Educational Smart Attendance System Using Face Recognition

Sahar Kamal [1], Mohamed Amr Madkour[2], Mohamed Salah Eldin Abdel Qader [3]

ABSTRACT

Abstract— In recent years, advancements in artificial intelligence and machine learning have significantly enhanced the accuracy and efficiency of face recognition technologies. This research explores the development of a Face Recognition Smart Attendance System leveraging deep transfer learning to optimize performance while reducing computational costs. The proposed system integrates pre-trained deep learning models to harness their robust feature extraction capabilities and applies fine-tuning techniques to adapt these models to the specific domain of attendance management. The system's architecture is designed to ensure high recognition accuracy under varying conditions, including changes in lighting, pose, and facial expressions. It features real-time detection and authentication, with a focus on scalability for deployment in educational institutions, corporate environments, and public events. Performance metrics such as precision, recall, and recognition speed are evaluated using a publicly available facial dataset and a custom dataset created for this study. Our results demonstrate the effectiveness of deep transfer learning in reducing training time while maintaining superior accuracy compared to conventional deep learning methods. Additionally, the system includes robust privacy and security measures to protect user data. This research contributes to the growing field of smart attendance solutions and showcases the potential of transfer learning in practical, real-world applications.

Keywords: Face Recognition, Smart Attendance System, Deep Transfer Learning, Artificial Intelligence.

I. INTRODUCTION

Organizations require attendance management systems to efficiently record staff attendance, whether manually or automatically. In educational settings, tracking students' daily attendance is crucial for performance assessment and quality assurance. Traditional methods, such as roll-calling or signing attendance sheets, are widely used but are often time-consuming and lack security [1]. Conversely, many automated human identification systems rely on conventional approaches like fingerprint scans, passwords, or ID cards. While these methods are commonly employed, they present significant drawbacks, including forgotten passwords and misplaced ID cards. A smart face recognition system offers a more reliable and secure alternative, capable of ensuring data integrity and maintaining historical records [2]. As a rapidly evolving field, face recognition technology has become integral to enhancing security due to its accuracy in identifying and verifying individuals [3] [4].

Tra Facial recognition technology is used to identify or verify an individual's identity from either a digital image or continuous frames from a video source. Face recognition systems operate through various methods, primarily by comparing facial data stored in a database with the facial features captured during recognition [5].

[1] Communication and Computers Engineering Department Al Shorouk Academy, Cairo, Egypt

[2] Communication and Computers Engineering Department Al Shorouk Academy, Cairo, Egypt

[3] Communication and Computers Engineering Department Al Shorouk Academy, Cairo, Egypt

This approach serves as one of the fastest and most intelligent time and attendance management solutions, leveraging facial recognition [6] to efficiently track and display attendance records. Additionally, face recognition [6] is quicker than other techniques and effectively minimizes the risk of attendance fraud's learning is a machine learning approach in which a model designed for a specific task is repurposed and adapted for a new task, serving as a starting point for further modifications. In deep learning, it is commonly employed as a pre-trained model in fields such as computer vision and natural language processing to build neural network models for these applications [7].

This technique is particularly valuable in deep learning, where real-world problems often involve massive datasets with billions of labeled samples, necessitating complex models [6]. Transfer learning optimizes the training process, saving time and enhancing performance. It enables developers to integrate multiple applications into a single framework and efficiently train new models for intricate tasks. Additionally, transfer learning is highly effective for boosting the accuracy of computer vision models [7].

II. RELATED WORK

Numerous studies have focused on the development of automated attendance systems in educational settings using facial recognition and deep learning techniques. These systems aim to enhance accuracy, reduce human intervention, and monitor student behavior in the classroom.

Arsenovic et al. [8] introduced a face recognition-based attendance system utilizing advanced deep learning techniques. Their approach consists of a sequence of key components, including a CNN-based cascade for accurate face detection and a CNN model for generating robust facial embeddings. The system was evaluated on a limited dataset comprising real-world facial images of employees, achieving an overall accuracy of 95.02%. Despite being tested on a small-scale dataset, the results indicate that the model holds strong potential for broader applications in various automated attendance and identity verification systems.

