ABUBAKAR SIDDIQ

240-4813766 | abubakarsiddiqpalli99@gmail.com | LinkedIn | GitHub | Website

SUMMARY

AI/ML Engineer specializing in Agentic AI, Deep Learning, and Computer Vision. Experienced in building and deploying Agentic AI systems, transformer-based architectures, multimodal, and fine-tuned LLM systems that drive real-world impact. Skilled in full-stack MLOps (PyTorch, TensorFlow, AWS) and integrating scalable AI solutions from model design to production deployment. Passionate about creating intelligent, autonomous systems that bridge perception and reasoning.

EDUCATION

University of Maryland | College Park, MD Aug 2023 – May 2025

Master of Engineering (M.Eng.) – Robotics | *Multi-Modal Models, Deep Learning, Perception, Autonomous vehicles, Localization, Path Planning*

SKILLS

Deep Learning & Machine Learning – Transformers, CNNs, Object Detection, LLMs, NLP, Multimodal Learning, Generative Models, MCP, ADK, A2A

Cloud and Deployment (AWS CERTIFIED) - AWS EC2, S3, Lambda, SageMaker, Cloud formation, CloudWatch, ECR, ECS, EB, Docker

Programming Languages – Python, C++, SQL, MATLAB

Frameworks & Libraries – PyTorch, TensorFlow, OpenCV, PEFT, LangChain, LangGraph, Open3D, Hugging Face, TIMM, Open AI SDK, RAG

Optimization & Tools – CUDA, SciPy, NumPy, Pandas, TensorRT, ONNX, Statistical Methods, Data Visualization, Exploratory Data Analysis.

EXPERIENCE

AI Engineer | Easybee AI August 2024 - Present

- Designed and deployed multi-agent system using LangGraph and CrewAI, assigning specialized agent roles to automate complex, interconnected workflows and boost operational efficiency. Implemented scalable deployment on AWS ECS for automatic load management.
- Engineered a custom web crawler and data ingestion pipeline to create client-specific Knowledge Bases for Retrieval-Augmented Generation (RAG), incorporating advanced chunking, data enrichment, and indexing techniques using Pinecone vector DB for optimized Agentic AI access.
- Integrated WhatsApp and Twilio messaging channels using Model Context Protocol (MCP) to enable seamless and continuous communication between customers and the AI system across different platforms.

Deep Learning Engineer | UMD PAL March 2024 – May 2025

- Developed a custom vision model based on Masked Autoencoders (MAE) and trained it on RGB orthomosaics from drones to estimate Leaf Area Index (LAI), achieved a novel R2 score of 0.87 using only RGB data. Reduced reliance on multispectral hardware and simplified data acquisition.
- Designed a ViT based multispectral pipeline that takes Digital Surface and Terrain Models (DSM & DTM) as inputs and trained for plant height estimation. The model maps features to ground truth values, achieving an RMSE of 0.26m, outperforming CNN baselines.
- Automated large-scale aerial data preprocessing and normalization pipelines, integrating OpenCV, rasterio, and PyTorch for scalable experimentation.

Deep Learning Engineering Intern | Vyorius May 2021 – Nov 2021

- Built YoLO-based video analytics modules for UAVs, enabling real-time object detection, tracking, and classification from aerial streams and enhanced autonomous surveillance coverage by 60% across remote deployments.
- Used SOTA segmentation models and integrated them into Vyorius's DataSync Intelligence pipeline and enhanced vision-based landing, enabling 40% faster detection of anomalies and ROI's and contributed to 25% higher precision in GPS-denied landings.
- Partnered with engineering to integrate and deploy these computer vision modules onto the Vyorius cloud platform, directly enabling scalable, AI-driven analytics for multi-robot commercial missions.

PROJECTS

AV Driving descriptor using fine-tuned LLaMA and Multimodal context | Link Feb-March 2025

- Fine-tuned LLaMA to generate human-readable explanations of autonomous vehicle behavior by integrating BLIP-2 scene descriptions, sensor data, and road rules via RAG.
- Developed a multimodal preprocessing pipeline that fuses image captions, sensor metadata, and retrieved traffic regulations to create rule-aware, grounded outputs.
- Enabled interpretable driving action summaries for AV systems, showcasing multimodal fusion, LLM adaptation, and retrieval-augmented reasoning.

Personalized Audio Learning Generator | MIT Global AI Hackathon 2025 | Link May 2025

- Built a RAG app that turns natural language prompts into personalized 5-minute audio lessons using Gemini with google genkit, Tavily web search, and ElevenLabs TTS.
- Designed a real-time pipeline with session-aware topic tracking, adaptive content expansion, and robust error handling; deployed seamlessly on AWS. Delivered dynamic audio playlists with context-enriched scripts, ranking Top 10 in the hackathon.

Multimodal Alignment model for LiDAR and Image Data | Link Oct-Dec 2024

- Engineered a novel multimodal model for LiDAR-RGB alignment, enabling fusion of 3D spatial data and 2D visual cues.
- Developed a custom Query-Based Transformer (Q-Former) model, leveraging ViT and PointNet for cross-modal feature extraction.
- Improved cross-modal feature matching by 94%, which directly enhanced the model's 3D object detection accuracy (0.61 mAP) for more reliable obstacle identification in autonomous driving simulations.

Transformers for 3D Object Detection in LiDAR Point Clouds | Link Sept-Nov 2024

- Designed a custom transformer-based 3D object detection model for LiDAR point clouds, trained using KITTI dataset.
- Integrated PointNet++ embeddings with custom transformer encoders and designed a novel loss function, enhancing 3D detection and achieved a mAP IoU of 0.67 for 3D bounding box detection on par with existing benchmarks.