B.Sc. in Computer Science and Engineering Thesis

Enhancing Workforce Attendance Systems in Garment Manufacturing Through a Comparative Analysis of Advanced Face Detection and Recognition Models

Submitted by

Mohammad Rahim

201905031

Kh Md Ibtihal

201905055

Supervised by

Dr. Md. Mostofa Akbar

Department of Computer Science and Engineering

Bangladesh University of Engineering and Technology

Dhaka, Bangladesh

March 2025

Candidates' Declaration

We hereby declare that the work presented in this thesis, titled *"Enhancing Workforce Attendance Systems in Garment Manufacturing Through a Comparative Analysis of Advanced Face Detection and Recognition Models"*, is the outcome of the research carried out by us under the supervision of Dr. Md. Mostofa Akbar. We confirm that neither this thesis nor any part of it has been submitted elsewhere for any degree or qualification.

Mohammad Rahim

201905031

Kh Md Ibtihal

201905055

ACKNOWLEDGEMENT

First and foremost, we express our heartfelt gratitude to Allah (SWT) for granting us the strength and perseverance to complete this work. We are deeply thankful to our supervisor, Associate Professor Dr. Md. Mostofa Akbar, for his invaluable guidance and constant support throughout this thesis. His insightful feedback, constructive criticism, and patient listening have been crucial in shaping this work.

We would also like to acknowledge the contributions of our seniors, Pritam Saha (201805102) and Yeasir Khandaker (201805105), whose previous research, *"Enhancing Workforce Management with OpenCV-Based Attendance System in Garment Manufacturing,"* provided valuable insights and inspiration for our work. Their study laid a strong foundation for exploring the applications of computer vision in practical scenarios, which has greatly influenced our approach.

Dhaka — Mohammad Rahim

March 2025 — Kh Md Ibtihal

Contents

List of Figures

List of Tables

Abstract

This paper presents a comparative analysis of advanced face detection and recognition models for developing an efficient workforce attendance system in the garment industry. Traditional attendance methods often suffer from inefficiencies, inaccuracies, and manual labour dependency, highlighting the need for a more robust and automated solution. This study evaluates both traditional machine learning and deep learning-based face detection and recognition approaches to assess their performance under real-world constraints. The analysis is conducted on a self-curated dataset comprising video clips and images collected from online sources, covering diverse conditions such as varying lighting, camera distance, and occlusions. Additionally, segmented evaluations on specific subsets, such as low-light scenarios, provide deeper insights into model effectiveness. The findings highlight the trade-offs between accuracy, computational efficiency, and adaptability, offering a comprehensive guideline for selecting the most suitable approach for large-scale attendance monitoring in garment manufacturing.

Chapter 1

Introduction

An effective workforce attendance system requires an accurate and efficient face recognition model capable of handling real-world challenges. Traditional attendance systems, such as those based on manual sign-ins, RFID, or fingerprint scanners, have proven inefficient and prone to inaccuracies. These systems often suffer from manual errors, fraud (e.g., buddy punching), and maintenance challenges, leading to inefficiencies and delays. The performance of modern attendance systems, however, depends on the choice of face detection and recognition algorithms, as different models vary in accuracy, speed, and robustness under challenging conditions such as lighting variations, occlusions, and varying camera distances. Face detection, the first step in the recognition pipeline, faces challenges like detecting faces in crowded environments, at different angles, and under poor lighting. Face recognition, while effective in controlled environments, struggles with issues like facial variations due to ageing, disguises, or facial hair, as well as handling multiple individuals in complex scenes. This study systematically evaluates and compares different models to determine the most suitable approach for large-scale workforce attendance tracking.

Face detection methods include traditional machine learning approaches like Haar Cascades and Histogram of Oriented Gradients (HOG) with Linear SVM, which are lightweight and efficient but often struggle in complex environments. In contrast, deep learning-based detectors such as Single Shot MultiBox Detector (SSD) MobileNet and Multi-task Cascaded Convolutional Networks (MTCNN) provide higher accuracy, especially in challenging scenarios, though at the cost of increased computational complexity.

