

EDUCATION

- Indian Institute of Information Technology, Design and Manufacturing, Kancheepuram, India** 2018 — 2022
- Bachelor of Technology in Computer Science and Engineering (Distinction)
 - CGPA: 9.17/10.00
 - Relevant Coursework:** High Performance Computing, Advanced Data Structures and Algorithms, Pattern Recognition, Computer Architecture, Computer Organization, Operating Systems, Linear Algebra, Probability Theory, Differential Equations
- SBOA School and Junior College, Chennai, India**
- Class XII - 475/500 (95.0%) 2018
 - Class X - 10.0/10.0 CGPA 2016

WORK EXPERIENCE

- NVIDIA** July 2022 — Present
Systems Software Engineer Pune, India
- Working as a DevTools Engineer on the CUPTI API in the Cuda Profiling Tools Development Team.
 - Analyzed the performance overheads of tracing and profiling in AI/HPC Workloads using CUPTI Injections.
 - Contributed to enabling CUPTI support in the bringup phase of upcoming GPUs.
- KLA-Tencor** Dec 2021 — Mar 2022
Software Intern Chennai, India
- Migrated serial algorithms used in semiconductor defect-detection to GPUs and parallelized them using CUDA.
 - Profiled CUDA code using NSight Compute and NSight systems to perform co-optimization.

RESEARCH EXPERIENCE

- Indian Institute of Technology, Roorkee (IIT-R) - Bachelor's Thesis** Nov 2021 - June 2022
Research Intern (Guide : Dr. Debiprasanna Sahoo) Roorkee, India
- Studied the design and micro-architecture of GPUs - SIMT Cores, Warp Schedulers and SIMT pipelines, with GPGPU-Sim's model as reference.
 - Formalized the Warp Issue Scheduler and verified the safety and liveness properties of the formal model using SAL.
- Indian Institute of Science (IISc), Bangalore - Report, Github** May 2021- Oct 2021
Research Intern (Guide : Dr. R. Govindarajan) Bangalore, India
- Recipient of the Indian Academy of Sciences' Summer Research Fellowship SRF '21.
 - Constructed pipelined CNNs with multiple GPUs for parallel training using Tensorflow Lingvo and GPipe.
 - Analyzed the performance and memory tradeoffs between model-parallel, pipeline-parallel and hybrid-parallel training in CNNs across multiple GPUs.
- Indian Institute of Technology, Madras (IIT-M) - Github Links : MaxFlow, SSSP** Mar 2020 — Oct 2020
Research Intern (Guide : Dr. Rupesh Nasre) Chennai, India
- Implemented parallel algorithms to compute maximum network flow on GPUs using CUDA.
 - Experimented with fundamental graph problems like parallel BFS, parallel Bellman-Ford SSSP on GPUs with CUDA.
- HPRCSE Labs, IIITDM Kancheepuram** Dec 2019 — Jan 2020
Intern (Guide : Dr. Noor Mahammad Sk) Chennai, India
- Conducted literature surveys on parallel computing taxonomies and gave a talk in the Computer Science Club's High Performance Computing Track (Slides).
 - Gained understanding of parallel programming in OpenMP and MPI and explored profiling tools like Valgrind and Gprof.

SKILLS

- Programming Languages** C, C++, Python
- Frameworks/Libraries** CUDA, Tensorflow, PyTorch, OpenMP, OpenMPI
- Tools** NSight Compute, NSight Systems, Bash, Git
- Other** MySQL, AWS, GPU Architecture
- Interests** High Performance Computing, Scientific Computing, Artificial Intelligence, Computer Graphics

ACHIEVEMENTS

- Recipient of the Indian Academy of Sciences' Summer Research Fellowship SRFP '21. May 2021 — Oct 2021
- ICPC 2020 Regionalist (Gwalior-Pune) - rank 222, Regionalist (Amritapuri) - rank 342 Oct 2021
- Winner, Special Mention (Recognition Team Award) at Startup Weekend 2k19. Feb 2019

PROJECTS

CUDA Maxflow Solver

- Implementation of parallel maximum-flow in CUDA using the parallel push-relabel algorithm.
- Asynchronous push-relabel works on static flow networks with non-negative edge capacities.

CUDA SSSP Solver

- Implementation of Single Source Shortest Path in CUDA using the parallel Bellman-Ford Algorithm.
- Edge-centric BFS traversal is used.
- Runs on 0.5ms on [bitcoin-otc \(SNAP dataset\)](#) on a MX150 2GB GPU.

POSIX PathTracer

- A primitive multi-threaded path tracer built in C++ and PThreads, based on [smallpt](#).
- Can render 200 spp in less than 25 minutes.

Pipeline Accelerated CNNs

- A modified fork of [Tensorflow Lingvo](#), with added definitions of AlexNet and VGG16, pipelined with [GPipe](#).
- [Experiments conducted](#) to analyze performance-memory tradeoffs across pipeline/model/hybrid parallel training, on multiple GPUs on AWS.

Garden Buddy - Plant Species Identifier Web App

- A Machine Learning based Gardener Assistance App that identifies the plant species from a picture of the leaves.
- Performed image augmentation, ensembling and stratified k-fold validation to achieved 96.8% classification accuracy.
- Trained using EfficientNet-B5s on the LeafSnap dataset. Deployed using PyTorch and Streamlit.

ACTIVITIES

- Lead Core (2020) / Joint Core (2019), EdITH (Education in IT & Hardware), Computer Science Club** 2019 — 2021
 - Led a team of 20+ juniors to organise numerous CS-related events, workshops and intra-campus competitions.
 - Organized bi-weekly sessions on various arenas in computer science.
- High Performance Computing (HPC) Track Lead, Computer Science Club** 2020 — 2021
 - Presented a [Talk](#) on Parallel Computing Taxonomies.
 - Organized workshops and sessions on HPC and Parallel Programming.
- Organizer - Vashisht 2020 (Inter-Collegiate Technical Fest)** 2020 — 2021
 - Conducted and organized coding competitions and CS-related talks.
 - Coordinated a team to raise funds via alumni network.
- Core member, Institute Innovation Council (IIC), MHRD's Innovation Cell (MIC)** 2018 — 2020
 - Organized design ideathons, and summer industry open houses (EHIPASSIKO).
- Coordinator, Music Club, IIITDM** 2019 — 2020
 - Played the drums in the institute band as part of several shows and cultural.
 - Conducted music events in SAMGATHA (Inter-collegiate Cultural Fest).
- Coordinator, Tamil Saalaram IIITDM** 2019 — 2020
 - Organized Tamil Language-related competitions in SAMGATHA (Inter-collegiate Cultural Fest).