A project Report

Smart Vision: Attendance marking System using CNN

by

Pranshu Saini
Pratham Agarwal
Saloni Tyagi

Under the Supervision of

Prof. saurabh kumar Shrivastav

Bennett University Greater Noida, Uttar Pradesh

FEB-MAY 2025

I. Introduction

Traditional attendance-taking techniques, such as calling names, using ID cards, or fingerprint scanners in workplaces, schools, and colleges, are frequently laborious, prone to mistakes, and easy to tamper with. These systems require additional effort and time and can be misused through proxy attendance, where someone marks attendance for a friend.

Many of these traditional methods rely on specialized equipment, such as RFID systems or biometric scanners, which can be expensive to install and maintain. Additionally, physical contact is necessary for ID card swipes and fingerprint scanners, raising hygiene concerns, especially in the wake of the COVID-19 pandemic.

A smart and automated method for tracking attendance is offered by a Face Recognition-Based Attendance Management System using CCTV cameras. This system enhances efficiency, saves time, and prevents cheating.

A. Ongoing Work in Deep Learning-Based Face Recognition

Deep learning techniques have significantly advanced face recognition (FR) systems, enabling more accurate and robust identification. Deep convolutional neural networks (CNNs) have revolutionized FR by leveraging multiple layers of feature extraction, learning hierarchical representations of facial features. These techniques, collectively termed **Deep Face Recognition (Deep FR)**, have evolved primarily through advancements in **network architectures and loss functions**.

One of the major challenges in FR is training complex models with limited data. To address this, **transfer learning** is widely adopted, where pre-trained models on large datasets are fine-tuned for specific tasks. Several loss functions have been proposed to enhance feature discrimination, including:

- **ArcFace** – Introduced an **additive angular margin loss** to improve class separability.
- **SphereFace** – Used **ResNet-64 with angular softmax loss** for better intra-class compactness and inter-class separability.
- **FaceNet** – Employed a **triplet loss function** and a large dataset to train a GoogleNet architecture.
- **CosFace** – Designed a **cosine-margin-based loss function** to enhance decision boundaries in the angular space.

These deep learning models achieve **high accuracy (98%–100%)** on benchmark datasets, demonstrating their effectiveness.

B. Key Components of Face Recognition Systems

Face recognition typically involves three major components:

1) **Face Detection** – Identifies and localizes faces in images or videos, ensuring robustness to pose, lighting, and occlusions.
2) **Facial Landmark Extraction** – Detects key facial points (e.g., eyes, nose, mouth) for **face alignment** and normalization.
3) **Feature Embedding & Classification** – Encodes identity information and computes similarity scores for face identification.

C. Recent Developments and Challenges

- **Real-time Optimization** – Researchers are working on **lightweight CNN models, quantization, and edge computing** for real-time applications.
- **Security Enhancements – Liveness detection, multi-factor authentication, and blockchain-based storage** are being explored to counter spoofing attacks and ensure data integrity.
- **Privacy-Preserving FR** – Techniques like **federated learning and homomorphic encryption** are being developed to comply with privacy regulations (e.g., GDPR).

The continuous advancements in deep learning are making AI-based attendance and authentication systems more **accurate, secure, and scalable**, transforming traditional attendance methods in educational and corporate environments.

II. Methodology

In modern educational institutions and workplaces, traditional attendance systems are often time-consuming and prone to human error. This project implements an automated attendance system using CCTV cameras integrated with Deep Neural Networks (DNNs), Multi-task Cascaded Convolutional Networks (MTCNN), and Haar Cascade techniques. The system captures real-time images from video feeds, processes them using machine learning models, and records attendance in an Excel file.

The methodology consists of the following key steps, ensuring accuracy, efficiency, and ease of use.

A. Data Collection & Preprocessing

The first and most crucial step in developing an AI-based attendance system is collecting a high-quality dataset of facial images. A well-structured dataset significantly enhances the accuracy and reliability of the face recognition model. Since individuals may appear in different lighting conditions, angles, and facial expressions, multiple images of each person are collected to improve the model's generalization.

