

## EXPERIENCE

- MediumAI** · **Software Development Engineer — Hybrid, Toronto, Canada** · Freelance Jan. 2026 – Present
- Build API integrations for a voice-LLM powered medical scribe agent, enabling secure multilingual clinical records.
- iQIYI.com** · **Data Analyst — Remote, Beijing, China** · Internship Oct. 2025 – Dec. 2025
- Combined three-arm A/B testing with causal forests and decision-tree rules for deployable personalization, lifting renewal rates from 37% to 54.06%.
- Qualcomm Canada** · **RFIC Design Verification Engineer — Onsite, Markham, Canada** · Permanent Full-time Jun. 2024 – Jul. 2025
- Verified UWB RX analog front-end and WLAN CP-PLL across 500+ channels using UVM; built GRU/MLP models for analog gain and VCO tuning, replacing manual calibration across 1000+ targets.
- University of Toronto – WIRLab** · **Research Assistant — Remote, Toronto, Canada** · Contract Part-time Jul. 2021 – Sep. 2021
- Applied PCA-based embedding visualization to investigate semantic clusters for **Transformer-based NLP tasks**. [Code]
- Alphawave Semi** · **SerDes Digital Verification Engineer — Remote, Toronto, Canada** · Internship May 2020 – Jun. 2021
- Built UVM testbenches for SerDes subsystems (clocking, datapath, SRAM), expanding coverage to 50+ scenarios; Scaled Jenkins CI/CD regression from 4 to 60+ projects, enabling 15× throughput growth.

## EDUCATION

- University of Toronto** Sep. 2017 – Jun. 2022
- Bachelor of Applied Science, Electrical Engineering*
- Capstone:** CNN NPU overlay + compiler for Intel Stratix 10 NX; custom CONVU + DLS compiler for MobileNetV1, achieving 4.3× speedup over Brainwave. *Supervisors: Prof. Vaughn Betz, Andrew Boutros.*
- University of Toronto** Sep. 2022 – Jun. 2024
- Master of Engineering, Electrical & Computer Engineering*
- MEng Thesis:** Pilot Training for AoA and Channel Estimation in 5G MIMO-OFDM; evaluated LMMSE vs. Matrix-Pencil methods under 3GPP-compliant multicell models, showing **higher AoA accuracy** with limited antennas. *Supervisor: Prof. Raviraj Adve.*

## PUBLICATIONS | PEER-REVIEWED CONFERENCES

- Arash Ahmadian, Louis S.P. Liu, **Yue Fei**, Konstantinos N. Plataniotis, Mahdi S. Hosseini. *Pseudo-Inverted Bottleneck Convolution for DARTS Search Space*. IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH AND SIGNAL PROCESSING (ICASSP), 2023.
- Abnash Bassi, **Yue Fei**, Gilead Posluns, Mark C. Jeffrey. *Optimized Priority Scheduling for Faster Scalable Belief Propagation*. PROBABILISTIC GRAPHICAL MODELS (PGM), [In Submission], 2026.

## SELECTED ACADEMIC PROJECTS

### COMPUTER ARCHITECTURE & DIGITAL DESIGN SYSTEMS

- Parallel Beamforming & Cache Coherence (ECE1755):** Accelerated ultrasound beamforming by 7× (17 s → 2.5 s) using 16-thread SIMD parallelism and memory optimizations (`restrict`); designed and verified a 3-hop MSI/MESI directory protocol in Murφ with an Exclusive (E) state to reduce bus traffic. [Code]
- Parallel Graph Inference Scheduler (ECE1755 Project):** Built a multi-threaded asynchronous scheduler for LDPC, Ising, and Potts models using a **Stealing Multi-Queue (SMQ)**; optimized sparse graph traversal via **CSR layouts**, improving cache efficiency on 48-thread manycore systems.
- Computer Architecture (ECE552):** Implemented MIPS 6-stage pipelined CPU with hazard detection, Tomasulo OoO execution, perceptron branch prediction, MSI directory coherence, and a **Bouquet L1 prefetcher** (IPCP+DPT+GHB); evaluated area/latency tradeoffs using CACTI (~24% area, < 1 ns access).
- CNN NPU Overlay & Compiler (Capstone):** Designed a **VLIW-based Python compiler** packing 130-bit instruction chains with dependency tracking and NOP insertion; engineered a SystemVerilog RTL scheduler with **credit-based flow control**, achieving 450 MHz Fmax on Intel Stratix 10 NX.

### ERROR-CORRECTING CODES & GRAPHICAL INFERENCE

- Graph-Based Error-Correcting Codes:** Implemented LDPC, Polar, and Fountain/LT encoders/decoders over erasure channels, with custom simulators in Julia and MATLAB. [Code]
- Error Control Coding:** Implemented convolutional and BCH codes with Viterbi decoding and finite-field arithmetic over  $GF(2^m)$ .

## AWARDS & TALKS

- Edward S. Rogers Sr. Dept. Betz Entrance Scholarship (\$5,000) 2017
- Certificate in Engineering Business Jun. 2022
- Qualcomm Panel: *Demystifying Machine Learning* Mar. 2025

## TECHNICAL SKILLS

**Software & Tools:** C/C++, Python (NumPy/Keras), Julia, SystemVerilog, SIMD, MLIR/LLVM, ISA simulation, Graph IR, Git, Jenkins CI/CD, Quartus Prime, Cadence Virtuoso, Synopsys VCS