

SKILLS

- **Programming:** C++, CUDA, Python, Java, Solidity
- **Machine Learning:** PyTorch, MLflow, LangChain, LangGraph, FastMCP, Scikit-learn, OpenCV
- **Web & Systems:** React.js, Next.js, Flask, Tailwind CSS, Prisma, MongoDB, MySQL, IPFS

PUBLICATIONS

- **Parikshit Gehlaut** and Nikhil Hegde, "*Estimating Optimal GPU Occupancy Levels for Performant Matrix Multiplication Codes*," accepted in **SRS, IEEE-HiPC 2025**, IEEE Xplore (to appear).

WORK EXPERIENCE

Yethi Consulting | AI Engineer Intern | Bengaluru | Jan 2026 - Present

- Evaluated Kairo Copilot (Yethi's in-house AI assistant) using an MLflow agent-as-a-judge framework, Extended the Tenjin MCP server by implementing automation tools

DoctorSpectra | Machine Learning Intern | Remote | February 2025 - March 2025

- Trained a U-Net image segmentation model (PyTorch, OpenCV) on the Carvana dataset achieving a 98% Dice score, and optimized the inference latency-accuracy trade-off for real-time production deployment; expanded the labeled dataset using CVAT.

EDUCATION

- **IIT DHARWAD** | Bachelor of Technology in Computer Science and Engineering | CGPA **8.00/10.0**

PROJECTS

Performance Modelling of CUDA Programs | R&D Project, IIT Dharwad | August 2025 - Dec 2025

- Developed three microkernels to empirically measure arithmetic latency, memory latency, and memory throughput on GPU hardware.
- Evaluated Volkov's analytical model across modern NVIDIA GPU architectures using synthetic kernels and matrix-multiplication operations that utilize these three measured parameters.

DeepFake Detection DApp ([Source Code](#)) | React.js, Flask, Solidity | Feb 2025 - April 2025

- Built a decentralized media verification system combining deepfake detection with blockchain-based metadata storage.
- Developed a React and MetaMask frontend communicating with Flask APIs running TensorFlow models, utilizing Solidity smart contracts to immutably record media hashes and IPFS links on-chain.