F U et al. [9] proposed an automated classroom attendance system that integrates two deep learning algorithms: the Multi-Task Cascaded Convolutional Neural Network (MTCNN) for face detection and the Centre-Face algorithm for face recognition. Their system is capable of identifying and reporting three types of attendance violations: absence, lateness, and early departure. Based on extensive experimental results, the system demonstrated real-time face recognition capabilities, requiring only 100 milliseconds per frame while maintaining high accuracy. The model achieved an accuracy rate of 98.87%, with a true positive rate below 1 in 1,000 and a false positive rate of 93.7%. Moreover, attendance records and learning statuses are automatically compiled and stored after each class session. Zulfiqar et al. [10] developed a face recognition system based on convolutional neural networks (CNNs), where the Viola-Jones algorithm [11] was employed for initial face detection. Following detection, facial features were extracted using a pre-trained CNN model. To improve recognition performance, the authors curated a large dataset of facial images, applying augmentation techniques to simulate various lighting and noise conditions. Their approach included the optimization of hyperparameters and selection of a robust pre-trained CNN architecture for deep face recognition. The system demonstrated strong performance in biometric authentication tasks, achieving an overall accuracy of 98.76%, thus validating the effectiveness of CNN-based approaches in real-world attendance applications.

III.PROPOSED MODEL

The proposed automated attendance management system relies on a facial recognition algorithm. This involves capturing multiple photos of the student using a camera installed in the classroom at different times to identify the faces of the students present. The proposed approach involves multiple stages: data collection, preprocessing, augmentation, CNN training and validation, followed by system testing. Dataset of students is stored in database and attendance reports is generate by web application and mobile application for students and instructor.

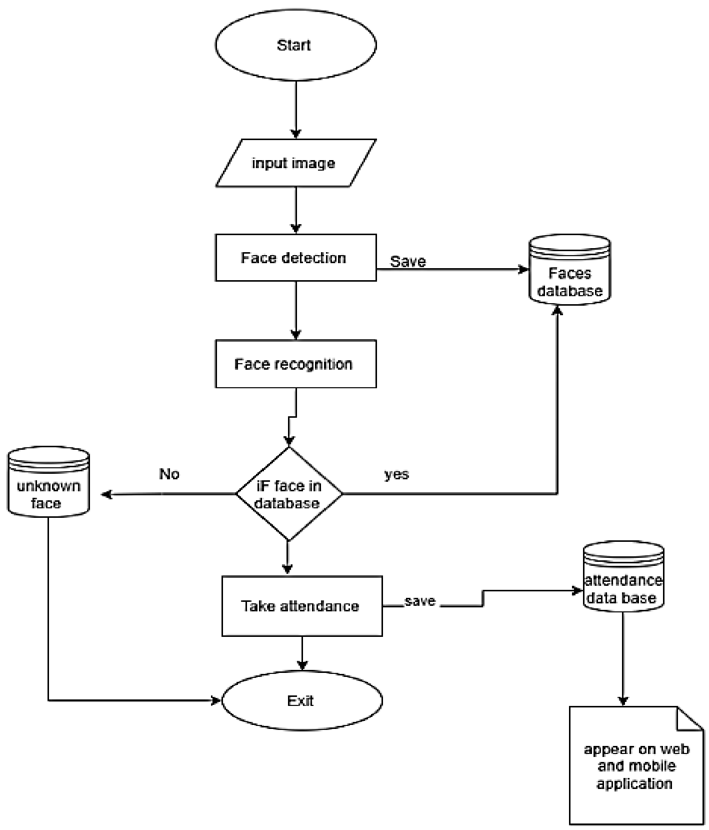


Figure. 1. Proposed model

A. Data collection

Our dataset comprises 300 images captured using the front-facing camera of an iPhone 12, which features a 12-megapixel, f/2.2 lens. The images are organized into 15 classes, each representing a different individual with 20 images per class. These 15 individuals include both male and female subjects, and save images for each student in a folder named with student ID as illustrated in Fig. 2.

