AALIYAN AHMED

AI ENGINEER

+92 3130457996 | AALIYANAHMEDRAJPUT@GMAIL.COM | LINKEDIN

GITHUB

BACKGROUND

AI Engineer with hands-on experience in Computer Vision (CV), Machine Learning (ML), Deep Learning (DL), Generative AI (GenAI), Agentic AI, and Natural Language Processing (NLP). Adept at developing, fine-tuning, and deploying end-to-end AI solutions, including data preparation, model training, real-time inference, and system integration. Expertise in creating intelligent, autonomous workflows using Agentic AI frameworks that enable decision-making, multi-step reasoning, and task automation. Skilled in designing and deploying Generative AI (GenAI) models for a variety of applications such as image captioning, content generation, and multimodal tasks. Proven ability to deliver robust, scalable, and production-ready AI systems with a strong focus on execution, efficiency, and impactful real-world applications.

EDUCATION

BACHELOR OF SCIENCE IN ARTIFICIAL INTELLIGENCE.

- University of Management and Technology, Lahore.

Experience

- **Freelance AI Engineer | Fiverr**
  Remote | 2022 – Present
- Deliver custom, production-ready AI solutions for a diverse global client base with expertise in Computer Vision, LLM-based Chatbots, and Agentic AI systems. Handle end-to-end AI pipelines, from requirement analysis to model deployment on cloud and multi-client processing (MCP) servers.
- **Computer Vision**:
  Developed real-time object detection and tracking systems using **YOLO, Deep SORT**, and custom-trained **CNNs**. Utilized **ViT (Vision Transformer)** and **DINOv2** for visual representation learning. Built solutions for **vehicle recognition** using frameworks like **OpenCV**, **TensorFlow**, **PyTorch**, and **Keras**. Optimized models for deployment on edge devices, ensuring low-latency performance for real-time streaming applications.
- **Generative AI & Multimodal Systems**: Built multimodal AI systems integrating image+text for rich captioning and visual reasoning using models like Florence-2, BLIP, GIT, and DINOv2. Fine-tuned models using LoRA, QLoRA, RAG workflows, and vector search (FAISS, Chroma).
- **LLMs & Chatbots**: Designed and deployed custom chatbots using LangChain, OpenAgents, and agentic AI frameworks. Integrated multi-step reasoning, memory, and tool usage in intelligent agents (AutoGPT, BabyAGI). Implemented RAG for personalized responses.
- **Agentic AI & Automation**: Developed intelligent multi-step agents and integrated them with workflow automation tools like N8n, Make.com, and Zapier to automate tasks and enhance operational efficiency.
- **Model Development & Training**: Built end-to-end machine learning and deep learning pipelines, specializing in transfer learning, model fine-tuning, and supervised learning with PyTorch and TensorFlow. Delivered custom solutions based on clients' datasets.

- **MCP Server Deployments**: Deployed AI systems on multi-client servers, optimizing for real-time inference, GPU resource management, and parallel processing to ensure scalable performance in production environments.
- **API & Backend Engineering**: Developed FastAPI-based RESTful APIs to serve AI models efficiently. Implemented secure, fault-tolerant services with robust logging and async handling for production-grade environments.
-

Skills

- **Generative AI (GenAI)**:
  - **Text**: LLM pipelines, custom story generators, summarizers, paraphrasers using GPT-J, Mistral, LLaMA
  - **Vision**: Florence-2, BLIP, GIT, ViT, DINOv2 for image captioning, visual reasoning, and multimodal tasks
  - **Multimodal**: Integrated image+text systems for rich captioning, scene understanding, and AI-generated content
  - Fine-tuning and optimization using LoRA, QLoRA, RAG workflows, and vector search (FAISS, Chroma)
- **Computer Vision (CV)**:
  Object detection and tracking (YOLOv5/v8, DeepSORT), facial attribute analysis (age, gender, race, emotion), vehicle recognition, license plate detection, real-time surveillance, motion-triggered processing
- **LLMs & Chatbots**:
  Custom chatbot development with LangChain, OpenAgents, memory integration, Retrieval-Augmented Generation (RAG), agentic reasoning, and tool usage
- **Agentic AI & Workflow Automation**:
  Built intelligent multi-step agents using CrewAI, LangGraph, AutoGPT, BabyAGI, integrated into tools like N8n, Make.com, Zapier
- **Model Development & Training**:
  End-to-end ML/DL pipelines with PyTorch, TensorFlow, custom dataset training, transfer learning, supervised learning, model evaluation
- **MCP Server Deployments**:
  Scalable deployment of AI systems on multi-client MCP servers with real-time inference, GPU resource optimization, and parallel processing
- **API & Backend Engineering**:
  FastAPI-based RESTful APIs, async handling, logging, secure production-grade service deployment
- **Tools & Platforms**:
  Hugging Face, Weights & Biases, Docker, Git, Google Colab, Jupyter Notebooks
- **Programming & Languages**:
  Python (expert), SQL (intermediate), Bash (basic)

