GOKULKRISHNA MUTHUSAMY

[Address] | [Phone Number] | gokulkrishna98@gmail.com

Education

New York University, New York, NY — May 2025
Master of Science in Computer Science

National Institute of Technology (NIT), Tiruchirappalli, TN — Jun 2020
Bachelor of Technology in Computer Science and Technology — 8.6/10
GRE – 326, TOEFL – 109

Experience

Senior Engineer, *Samsung Research India*, Bangalore, KA — Apr 2022 – Aug 2023
CPU, GPU and Compiler team for Neural Acceleration.

- Programmed kernels for complex neural layers like group normalization and modified MLIR to support TfLite model conversion to enable complete execution of Stable Diffusion model on Mobile GPUs, resulting in 53.3% improvement in performance over CPU.
- Designed an automated recompilation & caching tool for OpenCL kernels on GPUs to reduce caching overhead and prevent crash during GPU driver updates, which resulted in _ % reduction in crash report during testing and _% improvement in performance.
- Accelerated __ use cases on Samsung flagships (S23 devices) to improve user experience, achieved _ % improvement in load time and _ % in execution time.

ML Software Engineer, *Samsung Research India,* Bangalore, KA — Jan-2021 – Apr 2022
CPU, GPU and Compiler team for Neural Acceleration.

- Engineered Profiling tool for layer wise performance data in ArmNN during ML model execution to get better diagnosis of performance degradation. This improved overall quality of life of team and use case engineers for optimizing the models and kernels.
- Collaborated with Galaxy RAW team to solve greenish tinge issues on RAW images by implementing new normalization methods to support float16 quantized precision execution. This improved performance by 34% over previous method.
- Directly helped in enabling of over _ USP camera and gallery features deployed on Samsung flagship S22 and S23 devices. Contributing to over 15x optimizations for speed, memory and battery.

Intern, ***Samsung Research India,*** Bangalore, KA — May 2019 – July 2019
Keyboard Intelligence for OnDevice AI

- Evaluated the improvement in emoji prediction by incorporating keystroke statistics in NLP sentiment analysis algorithm, resulting in 10% improvement in accuracy on personalized dataset.

Patents & Publications

A1 graded patent by Samsung Research HQ — Oct 11, 2022
Main inventor of Systems and Methods for On-Device Validation of a Neural Network Model, M Gokulkrishna, Siva Kailash, R. Prasanna, Rajath Elias, Ashok Kumar, Praveen Doreswamy. Pending Patent:

Contributions & Activities

- Active contributor to TensorFlow open-source community.
- Lead the design team for Vortex, Computer Science department in NIT Tiruchirappalli.
- Volunteered for awareness activities on sustainable farming practices for National Service Scheme (India).
- Received Full Stack Web Developer Nanodegree from Udacity [credentials]

Technical Skills

Programming: Proficient in C++, Python, Shell, Rust, Android NDK, Git, Docker, Flatbuffer
Deep Learning: TensorFlow, Keras, ArmNN, Pytorch, ONNX, Apache TVM, CNN, Transformers, Generative AI. Quantization.
System & Performance: OpenCL, MLIR, CUDA, Perfetto, Gdb, Valgrind, Asan, Hwasan, TensorFlow Lite (GPU Delegate).