For face recognition, this study compares various models implemented using dlib, including dlib's ResNet-based embeddings, FaceNet, and Python's face recognition library. While traditional models offer efficiency in controlled settings, deep learning-based models like FaceNet provide superior accuracy and adaptability to real-world environments with variable lighting and occlusions.

To further enhance system reliability, we introduce an error detection API that dynamically updates detected faces based on confidence scores, ensuring the most accurate embeddings are stored. This allows for improved verification by presenting maintainers with stored embeddings alongside detected faces, simplifying manual confirmation. This manual confirmation can also be used to identify unidentified faces. The study aims to provide a clear guideline for selecting the best face detection and recognition combination, optimizing accuracy, efficiency, and real-world applicability for workforce attendance in large-scale industrial environments.

1.1 Motivation

The rapid growth of industries reliant on large-scale workforce management, such as the garments and manufacturing sectors, highlights the necessity for accurate and efficient attendance tracking systems. Traditional methods of attendance monitoring, including manual logging and RFID-based systems, are increasingly proving inadequate due to issues like human error, security vulnerabilities, and environmental interference. These methods often result in payroll discrepancies, inefficiencies in workforce allocation, and reduced productivity, underscoring the urgent need for more advanced solutions.

Our motivation lies in addressing these challenges by exploring the potential of computer vision and machine learning, specifically focusing on face detection and face recognition technologies. Facial recognition offers a promising alternative to conventional attendance systems, providing a more reliable, contactless, and scalable approach to attendance management. The ability to accurately identify individuals in real-time, even in varied lighting conditions and crowded environments, is a key advantage over traditional systems.

With the proliferation of such technologies, we aim to conduct a comparative analysis of different face detection and recognition models to determine which offers the highest accuracy and efficiency. By evaluating models like Haar Cascades, HOG with Linear

SVM, and deep learning-based methods, we aim to identify the most suitable solution for real-world applications, ensuring that it can operate in diverse conditions without compromising performance.

In addition to improving accuracy, our research also seeks to explore the scalability, cost-effectiveness, and computational complexity of these models. A system capable of handling a large number of faces while remaining easily maintainable and scalable is crucial for deployment across various industries, from small factories to large enterprises.

Ultimately, our motivation is driven by the desire to revolutionize the way attendance is tracked in various sectors, especially in the garment industry, enhancing productivity and fairness through the application of cutting-edge technology. By developing a robust and efficient face recognition-based system, we aspire to contribute to improving operational workflows, reducing manual intervention, and ensuring a more streamlined and secure attendance management process.

1.2 Problem Definition

The garments sector, a key industry employing millions globally, faces persistent challenges in workforce attendance management. Traditional methods such as manual entry and RFID systems are widely used but suffer from significant drawbacks. Manual entry is prone to human errors and manipulation, leading to inaccurate records and payroll discrepancies. RFID systems, while an improvement, remain susceptible to misuse and require substantial infrastructure and maintenance.

Face recognition-based attendance systems offer a promising alternative, eliminating physical cards and manual tracking. However, their deployment in garment factories presents several challenges. The demanding industrial environment—characterized by high levels of dust, noise, and fluctuating lighting conditions—can significantly impact the accuracy of face detection and recognition models. Harsh lighting can lead to overexposed or underexposed images, while occlusions caused by masks, scarves, or head coverings commonly worn in factories can reduce recognition accuracy. Additionally, the high employee turnover and the need to process thousands of faces in real-time demand a system that is both scalable and computationally efficient.

Furthermore, balancing recognition speed and accuracy is crucial—delays in real-time identification can disrupt workflow, while high false positive or false negative rates can undermine trust in the system. Many existing models struggle to generalize well across varying environmental conditions, making it essential to evaluate and optimize different approaches for robustness.

To address these issues, this study focuses on enhancing workforce attendance systems in garment manufacturing through a comparative analysis of advanced face detection and recognition models. By evaluating different models under real-world conditions, we aim to develop a robust system capable of handling large-scale, real-time attendance tracking while ensuring high accuracy, adaptability to environmental factors, and efficient computational performance.

Chapter 2

Background and Related Works

2.1 Face Detection and Recognition

Face detection and recognition are critical components of automated attendance systems. Face detection involves identifying and locating human faces in images or video frames, while face recognition extends this process by verifying or identifying individuals based on their facial features.