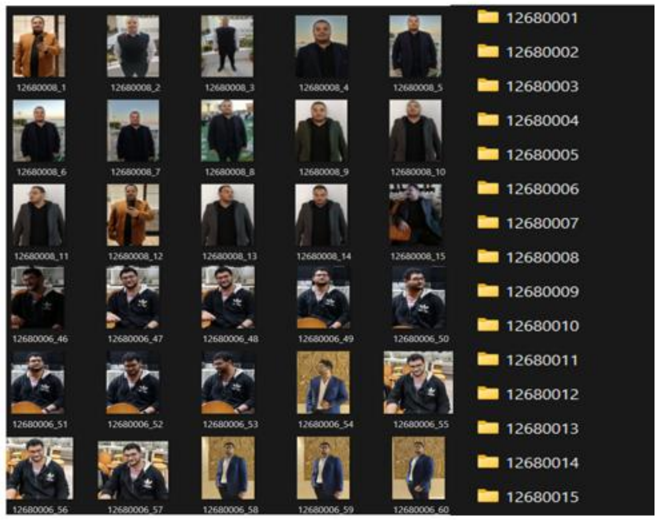


Figure. 2. Dataset

B. Data Formatting

The collected data is stored in JPG format, with image sizes ranging from 3.00 MB to 4.00 MB. Since deep learning algorithm requires a specific input size, the images were resized accordingly: 160 × 160 as in Figure. 3

Figure. 3. resize image

C. Data Augmentation

Data augmentation is a technique used to expand the dataset by creating modified versions of existing data or generating synthetic data based on it. This process helps regularize the model and reduce overfitting during training. In deep learning, common augmentation methods include geometric transformations such as flipping, color adjustments, cropping, rotation, noise injection, and random erasing [12].

In our trained networks, data augmentation was applied by capturing multiple images from various angles, environments, lighting conditions, orientations, and positions. After importing the data into the network, we implemented four specific augmentation techniques: rotation, scaling, cropping and, noise injection.

D. spilt dataset

Split images into 80% for training and 20% for testing as in Figure. 4

```
Class 12680001: 24 images (19 train, 5 val)
Class 12680002: 16 images (12 train, 4 val)
Class 12680003: 16 images (12 train, 4 val)
Class 12680004: 22 images (17 train, 5 val)
Class 12680005: 20 images (16 train, 4 val)
Class 12680006: 60 images (48 train, 12 val)
Class 12680007: 61 images (48 train, 13 val)
Class 12680008: 65 images (52 train, 13 val)
Class 12680009: 24 images (19 train, 5 val)
Class 12680010: 24 images (19 train, 5 val)
Class 12680011: 33 images (26 train, 7 val)
Class 12680012: 25 images (20 train, 5 val)
Class 12680013: 45 images (36 train, 9 val)
Class 12680014: 22 images (17 train, 5 val)
Class 12680015: 24 images (19 train, 5 val)
Splitting completed successfully!
```

Figure. 4. Dataset spilt

E. Face detection

In this paper, we propose a robust, real-time automated student attendance system designed for educational environments. The system integrates YOLOv11 and Face Net for precise face detection and for fast and accurate face recognition. This two-stage approach ensures both high detection reliability and real-time recognition performance, making it suitable for classroom and remote learning contexts [13].

This section outlines the design and implementation of a real-time face detection module based on **YOLOv11** as part of a smart classroom attendance system. The objective is to automate the student attendance process by detecting faces in live classroom footage efficiently and accurately under real-world conditions, including variable lighting, student pose variation, and partial occlusions. The proposed system consists of three main components: Video Input Stream from a classroom-installed camera. Face Detection Module using YOLOv11 [14]. Face Recognition Module (e.g., FaceNet, described in a separate section). YOLOv11 is selected for its advanced object detection capabilities and real-time performance. Unlike traditional face detection methods, YOLOv11 processes the entire image in a single forward pass, identifying multiple faces with high precision and minimal latency [15].

1. **Detection Pipeline**
   - **Preprocessing**:
     - Input frames are resized to 640×640 resolution.
     - Images are normalized and batched before feeding into the model.
   - **Face Detection using YOLOv11:**
     - The model outputs bounding boxes with class labels and confidence scores.
     - Non-Maximum Suppression (NMS) is applied to remove duplicate or overlapping detections.
   - **Face Cropping:**
     - Detected face regions are extracted and passed to the recognition module for identity verification.
   - **Integration with Attendance Module:**
     - Once recognized, student ID is matched with the database.
     - Timestamped attendance is recorded automatic

A YOLOv11 model pre-trained on the COCO dataset [16] is fine-tuned using a custom dataset of annotated classroom images. This dataset includes: Varying lighting conditions Different classroom angles. Students with partial occlusions (e.g., masks, hands). Multiple students per frame. Test result shown in figure 5.

