Osama Altaf

AI Researcher & Engineer | Vision-Language Models, Medical Imaging, Computer Vision

Pakistan (open to relocation) | (+92) 318 7661096 | osamaaltaf.pk@gmail.com
github.com/osamaaltaf-pk | linkedin.com/in/osamaaltafpk | osama-altaf.vercel.app

RESEARCH INTERESTS

Vision-language models for medical imaging, with a focus on visual grounding: producing answers that can be traced back to the region of the scan they came from, rather than captions for a whole volume. Related interests in segmentation, 3D medical imaging, multimodal representation learning, and the construction and validation of benchmarks that hold up under distribution shift across sites and acquisition conditions.

EDUCATION

BS Physics - Computational Physics & Computer Programming

Khwaja Fareed University of Engineering & IT (KFUEIT), Rahim Yar Khan, Pakistan 2019 - 2023

CGPA: 3.63 / 4.00 - 134 / 134 credit hours - overall 90 percent - EQF Level 6. Final-year semester GPAs of 3.82 and 3.94.

Thesis: Object Detection Transformer for disease detection on cotton leaves. Designed a transformer-based detector coupled with an AI-modelled digital microscope to identify cotton-crop disease from visual symptoms; outperformed CNN baselines on classification accuracy.

Relevant coursework: Introduction to Computational Physics (A), Linear Algebra (A), Probability and Statistics (A), Mathematical Methods of Physics I-II, Introduction to Computer Programming and Applications (A+), Digital Electronics (A+), Statistical Mechanics (A), Solid State Physics I-II (A), Nuclear Physics (A+), Introduction to Nanoscience and Nanotechnologies (A), Technical Writing and Presentation Skills (A).

Additional training: Machine Learning, MLOps and LLM Zoomcamps - DataTalks.Club (2023).

Pre-university: F.Sc. Pre-Medical, BISE Bahawalpur (89 percent). Matriculation, BISE Bahawalpur (95 percent).

RESEARCH EXPERIENCE & PROJECTS

Vision-Language Fine-Tuning on Radiography

Independent research project - multimodal ML 2025 - 2026

- Fine-tuned Llama 3.2 Vision with Unsloth and LoRA on radiography imagery, adapting a general vision-language model to medical image understanding.
- Observed that report-like fluency emerges long before verifiable grounding: the model describes findings confidently while the supporting evidence stays untraceable to a region. This motivates my current interest in grounded medical VLMs.

Visual Grounding Pipeline for Chest Radiographs

Independent project - multimodal ML 2025

- Built a five-stage pipeline that forces a model to commit to a location: multimodal analysis, grounded retrieval, normalised bounding-box emission, box rendering onto the source image, and a clinician-facing discussion stage.
- Demonstrates region-level accountability in 2D, and makes explicit how much scarcer grounded supervision is than descriptive supervision.

Airport Vehicle Detection: Vision versus LiDAR

Freelance research study for a UK-based researcher 2025

- Benchmarked a range of computer-vision detection and depth-estimation models against LiDAR for vehicle detection at airports, across varying lighting, seasonal and weather conditions.
- The central finding was the instability of model ranking under distribution shift: aggregate accuracy was a poor predictor of behaviour once acquisition conditions changed.

Edge-AI Vision on NVIDIA Jetson

Embedded computer vision 2024 - 2025

- Deployed quantised YOLOv8 detection to a Jetson Orin Nano at sub-40 ms fully offline inference for low-bandwidth environments.

Physics Instrumentation & Remote Sensing

Undergraduate hardware projects 2021 - 2023

- Portable Raspberry Pi radar for remote sensing with real-time signal processing for distance and velocity estimation; IoT safety systems including a blind-spot accident-prevention module and a gas-leakage detector.

PROFESSIONAL EXPERIENCE

Founder & CEO - Wisko (wiskoai.com)

Multi-tenant agentic AI SaaS - whiteboard animation Current

- Lead a platform where an agentic pipeline turns a prompt into a finished animation; own the architecture and the multimodal generation stack.

AI Engineer - LLM Systems Research & Analyst

Contract, remote - London, United Kingdom Jul 2025 - Jan 2026

- Deployed and scaled LLMs with vLLM for high-throughput, low-latency inference and efficient KV-cache use.
- Fine-tuned models on proprietary corpora; built FastAPI services with LMCache and Redis caching.

AI Voice Agent Developer & Research Analyst

Confidential GovTech project, remote - Morocco Feb 2025 - Jun 2025

- Co-developed real-time bi-directional translation across 30+ languages at sub-1500 ms latency, integrating Whisper ASR with custom speech pipelines and context-aware memory.

AI Engineer - Voice AI Agent Platform

US project, remote - Miami, United States Nov 2024 - Feb 2025

- Designed a telephony-scale voice platform with SIP routing and LangGraph agents for multi-turn conversation.

Junior AI Engineer (Intern)

iAxon Software House, Pakistan Apr 2023 - Feb 2024

- Cut model-training compute time 30 percent through feature engineering; improved forecasting accuracy 15 percent.

TECHNICAL SKILLS

Vision & multimodal:	Vision transformers, object detection, visual grounding, vision-language models, depth estimation, PyTorch, OpenCV
LLM & training:	LoRA / QLoRA, Unsloth, GRPO, quantisation, KV cache, RAG, Hugging Face, TRL
Inference & MLOps:	vLLM, ONNX Runtime, Docker, FastAPI, edge deployment (Jetson Orin Nano), managed cloud training and evaluation
Data & retrieval:	Qdrant, Pinecone, Weaviate, Redis, PostgreSQL
Scientific foundations:	Computational physics, mathematics, statistics, signal processing

ACADEMIC SERVICE & LEADERSHIP

- Conference organiser: ICPSE-2021 and ICPSE-2022 (International Conferences on Physical Sciences and Engineering); ICSMAND (9th International Conference on Semiconductor Materials and Nano-Devices).
- President, Natural Resources Defence Council and Society, KFUEIT.
- Organised the 1st South Punjab Science Mela (2021); student proctor at the 1st KFUEIT Convocation.

LANGUAGES & CERTIFICATIONS

English:	IELTS 7.5 overall - C2 Listening and Reading, C1 Speaking and Writing.
Urdu:	Native.
Certifications:	DeepLearning.AI - Fine-Tuning LLMs; LLMOps: Deployment and Inference; Building AI Voice Agents for Production; Practical Multi-AI Agents with CrewAI; AI Agentic Design Patterns with AutoGen; Evaluating AI Agents; Vector Databases; LangChain for LLM App Development. edX / IBM - Generative AI: Agents with RAG and LangChain. DataTalks.Club - MLOps and LLM Zoomcamp.