

MUHAMMAD AHMAD

AI/ML Engineer | Computer Vision & Generative AI Specialist

Lahore, Pakistan | ahmadaiwizard42@gmail.com | +92 309 2139073 | [LinkedIn](#) | [Portfolio](#)

PROFESSIONAL SUMMARY

AI/ML Engineer with 4+ years of experience designing, training, and deploying production-grade deep learning, computer vision, NLP, and generative AI systems. Proven track record delivering end-to-end ML solutions — from model development in PyTorch and TensorFlow through backend integration (Django, Flask, FastAPI) to full deployment — including computer vision models achieving 99.5%+ accuracy. Hands-on expertise building Retrieval-Augmented Generation (RAG) pipelines and multi-agent systems with LangChain, LangGraph, and CrewAI, plus diffusion and VAE-based generative models. Skilled at translating research into scalable, cost-efficient AI products for real-world business use cases.

CORE SKILLS

- **Programming & Databases:** Python, SQL, MongoDB, PostgreSQL, MySQL
- **Machine Learning & Deep Learning:** PyTorch, TensorFlow, Scikit-learn, ANN, CNN, Transfer Learning, Model Deployment
- **Computer Vision:** OpenCV, MediaPipe, YOLOv8, MobileNetV2/SSD, OCR (PaddleOCR, Tesseract, AWS Textract, Google Vision AI)
- **Generative AI & LLM Engineering:** RAG, LangChain, LangGraph, CrewAI, Diffusion Models, GANs, VAE, Prompt Engineering, LLM, Agentic AI, Generative AI
- **Backend & Deployment:** Flask, FastAPI, Django, Gradio, Streamlit, REST APIs, Docker, Git, AWS, GCP, Azure
- **Automation & Integration:** N8N, Webhooks, Workflow Automation, Pushover, AI Automation

PROFESSIONAL EXPERIENCE

AI/ML Engineer | Arhamsoft (Pvt) Ltd, Lahore, Pakistan 2022 – Present

- Lead end-to-end delivery of AI/ML projects spanning object detection, speech-to-video generation, generative modeling, and OCR pipelines, taking products from prototype through to production deployment.
- Architect and deploy Retrieval-Augmented Generation (RAG) and multi-agent AI systems using LangChain, LangGraph, and CrewAI to automate complex business workflows.
- Engineer image-to-video and audio-to-video generation pipelines using diffusion models and Variational Autoencoders (VAEs) to expand content-generation capabilities for client products.
- Collaborate with backend and product teams to integrate trained AI models into scalable REST API services built on Flask, FastAPI, and Django.

KEY PROJECTS

American Sign Language (ASL) Recognition System

- Designed and built 3 integrated AI models — pose prediction, an LLM-based chatbot, and digit/alphabet recognition (MobileNetV2) — plus a video-level ASL recognition pipeline.
- Achieved 99.5%+ accuracy on digit and alphabet recognition, enabling reliable real-time sign language translation.

PDF AI Chatbot (RAG)

- Built a Retrieval-Augmented Generation chatbot powered by local LlamaCpp inference, eliminating API costs and ensuring full data privacy for document Q&A.

Gemini AI Chatbot with Vision

- Developed a Flask-based multimodal chatbot integrating the Google Gemini API with RAG to handle combined text-and-image queries.

N8N-Powered AI Chatbot

- Built a web-based chatbot connected to an N8N webhook to automate conversational logic and backend workflows.

Personal AI Agent

- Built a CrewAI-based autonomous agent with function calling for task automation and real-time notification delivery.

VigiCount — Real-Time Security & People-Counting System

- Developed a real-time people-counting system using MobileNet SSD and YOLOv8 for security and occupancy-monitoring use cases.

PaddyLeafShield — Crop Disease Classifier

- Trained a PyTorch CNN model to classify 6 paddy leaf diseases and deployed it via a web-based interface for accessible field diagnostics.

Face Mask Detection

- Built a computer vision-based face mask detector using OpenCV, deployed with a Streamlit interface.

AI Background Remover

- Developed a browser-based background removal tool using the rembg library.

Unstructured-to-Structured OCR Pipeline

- Researched and benchmarked PDF/Excel table-extraction approaches across PaddleOCR, Camelot, AWS Textract, and Google Vision AI to identify the most accurate extraction pipeline.

RESEARCH & DEVELOPMENT

- **Unstructured Data Parsing:** Conducted in-depth research on unstructured table conversion; identified LlamaIndex as the most robust framework for efficient parsing and contextual data extraction.
- **Sign Language Recognition:** Investigated ASL recognition approaches through comparative analysis of pose-estimation datasets and deep learning architectures for production-level accuracy.
- **Dimensionality Reduction:** Explored high-dimensional data transformation techniques, identifying MST++ as the optimal method for converting 3D RGB images into a compact 31D feature representation.

EDUCATION

Bachelor of Science in Computer Science

University of Management and Technology, Lahore

- Final Year Project: American Sign Language Recognition using AI and Deep Learning (3 integrated AI models).