# Facial Recognition System

## (a) Problem Description

This project implements a facial recognition system designed to accurately identify and verify human faces in real-time. The system addresses the challenge of secure and efficient biometric authentication by utilizing deep learning techniques to create unique facial representations (embeddings) that can be compared for identity verification.

The specific problems that this facial recognition system aims to solve include:

1. Creating a user-friendly interface for face registration and verification

2. Accurately detecting faces from webcam input

3. Generating consistent and discriminative facial embeddings

4. Securely storing facial embeddings for future verification

5. Providing real-time verification with high accuracy

The system supports multiple users and allows for straightforward management of registered identities, making it suitable for access control, attendance systems, and other authentication applications.

## (b) Dataset Description

The model was trained on the CASIA-WebFace dataset, a large-scale face recognition dataset containing 494,414 images across 10,575 different identities. This dataset is widely used in the research community for facial recognition tasks.

Dataset characteristics:

**Number of images:** 494,414 facial images

**Number of identities:** 10,575 unique individuals

**Image characteristics:** Various poses, expressions, and lighting conditions

**Source:** CASIA (Chinese Academy of Sciences' Institute of Automation)

https://www.kaggle.com/datasets/debarghamitraroy/casia-webface

For training efficiency, a subset of the dataset was used, limiting to 100 images per identity. This approach balanced training quality with computational constraints while ensuring sufficient examples of each identity were available for learning robust feature representations.

The CASIA-WebFace dataset was selected for this project because:

1. It contains a large number of identities, which helps the model learn discriminative facial features

2. The images capture real-world variations in pose, expression, and lighting, making the model more robust

3. It is a publicly available dataset designed specifically for facial recognition research

4. Its size and quality allow for effective training without requiring the computational resources needed for even larger datasets like MS-Celeb or VGGFace2

## (c) Machine Learning Algorithm Selection

The project implements a deep learning approach for facial recognition using the FaceNet framework, specifically with an InceptionResnetV1 architecture for embedding generation.

### Selected Machine Learning Algorithms:

1. **Face Detection**: Multi-Task Cascaded Convolutional Networks (MTCNN)

- Three-stage cascade architecture for face detection and alignment
- Provides bounding box coordinates and facial landmarks
- Ensures only properly aligned faces are processed for recognition

2. Facial Embedding Generation: InceptionResnetV1

- Deep convolutional neural network architecture
- Combines Inception modules with residual connections
- Outputs 512-dimensional facial embeddings
- Trained using CrossEntropyLoss for classification during training

3. **Face Verification**: Cosine Similarity Comparison

- Compares embeddings using cosine similarity metrics
- Normalizes embeddings before comparison
- Uses a threshold of 0.6 to determine match/non-match

### Justification for Algorithm Selection:

1. InceptionResnetV1 was selected because:

- It achieves state-of-the-art accuracy in facial recognition tasks
- The inception modules efficiently extract features at multiple scales
- Residual connections help with gradient flow during training
- The 512-dimensional output provides a compact yet discriminative facial representation
- The architecture balances computational efficiency with accuracy

2. MTCNN was chosen for face detection because:

- It provides accurate face detection even in challenging conditions
- It performs face alignment, which improves embedding consistency
- It is computationally efficient for real-time applications
- It has proven effectiveness in facial recognition pipelines

3. Cosine similarity was selected for verification because:

- It is invariant to scaling, focusing on the direction of the embedding vectors
- It works well with normalized facial embeddings
- It provides a simple similarity score between 0 and 1
- It is computationally efficient for real-time comparisons

The hybrid approach of using CrossEntropyLoss during training and cosine similarity during inference allows the model to learn discriminative features while providing a simple, interpretable similarity metric for verification.

## (d) Performance Evaluation Results

The model was trained for 5 epochs on a subset of the CASIA-WebFace dataset using Google Colab's T4 GPU, with training taking approximately 6 hours.

### Training Performance:

**Final Accuracy:** 84% after 5 epochs

**Training Duration:** 6 hours on Google Colab's T4 GPU

**Loss Function:** CrossEntropyLoss

### Verification Performance:

During testing, the system demonstrated:

- Successful face detection in various lighting conditions
- Accurate identity verification with properly registered faces
- Real-time performance suitable for interactive applications

The system implements a multi-sample registration approach, where multiple images of the same person are captured and their embeddings averaged. This enhances robustness by accounting for slight variations in pose, expression, and lighting.

For verification, a similarity threshold of 0.6 was established based on empirical testing. This threshold balances:

**False Accept Rate (FAR):** Minimizing incorrect authentications

**False Reject Rate (FRR):** Ensuring legitimate users are correctly recognized

The combination of MTCNN for precise face detection and alignment with InceptionResnetV1 for embedding generation proved effective in creating a robust facial recognition system. The approach of averaging multiple sample embeddings during registration further enhanced the system's reliability by creating more representative reference templates for each identity.