Oliver Cassidy

ollyj.cassidy@gmail.com | P: +44 7484 232315 | https://github.com/ollycassidy13 | https://olly-cassidy.vercel.app/ | www.linkedin.com/in/oliver-cassidy-286ba3235

EDUCATION

Imperial College London

London, UK

MEng in Electronic and Information Engineering

Expected May 2027

Predicted First-Class Honours; Dean's List 2024; Ranked 2nd in the year

Relevant coursework: Pipelined ray-tracing engine implemented in RTL on PYNQ, pipelined RISC-V CPU in SystemVerilog, C90 compiler, FPGA/AWS based duel game system, WiFi controlled rover

The Manchester Grammar School

Manchester, UK

A Levels – Mathematics A*, Further Mathematics A*, Physics A*, Electronics A*; Winner of the Paton Electronics Prize

Jun 2023

PROFESSIONAL EXPERIENCE

Cloud Engineering Intern

London, UK

Jun 2025-Sep 2025

T. Rowe Price
Implemented a research environment for OpenAI agents in Azure with Bing search and MCP servers securely

- Implemented a research environment for OpenAr agents in Azure with onig search and MCF servers securely
- Coordinated with Microsoft and global teams to develop Bicep frameworks to support rapid and consistent prototyping
- Implemented crucial roles in AWS to limit costs and protect the deployment services

Private 1:1 Tutoring

Jan 2021-Sep 2024

- Used social media to market my own tutoring business and attract clients, leading to a full client roster and a waitlist
- Tutored over fifteen students at GCSE and A level in preparation for their examinations, leading to an increase in achieved grades

CONFERENCE PUBLICATIONS

ReducedLUT: Table Decomposition with "Don't Care" Conditions

London, UK

Paper published in the ACM/SIGDA International Symposium on Field-Programmable Gate Arrays 2025

Aug 2024-Feb 2025

- Lead-author of a paper focused on reducing the physical lookup table (P-LUT) utilisation of L-LUT based neural network (NN) models for ultra-low latency applications by introducing extra similarities within the data to allow for more effective decomposition
- Reduced the P-LUT utilisation by over 38% with a maximum test accuracy drop of 0.01%: https://doi.org/10.1145/3706628.3708823
- Presented this paper to leading academics at both ISFPGA 2025 and to the CAS research group at Imperial College London

RESEARCH EXPERIENCE

Ultra-Low Latency FPGA ML Reasearch

Jun 2024 – Present

- Adapted the open-source NeuraLUT toolflow to integrate Verilator testing, CUDA for improved inference and Vivado for synthesis of the model using Vivado along with modifying various L-LUT compression techniques.
- Implemented a new toolflow to integrate the use of the lossless CompressedLUT, and then ReducedLUT to NeuraLUT models
- Developed an imask-optimized input buffer for NeuraLUT-Assemble models, cutting latency to one-third of the previous best
- Implemented L-LUTs for an input quant layer to allow for end to end testing of models

FPGA-Based Brain Simulation

Mar 2025

Worked with a team from TU Delft to map gate activation and timestep functions to L-LUTs and applied decomposition techniques

ACADEMIC PROJECTS

CMATMUL - Cache Based Matrix Multiplication Kernel

Feb 2025-Mar 2025

- Developed an effective C++ based matrix multiplication kernel leading to over a 100x increase in throughput from a naïve solution
- Optimised the throughput by implementing cache-aware tiling, register blocking, an AVX2 microkernel and OpenMP parallelism

Collabify Mar 2024-Aug 2024

- Designed and created a website which introduced more collaboration to Spotify
- Used a release date weighted cosine similarity model in Python and the Spotify API to suggest new music personalised to the user
- Initialised and maintained a database using SQL to store account details and the users' liked songs to allow for a shared liked songs playlist

Remote Control Car from Logic

Dec 2022-Mar 2023

- Designed and built a remote control car using a RF transmitter/receiver pair, logic gates, counters and motor drivers
- Built the RF unit using a crystal oscillator and filter to obtain the carrier wave, and ASK modulation to transmit the information

ADDITIONAL

Programming Proficiencies: Advanced in C++, Python, PyTorch, HTML/CSS, CUDA, Verilator, Git; Proficient in C, JavaScript, SQL, Java, Verilog, SystemVerilog, React, Basic, Assembly languages, Flask, Tcl, Terraform

Awards: Dean's List (2024), Paton Electronics Prize (2023), Gold Crest Award, Gold Kangaroo in the Senior Maths Challenge, Silver in the Physics Olympiad and Silver Industrial Cadets Award (2022)

Interests: 1500m competitive track running for The Thames Valley Harriers (2023-Present), Raced nationally for junior cycling development teams and trained weekly with British Cycling (2020-2023)