1) CCTV Camera Feeds: One of the primary sources of images for the attendance system is live CCTV camera feeds. The system continuously extracts frames from real-time video streams, identifying and capturing individual faces automatically. This method eliminates the need for manual image collection and ensures that images are obtained in natural environments.

- **Automated Image Capture:** The system runs continuously, detecting and saving facial images from video feeds without human intervention.
- **Realistic Data Collection:** Unlike manually captured images, CCTV-based data collection ensures natural facial expressions and postures, reducing overfitting to specific poses.
- **Multiple Angles & Lighting Conditions:** The system captures faces in different orientations and lighting, improving recognition accuracy in dynamic settings.

By utilizing CCTV cameras, organizations can effortlessly gather a comprehensive dataset without disrupting normal activities.

2) Manual Image Uploads: To further enhance the dataset, administrators can manually upload images. This method is useful when certain individuals are not detected properly by the CCTV system or require additional images for better recognition.

- **Admin-Controlled Uploads:** Administrators can add new students, employees, or visitors by uploading their photos manually.
- **Multiple Images for Training:** If an individual's face is not frequently captured in CCTV footage, additional images can be provided to improve recognition accuracy.
- **Ensuring High-Quality Data:** Manual uploads allow for clear and well-lit images, improving model performance.

This hybrid approach, combining automated CCTV image extraction with manual uploads, helps build a well-balanced dataset with minimal gaps.

3) Image Preprocessing: Once facial images are collected, they undergo several preprocessing steps to ensure they are suitable for training a machine-learning model. Raw images from video feeds or manual uploads may contain unnecessary background information, varying resolutions, or inconsistent lighting, which can affect model performance.

- **Grayscale Conversion:** - Color information is often unnecessary for face recognition, and removing it reduces computational complexity. - Each image is converted from RGB to grayscale, preserving only intensity information while reducing memory usage and processing time. - This results in faster processing and reduced storage requirements without significantly affecting recognition accuracy.
- **Face Detection & Cropping:** - The raw image may contain unwanted background elements that interfere with recognition. - The system applies face detection algorithms such as Haar Cascade or MTCNN to identify faces. Once detected, faces are cropped to remove unnecessary parts of the image. - This helps the model focus only on facial features, improving recognition accuracy.
- **Resizing & Normalization:** - Different image sizes can affect model performance, making it difficult to learn consistent features. - All images are resized to a standard dimension (e.g., 128x128 pixels) to maintain uniformity. - Pixel values are normalized (scaled between 0 and 1) to standardize intensity variations across different images. - This ensures that all images have the same format, making training more efficient and improving model convergence.

Through these preprocessing steps, the dataset is refined, making it well-structured and optimized for deep learning models. This ensures that the system operates with high efficiency and accuracy, even in real-world scenarios.

B. 2.2 Model Training for Face Recognition

Once the dataset has been collected and preprocessed, the next critical step is training a face recognition model capable of distinguishing between different individuals. The accuracy and efficiency of the system depend on the choice of recognition method and the training process. The model must be robust enough to recognize faces even under variations such as different lighting conditions, head orientations, facial expressions, and partial occlusions (e.g., glasses or masks).

2.2.1. Face Recognition Methods: The system supports multiple face recognition techniques, each with its own strengths and use cases. Depending on the computational resources and real-time requirements, organizations can select the most suitable method:

*2.2.1.1. Deep Neural Networks (DNNs)

Overview: Deep learning-based face recognition models leverage large-scale datasets and neural networks to achieve highly accurate facial identification.

How It Works:

- The system uses **Convolutional Neural Networks (CNNs)** to extract deep facial features.
- The network learns distinctive features such as eye shape, nose position, and jawline structure.
- A classification layer at the output assigns a label (identity) to each detected face.

Advantages:

- High accuracy, even with variations in lighting, angle, and expression.
- Adaptable to large datasets, making it scalable for institutions with many individuals.
- Can integrate with advanced deep learning models such as **FaceNet** or **ResNet** for even better performance.

Limitations: Requires significant computational resources (GPU/TPU) for training and inference.

*2.2.1.2. MTCNN (Multi-task Cascaded Convolutional Network)

Overview: MTCNN is a specialized deep learning model designed for face detection and alignment, making it highly efficient for real-time applications.