Projects

Virtual Try-On System (OOTDiffusion)

Implemented a state-of-the-art virtual try-on system using OOTDiffusion, a diffusion-based model for realistic outfit transfer. The system supports swapping user-uploaded clothing onto a person image using pose estimation and cloth warping techniques. It includes preprocessing pipelines, pose conditioning, and garment warping stages. Ideal for e-commerce virtual fitting rooms.
Key Features:
Swaps garments with high-fidelity preservation of pose and texture
Automated inference pipeline with user image and cloth input
Tech Stack: PyTorch, Diffusers, OpenCV, Pose Estimation (OpenPose/MediaPipe)

Text-to-Speech (TTS) API with Polly + React Frontend

Built a custom API using AWS Polly for neural text-to-speech (TTS), deployed with FastAPI for backend and integrated into a React frontend. Users can input text, choose voice types, and download speech in real-time.
Key Features:
React UI with real-time audio feedback
Backend handles AWS Polly voice generation
Audio caching and download support
Tech Stack: FastAPI, AWS Polly, React.js, HTML5 Audio API

Lead Generation Automation (n8n + LLM Response Agent)

Created an automated lead generation pipeline using n8n. The workflow scrapes or receives lead data via APIs or forms, processes it through filters, then forwards it to an integrated local LLM-based agent (e.g., LLaMA/phi) for tailored outreach email/message generation.
Key Features:
Fully automated: From lead capture → LLM response → CRM input
Uses n8n for visual workflow design and integration
LLM responses customized based on lead context
Tech Stack: n8n, Node.js, Custom REST APIs, Local LLMs (LLaMA, phi), JSON Webhooks

Facial Recognition and Demographic Analysis

Developed a facial recognition system that predicts age, gender, and race using a custom-trained ResNet model on the FairFace dataset. The application includes a backend built with FastAPI, a UI implemented in Flask, and is containerized with Docker for CI/CD. Utilized PyTorch for model training, dataset preparation, and data augmentation, ensuring high accuracy and scalability.

Real-Time Entry/Exit Person Counting

Created a real-time person counting system that tracks the entry and exit of individuals through a camera. Leveraged YOLO and OpenCV for accurate object detection and motion tracking. This system enables precise monitoring of people flow in various environments, enhancing security and operational efficiency.

Real-Time ROI-Based Object Detection

Developed a system that detects objects only inside a defined Region of Interest (ROI). Integrated motion detection to trigger alerts efficiently, ensuring precise detection within designated zones.
Libraries/Frameworks: Python, YOLO, OpenCV, NumPy, Motion Detection

Boundary Line Crossing Detection (Outside ROI)

Built a system to detect objects outside a defined boundary line and track when they cross into restricted areas. Used motion-based event triggering for efficient processing and alerting.
Libraries/Frameworks: Python, YOLO, OpenCV, Motion Detection
Background Filter Application

Developed a real-time background-changing application that can blur or replace the user's background with a custom image. Integrated with Zoom via OBS virtual camera for seamless virtual meeting use.
Libraries/Frameworks: Python, OpenCV, MediaPipe, Pillow, NumPy

Vehicle Parts Quality Classification

Built a classification model for assessing the quality of vehicle parts, aiding in meeting safety standards and minimizing mechanical failures.
Libraries/Frameworks: Python (Tkinter GUI), PyTorch, TensorFlow, OpenCV, Pillow, NumPy
Real-Time Surveillance System
Designed a high-performance surveillance system for real-time object detection and visual analytics using streaming architectures.
Libraries/Frameworks: Python, YOLO, PyTorch, TensorFlow, RabbitMQ, FastStream, Apache Kafka

Real-Time Accident Detection

Developed an AI-powered surveillance module using a custom-trained YOLOv11n model to detect traffic accidents in real-time. Alerts are triggered via Twilio or custom APIs for emergency response.

Technologies: Python, PyTorch, YOLO, Data Labeling

Weapon Detection Model

Created a real-time firearm detection model for surveillance footage, using YOLO and a custom dataset. Optimized for low-latency inference on edge devices, enabling real-time alerts to security teams.

Libraries/Frameworks: Python, YOLOv5, OpenCV, PyTorch, TensorFlow

Real-Time Class Attendance System

Implemented an automated attendance system using facial recognition powered by a YOLOv11n model, enabling accurate and real-time attendance tracking in classrooms.

Technologies: Python, PyTorch, YOLO, Data Labeling

Customer Support Chatbot (LLM-Based)

Built a lightweight customer support assistant powered by a local LLaMA model, capable of retrieving and summarizing relevant database responses in real-time.

Libraries/Frameworks: Python, LLaMA, SQLite, Flask, FastAPI

GenAI Image Captioning and Visual Reasoning

Developed a multimodal image captioning system using BLIP and Vision Transformers to generate high-quality, context-aware image descriptions for accessibility and content tagging.

Libraries/Frameworks: Python, BLIP, ViT, PyTorch, Hugging Face Transformers