

Krishna Shravan

India | Krishna18698@gmail.com | <https://www.krishna-portfolio.in/>
<https://linkedin.com/in/krishna-shravan/> | <https://github.com/Krishna18698>

Professional Summary

Software Engineer with 3.5 years of experience building production backend applications and AI-powered software systems. Experienced in Python, C#/.NET, FastAPI, PostgreSQL, Azure, and cloud-native development, with expertise in computer vision, RAG, and edge AI deployment. Improved inference throughput from 20 to 30–35 FPS, increased model accuracy from 50–60% to 80–85%, and delivered \$120K+/year in direct labor savings through production software solutions.

Experience

Software Engineer Wipfli India LLP – Bengaluru, India Aug 2022 – Feb 2026

- Independently designed and developed a C#/.NET healthcare reporting add-in covering 60+ hospitals and clinics across 3 facility types, reducing multi-hour reporting workflows to 1–2 minutes and delivering \$120K+/year in direct labor savings.
- Developed a C#/.NET Excel AI Visualizer that converted natural-language input into charts and visualizations, reducing manual charting and analysis effort by 80% across 50+ internal stakeholders.
- Increased real-time YOLO inference throughput from 20 FPS to 30–35 FPS on OAK-D edge cameras by converting models to ONNX/OpenVINO and applying quantization and batched inference for Intel Myriad X VPUs.
- Reduced Real-ESRGAN super-resolution processing time for airport security X-ray images from 20–25 seconds to 7–9 seconds by deploying GPU-backed inference on GCP and optimizing the serving pipeline with TensorRT.
- Improved factory computer-vision accuracy from 50–60% to 80–85% for zone occupancy and from under 50% to 70–80% for PPE detection by collecting and hand-labeling custom CCTV datasets and training YOLOv5/v8 models.
- Led technical knowledge-sharing sessions on Azure AI, VSTO add-ins, and AI-assisted development for a 15–20 member innovation group spanning engineers, senior staff, managers, and partners.

Projects

- Codeward – AI-Powered Interview Prep Platform [Live] [GitHub]**: Architected and developed a full-stack interview preparation platform with Next.js, REST APIs, NextAuth, PostgreSQL, Prisma, Redis-backed rate limiting, and sandboxed multi-language code execution. Integrated RAG-powered AI mentoring using pgvector, Voyage AI embeddings, and Groq across DSA, system design, code review, debugging, and staged LLD workflows.
- FolliScan – On-Device ML Hair-Loss Classifier [GitHub]**: Engineered a React Native + TensorFlow Lite application for fully offline Norwood 1–7 classification using quantized MobileNet models, scalp validation, automated image-quality checks, GPU delegates, and CPU fallback.
- Lifestyle Tracker – Adaptive Android Fitness Platform**: Built a Kotlin/Jetpack Compose app with Health Connect, barcode/OCR nutrition capture, adaptive TDEE and trend analysis, Room/WorkManager offline architecture, and 750+ automated tests.
- Pit Wall – Live Formula 1 Data Dashboard [Live] [GitHub]**: Architected and developed a server-rendered Next.js dashboard with server-side API routes, cached data pipelines, real-time telemetry integration, live driver tracking, automated standings, RSS aggregation, and responsive visualizations for Formula 1 race weekends.

Skills

- Languages**: Python, C#/.NET, TypeScript, JavaScript, SQL
- Backend, Web & Cloud**: FastAPI, Flask, REST APIs, Next.js, React, React Native, PostgreSQL, Prisma, Neon, Supabase, Azure AI Services, Azure Functions, GCP GPU VMs, Upstash Redis
- GenAI & LLM**: RAG, vector search, pgvector, Voyage AI embeddings, Gemini, Groq
- ML Infrastructure & Optimization**: ONNX, ONNX Runtime, OpenVINO, TensorRT, Docker, quantization, batched inference, GPU delegates, edge deployment
- ML & Computer Vision**: TensorFlow, TensorFlow Lite, Keras, PyTorch, YOLOv5/v8, MobileNetV2/V3, Real-ESRGAN, OpenCV, MediaPipe, scikit-learn

Education

Bachelor of Engineering, Computer Science

K.V.G. College of Engineering, Sullia, Karnataka – 2022

Certifications

Microsoft Certified: Azure AI Engineer Associate (AI-102) | Azure AI Fundamentals (AI-900)