

## EDUCATION

<b>M.Sc.</b> in Electrical Engineering <b>University of Pennsylvania</b> Overall GPA: 3.85/4.0, Track: Circuits and Computer Engineering	09/2021-05/2023 PA, United States
<b>B.Eng.</b> with Honors of the First Class in Electronics and Electrical Engineering <b>University of Glasgow</b> (UoG)	09/2016-06/2020 Glasgow, United Kingdom
<b>B.Eng.</b> in Electronic Information Engineering <b>University of Electronic Science and Technology of China</b> (UESTC) Overall GPA: 3.7/4.0, Outstanding Graduate in UESTC (Distinction)	09/2016-06/2020 Chengdu, China
<b>Summer Program</b> in Electrical and Computer Engineering, <b>University of British Columbia</b>	07/2018-08/2018 Vancouver, Canada

## RESEARCH INTERESTS

Programming Languages, Compilers, Software Platforms for Hardware Accelerators; Heterogeneous/Parallel Computing

## PUBLICATIONS

N<sup>3</sup>H-Core: Neuron-designed Neural Network Accelerator via FPGA-based Heterogeneous Computing Cores (1<sup>st</sup> Author)

2022 ACM/SIGDA International Symposium on Field-Programmable Gate Arrays (FPGA)

IoT Enabled Smart Security Framework for 3D Printed Smart Home (1<sup>st</sup> Author)

2020 IEEE International Conference on Smart Internet of Things (SmartIoT)

High- and Low-Level Feature Enhancement for Medical Image Segmentation

2019 International Workshop on Machine Learning in Medical Imaging (MLMI)

## EMPLOYMENT

<b>GPU Design Verification Intern</b> , NVIDIA Corporation (US)	05-08/2022
<ul style="list-style-type: none"><li>Summer Intern at Streaming Multiprocessor (SM) Design Verification Team focusing on testlist optimization.</li><li>Designed and deployed Baysian Optimization flow on testlists to improve test coverage. The algorithm explores the knob space and tunes knob values with a test result feedback loop to improve the hitrate on specific targets.</li><li>Implemented a NN trained with random tests, which accelerates the BayOpt flow to generate optimized testlists.</li></ul>	
<b>Research Assistant</b> , advised by Prof. Li Jiang, Shanghai Jiao Tong Univ. & Shanghai Qi Zhi Institute	09/2020-07/2021
<ul style="list-style-type: none"><li>Researched the neural network (NN) compression and design space exploration on the FPGA-based NN accelerator.</li><li>Full-time worked in the Advanced Computer Architecture (ACA) lab at Shanghai Jiao Tong University (SJTU).</li></ul>	
<b>Teaching Assistant</b> , UESTC	09/2019-12/2019
<ul style="list-style-type: none"><li>Assisted instructor in designing labs, holding office hours, and grading homework for course Electronic Devices.</li></ul>	

## EXPERIENCES

<b>Design of Configurational Logic Block (CLB) in FPGA (in progress)</b> , UPenn	11-12/2022
<ul style="list-style-type: none"><li>Design the circuit-level implementation of a CLB with a 16:1 Look-up Table (LUT) using pass transistor logic for multiplexers optimized for the delay, with control logic to load the 6-T SRAM cells holding a custom truth table.</li></ul>	
<b>Processor Design with Superscalar Pipelined Datapath</b> , UPenn	02-05/2022
<ul style="list-style-type: none"><li>Designed a superscalar with two in-order pipelines for LC4 ISA, with multi-ported and bypassed register file.</li><li>Designed the stalling logic, squash logic, and bypassing logic for pipeline switch. Tested the processor on Zedboard.</li></ul>	

- Accelerating VGG16 DCNN with a cloud FPGA**, advised by **Prof. Jing (Jane) Li**, UPenn 11-12/2021
- Designed a VGG16 accelerator under heterogeneous computing framework OpenCL with AWS EC2 F1 Instance.
  - Applied General Matrix Multiplication (GEMM) algorithm in C++ OpenCL host code binded with Pytorch and systolic array kernel for matrix multiplications on FPGA. Optimized the kernel with HLS pragmas like array partition.
- Data Compressor Design and Acceleration on the Heterogeneous Platform Ultra96**, UPenn 10-12/2021
- Designed a compressor receiving data in real-time and reducing data size by identifying and reducing redundancy.
  - Implemented the flow with content-defined chunking, SHA-384 deduplication, and LZW compression steps.
  - Achieved C solution on the Single ARM Processor first. Identified performance bottlenecks and mapped the bottleneck functions on the FPGA logic. Optimized the throughput with data streaming, loop unrolling, pipeline, etc.
- N<sup>3</sup>H-Core (FPGA '22)**, advised by **Prof. Li Jiang**, Advanced Computer Architecture (ACA) Lab at SJTU 09/2020-07/2021
- Designed a heterogeneous DNN accelerator architecture called N3H-Core that consists of DSP- and LUT-centric computing units (aka. DSP-core and LUT-core) to fully exploit the on-chip resource of the target FPGA.
  - Constructed scalable and adaptable cost model across different DNNs and FPGA to precisely estimate the resource utilization, inference latency, and other metrics-of-interests for further design space exploration.
  - Applied the Reinforcement Learning (RL) technique to build the end-to-end optimization framework that automatically generates the architecture configurations (resource), dataflow (both cores), and DNN respectively.
- IoT Smart Home Design for Security (SmartIoT '20)**, advised by **Dr. Qammer H. Abbasi**, UoG 12/2019-04/2020
- Designed an MCU-based smart home with a visual-intelligent surveillance system supported by the Raspberry Pi.
  - Correlated the surveillance system with the door lock system innovatively to further enhance smart security.
- Medical Image Segmentation (MLMI '19)**, advised by **Prof. Guotai Wang**, UESTC 09-12/2019
- Implemented different fully convolutional networks (FCNs) as baseline models and compared their model size and accuracy on two challenging medical image segmentation datasets – skin lesions and spleen.
  - Proposed HLE-Net that utilizes the attention mechanism to selectively aggregate the optimal feature information and uses global semantic information to enhance the semantic consistency of high- and low-level features.

## HONORS & AWARDS

Outstanding Graduate in UESTC (Distinction)	12/2019
Outstanding Student Leadership Scholarship	11/2019
Outstanding Student Scholarship in UESTC (Top 10%)	10/2017 10/2018 10/2019
Honorable Mention of 2019 Mathematical Contest in Modeling (MCM)	04/2019

## SKILLS

**Knowledge:** FPGA accelerator design from chip architecture, physical layout to applications, Heterogeneous and parallel computing with High-Level Synthesis (HLS), Full-stack computer system from cmos devices to applications, System-on-Chip Architecture Design and Verification, Machine Learning and Optimization

**Programming:** C++, Python, Verilog/System Verilog, HLS, Linux Shell, Matlab

**Frameworks:** Pytorch, Quartus, Vivado, ModelSim, Vitis, Vitis HLS, Vitis Analyzer, OpenCL, Cadence, LTspice

**Hardware Platform:** Low Power: Xilinx PYNQ, Ultra96, Zedboard, Intel DE1-SoC board; High Perf: Amazon EC2 F1 Instances

**Operating System:** Linux-Ubuntu, Mac OS, Windows