

Ronak Chhatbar

☎ +1716-507-2419 | @ ronakchhatbar@gmail.com | 🔗 LinkedIn | 🐙 GitHub | 📁 Portfolio | 📍 Buffalo, New York

PROFESSIONAL SUMMARY

AI Engineer with **5+ years** developing production AI systems, specializing in **computer vision and conversational AI integration**. Currently building real-time monitoring platforms that combine video analysis, IoT alerts, and multi-agent orchestration, achieving **33% cost reduction** and **67% faster incident response**. Expert in GPU optimization, video processing pipelines, and deploying scalable AI systems that drive measurable operational efficiency.

EXPERIENCE

- **Centific** Remote
AI Engineer *Dec 2024 - Present*
 - Reduced facility monitoring costs by **33%** by building automated computer vision system using **NVIDIA VSS and Deepstream**, processing **50+ camera streams** and eliminating need for round-the-clock human surveillance
 - Decreased incident response time from **15 minutes to 5 minutes** by implementing real-time alert system with unified dashboard, achieving sub-second detection across facility monitoring
 - Automated **80% of facility workflows** by deploying **conversational AI system with LangGraph orchestration**, enabling voice/text interactions for incident reporting, maintenance scheduling, and compliance reports
 - Cut infrastructure costs by **20%** by optimizing AI model serving with **TensorRT-LLM and distributed deployment**, maintaining real-time performance while reducing compute requirements
- **Tensorgo Technologies** Hyderabad, India
Computer Vision Engineer *Sep 2020 - Aug 2022*
 - Enhanced deep learning models (eye-gaze, emotion) performance by **40% through TensorRT and Deepstream optimization**, increasing system throughput by **25%**.
 - Developed computer vision-based heart rate estimation system, improving prediction accuracy by **8%** using datasets of over **20,000+ images** across diverse demographics.
 - Refined heart rate estimation system addressing demographic diversity using **BP4D+, UBFC-1, and UBFC-2 datasets**, achieving enhanced accuracy for inclusive biometric applications across varied skin tones and lighting conditions.
 - Led development of compliance software with ASR integration, enhancing meeting analytics accuracy by **16%** through real-time speaker segmentation.
 - Automated pipeline processes for deep learning applications, reducing manual intervention by **60%** and increasing deployment efficiency by **30%**.
 - Deployed AI models on **Jetson-NX and Nano**, achieving **30-40 FPS** real-time processing in field applications through advanced memory optimization using **TensorRT and mixed precision**.
- **Wavelabs Technologies** Hyderabad, India
Machine Learning Engineer *May 2019 - Aug 2020*
 - Implemented weapon detection system on Jetson-Nano, achieving **sub-2-second threat identification** with **30-40 FPS** processing, integrated with mobile apps (iOS/Android).
 - Handled image labeling and bounding box annotation for weapon detection dataset, creating **1,50,000+ manually labeled training images** with data augmentation techniques to generate adequate dataset for model training.
 - Conducted real-time sentiment analysis on **5,000+ customer interactions monthly**, resulting in **6%** improvement in customer service quality using ULMFit language modeling.
 - Developed dynamic pricing algorithms for financial services, impacting **10,000+ transactions** and contributing to **15% revenue increase** through optimized pricing strategies.
 - Streamlined AI model deployment using **Docker** and AWS SageMaker, reducing resource utilization by **35%** and deployment time by **20%**.
- **Wavelabs Technologies** Hyderabad, India
Computer Vision Research Intern *Nov 2018 - Apr 2019*
 - Created facial recognition system with **95% accuracy** for employee identification, incorporating age and gender analysis using ResNet50 and HOG features.
 - Led image classification optimization project, **improving model accuracy by 16%** through systematic testing of architectures and learning parameters in TensorFlow.

- **Spatial AI & Robotics Lab**

- *Graduate Research Assistant*

University at Buffalo

Dr. Chen Wang — May 2023 - Dec 2024

- Developed C++ plugins for visual odometry optimization, **successfully integrating with TensorRT backend** and achieving **33% increase in inference efficiency**.
 - Enhanced optical flow estimation using specialized algorithms, marking substantial advancement in accuracy and speed of real-time data processing.
 - Spearheaded backend development of robotranking.org platform for robotics research assessment, contributing valuable resource to international research community.

- **GPU Computing Research Portfolio**

- *Advanced CUDA Development*

- **Hessian Matrix Inversion:** Achieved **526x GPU speedup** implementing Hessian matrix inversion via LU decomposition with cuSOLVER integration and Python bindings.
 - **Advanced Performance Profiler:** Developed comprehensive CUDA profiling framework using Nsight Compute integration, automated CSV-based metrics collection, and interactive Streamlit visualization dashboard for systematic kernel optimization.
 - **Neural Network Framework:** Implemented complete GPU-accelerated library with custom kernels for dense layers, softmax, and attention mechanisms, featuring end-to-end training pipelines and ONNX export.
 - **Convolution Analysis:** Published comprehensive performance analysis with **18 optimization metrics**, mathematical foundations, and advanced profiling using Nsight Compute with interactive **Streamlit** visualization.
 - **Heat Equation Solver:** Created distributed numerical computing solution using CUDA and OpenMP with NVTX profiling, featuring multi-node optimization and cross-platform deployment.

EDUCATION

- **University at Buffalo, The State University of New York**

Buffalo, NY

- *Masters in Computer Science; GPA: 3.4/4.0*

Aug 2022 - Jan 2024

- *Courses: Operating Systems, Analysis Of Algorithms, Biometrics Image Analysis, Reinforcement Learning, Computer Vision.*

- **Jawaharlal Nehru Technological University Hyderabad**

Hyderabad, India

- *Bachelor of Computer Science; GPA: 3.6/4.0*

Aug 2015 - May 2019

- *Courses: Machine Learning, Cloud Computing, DSA, Computer Networks, Probability, Statistics, Mathematics, Compiler Design.*

SKILLS SUMMARY

- **Agentic AI & Advanced Systems:** LangChain/LangGraph, Model Context Protocol (MCP), Multi-Agent Orchestration, Agent State Management, PostgreSQL, Vector Embeddings
- **GPU Computing & Performance Optimization:** CUDA Programming, TensorRT, NVIDIA VSS/NIMs, GPU Memory Optimization, Performance Profiling (Nsight Compute), OpenMP
- **Computer Vision & Machine Learning:** PyTorch, TensorFlow, Computer Vision Pipelines, ASR Integration, Real-time Processing, ONNX, Mixed Precision
- **Production Systems & Development:** C/C++, Python, Docker, Kubernetes, AWS, Jetson (NX/Nano), Deepstream, Microservices, Linux