Nikhil Pratap Ghanathe

PhD Candidate at University of British Columbia

nikhilghanathe@ece.ubc.ca nikhilghanathe.github.io nikhilghanathe nikhil-pratap-ghanathe

Education

University of British Columbia (UBC) — *Vancouver, BC, Canada*

Doctor of Philosophy in Electrical and Computer Engineering — *September2019-PRESENT*

University of Florida (UFL) — *Gainesville, FL, USA*

Master of Science in Electrical and Computer Engineering — *August2014-May2016*

Visvesvaraya Technological University — *Bangalore, India*

Bachelor of Engineering in Electronics and Communication — *September2010 - July2014*

Work Experience

University of British Columbia — *Vancouver, Canada*

Graduate Research Assistant — *September 2019 - PRESENT*

- Devised novel on-chip monitoring and diagnostic solutions to preserve model reliability in resource-scarce setting
  - Resource-efficient uncertainty estimation mechanism outperforming prior works for corruption/out-of-distribution detection
  - Hyper-dimensional computing-inspired on-device diagnostics for pinpointing the causes of TinyML model failures
- Improved energy-efficiency of tiny neural networks through dynamic inference post-deployment
  - TinyML-optimized early-exit architecture with novel early-exit distillation, enabling average 33% reduction in FLOPS
- Developed a compiler for accelerating ML on resource-scarce low-power FPGAs
  - Enhanced ease of deployment for edge inference of classical ML algorithms on FPGAs, achieving a 2.5× speedup
- Collaborating with University of Victoria for supporting large language models at the edge through hierarchical inference

Microsoft Research — *Bangalore, India*

Research Fellow — *Aug. 2017 - Aug. 2019*

- Developed and integrated a FPGA-backend for a Domain Specific Language to run KB-sized ML models on low-cost FPGAs for IoT
  - Improved productivity and performance by achieving 10× speedup compared to Vivado High-level Synthesis Compiler and 211× better than microcontroller implementations
- Collaborated with University of Toronto to explore training of deep Neural Networks (LSTMs) on FPGAs

University of Florida — *Gainesville, FL-USA*

Research Engineer + Graduate Reseach Assistant — *May2015 - July2017*

- Co-developed Software & Firmware (Verilog) models for the CMS Level-1 Muon Track finder at the Large Hadron Collider at CERN
  - Led a team to investigate & formulate high-level synthesis techniques to boost productivity of FPGA-firmware development
  - Developed novel techniques and code optimizations for latency control, improving throughput and scheduling functions
  - Code-base integrated into Comapact Muon Solenoid Software framework (CMSSW)
  - Achieved improvement in latency (by one clock cycle) and resource consumption (~15%) using methods developed

Indian Space Research Organization — *Bangalore, India*

engineering intern — *Feb2014 - May2014*

- Designed a system to detect and estimate relative speed between star sensor and celestial bodies
  - Developed FPGA-based image acquisition hardware on a rad-hard FPGA; Novel algorithm reduced resource usage by 50%

Publications

DEBUG-HD: Debugging TinyML models on-device using Hyper-Dimensional computing

Nikhil P Ghanathe, Steven Wilton

Machine learning for Systems Workshop at NeurIPS 2024

QUTE: Quantifying Uncertainty in TinyML with Early-exit-assisted ensembles for model-monitoring

Nikhil P Ghanathe, Steven Wilton

Preprint Version, Under Review.

Enabling Risk Management of Machine Learning Predictions for FPGA Routability

Andrew David Gunter, Maya Thomas, **Nikhil P Ghanathe**, Steven Wilton

Proceedings of the 2024 ACM/IEEE International Symposium on Machine Learning for CAD

T-RECX: Tiny-Resource Efficient Convolutional neural networks with early-eXit (*BEST PAPER AWARD*)

Nikhil P Ghanathe, Steve Wilton

Proceedings of the 20th ACM International Conference on Computing Frontiers (CF '23), , pp. 123–133, 2023

MAFIA: Machine Learning Acceleration on FPGAs for IoT Applications

Nikhil P Ghanathe, Vivek Seshadri, Rahul Sharma, Steve Wilton, Aayan Kumar

In 2021 31st International Conference on Field-Programmable Logic and Applications (FPL) (pp. 347-354). IEEE.

Compiling KB-sized machine learning models to tiny IoT devices

Sridhar Gopinath, **Nikhil P Ghanathe**, Vivek Seshadri, Rahul Sharma

In 40th ACM SIGPLAN Conference on Programming Language Design and Implementation (PLDI 2019). ACM, New York, NY, USA, 79-95

Software and firmware co-development using high-level synthesis

Nikhil P Ghanathe, Alexander Madorsky, Herman Lam, Darin Acosta *et al.*

Presented at the Topical Workshop on Electronics for Particle Physics (TWEPP2016)

Published in the Journal of Instrumentation, Volume 12, January 2017

Presentations and Non-peer reviewed Contributions

Improving battery-life in ultra-low-power devices using early-exit networks — *Vancouver, canada*

Talk | UBC ECE Research Day — *February2024*

Accelerating edge machine learning algorithms on low-cost FPGAs — *Vancouver, canada*

Talk | NSERC COHESA Annual general meeting — *October2022*

Compiling Machine learning to low-cost Hardware — *Redmond, Washington-USA*

Industry Presentation | Microsoft TechFest — *February2019*

Endcap Muon Trigger firmware and software co-development using Vivado HLS — *CERN, Geneva-Switzerland*

Invited talk | European Organization for Nuclear Research (CERN) — *April2016*

High-level synthesis for CMS Endcap Muon Trackfinder — *Batavia, Illinois-USA*

Invited talk | Fermi National Accelerator Laboratory — *March2016*

Systolic Array Based SGEMM on Arria10 — *Austin, Texas-USA*

Presentation | SuperComputing Conference — *November2015*

Leadership and Volunteer

- ***UBC Engineering Mentor*** *(mentored two students)* — *Oct2023- April2024*
- ***UBC ECE Graduate Student Lab Ambassador*** — *Sept2022- Aug2024*
- ***UBC Graduate Student Well-being Ambassador*** — *Sept2022- Aug2023*
- ***Collaborator*** *, Workshop on Machine Learning on Constrained Hardware by* ***Microsoft Research****, India* — *Aug2018-Oct2018*
- ***Collaborator*** *, Center for High-performance and Reconfigurable Computing Annual Workshop at* ***Kennedy Space Center, NASA*** — *Dec2015*

Achievements and Extracurriculars

- ***Distinguished Lightning Talk Award*** *at UBC ECE Research Day*
- ***ACM Best Paper Award*** *at International Conference on Computing Frontiers (CF'23)*
- ***College of Engineering Achievement Award Scholarship****, University of Florida*
- ***Academic Excellence Award*** *by the* ***Indian Space Research Organization***