How It Works:

- Uses a three-stage deep learning architecture to detect and align faces.
- Identifies facial landmarks such as eyes, nose, and mouth positions to enhance recognition accuracy.
- Ensures that the detected face is well-aligned before passing it to the recognition model.

Advantages:

- Highly accurate in detecting faces under different orientations.
- Performs **face alignment**, improving recognition accuracy.
- Works well on both still images and video frames.

Limitations: Slower than traditional methods like Haar Cascade but provides better accuracy.

*2.2.1.3. Haar Cascade Classifier

Overview: Haar Cascade is a rule-based face detection method that relies on edge and texture patterns to identify faces.

How It Works:

- Uses a **cascade of classifiers** that analyze different parts of an image.
- Based on pre-trained Haar-like features, it identifies patterns such as eyes, nose, and mouth.

Advantages:

- Extremely fast and lightweight, making it ideal for real-time applications.
- Works efficiently on low-power devices and embedded systems.

Limitations:

- Less accurate than deep learning-based models.
- Can produce false positives (detecting faces where none exist).
- Struggles with occlusions and lighting variations.

2.2.2. Training Process: Once the appropriate face recognition model is selected, the training process begins. This involves multiple stages to ensure that the model can effectively distinguish between different individuals.

*2.2.2.1. Feature Extraction Using Convolutional Neural Networks (CNNs)

Why It's Important: Raw image data contains millions of pixels, but only certain facial features are relevant for recognition. CNNs help extract meaningful patterns.

How It Works:

- The first layers of the CNN detect basic features such as edges and corners.
- Deeper layers recognize complex patterns like eyes, nose, and jaw structure.
- The final layer represents the face as a unique feature vector (numerical representation).

Outcome: The system generates **high-dimensional feature embeddings** that uniquely represent each individual.

*2.2.2.2. Training with Backpropagation to Minimize Loss

Why It's Important: Training the model requires adjusting the internal weights to minimize errors in recognition.

How It Works:

- The model initially makes predictions based on random weights.
- The **loss function** (such as cross-entropy loss) measures how far the prediction is from the actual identity.
- The system updates weights using **backpropagation** and an **optimizer** (such as Adam or Stochastic Gradient Descent).
- This process repeats over multiple iterations (epochs) until the model achieves high accuracy.

Outcome: The model becomes increasingly accurate in distinguishing faces.

*2.2.2.3. Saving the Trained Model for Real-Time Inference

Why It's Important: Once trained, the model must be saved in a format that allows real-time face recognition.

How It Works:

- The trained model is stored in a file format such as **.h5 (HDF5)** or **.pkl (Pickle)**.
- This allows the system to **load the model quickly** for real-time recognition.
- The trained model can also be deployed on cloud-based services for remote access.

Outcome: The model is now ready to perform **real-time face recognition** on live CCTV feeds.

Training the face recognition model is one of the most crucial steps in building an AI-powered attendance system. By using advanced deep learning techniques such as CNNs and MTCNN, the system achieves high accuracy and robustness. The combination of feature extraction, backpropagation, and model saving ensures that the system is efficient and scalable for large organizations.

C. Attendance Marking System

1) Introduction: In modern educational and workplace environments, the process of attendance marking plays a crucial role in ensuring accountability and efficiency. Traditional attendance-taking methods are often time-consuming and prone to errors. This section details the automated attendance marking system using facial recognition technology integrated with CCTV cameras. The system captures real-time video frames, detects and recognizes individuals, and records their attendance in an organized manner. The implementation aims to reduce manual intervention, minimize errors, and enhance overall efficiency.

2) Workflow of the Attendance Marking System: The attendance marking process is structured into several key steps to ensure accuracy and reliability. These steps include face detection, feature extraction, identity matching, and attendance logging.

a) Step 1: Face Detection from Live Video Feeds:

- The system continuously receives real-time video feeds from CCTV cameras installed in designated areas such as classrooms or offices.
- Face detection is performed using models like MTCNN (Multi-task Cascaded Convolutional Networks) or Haar Cascade Classifier, which identify and extract faces from the video frames.
- Detected faces are cropped and prepared for further processing.