In this paper we test model of detection in different conditions, **easy** (Frontal faces, good lighting), **medium** (Side angles, medium lighting) and **hard** (Low light or partially occluded faces), result shown in Figure 6. Data augmentation techniques such as random cropping, brightness adjustment, and horizontal flipping were used to enhance generalization.

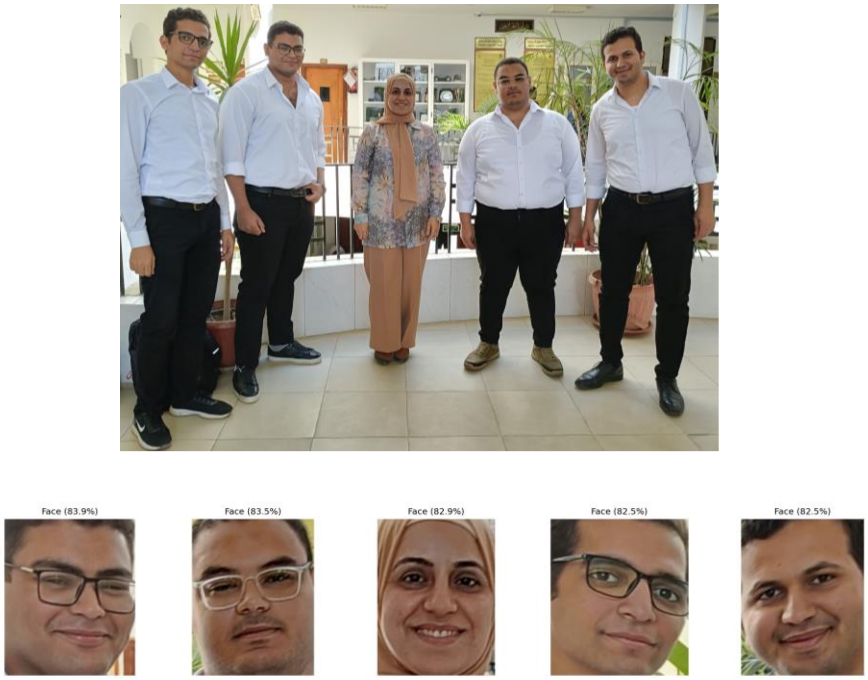


Figure. 5. Example for detection using YOLLO v11

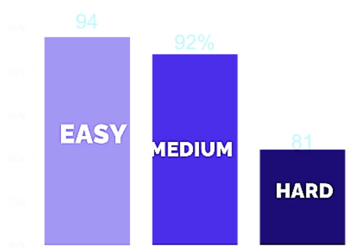


Figure 6: Accuracy results across different scenarios.

F. Face recognition

the face recognition component of the proposed smart classroom attendance system, which uses the FaceNet model to verify the identity of each detected student. Once faces are detected using YOLOv11, the FaceNet model processes each face and generates unique embeddings for recognition, enabling accurate and automated attendance logging.

1. Overview of FaceNet

FaceNet is a deep convolutional neural network architecture that maps facial images into a compact **128-dimensional embedding space**, where the Euclidean distance directly reflects facial similarity. Unlike traditional classification-based methods, FaceNet learns to distinguish identities through a **triplet loss function**, ensuring that embeddings of the same person are closer together than those of different individuals [17].

2. Recognition Pipeline

The face recognition pipeline consists of the following stages

- **Input from Face Detection**

- Cropped face images obtained from the YOLOv11 detection module are passed to the FaceNet model.

 Preprocessing

- Faces are aligned based on eye positions (optional for improved accuracy).
- Images are resized (e.g., 160×160 pixels) and normalized for input to FaceNet.

 Embedding Generation

- Each face is transformed into a 128-D embedding vector.
- This embedding represents the unique features of the face as shown in Figure 7.

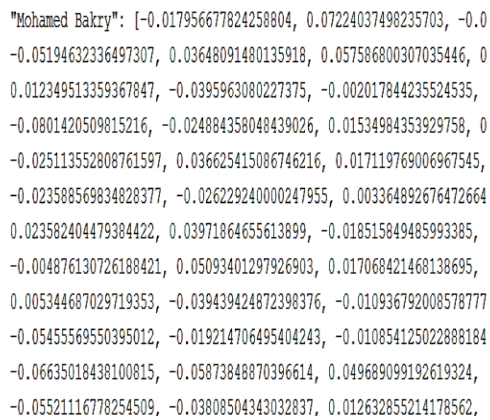


Figure. 7. Embedded features.

