

Shobhit Pal

AI Engineer — Generative AI & Computer Vision — Python — FastAPI

Phone: +91 91703 36377 — palshobhit318@gmail.com — LinkedIn — GitHub

Professional Summary

AI Engineer with **nearly 3 years of experience** in a fast-paced startup environment. Hands-on experience building **Generative AI applications** including **RAG-based chatbots**, vector search systems, and backend APIs using **Python, FastAPI, and FAISS**. Strong foundation in **backend development and data processing**, with practical experience developing **real-time computer vision systems** using YOLOv8 and OpenCV.

Technical Skills

Generative AI: RAG, FAISS, Ollama, Mistral LLMs, LangChain (basic)

Computer Vision: YOLOv8, OpenCV, Real-time Object Detection

OCR & Data Processing: PaddleOCR, PyTesseract, Pandas, Regex

Backend & APIs: Python, FastAPI, Flask, REST APIs

Databases: MongoDB

Automation: Selenium, BeautifulSoup

Tools: Docker, Git, WebRTC, GCP

Experience

FOURBRICK TECHNOLOGY

Apr 2023 – Present

ML / AI Engineer

Jun 2024 – Present

- Designed and developed a **RAG-based Generative AI chatbot** for document-driven question answering.
- Built end-to-end pipelines including document ingestion, chunking, embeddings, vector search, and response generation.
- Developed **real-time computer vision solutions** using YOLOv8 and OpenCV for object detection use cases.
- Exposed AI and CV inference pipelines via **FastAPI-based REST APIs** for production use.

Software Engineer

Apr 2023 – May 2024

- Developed scalable backend services using **Python, Flask, FastAPI, and MongoDB**.
- Built OCR-based data extraction and parsing pipelines using PaddleOCR, PyTesseract, Pandas, and Regex.
- Implemented automation and web data extraction workflows using Selenium and BeautifulSoup.

Projects

AiVerse — Generative AI Chatbot Platform

Python, FastAPI, RAG, FAISS, Ollama, Mistral LLMs, MongoDB

- Built a **ChatGPT-style conversational AI platform** with real-time streaming responses.
- Implemented **Retrieval-Augmented Generation (RAG)** for PDF ingestion and context-aware question answering.
- Designed and integrated a **FAISS-based vector search** layer for efficient semantic retrieval.
- Implemented secure authentication and user-specific conversation management using JWT.
- Architected a **modular and scalable backend** to support future feature expansion.

AutoVision Alert — Computer Vision Safety System

Python, YOLOv8, OpenCV, Flask, WebRTC, MongoDB

- Developed a **real-time object detection system** to monitor industrial machines and worker safety.
- Implemented a **distance-based hazard detection mechanism** to identify unsafe proximity scenarios.
- Integrated OpenCV-based video pipelines for continuous real-time inference on live camera feeds.
- Deployed the system using Flask, exposing REST APIs for monitoring and alert generation.

Education

B.Tech — Computer Science

2021

Dr. A.P.J. Abdul Kalam Technical University (AKTU), Lucknow