Early face detection methods include Haar Cascades and Histogram of Oriented Gradients (HOG). Haar Cascades, introduced by Viola and Jones, employ a cascade of classifiers trained using AdaBoost to detect faces efficiently. HOG, another classical approach, extracts gradient-based features and applies a classifier for face detection. While effective in constrained environments, these methods struggle with variations in lighting, pose, and occlusion.

Traditional feature extraction techniques for face recognition include Principal Component Analysis (PCA) used in Eigenfaces and Linear Discriminant Analysis (LDA) used in Fisherfaces. These techniques focus on reducing dimensionality while preserving essential facial features. However, they are highly sensitive to environmental changes, particularly lighting and pose variations, making them unsuitable for dynamic industrial settings. A foundational work in this area is by Turk and Pentland [1], who introduced the concept of Eigenfaces for face recognition, demonstrating a practical application of PCA.

The emergence of deep learning has revolutionized face detection and recognition systems, providing improved robustness and accuracy under varying conditions. Convolutional Neural Networks (CNNs) have become the state-of-the-art, with methods such as Multi-task Cascaded Convolutional Networks (MTCNN) and Single Shot MultiBox Detector (SSD) achieving high precision in face detection. The evaluation of Haar Cascade Classifiers for face detection has also been explored by Padilla et al. [4], emphasizing its effectiveness in constrained environments.

Modern face recognition relies on deep learning-based feature extraction, where CNNs learn hierarchical facial representations. Notable models include VGG-Face, FaceNet, and ArcFace, which generate high-dimensional embeddings that encode unique facial characteristics. Schroff et al.[2] introduced FaceNet, a unified deep learning framework that uses deep metric learning with triplet loss to achieve state-of-the-art performance in face recognition and clustering. Additionally, Zhang et al.[3] demonstrated the implementation of Dlib's deep learning face recognition technology, further enhancing the robustness of face recognition systems.

These embeddings are typically compared using distance metrics such as Euclidean distance or cosine similarity. Furthermore, deep metric learning techniques employing contrastive loss and triplet loss have enhanced the ability to differentiate between individuals effectively. Such advancements have led to more robust and scalable systems applicable to real-world scenarios.

2.2 Literature Review on Face Recognition-Based Attendance Systems

Several studies have explored face recognition-based attendance systems across various domains. Lukas et al. [5] implemented an Eigenface-based recognition system within a classroom setting, demonstrating feasibility but encountering limitations related to pose and lighting variations. The system's reliance on PCA made it sensitive to environmental changes, which may not be suitable for dynamic industrial environments.

Sawhney et al. [6] developed a real-time attendance system that integrates face recognition with video processing. While the system demonstrated good performance in controlled environments, it required high computational resources to achieve accuracy, making it less ideal for large-scale deployment.

Wagh et al. [7] proposed a PCA-based recognition system designed for classrooms but encountered difficulties in handling real-world variations such as illumination changes and occlusion. These limitations highlight the inadequacies of classical methods when applied to dynamic environments like garment manufacturing.

Yang and Han [9] introduced a real-time face recognition attendance system optimized for video-based tracking. While their approach improved processing efficiency, it was tailored primarily for small-scale environments and failed to address scalability concerns in large industrial settings.

Wheeler et al. [8] aimed to enhance face recognition at a distance for surveillance applications. Their work provided valuable insights into dealing with challenging conditions, such as low resolution and varying illumination, which are relevant when considering uncontrolled factory environments.

The studies reviewed demonstrate a progression from classical approaches like PCA to deep learning-based methods. However, most research has been conducted in controlled or small-scale environments, with limited focus on industrial-scale deployments. Additionally, the reliance on high computational resources remains a significant challenge.

Chapter 3

Model Utilization & Theoretical Analysis

3.1 Models and Their Role in the Study

In this study, four different face recognition models were used to detect and recognize faces. Each model consists of a face detection component and a face recognition method. These models help in identifying individuals by extracting unique facial features and comparing them to a stored database. The goal of using multiple models was to evaluate their effectiveness in different conditions, such as varying lighting, angles, and image quality.