Challenges:

- Multiple faces in a single frame may lead to overlapping detections.
- Environmental factors like poor lighting or motion blur may impact detection accuracy.

b) Step 2: Feature Extraction for Identification:

- The detected facial regions are processed through deep learning models to extract key facial features.
- The extracted features are converted into numerical embeddings that uniquely represent each individual.
- These embeddings are then compared with a pre-stored database of enrolled individuals for identification.

c) Step 3: Matching Detected Faces with Stored Data:

- The system compares the extracted face embeddings with the database using similarity metrics such as cosine similarity or Euclidean distance.
- If the similarity score surpasses a predefined threshold (e.g., 85% match confidence), the system successfully identifies the individual.
- If no match is found, the system either ignores the detection or marks it as an "Unknown" entry.

Challenges:

- Facial variations due to different expressions, angles, or accessories can affect recognition accuracy.
- Individuals wearing masks or glasses may require additional processing techniques.

d) Step 4: Threshold-Based Decision for Attendance Logging:

- If the recognition confidence score is above the set threshold, the system logs the attendance of the identified individual after clicking the mark attendance button.
- A timestamp, including date and time, is recorded to ensure attendance is only marked once per session.
- Additional business rules are applied, such as:
  - **Time-based restrictions:** Attendance is logged only within a defined time window.
  - **Minimum presence duration:** Individuals must remain within the frame for a specified time before being marked present.

3) Features of the Attendance Marking System: To enhance reliability and functionality, the system incorporates several advanced features:

1) **Real-Time Processing**
   - The system operates continuously, detecting and recognizing individuals in real time without requiring manual intervention.
2) **Multi-Person Detection**
   - The model is capable of detecting and recognizing multiple individuals simultaneously in a single frame.
   - Suitable for classrooms, office meetings, and large gatherings.
3) **Automatic Re-Identification**
   - If an individual temporarily moves out of the camera's view, the system re-identifies them upon return.
   - Prevents unnecessary reprocessing and improves system efficiency.
4) **Customizable Attendance Rules**
   - Organizations can customize attendance criteria based on their requirements, such as:
     - Setting a minimum confidence threshold for recognition.
     - Restricting attendance marking within a specific time window.
     - Generating absence alerts for missing individuals.
5) **Integration with Other Systems**
   - The system seamlessly integrates with:
     - Institutional databases for attendance tracking and management.
     - Payroll systems for employee work-hour calculations.
     - Notification services to send alerts regarding attendance status.

4) Challenges and Solutions: While the attendance system is designed to be robust and accurate, certain challenges may arise during implementation. The following are common challenges and their corresponding solutions:

1) **Handling Low-Quality Images**
   - **Problem:** Poor-quality images may result in inaccurate recognition.
   - **Solution:** Image enhancement techniques, such as noise reduction and resolution improvement, are implemented to improve clarity.
2) **Occlusion and Partial Faces**
   - **Problem:** Individuals wearing masks, glasses, or partially obscured faces may not be accurately recognized.
   - **Solution:** Training models with diverse datasets, including masked faces, and focusing on unique facial features like eyes and forehead.

3) **Variable Lighting Conditions**
   - **Problem:** Poor lighting may impact the accuracy of detection and recognition.
   - **Solution:** Adaptive contrast enhancement techniques, such as histogram equalization, are used to improve facial visibility.

5) Conclusion: The automated attendance marking system provides a seamless and efficient approach to attendance tracking using real-time face recognition. By integrating deep learning models, real-time feature extraction, and threshold-based decision-making, the system ensures accurate and efficient attendance logging. Additional features, including multi-person detection, customizable rules, and database integration, make it a scalable solution for educational institutions and corporate environments. Addressing challenges such as lighting variations, occlusions, and image quality enhancements further improves system robustness. With these capabilities, the system significantly reduces manual effort, minimizes errors, and enhances overall operational efficiency.

D. Attendance Marking System

1) Introduction: In modern educational and workplace environments, the process of attendance marking plays a crucial role in ensuring accountability and efficiency. Traditional attendance-taking methods are often time-consuming and prone to errors. This section details the automated attendance marking system using facial recognition technology integrated with CCTV cameras. The system captures real-time video frames, detects and recognizes individuals, and records their attendance in an organized manner. The implementation aims to reduce manual intervention, minimize errors, and enhance overall efficiency.

