Gopi Krishna Tummala

 Fremont, CA —   +1-614-620-7960 — tummalag.cseosu@gmail.com

 linkedin.com/in/gopi-krishna-tummala — gopikrishnatummala.com — Google Scholar

Visa Status: Green Card

SUMMARY

Machine Learning Engineer with **12+ years of experience** designing and deploying scalable ML systems for **Generative AI** and **Autonomous Driving** (L3-L5). Expert in developing **AV prediction models** and building production training infrastructure, custom GPU kernels, and multi-modal data pipelines. Proven track record of bridging research and production, with a Ph.D. in Computer Science, 10+ patents, and award-winning publications. Currently architecting data infrastructure for Adobe Firefly.

TECHNICAL SKILLS

- **Languages:** Python (Expert), C++, CUDA, SQL, Bash, TypeScript, Java, HTML/CSS.
- **ML & GenAI:** PyTorch, TensorFlow, JAX, Transformers, Diffusion Models (DiT/VAE), Vision-Language Models (VLMs), RAG, LangChain, FlashAttention, vLLM.
- **Infrastructure:** Distributed Training (FSDP, DeepSpeed), AWS (S3, DynamoDB, EC2), Docker, Kubernetes, Ray, Apache Arrow/Parquet.
- **Domain Expertise:** Autonomous Vehicles (Prediction/Planning), Computer Vision, Sensor Fusion, Camera Calibration, MLOps.

PROFESSIONAL EXPERIENCE

Adobe — San Jose, CA
Senior Machine Learning Engineer — Jan 2024 – Present

- **GenAI Data Marketplace:** Architecting a high-performance Data Marketplace (internal Hugging Face) to support **Adobe Firefly**, serving distinct data needs across research and product teams.
- **Production Dataloader:** Developed/Maintained a high-throughput dataloader for production training, handling >**1M calls**/**month** with optimized data streaming and caching.
- **Scalable Storage:** Managed a global training feature store on a **DynamoDB cluster (20B items)**, ensuring high availability and low-latency access for distributed training jobs.
- **MLOps & Profiling:** Engineered runtime profiling suites and CI/CD automation for high-throughput dataloaders and inference systems, optimizing resource utilization and accelerating deployment velocity.

Zoox — Foster City, CA
Software Engineer - Prediction — Feb 2022 – Jan 2024

- **Trajectory Prediction:** Designed and shipped multi-agent prediction models for L3-L5 autonomous fleets using **Transformers and Graph Neural Networks (GNNs)**.
- **Training Frameworks:** Co-designed a production training framework integrating **closed-loop reasoning**. This directly improved vehicle safety and passenger comfort metrics in dense urban environments.
- **Infrastructure Optimization:** Developed specialized dataloaders to handle high-fidelity sensor streams, reducing model training epoch time by ∼30%.
- **Legacy Migration:** Led the migration of legacy TensorFlow training pipelines to a modern **PyTorch** stack.

Qualcomm Research — San Diego, CA
Senior Systems Engineer — Feb 2020 – Feb 2022

- **Snapdragon Auto Stack:** Developed intent-based ML models for the Snapdragon Autonomous Driving Platform, focusing on discrete intention prediction and trajectory generation.
- **KPI Analysis:** Led the "Prediction KPI" initiative, building an automated analysis pipeline to mine petabytes of driving logs for edge-case failures and verify model performance against safety standards.

Qualcomm Research   San Diego, CA
Senior System Integration & Test Engineer   Aug 2018 – Feb 2020
  - Translated research-grade ML algorithms into efficient, modular C++ implementations for embedded deployment on Snapdragon hardware.
  - Developed comprehensive testing and automation pipelines for the behavior prediction module.

Standard Chartered Scope International   Chennai, India
Software Analyst   Jun 2012 – Jul 2013
  - Developed financial software systems and data analysis tools for banking operations.

RESEARCH INTERNSHIPS

Qualcomm Research   San Diego, CA
Interim Engineering Intern   Summer 2018
  - Optimized modular software implementations for the AV stack, bridging the gap between Python research models and C++ production code.

Microsoft Research   Bangalore, India
Research Intern   Summer 2016
  - **Project AutoCalib:** Designed software for automatic traffic camera calibration. Achieved speed estimation errors of $< 10\%$ without manual intervention.
  - Published results in *ACM BuildSys 2017* (Won Best Paper Award).

Tata Elxsi   Chennai, India
Project Intern   Summer 2011

EDUCATION

The Ohio State University   Columbus, OH
Ph.D. in Computer Science & Engineering   2013 – 2018
M.S. in Computer Science & Engineering   2013 – 2017

IIT Madras   Chennai, India
B.Tech in Electrical Engineering (Minor in Mathematics)   2008 – 2012

SELECTED PUBLICATIONS

  - **AutoCalib: Automatic calibration of traffic cameras at scale.**
    ACM Transactions on Sensor Networking (TOSN) 2018 & *ACM BuildSys 2017*.
    Awards: Best Paper Award & Best Demo Award.
  - **SmartDashCam: Automatic Live Calibration for DashCams.**
    ACM IPSN 2019.
  - **RoadView: Live View of On-Road Vehicular Information.**
    IEEE SECON 2017.
  - **CaneScanner: Obstacle Detection for People with Visual Disabilities.**
    IEEE MiSeNet 2018.

PATENTS

10+ patents granted/pending in AV prediction, behavior trees, and calibration.
  - **US Patent 10,580,164:** Automatic Camera Calibration (Microsoft).
  - **US Patent 10,032,370:** Methods for enabling Mobile communication device based Secure Interaction (Honda).
  - **US Patent App 17/455,853:** Managing Vehicle Behavior Based On Predicted Behavior of other vehicles (Qualcomm).
  - **US Patent App 17/352,886:** Tree based behavior predictor (Qualcomm).

HONORS & SERVICE

- **Awards:** Best Paper (ACM BuildSys 2017), Best Demo (IEEE MiSeNet 2018), NSF Travel Awards (2017, 2018).
- **Scholastic:** All India Rank 274 (Top 0.1%) in IIT-JEE; Top 1% in National Physics/Math Olympiads.
- **Service:** Associate Editor for *IEEE RA-L*; Reviewer for *ACM TECS*, *IEEE TMC*, *ToSN*.
- **Technical Blog:** Author of gopikrishnatummala.com, publishing in-depth technical deep dives on generative AI and model scaling, advanced MLOps and production-grade infrastructure, autonomous systems engineering, and agentic AI design and deployment.