Face recognition is widely used in security, surveillance, and authentication systems. By analyzing the effectiveness of different models, it is possible to understand their strengths and weaknesses, which can guide their practical application. The models chosen in this study represent a mix of traditional and deep learning approaches, providing insight into the trade-offs between computational efficiency and recognition accuracy.

3.2 Principles of the Models

Each model works based on different principles of face detection and recognition. Below is a detailed explanation of each model.

3.2.1 Haar Cascade with Face Recognition

Haar Cascade detects faces using edge and texture-based features. It applies a cascade of classifiers trained on different facial features, allowing it to identify faces quickly. However, it is sensitive to variations in lighting and pose, which can impact accuracy. The recognition component utilizes dlib's face encoding method, which converts facial features into numerical vectors for comparison.

One of the main advantages of this model is its speed, making it suitable for real-time applications with limited computational power. However, its reliance on simple feature detection makes it prone to false positives and false negatives, especially in images with complex backgrounds or poor lighting conditions.

3.2.2 HOG with Linear SVM and Dlib ResNet

This model detects faces using the Histogram of Oriented Gradients (HOG) method, which analyzes gradient-based features in an image. A Support Vector Machine (SVM) classifier is then used to differentiate faces from non-faces. The face recognition component uses dlib's ResNet-based face descriptor, which maps faces into a lower-dimensional space to improve classification accuracy.

HOG-based detection is robust against variations in illumination and minor pose changes. However, it struggles with occlusions and extreme angles. The combination of HOG and SVM offers a balance between accuracy and efficiency, making it a suitable choice for applications that require reliable face detection without deep learning overhead.

3.2.3 SSD with MobileNet and Dlib ResNet

SSD (Single Shot MultiBox Detector) with MobileNet is a deep learning-based method that detects faces in images efficiently. SSD predicts multiple bounding boxes in a single forward pass, making it suitable for real-time applications. MobileNet is a lightweight model optimized for mobile and embedded devices. Dlib's ResNet-based face embedding technique is used for recognition, ensuring high accuracy in distinguishing faces.

This model offers fast and accurate face detection, making it useful for applications where speed is critical, such as surveillance systems. However, it may struggle with very small faces or highly cluttered backgrounds. The MobileNet backbone ensures efficiency, but its performance may decline when dealing with extreme facial variations or partial occlusions.

3.2.4 MTCNN with FaceNet

MTCNN (Multi-task Cascaded Convolutional Networks) applies a cascade structure for face detection, refining the detection process through multiple stages. This method improves accuracy by progressively filtering non-face regions. FaceNet is used for recognition, utilizing deep neural networks with triplet loss optimization to generate compact and discriminative face embeddings. This approach results in high recognition accuracy but is computationally expensive.

MTCNN provides superior detection performance, particularly in challenging conditions such as varying lighting, poses, and partial occlusions. FaceNet further enhances recognition accuracy, making this model one of the most reliable but resource-intensive choices.

3.3 Application and Limitations

Each model has its own advantages and limitations.

3.3.1 Haar Cascade with Face Recognition

Haar Cascade is efficient for simple face detection tasks and requires minimal computational resources. However, it struggles with lighting variations, occlusions, and side profiles, leading to false negatives. The recognition process is effective but depends on the quality and diversity of training images. Its limited ability to detect faces in complex environments makes it less suitable for real-world scenarios where lighting and angles are inconsistent.

3.3.2 HOG with Linear SVM and Dlib ResNet

HOG with Linear SVM performs well in structured environments with clear facial features. However, it may fail in complex backgrounds or with extreme variations in facial expressions. The dlib ResNet-based face embedding enhances recognition but can be affected by image noise and misalignment. This model is more reliable than Haar Cascade but may still struggle in highly dynamic settings where faces appear in various orientations.

3.3.3 SSD with MobileNet and Dlib ResNet

SSD with MobileNet provides fast and efficient face detection, making it ideal for real-time applications. However, it may miss small or partially obscured faces. The recognition process is robust due to ResNet-based embeddings, but performance can degrade if faces are detected inaccurately. This model is best suited for applications where speed is essential, but accuracy may suffer under challenging conditions.