2) Workflow of the Attendance Marking System: The attendance marking process is structured into several key steps to ensure accuracy and reliability. These steps include face detection, feature extraction, identity matching, and attendance logging.

a) Step 1: Face Detection from Live Video Feeds:

- The system continuously receives real-time video feeds from CCTV cameras installed in designated areas such as classrooms or offices.
- Face detection is performed using models like MTCNN (Multi-task Cascaded Convolutional Networks) or Haar Cascade Classifier, which identify and extract faces from the video frames.
- Detected faces are cropped and prepared for further processing.

Challenges:

- Multiple faces in a single frame may lead to overlapping detections.
- Environmental factors like poor lighting or motion blur may impact detection accuracy.

b) Step 2: Feature Extraction for Identification:

- The detected facial regions are processed through deep learning models to extract key facial features.
- The extracted features are converted into numerical embeddings that uniquely represent each individual.
- These embeddings are then compared with a pre-stored database of enrolled individuals for identification.

c) Step 3: Matching Detected Faces with Stored Data:

- The system compares the extracted face embeddings with the database using similarity metrics such as cosine similarity or Euclidean distance.
- If the similarity score surpasses a predefined threshold (e.g., 85% match confidence), the system successfully identifies the individual.
- If no match is found, the system either ignores the detection or marks it as an ”Unknown” entry.

Challenges:

- Facial variations due to different expressions, angles, or accessories can affect recognition accuracy.
- Individuals wearing masks or glasses may require additional processing techniques.

d) Step 4: Threshold-Based Decision for Attendance Logging:

- If the recognition confidence score is above the set threshold, the system logs the attendance of the identified individual after clicking the mark attendance button.
- A timestamp, including date and time, is recorded to ensure attendance is only marked once per session.

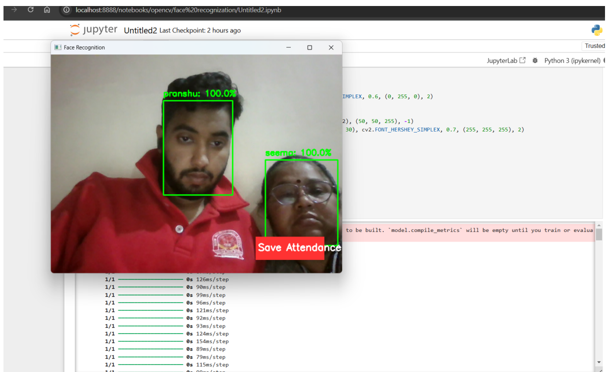

Fig. 1. faces recognized by model

Additional business rules:

- Time-based restrictions: Attendance is logged only within a defined time window.
- Minimum presence duration: Individuals must remain within the frame for a specified time before being marked present.

3) Features of the Attendance Marking System:

- **Real-Time Processing:** The system operates continuously, detecting and recognizing individuals in real time without requiring manual intervention.
- **Multi-Person Detection:** The model is capable of detecting and recognizing multiple individuals simultaneously in a single frame.
- **Automatic Re-Identification:** If an individual temporarily moves out of the camera's view, the system re-identifies them upon return.
- **Customizable Attendance Rules:** Organizations can customize attendance criteria based on their requirements.
- **Integration with Other Systems:** The system seamlessly integrates with institutional databases, payroll systems, and notification services.

4) Challenges and Solutions:

- **Handling Low-Quality Images:** Image enhancement techniques, such as noise reduction and resolution improvement, are implemented to improve clarity.
- **Occlusion and Partial Faces:** Training models with diverse datasets, including masked faces, and focusing on unique facial features like eyes and forehead.
- **Variable Lighting Conditions:** Adaptive contrast enhancement techniques, such as histogram equalization, are used to improve facial visibility.

5) Conclusion: The automated attendance marking system provides a seamless and efficient approach to attendance tracking using real-time face recognition. By integrating deep learning models, real-time feature extraction, and threshold-based decision-making, the system ensures accurate and efficient attendance logging.

