

# Gopi Krishna Tummala

📍 Fremont, CA — 📞 +1-614-620-7960 — [tummalag.cseosu@gmail.com](mailto:tummalag.cseosu@gmail.com)

🔗 [linkedin.com/in/gopi-krishna-tummala](https://linkedin.com/in/gopi-krishna-tummala) — 🌐 [gopikrishnatummala.com](http://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 – Aug 2024
- **Vehicular Prediction Models:** Re-engineered ground-truth labeling for intent map points using a stable algorithm that jointly considers agent direction and road target orientation (TNT-inspired target-driven approach), replacing a flickery closest-distance heuristic. Delivered **2–3 production model releases** with significant improvements in both intent/target prediction accuracy and downstream trajectory quality.
  - **VRU Model Overhaul:** Migrated pedestrian/cyclist/motorcycle prediction from legacy CNN + top-down rendered images + MLP architecture to **GNN-based multi-intent/trajectory** architecture. Solved hard problems including generating intent goal points for off-road agents and ambiguous scenarios. Achieved model parity and outperformed mature baseline; owned end-to-end deployment (PyTorch → ONNX → C++ runtime plugins).
  - **PyTorch Training Framework:** Built config-driven module registry from scratch (YAML/Proto) to instantiate MLPs, CNNs, RNNs/LSTMs, GNNs, Transformers, and top-down UNet backbones. Wrapped on **PyTorch Lightning** with wandb/CometML integration. Achieved near-parity between TensorFlow and PyTorch for vehicular + VRU models, enabling framework migration.
  - **Data Infrastructure & Evaluation:** Engineered PyTorch-compatible dataloaders from scratch, converting legacy TF Records to NPZ format. Scaled evaluation infrastructure to run **~40,000 autonomous driving scenarios**

per batch. Authored novel intent & trajectory prediction metrics, aggregated results into Streamlit dashboard for cross-team visibility.

- **Prediction Visualization Tools:** Built prediction visualization plugin on Argus (internal ROS/RViz-like framework) using TypeScript/React, Python, and C++ to debug probabilistic multi-trajectory outputs across perception → prediction → planning pipeline.

#### Qualcomm Research

San Diego, CA  
Feb 2020 – Feb 2022

##### Senior Systems Engineer

- **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  
Aug 2018 – Feb 2020

##### Senior System Integration & Test Engineer

- 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  
Jun 2012 – Jul 2013

##### Software Analyst

- Developed financial software systems and data analysis tools for banking operations.

## RESEARCH INTERNSHIPS

#### Qualcomm Research

San Diego, CA  
Summer 2018

##### Interim Engineering Intern

- Optimized modular software implementations for the AV stack, bridging the gap between Python research models and C++ production code.

#### Microsoft Research

Bangalore, India  
Summer 2016

##### Research Intern

- **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  
Summer 2011

##### Project Intern

## EDUCATION

#### The Ohio State University

Columbus, OH

2013 – 2018

##### Ph.D. in Computer Science & Engineering

2013 – 2017

##### M.S. in Computer Science & Engineering

Chennai, India

#### IIT Madras

2008 – 2012

##### B.Tech in Electrical Engineering (Minor in Mathematics)

## 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](http://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.