3.3.4 MTCNN with FaceNet

MTCNN offers high accuracy in face detection by refining results through multiple stages. However, it requires more computational power and can be slower than other models. FaceNet provides excellent recognition performance, but its sensitivity to low-quality images and occlusions can reduce accuracy in challenging conditions. Despite its computational demands, this model is highly effective in security applications where precision is a priority.

Understanding these limitations helps in selecting the most suitable model based on the application requirements. By comparing these models, it is possible to determine which approach works best under different conditions, balancing speed, accuracy, and computational efficiency. In real-world implementations, a trade-off between accuracy and efficiency must be carefully considered to achieve optimal performance in face recognition systems.

Chapter 4

Model Testing & Performance Evaluation

4.1 Experimental Setup & Test Scenarios

4.1.1 Setup Overview and Model Selection

This study evaluates various face detection and recognition models to assess their performance in a workforce attendance system. The models selected for testing were chosen based on compatibility with one another, without a specific reason behind their combination. Both traditional and deep learning-based methods were tested to gauge their effectiveness under real-world conditions typical of garment manufacturing environments.

The dataset used for this evaluation was custom-built, consisting of facial images extracted from publicly available YouTube and Google videos. Although relatively small, the dataset was designed to simulate challenges faced in manufacturing settings. All images underwent preprocessing, where resizing was applied to maintain consistency across the models.

4.1.2 Testing Conditions and Scenario Design

To ensure the models' robustness, they were exposed to different conditions, reflecting both controlled and uncontrolled environments. Two primary factors were considered in this testing: lighting conditions and distance from the camera. The dataset was divided into subsets that represent well-lit and poorly-lit environments, as well as different distances (close-range and far-range faces).

Though real-time attendance tracking was not implemented in this study, models were evaluated using static images and video frames. One critical factor considered was the models' ability to handle multiple faces in a frame, an essential feature for workforce attendance systems. While specific variables such as facial expressions, pose, and accessories (masks, glasses, etc.) were not individually tested, the overall recognition accuracy was used as an indirect measure of how well the models perform under these variations.

4.2 Performance Metrics & Evaluation

In this study, four key performance metrics were used to evaluate the models: **Accuracy**, **Precision**, **F1-score**, and **Processing Time**. These metrics were applied across both face detection and recognition tasks, without separating their individual contributions. The results of these evaluations are presented in a series of tables to provide a comprehensive overview of each model's performance.

4.2.1 Overall Evaluation Results

The first table presents the overall results for all models, summarizing the performance metrics (accuracy, precision, F1-score, and processing time) across all testing conditions.

Table 4.1: Overall Results for All Models

Model	**Accuracy (%)**	**Precision (%)**	**F1-Score (%)**	**ETPF (ms)**
Model 1	75	66	78	199
Model 2	69	56	65	291
Model 3	73	62	72	275
Model 4	80	72	82	483

Model 1 → Haar Cascades & Python's Face Recognition

Model 2 → HOG with Linear SVM & Dlib's ResNet

Model 3 → SSD with MobileNet & Dlib's ResNet

Model 4 → MTCNN & Dlib's FaceNet

ETPF → Elapsed Time per Frame

4.2.2 Scenario-Specific Performance

To assess how each model performed under different conditions, the following tables present the results for specific scenarios:

Table 4.2: Results Under Light and Dark Conditions

Model	Accuracy (%)	Precision (%)	F1-Score (%)	ETPF (ms)
Model 1	77 / 69	66 / 54	75 / 60	229 / 184
Model 2	70 / 66	57 / 54	66 / 61	213 / 242
Model 3	74 / 70	62 / 56	70 / 67	224 / 260
Model 4	83 / 80	74 / 70	84 / 75	473 / 511

Table 4.3: Results for Close and Far Range Faces

Model	Accuracy (%)	Precision (%)	F1-Score (%)	ETPF (ms)
Model 1	76 / 52	62 / 48	68 / 53	288/298
Model 2	69 / 55	66 / 51	72 / 61	318/241
Model 3	74 / 65	63 / 56	70 / 65	246/205
Model 4	86 / 77	74 / 71	82 / 76	533/459

These scenario-specific tables allow for a more detailed comparison of the models' strengths and weaknesses in varying real-world conditions, such as different lighting and distances.