E. Excel-Based Record Keeping

1) Data Storage Structure: Each attendance entry is logged with the following key attributes:

- **Student Name:** The name of the identified individual based on facial recognition.
- **Confidence Score:** A numerical value indicating the model's certainty in recognizing the face.
- **Timestamp:** The exact date and time when the attendance was marked.

2) Advantages of Excel-Based Record Keeping: Using Excel as the primary format for storing attendance data offers several benefits:

- **Ease of Access:** Administrators can open and review records without requiring specialized database software.
- **Sorting and Filtering:** Attendance data can be sorted by date, name, or confidence score for quick analysis.
- **Integration with Other Tools:** The data can be imported into visualization tools for generating attendance reports.
- **Backup and Security:** Excel files can be periodically backed up, and access can be restricted to authorized users.

3) Automation of Record Updates: The system automatically updates the Excel sheet in real time whenever a student's attendance is recorded. The automation process involves:

- **Opening the Attendance File:** The system loads the existing Excel file where attendance records are maintained.
- **Appending New Entries:** Each new recognition result is appended to the sheet.
- **Saving the File:** Once the new records are added, the updated file is saved to ensure data persistence.

4) Error Handling and Data Validation: To prevent incorrect entries or data corruption, the system implements error handling mechanisms:

- **Duplicate Check:** Ensuring the same student is not marked twice within a short time frame.
- **Data Type Validation:** Verifying that names are stored as text, confidence scores are numerical, and timestamps are in the correct format.
- **Exception Handling:** Detecting and handling errors such as missing or inaccessible files.

5) Future Enhancements: To further improve record-keeping capabilities, the following enhancements can be considered:

- **Cloud Integration:** Storing attendance records on cloud platforms for remote access.
- **Database Migration:** Transitioning from Excel to a relational database for enhanced querying and scalability.
- **Automated Analytics:** Implementing built-in attendance trend analysis and report generation.

F. User Interface & Functionalities

A well-designed user interface (UI) is crucial for ensuring ease of use, especially in an automated attendance system that may be operated by non-technical users. The system provides a simple and intuitive UI that allows administrators to perform key actions efficiently. The interface is web-based and developed using Django, ensuring accessibility from various devices.

1) Key Functionalities: The system's UI provides the following core functionalities:

1) **Start Attendance**
    - Users can initiate the attendance-taking process with a single button click.
    - The system activates the connected CCTV camera, captures real-time video frames, and applies face recognition to mark attendance.

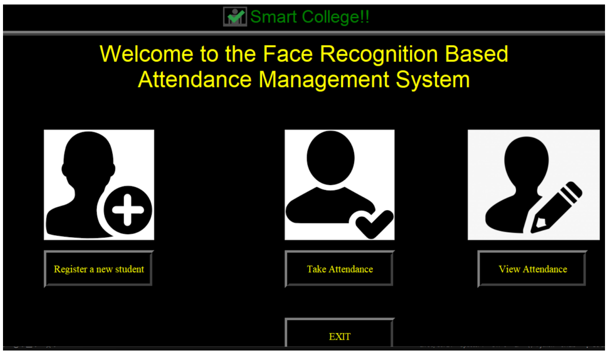

Fig. 2. User Interface

- A notification or success message confirms when the task is completed.

2) **Register New Students**
   - A dedicated section allows administrators to register new students.
   - Users can upload images of new students, which are then processed and added to the training dataset.
   - The system updates the recognition model accordingly to ensure accurate identification.

3) **View Attendance Records**
   - Users can view stored attendance logs in a structured tabular format.
   - The table displays student names, confidence scores, and timestamps for each recorded entry.
   - Filtering and sorting options allow easy retrieval of specific records.

2) UI Design Considerations: The UI is designed with the following principles in mind:

- **Simplicity:** A clean layout ensures ease of navigation for users with minimal technical knowledge.
- **Responsiveness:** The web-based interface adapts to different screen sizes, allowing use on desktops, tablets, and mobile devices.
- **Real-Time Updates:** Attendance results are displayed immediately after processing, minimizing delays.
- **Error Handling:** Clear error messages and validation checks prevent invalid data entry or system malfunctions.

