Facial Recognition System - Project Report

# 1. Introduction

Facial recognition is a biometric technology that uses distinctive facial features to identify and verify individuals.   
This project aims to develop a robust facial recognition system using Python. The system is built using a dataset of facial images represented as numerical feature vectors (arrays), instead of traditional image files. The model compares these vectors to accurately identify individuals.

# 2. Objective

The primary objective of this project is to create a reliable and efficient facial recognition system that can identify and verify individuals using numerical image data. Unlike typical face recognition models that depend on direct image inputs, we worked with pre-extracted facial vectors that describe the face numerically.

The key objective of this project is to verify the identity of users based on their facial features. We wanted to ensure that the system could distinguish between known and unknown users using efficient similarity-based matching.  
The system is designed to work in real-time scenarios, ensuring high accuracy and adaptability under varying environmental conditions.

# 3. Dataset Description

The dataset used in this project comprises numerical feature representations of facial images. Each entry corresponds to a person’s face converted into a high-dimensional vector. These vectors are used to train the machine learning model to distinguish between different individuals.

# 4. Data Preprocessing

Before feeding data into the model, it undergoes several preprocessing steps such as normalization, dimensionality reduction (using PCA), and encoding labels. This step ensures that the input data is clean, consistent, and optimized for model training.

# 5. Data Visualization

Data visualization plays a vital role in understanding the distribution of facial vectors and the performance of the model. I used plots for visualization like Distribution of Plots. Top 10 features with Higher Variance etc. Dimensionality reduction techniques like PCA are used to visualize high-dimensional data in 2D or 3D space for better understanding.

# 6. Model Training

We experimented with multiple algorithms: K-Nearest Neighbors (KNN), Support Vector Machine (SVM), Logistic Regression, and Random Forest. These models were trained on 80% of the dataset and tested on the remaining 20%.  
The training data, after preprocessing, is fed into the classifier, which learns to distinguish different individuals based on their unique facial vectors. The best performing model for this training dataset here is Logistic Regression.  
Evaluation metrics like accuracy, precision, recall, and confusion matrices helped us understand each model’s strengths and weaknesses.

# 7. Identity Verification

Once the model is trained, it is used to verify or predict the identity of new inputs by comparing the incoming face vector with the known identities in the database. The system outputs the predicted identity and the associated confidence score.

# 8. Deployment

The final model will be deploying using Streamlit, enabling an interactive and user-friendly interface for uploading vectors and performing identity verification. The deployment facilitates real-time face recognition and identity verification in a simple UI.

# 9. Challenges Faced

• Handling high-dimensional data without actual images was complex.  
  
• Ensuring model accuracy across varied environmental factors like lighting and angle.  
  
• Designing an intuitive user interface for non-technical users.  
  
• Managing performance and memory constraints for real-time execution.

# 10. Key Learnings

• Gained deeper understanding of facial feature extraction using vectors.  
  
• Learned how to implement dimensionality reduction techniques like PCA.  
  
• Acquired skills in deploying machine learning models with Streamlit.  
  
• Understood real-world challenges of face recognition technology in terms of accuracy, security, and scalability.

# 11. Conclusion

This project successfully demonstrates a working facial recognition system using feature vectors instead of raw images. The model achieved high accuracy and was effectively deployed in a real-time UI. With advancements in AI, such systems are becoming increasingly crucial for security, authentication, and surveillance applications.

I hope this presentation provided a clear understanding of how facial recognition systems can be implemented using numerical feature vectors.

Created By – Sufiyan Zamindar