4.3 Results & Discussion

The evaluation results provide valuable insights into the performance of the tested models under various conditions. The overall results presented in Table 1 show the general capabilities of each model across all testing scenarios, with Model 4 emerging as the highest-performing model. It achieved the best scores for accuracy (80%), precision (72%), and F1-score (82%), though it also exhibited the highest processing time (483 ms per frame). This indicates that while Model 4 offers robust performance, its computational demands may be a limiting factor in certain real-time applications.

4.3.1 Analysis of Scenario-Specific Performance

The results under light and dark conditions, as shown in Table 2, reveal notable variations in performance. In well-lit environments, Model 4 consistently outperforms the other models in terms of accuracy, precision, and F1-score. For instance, the accuracy in light conditions was 83%, whereas in dark conditions, it still maintained a strong performance at 80%. On the other hand, Model 1 and Model 2 show a more significant drop in accuracy when transitioning from light to dark settings, with Model 1 showing a decrease from 77% to 69%, and Model 2 from 70% to 66%. This suggests that models relying on more traditional methods may struggle in poorly lit environments, highlighting the importance of robust model design for varied lighting conditions.

When considering close-range and far-range faces (Table 3), the performance differences become even more evident. Model 4 continues to demonstrate a high degree of reliability, maintaining an accuracy of 86% for close-range faces and 77% for far-range faces. This suggests that Model 4 is better equipped to handle variable distances between the camera and subject. In contrast, Model 1 and Model 2 show a more significant decline in accuracy when faces are farther from the camera, with Model 1 experiencing a drop from 76% at close range to just 52% at far range. This indicates that some models may struggle with recognizing faces at a distance, which is a critical factor in real-world applications where individuals may not always be positioned at a fixed distance from the camera.

4.3.2 Insights on Model Choice for Real-World Applications

Based on these results, it is clear that deep learning-based models, like Model 4, generally perform better in challenging environments (i.e., varying lighting and distance). These models tend to have higher accuracy, precision, and F1-scores, especially under less-than-ideal conditions. However, the increased processing time of these models must be considered in scenarios where real-time attendance tracking is crucial, as delays in processing could hinder overall system performance.

Models such as Model 1 and Model 2, while potentially more efficient in terms of processing time, demonstrate lower overall performance in terms of accuracy, particularly under difficult conditions like low lighting or at greater distances. Therefore, they may be better suited for environments where lighting and distance remain relatively consistent, or where computational resources are more limited.

Chapter 5

Proposed Face Recognition-Based Attendance System

5.1 System Overview

The proposed system modernizes workforce attendance tracking in garment manufacturing by leveraging facial recognition technology. Unlike traditional methods (e.g., manual sign-ins, RFID, fingerprint scanning), it offers a contactless, automated, and efficient alternative.

Cameras positioned at factory entry and exit points continuously capture images and videos for real-time face detection and recognition. The system employs preprocessing techniques such as grayscale conversion and normalization to ensure robustness across varying lighting and distances. Multiple detection and recognition models are evaluated to determine the optimal combination for industrial environments.

To enhance accuracy, a manual confirmation API enables administrators to verify mismatches and refine employee embeddings, reducing false positives and improving long-term system performance. This scalable solution minimizes administrative overhead while ensuring reliable attendance tracking.

5.2 Face Detection and Recognition Process

5.2.1 Image Acquisition & Preprocessing

Cameras at factory entry and exit points continuously capture images or video frames, eliminating manual attendance tracking. Captured images undergo preprocessing to optimize recognition accuracy:

- **Grayscale Conversion:** Reduces computational complexity.
- **Normalization:** Adjusts lighting variations.
- **Resizing:** Ensures uniform processing across models.

5.2.2 Face Detection & Recognition

The system integrates multiple detection models to enhance accuracy:

- **Haar Cascade:** Efficient for detecting facial regions.
- **HOG & SVM:** Gradient-based feature extraction with classification.
- **SSD MobileNet:** Deep learning-based real-time detection.
- **MTCNN:** Multi-stage face detection for better accuracy.