3) Future Enhancements: To further improve user experience and functionality, potential enhancements include:

- **Mobile App Integration:** Developing a mobile-friendly version for on-the-go access.
- **Role-Based Access Control:** Implementing different user roles (e.g., administrators, teachers, students) with varying permissions.
- **Graphical Analytics Dashboard:** Adding visual representations of attendance trends for better data insights.

The user interface serves as the primary point of interaction between the system and administrators, ensuring that attendance management is both seamless and efficient.

G. Error Handling & Enhancements

A robust error-handling mechanism is essential for ensuring the reliability and accuracy of an AI-based attendance system. Since the system relies on real-time video processing, machine learning models, and database operations, various types of errors can occur, ranging from hardware failures to incorrect face recognition. To mitigate these issues, the system incorporates multiple error-handling techniques and planned enhancements.

1) Error Handling Mechanisms: To prevent system failures and improve usability, the following error-handling strategies have been implemented:

1) **Exception Handling for File Operations**
   - The system stores attendance records in an Excel file using the OpenPyXL library.
   - Proper exception handling ensures smooth file reading/writing, even if the file is missing or in use.
   - Error messages are displayed when an issue arises, preventing data corruption.

2) **Model Retraining & Data Update**
   - Over time, face recognition accuracy may decline due to changes in facial features (e.g., aging, new hairstyles, glasses).
   - A retraining option is included, allowing administrators to update the dataset and retrain the model for improved performance.
   - Users can add new images for training, ensuring the system remains adaptable.

3) **Multi-Model Support for Fault Tolerance**
   - The system supports three different face detection models:
     - Deep Neural Networks (DNN) for high-accuracy recognition.
     - MTCNN (Multi-task Cascaded Convolutional Network) for efficient detection and alignment.
     - Haar Cascade Classifier, a lightweight model for real-time processing.
   - If one model fails to recognize a face due to lighting conditions or occlusions, another model can be used as a fallback.

4) **Handling Camera & Hardware Issues**
   - If the CCTV camera malfunctions or is disconnected, the system provides an error message and logs the issue.

- Administrators receive alerts when the camera feed is unavailable, ensuring timely troubleshooting.
- The system can switch to an alternate camera if multiple devices are available.

5) **Network & Database Connectivity Issues**
   - If network connectivity is lost during data transmission, the system retries automatically to prevent data loss.
   - Database errors are handled with rollback mechanisms to maintain consistency in stored attendance records.

III. Results

The performance of the AI-based attendance system is evaluated using various metrics to assess the effectiveness of the face recognition model. The key evaluation components include the **Confusion Matrix**, **Training & Validation Loss and Accuracy Graph**, and the **ROC-AUC Curve**.

A. Confusion Matrix

The confusion matrix provides insight into the classification performance of the model by displaying the number of correctly and incorrectly identified faces. It helps in understanding false positives (incorrectly marked attendance) and false negatives (missed recognitions).

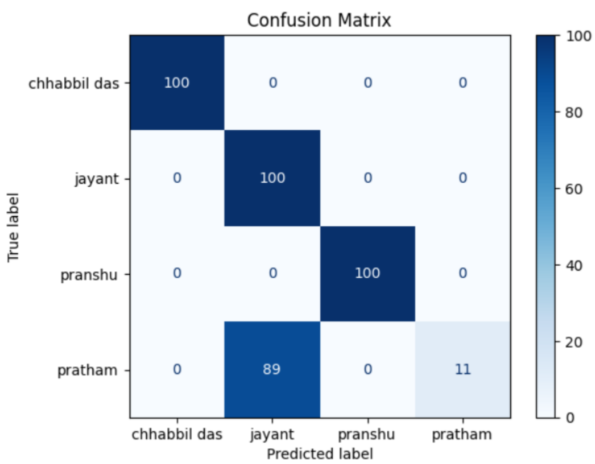

Fig. 3. Confusion Matrix

B. Training & Validation Loss and Accuracy Graph

The training and validation curves illustrate the model's learning process over epochs. These graphs help in analyzing:

- Whether the model is **overfitting** (training accuracy much higher than validation accuracy).
- Whether the model is **underfitting** (both training and validation accuracy are low).
- The effectiveness of the learning process through loss reduction and accuracy improvement over time.

