, ***k*Means Mask Transformer**

OCULOGYX (OX) — **Bentonville, AR**

Research Engineer — *May '21 – Sep '21*

- Leading the development of mapping warehouse floors with SKU-level info to ~1m accuracy.
- Involved in business decisions with the CTO and CEO of the company.
- Worked on developing **Ox Orion**, a near real-time computer vision recognition for groceries.
  - 2-pass KNN combining approximate nearest neighbors for faster lookup speeds
  - Deep learning one-stage one-shot object detection
  - Pipelined algorithm using SIFT features, RANSAC homography, and triplet loss for object recognition and geometric verification
- Developed **Ox Automapper** product from scratch, a pedestrian GraphSLAM algorithm mapping warehouse and supermarket store floors with SKU-level information
  - GraphSLAM for pedestrian data using inertial (IMU) odometry
  - Deep learning sensor correction and sensor fusion for natural pedestrian walk routines.

GEORGIA INSTITUTE OF TECHNOLOGY — **Atlanta, GA**

Graduate Researcher (AI/ML & Robotics) — *Aug '21 – Present*

- Conducting research under Prof. Sonia Chernova on semantic rearrangement: the ability for a robot/planner to organize a scene without explicit detailed human instruction.
  - Working with PDDLStream, Graph NNs, and pose graphs

Graduate Researcher (Computer Vision & Unsupervised Learning) — *Aug '20 – Present*

- Conducting research under Prof. Thad Starner on AI Through Symbiosis (wearable technology, unsupervised learning) specializing in computer vision and SLAM.
  - Researching a new HMM-based algorithm, utilizing its model capacity to recover event labels in a weakly supervised manner, used to train deep vision and time-series models.

Graduate Teaching Assistant — *Aug '21 – Dec '21*, *Jan '22 – May '22* ***(HEAD TA)***

- TA/Head TA of the *Mobile & Ubiquitous Computing* course (i.e. wearables, HCI)
  - focus on teaching applied research methods, conducting user studies, prototyping

United States Citizen

Continues on next page →

Undergraduate Teaching Assistant *Jan '20 – May '20*

- Lead teaching assistant for *Machine Learning* (CS 4641), a fourth-year level course

IBM **Littleton, MA**

Software Engineering Intern *Jun '20 – Aug '20*

- Worked on IBM Food Trust Blockchain Transparent Supply
- Significantly expanded open-source **Recall Assistant** capabilities
  - Worked directly with customers to support complex, real recall scenario types
  - Currently in use by real IBM Food Trust customers including Walmart for faster, more accurate recall assistance
- Developed IBM cloud solutions for improving internal production pipeline
- Led "Farming Insights" application, designed app stack (database, server, and front-end)

SALLIE MAE INC **Newton, MA**

Software Development Intern *May '19 – Aug '19*

- Spearheaded project developing **Chatbot Integration Manager** (JS/Python Web App), AI integration tool for moving to AWS Lex chatbot from human chat using NLP, from scratch.
  - NLP logic developed in Python, using word2vec to calculate semantic text-similarity of Q&A pairs and rank answers using a proprietary ranking algorithm based on PageRank.

Publications

Unifying exemplar and prototype models of categorization 2023
Zuo, M., Marupudi, V., & Varma, S. (under review). *Proceedings of the 45th Annual Cognitive Science Society Conference, Sydney, Australia.*

ATCON: Attention Consistency for Vision Models 2022
Mirzazadeh, A., Dubost, F., Pike, M., Maniar, K., **Zuo, M.**, Lee-Messer, C., & Rubin, D. (2022). *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (pp. 1880-1889).*

Efficient Exploration via First-Person Behavior Cloning Assisted Rapidly-Exploring Random Trees 2022
Zuo, M., Schick, L., Gombolay, M., & Gopalan, N. (2022). *HRI 2022 Workshop - MLHRC*

Awards & Achievements

GVU Distinguished Masters' Finalist '22
HackGT '21 – First place overall & best design
GT Highest Honors '21 – 4.00 GPA for BS in CS
GT Sports Innovation '20 – Winner, computer vision football analysis
HackGT '19– NSA: Secure Code Challenge Winner
MIT Blueprint 2017 – First place

Personal Projects

All: github.com/maxzuo

Hypercut (HackGT, Oct 2021) – Video summary generator

- Using sentence transformers MPNet and TextRank to reduce the content of a video while maintaining as much pertinent information as possible.
- Wav2Vec2 + CTC for offline transcription, Google Cloud Speech API for online transcription

Datalytics (GT Sports Innovation, Mar 2020) – Computer vision tool to automatically analyze football footage

- Yard line extraction, score information extraction, and team formation extraction
- Action segmentation (detects start and end of plays)

Skills

Software Development	Python, Java, Go, C, SQL, JavaScript, TypeScript, HTML, CSS
Libraries	OpenCV, NumPy, Keras, Tensorflow, PyTorch, Scikit-Learn Firebase, React, Flask, JQuery
Machine Learning	Computer vision, Object detection, Few/one-shot learning, Open-vocabulary detection, Convolutional Neural Networks, Graph Neural Networks, Transformers, HMMs, Autoencoders, SVM, Random Forests, Word2Vec, LSTM, Text/PageRank
Robotics	SLAM, Planning (PDDL/PDDLStream), Scene graphs, Learning from demonstrations, Inverse reinforcement learning
Foreign Languages	Fluent Mandarin, Spanish (National Spanish Exam 3 Bronze, NSE2 Silver)
Misc	JSON, Git, VSTS, Agile, Jenkins, IBM Cloud

United States Citizen