For recognition, the system evaluates:

- **HOG + CNN (Python's `face_recognition`):** Lightweight yet effective.
- **Dlib's ResNet:** High-accuracy feature extraction.
- **FaceNet:** Deep metric learning for robust recognition.

5.3 Database Management & Model Training

5.3.1 Database Management

The system stores employee feature vectors rather than raw images to optimize storage and retrieval. Key functionalities include:

- **Efficient Storage:** Encoded embeddings for quick lookup.
- **Real-Time Retrieval:** Fast matching against stored data.
- **Continuous Updates:** New employees and refined embeddings.

5.3.2 Model Training & Storage

Trained models are managed using Python's Pickle library, enabling seamless storage and retrieval while ensuring scalability.

5.4 Manual Confirmation API

5.4.1 Mismatch Detection and Confidence Score

The system utilizes a confidence score to assess recognition accuracy:

- **High Confidence Score (e.g., 0.87 and above):** The recognized face closely matches the stored face.
- **Low Confidence Score:** Indicates potential misidentification, requiring manual review.

5.4.2 Side-by-Side Face Verification

To facilitate manual verification, the system presents a **side-by-side comparison** of the stored and recognized faces, as shown in Figure 5.1. The administrator can take one of the following actions:

- **Confirm:** Accept the match, verifying the recognition.
- **Confirm & Add to Database:** Accept the match and update the database with the new face image.
- **Mismatch:** Reject the recognition, preventing incorrect attendance logging.

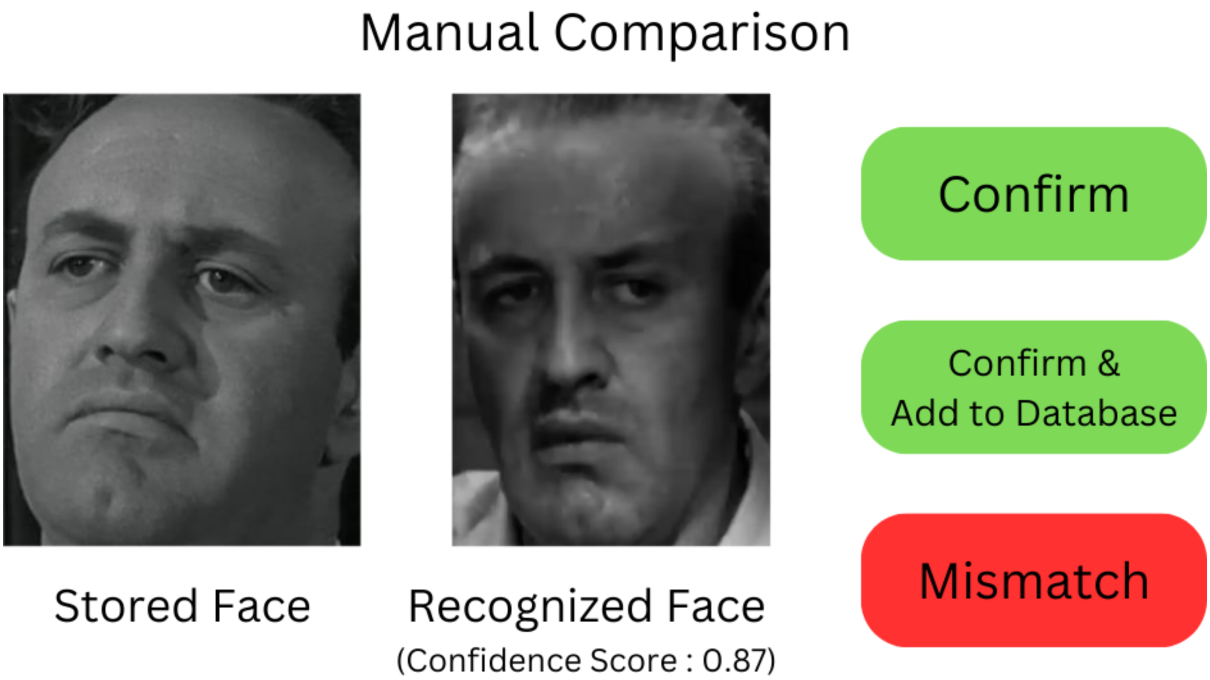

Figure 5.1: Manual comparison interface for verification of recognized faces.