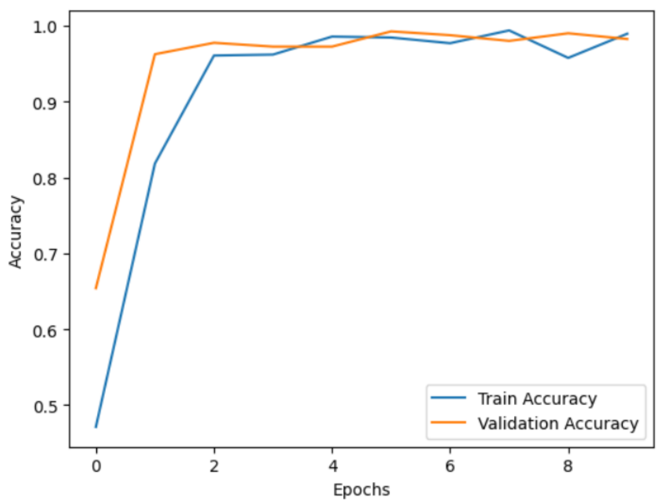

Fig. 4. Training & Validation Loss Curve of mtcnn

C. ROC-AUC Curve

The Receiver Operating Characteristic (ROC) curve and the **Area Under the Curve (AUC)** score help in evaluating the model's ability to distinguish between recognized and unrecognized faces. A higher AUC value (closer to 1) indicates better classification performance.

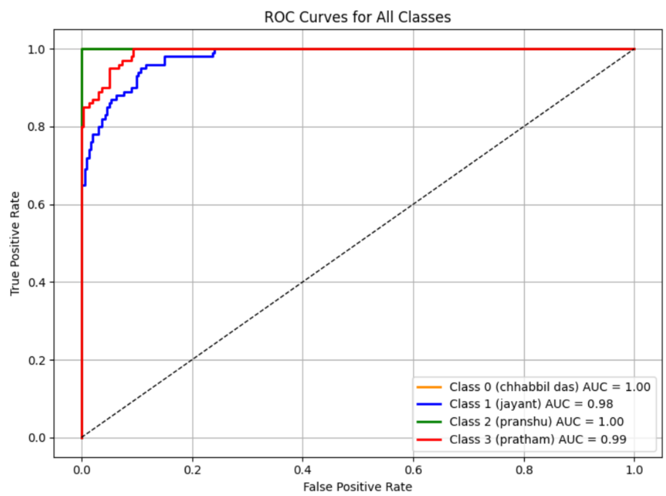

Fig. 5. ROC-AUC Curve

These evaluation metrics provide a comprehensive analysis of the system's effectiveness and highlight areas for further improvement. The results demonstrate the system's ability to automate attendance with high accuracy while maintaining robustness against variations in lighting, pose, and occlusion.

IV. Future Enhancements

To further improve system efficiency and reliability, the following enhancements are planned:

1) **RFID & Biometric Integration**

- Combining face recognition with RFID (Radio Frequency Identification) or fingerprint scanning for multi-factor authentication.
- Reducing false positives by requiring multiple verification methods.

2) **Cloud-Based Storage for Real-Time Remote Access**
   - Implementing cloud storage for attendance records, allowing administrators to access data from any location.
   - Integration with Google Drive, Firebase, or AWS for secure and scalable storage solutions.
3) **Liveness Detection to Prevent Spoofing**
   - Implementing liveness detection techniques to prevent fraudulent attendance marking using printed photos or videos.
   - Using depth analysis, blinking detection, and movement tracking to ensure real individuals are being recognized.
4) **Improved AI Algorithms for Faster & More Accurate Recognition**
   - Fine-tuning deep learning models for enhanced speed and precision.
   - Exploring transformer-based architectures (e.g., Vision Transformers - ViTs) for superior face recognition performance.

By continuously improving error handling and integrating advanced features, the system ensures reliability, security, and efficiency, making it a practical solution for educational institutions and workplaces.