
📘 Khoj AI: Missing Person Identification & Reunification System

---

1. Introduction

1.1 Problem Statement

Every year, thousands of people go missing across the world due to multiple reasons such as abduction, accidents, natural calamities, or simply losing contact with families. Locating missing individuals is often a **time-consuming, manual, and error-prone** process.

Traditional methods rely heavily on **paper notices, posters, and manual searches** which are slow and ineffective in today's digital age.

1.2 Objective

Khoj AI has been designed as a **technology-driven solution** to help in finding missing persons faster by combining:

- **AI-based face recognition** for identifying individuals.
- **Age progression** to predict how a person might look after several years.
- **Liveness detection** to prevent fraud and ensure authenticity.
- **AI chatbot assistance** for user support.
- **Data analytics and web scraping** to maintain updated case records.

The ultimate goal is to **reduce the time to reunite missing persons with their families**.

---

2. Project Architecture

2.1 System Overview

Khoj AI is built using a **full-stack architecture** with AI/ML integration:

1. **Frontend (User Interaction Layer)**
    - Developed using **React.js** with **TypeScript**, styled with **Tailwind CSS**.
    - Provides an **intuitive and responsive UI**.
    - Allows users to:
        - Upload images.

        - Use webcam for real-time detection.
        - Interact with chatbot assistant.
        - View case statistics and reports.

2. **Backend (Application & AI Layer)**
    - Built on **Python FastAPI** for high-performance API handling.
    - Integrates multiple **AI/ML models** including:
        - **InsightFace Antelope ONNX** for facial recognition and similarity matching.
        - **Liveness detection model** using eye-blink tracking.
        - **Age progression model** to generate future-aged appearances.
    - Implements **web scraping** pipelines to collect data from missing persons' registries and online reports.
    - Provides **REST APIs** for communication between frontend and backend.
3. **Database (Storage Layer)**
    - Stores **missing person records, images, and case histories**.
    - Maintains **analytics logs** for model predictions and user queries.
4. **Dataset**
    - Core dataset modified from **Kaggle**.
    - Preprocessed with techniques like **face alignment, augmentation, and labeling**.
    - Enhanced with **synthetic data** to improve performance in real-world cases.

---

3. Core Functionalities

3.1 Face Recognition

- Accepts an **uploaded image** or **live webcam feed**.
- Extracts facial embeddings using **InsightFace Antelope**.
- Matches embeddings against database of missing persons.
- Returns:

    - **Details of the matched person**.
    - **Similarity percentage (confidence score)**.

3.2 Webcam Integration

- Real-time face detection using webcam.
- Allows instant verification without needing image uploads.
- Useful for **NGO volunteers, police checkpoints, and search operations**.

3.3 Liveness Detection

- Prevents spoofing by verifying if the input is from a **real person**.
- Detects **eye blinks and subtle movements**.
- Ensures that printed photos or videos cannot fool the system.

3.4 Age Progression

- Accepts **younger photographs** of a missing person.
- Applies AI-based age-progression techniques.
- Generates **aged appearance** images to reflect current estimated look.
- Very useful in cases where people have been missing for **many years**.

3.5 AI Chatbot Assistant

- Provides **natural language interaction**.
- Assists in:
    - Searching for missing persons.
    - Explaining system functionality.
    - Answering FAQs (e.g., "How do I report a missing case?").
- Built with **AI conversational models** for accuracy and user-friendliness.

3.6 Web Scraping

- Collects data from:
    - Government missing person registries.
    - NGO portals.
    - Online news reports.
- Regularly updates the **backend database** with fresh cases.

---

4. Statistics & Analysis

The project includes an **analytics dashboard** that presents:

- **Model Accuracy**: Percentage of correct identifications.
- **Precision**: Ability to correctly identify actual matches.
- **Recall**: Ability to detect all possible missing persons in database.
- **Confidence Scores**: Similarity values for each prediction.
- **State-wise Missing Person Records**: Helps authorities track regional data.
- **Prediction Logs**: Stores history of all user queries and results.

This data helps in **system improvement, transparency, and research**.

---

5. Dataset

- **Source**: Kaggle dataset (modified).
- **Size**: Thousands of face images across different ages, genders, and ethnicities.
- **Preprocessing Steps**:
  - Face alignment and cropping.
  - Image augmentation (rotation, scaling, color adjustment).
  - Balancing across categories (children, adults, elderly).
- **Age Progression Support**: Augmented dataset includes **synthetic aged images** for training.
- **Validation**: Dataset split into training, validation, and testing sets.

---

6. Technologies Used

- **Frontend**:
  - React.js, TypeScript, Tailwind CSS, HTML5.
- **Backend**:
  - Python, FastAPI, REST APIs.
- **AI/ML Models**:

  - InsightFace Antelope (ONNX models).
  - Age Progression Model (GAN/Deep Learning based).
  - Liveness Detection (Blink-based ML).
- **Other Tools**:
  - ONNX Runtime.
  - Web Scraping (BeautifulSoup, Scrapy).
  - Data Visualization (Charts, Graphs).
- **Dataset**: Kaggle (modified and augmented).

---

7. Conclusion

Khoj AI is a **next-generation humanitarian technology project** that aims to make the process of finding missing persons **faster, smarter, and more reliable**.

Its **unique combination of AI-powered recognition, liveness detection, and age progression** sets it apart from conventional systems.

With further enhancements such as **real-time integration with government databases, cloud deployment, and multilingual chatbot support**, Khoj AI has the potential to become a **national and global platform** for missing person identification and reunification.
