ZHIYU LIANG

3412-403 Church St, Toronto, ON M4Y 0C9, Canada
(+1)647-936-8870 ⋄ edward.liang@mail.utoronto.ca

EDUCATION

University of Toronto - St. George Campus *September 2015 - December 2019*
Honours Bachelor of Science with High Distinction, Computer Science GPA: 3.84/4.00
Focus on Artificial Intelligence and Computer Vision

EXPERIENCE

Qualcomm *February 2020 - July 2021*
Machine Learning Software Engineer *Markham, ON*

- Identified severe accuracy drop (74.7% lower) for 8-bit quantized Transformer models on Machine Translation tasks
- Developed a post-training quantization solution to recover the accuracy by 97.7%
- Developed quantization-friendly training techniques for Transformer models which achieved nearly no loss in post-training quantization

Vector Institute *May 2019 - December 2019*
Research Intern *Toronto, ON*

- Crawled and prepared large-scale high-quality text data (12 billion word tokens)
- Implemented and trained large-scale Transformer-based language models (GPT-2) in Tensor2Tensor
- Coded scripts for distributed training setup, available for hundreds of researchers and sponsors to use
- Profiled distributed training performance on Vector training clusters which helped decide the multimillion-dollar purchase of additional training hardware

Qualcomm *May 2018 - April 2019*
Machine Learning Software Intern *Markham, ON*

- Collaborated on developing the Quantization-Friendly MobileNet, won the 1st prize in 2018 IEEE Low Power Image Recognition Challenge and published a NeurIPS Workshop paper
- Automated quantization process which reduced the time for the pipeline from 2 hours to 5 seconds
- Developed an Android App FaceBlock to protect privacy in photos and videos for Qualcomm 2018 HackMobile, featured in Qualcomm Developer Network and won the best hackathon project out of 60+ teams and 250+ participants.

TEACHING

CSC384 Introduction to Artificial Intelligence *January 2020 - May 2020*
CSC320 Introduction to Visual Computing *January 2019 - May 2019*
CSC343 Introduction to Databases *September 2017 - January 2018*

PUBLICATIONS

Low Power Inference for On-Device Visual Recognition with a Quantization-Friendly Solution
Chen Feng, Tao Sheng, Zhiyu Liang, Shaojie Zhuo, Xiaopeng Zhang, Liang Shen, et al
Neural Information Processing Systems 2018 MLPCD 2
Montreal, Canada

2018 Low-Power Image Recognition Challenge
Sergei Alyamkin, et al
Computer Vision and Pattern Recognition 2018 LPIRC
Salt Lake City, Utah, United States

ACHIEVEMENTS / AWARDS

New College Student Council In-Course Scholarship *June 2019*

Recognition of overall academic achievement

Qualcomm Hack Mobile First Place *July 2018*

Best hackathon project out of 60+ participating teams and 250+ participants. Presented to the CEO of Qualcomm and judged by a panel of Executive VPs

IEEE Low Power Image Recognition Challenge First Prize *June 2018*

Achieved the highest image recognition accuracy under 30ms on SDM835 platform

New College Council In-Course Scholarship *September 2017*

Recognition of overall academic achievement

Dean's List Scholar *September 2016 - December 2019*

Awarded for maintaining a Cumulative GPA higher than 3.5 every year

PROJECTS

FaceBlock

· Developed an Android app using Java, TensorFlow, Qualcomm SNPE to protect people's privacy in live video streams by detecting, tracking and blocking unwanted faces with a selected emoji in real time
· Won 1st place in Qualcomm Hack Mobile 2018 out of 60+ teams and 250+ participants

iOS Weather App

· Designed and implemented the app UI with automatic dark mode enabled
· Implemented core functionality following MVC and Delegate Design Pattern
· Implemented networking with OpenWeatherMap API and data parsing
· Implemented location based weather presenting

Fake News Detection

· Crawled news headlines as training data from Kaggle and other sources
· Applied lemmatization and Part-of-Speech tagging for feature extraction
· Extracted 27 other features useful for sentiment analysis for each headline
· Trained a Random Forest (RF), MLP and AdaBoost classifiers which vote to produce predictions
· Won the 1st place out of 200+ submissions in the Fake News Challenge of CSC411/2515

SKILLS

Languages	Python, Swift, C/C++, Matlab, Java, Shell, JavaScript, R
Frameworks	PyTorch, TensorFlow, NumPy, OpenCV, Fairseq, Tensor2Tensor