# YIREN (AARON) ZHAO

## yaz21@cam.ac.uk

UK Contact: (0044) 07547842218

#### EDUCATION

University of Cambridge, Cambridge, UK Expected Grad. 2020 PhD in Computer Science University of Cambridge, Cambridge, UK Grad. June 2017 Master of Philosophy in Advanced Computer Science, Award of Distinction Imperial College London, London, UK Grad. June 2016 Bachelor of Engineering in Electrical & Electronic Engineering Award of honors: First Class Honors; Decile: Ranked 2nd of the year; Selected Awards and Honors EPSRC International Doctoral Studentship joint University of Cambridge Computer Laboratory Premium Scholarship 2017 Fully funded phd scholarship for 3.5 years. Willis Jackson Metal and Prize 2016 For excellence in academic performance, one award per academic year. Dean's List 2016

For top-ranked students in the department.

# Faculty of Engineering UROP Award,

2015

Award to undergraduate Imperial College students with excellence in research projects.

### Papers & Thesis

### Pay Attention to Features, Transfer Learn Faster CNNs

K Wang\*, X Gao\*, Y Zhao, X Li, D Dou, X Gao, and C Xu, in submission

# Blackbox Attacks on Reinforcement Learning Agents Using Approximated Temporal Information

Y Zhao\*, I Shumailov\*, C Han, X Gao, R Mullins and R Anderson, in submission

#### Focused Quantization for Sparse DNNs

Y Zhao\*, X Gao, R Mullins and C Xu

Thirty-third Conference on Neural Information Processing Systems (NeurIPS 2019)

#### Sitatapatra: Blocking the Transfer of Adversarial Samples

I Shumailov\*, X Gao\*, Y Zhao\*, R Mullins, R Anderson and C Xu, in submission

#### Automatic Generation of Multi-precision Multi-arithmetic CNN Accelerators for FPGAs

Y Zhao\*, X Gao\*, X Guo\*, J Liu, E Wang, R Mullins, P Cheung, G Constantinides and C Xu in submission

#### The Taboo Trap: Behavioural Detection of Adversarial Samples

I Shumailov\*, Y Zhao\*, R Mullins and R Anderson, in submission

#### Characterizing Sources of Ineffectual Computations in Deep Learning Networks

M Nikolic, M Mahmoud, Y Zhao, R Mullins and A Moshovos

International Symposium on Performance Analysis of Systems and Software 2019 (ISPASS 2019)

#### Dynamic Channel Pruning: Feature Boosting and Suppression

X Gao\*, Y Zhao\*, R Mullins and C Xu

International Conference on Learning Representations 2019 (ICLR 2019)

# To compress or not to compress: Understanding the Interactions between Adversarial Attacks and Neural Network Compression

Y Zhao\*, I Shumailov\*, R Mullins and R Anderson

The Conference on Systems and Machine Learning 2019 (SysML 2019)

#### Mayo: A Framework for Auto-generating Hardware Friendly Deep Neural Networks

Y Zhao\*, X Gao\*, R Mullins and C Xu

2nd International Workshop on Embedded and Mobile Deep Learning (Workshop of Mobisys) (EMDL 2018)

# Redundancy-Reduced MobileNet Acceleration on Reconfigurable Logic For ImageNet Classification

J Su, J Faraone, J Liu, Y Zhao, D Thomas, P Leong and P Cheung

14th International Symposium on Applied Reconfigurable Computing (ARC 2018)

### An Efficient Implementation of Online Arithmetic

Y Zhao, J Wickerson and G Constantinides

2016 International Conference on Field-Programmable Technology (FPT 2016)

### Improving Compression Pipelines For Convolutional Neural Networks

Master thesis

\* indicates equal contribution

#### RESEARCH EXPERIENCE

### Microsoft Research New England, Research Intern

June 2019-Oct 2019

Supervisor: Dr. Nicolo Fusi

- Working on Project AutoML
- Design novel hardware-aware AutoML methods

### Microsoft Research Redmond, Research Intern

June 2018-Oct 2018

Supervisor: Dr. Daniel Lo and Dr. Eric Chung

- Working on Project Catapult and Project Brainwave
- Design and implement novel compression techniques for neural network training
- The developed techniques are approved to be in the next generation training hardware architecture at Microsoft
- More than 4 US patents in application

#### Microsoft Research Cambridge, Research Intern

June 2017-Oct 2017

Supervisor: Dr. Hitesh Ballani

- Working on Project Sirius
- Design and research on new routing methods between racks in a datacenter
- Implemented a Top of Rack switch prototype using FPGAs

#### Ocado Technology 10x, Machine learning researcher (Intern)

June 2016-Oct 2016

- Ocado Technology 10x is the research department of Ocado, the world's largest online only grocery retailer
- Focused on designing an autonomous robot used for delivering grocery orders using Reinforcement Learning

#### Machine learning consultant at Tangi0

July 2016-Sep 2016

- Tangi0 is a R&D start-up company focusing on manufacturing novel materials
- My consulting work includes exploiting machine learning techniques (DNN and CNN) on signal processing and also gesture recognition using the company's novel materials

# Radio Frequency Laboratory in University of Electronic Science and Technology of China Supervisor: Prof. Guangjun Wen June 2014 - Oct 2014

- Implementing RFID Analog Circuits on small scale PCB(printed circuit board) for a commercial product
- Final product is currently on the commercial market

#### SKILLS

**Programming:** Python (proficient), C++, C, R, ARM Assembly, Verilog (proficient), Matlab (proficient), Haskell, LATEX

Operating Systems: Windows, Mac OS X, Linux

Practical skills: Compilers, Embedded systems, RTL design and PCB design