5.4.3 Database Enhancement and Learning

The system continuously improves by integrating feedback:

- **Verified matches update embeddings** to refine accuracy.
- **Newly confirmed faces** are added to expand the dataset.
- **Rejected mismatches** help prevent recurring errors.

This approach balances automation with human oversight, ensuring a **reliable and adaptable attendance system**.

5.4.4 System Workflow

The system combines automated recognition with manual verification for accuracy and efficiency.

5.4.5 Face Detection & Recognition

Cameras capture employee faces, preprocess images, and detect faces using classifiers. Recognized faces are compared against stored embeddings, and attendance is logged if the confidence score meets a predefined threshold.

5.4.6 Manual Confirmation & Database Updates

If recognition confidence is low, the system flags the entry for manual review. Administrators can:

- **Confirm the match** (attendance is recorded).
- **Correct misidentifications** (update the database).
- **Reject unrecognized faces** (prevent unauthorized access).

Corrections refine stored embeddings, enhancing future recognition accuracy.

5.4.7 Attendance Recording & Reporting

Attendance records, including timestamps and confidence scores, are securely stored and exportable for workforce management.

Chapter 6

Conclusion and Future Work

This study compared various face detection and recognition models for enhancing workforce attendance systems in garment manufacturing. The results indicated that deep learning-based models generally outperformed traditional approaches, especially under challenging conditions like poor lighting or distant faces. However, issues such as real-time processing and variations in facial expressions or accessories still present challenges.

Looking forward, future work should focus on improving model robustness to factors such as pose variations, facial accessories, and lighting conditions. Expanding the dataset to include a wider range of scenarios and incorporating advanced models like hybrid biometric systems could further improve accuracy and reliability. Additionally, real-world testing in dynamic, real-time environments is essential to assess practical performance. Ethical concerns, such as privacy issues related to face recognition, should also be addressed to ensure the technology is implemented responsibly in workplace settings.

Bibliography

[1] M. Turk and A. Pentland, "Eigenfaces for recognition," *Journal of Cognitive Neuroscience*, vol. 3, no. 1, pp. 71–86, 1991.

[2] F. Schroff, D. Kalenichenko, and J. Philbin, "FaceNet: A unified embedding for face recognition and clustering," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2015.

[3] D. Zhang, J. Li, and Z. Shan, "Implementation of Dlib Deep Learning Face Recognition Technology," *2020 International Conference on Robots and Intelligent System (ICRIS)*, 2020.

[4] R. Padilla, C. F. F. Costa Filho, and M. G. F. Costa, "Evaluation of Haar Cascade Classifiers Designed for Face Detection," *International Journal of Computer and Information Engineering*, 2012.

[5] S. Lukas, A. R. Mitra, R. I. Desanti, and D. Krisnadi, "Student Attendance System in Classroom Using Face Recognition Technique," *2016 International Conference on Information and Communication Technology Convergence (ICTC)*, Jeju, Korea (South), 2016, pp. 1032-1035.

[6] S. Sawhney, K. Kacker, S. Jain, S. N. Singh, and R. Garg, "Real-Time Smart Attendance System using Face Recognition Techniques," *2019 9th International Conference on Cloud Computing, Data Science and Engineering (Confluence)*, Noida, India, 2019, pp. 522-525.

[7] P. Wagh, R. Thakare, J. Chaudhari, and S. Patil, "Attendance System Based on Face Recognition Using Eigenface and PCA Algorithms," *2015 International Conference on Green Computing and Internet of Things (ICGCIoT)*, Greater Noida, India, 2015, pp. 303-308.

[8] F. W. Wheeler, R. L. Weiss, and P. H. Tu, "Face Recognition at a Distance System for Surveillance Applications," *2010 Fourth IEEE International Conference on Biometrics: Theory, Applications and Systems (BTAS)*, Washington, DC, USA, 2010, pp. 1-8.

[9] H. Yang and X. Han, "Retracted: Face Recognition Attendance System Based on Real-Time Video Processing," *IEEE Access*, vol. 8, pp. 159143-159150, 2020.