3. Identity Matching

- The embedding is compared to a database of pre-stored student embeddings.
- A match is determined using Euclidean distance:

$$d(e_1, e_2) = \sqrt{\sum_{i=1}^{128} (e_1^i - e_2^i)^2}$$

- If the distance is below a defined threshold (e.g., 0.7), the face is considered a match

4. Attendance Logging

- Once a face is recognized, the system marks the corresponding student as “Present” along with a timestamp.
- Duplicate detections within a session are ignored using time-based filters.

5. Embedding Database Management

Each student’s face is enrolled once by capturing 20 face images under different conditions (angles, lighting). The average embedding is stored in the database as the reference for recognition. This approach ensures stability and reduces false positives.

Following the integration of the YOLOv11 and FaceNet models, the system first utilizes YOLOv11 to detect faces within the input frames. The detected facial regions are then processed by FaceNet, which generates 128-dimensional embeddings and compares them against a pre-enrolled student dataset. Upon a successful match, the system identifies the student and automatically updates their attendance status in the database with ID number as shown in Figure 8.

Figure 8. Recognize IDs for detected students.

G. System modeling and Front-End

The proposed Smart Attendance System is modeled as a single functional unit that interacts with three primary external entities: students, instructors, and administrators. Students interface with the system to authenticate, submit attendance requests, and receive real-time notifications. Instructors manage class sessions and access analytics-driven attendance reports. Administrators (or registrars) oversee academic structures, maintain student records, and assign instructors to courses. This interaction diagram illustrates the flow of information between users and the system while abstracting internal operational details. Front end of attendance system shown in figure 9.

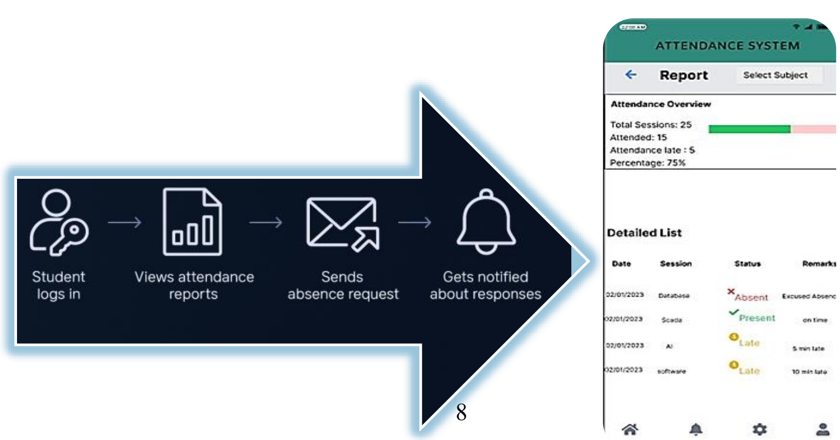
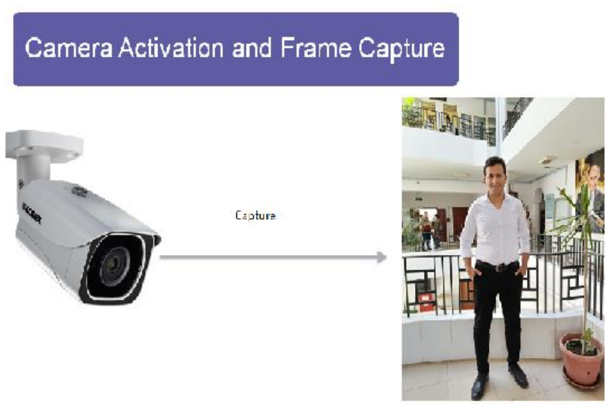


Figure 9. Mobile and web application for attendence system.

This workflow ensures that the system operates in real-time, with minimal manual intervention, and maintains high accuracy in student identification. The integration of YOLOv11 and FaceNet enables efficient face-based attendance in large classroom environments, significantly reducing administrative workload and increasing reliability



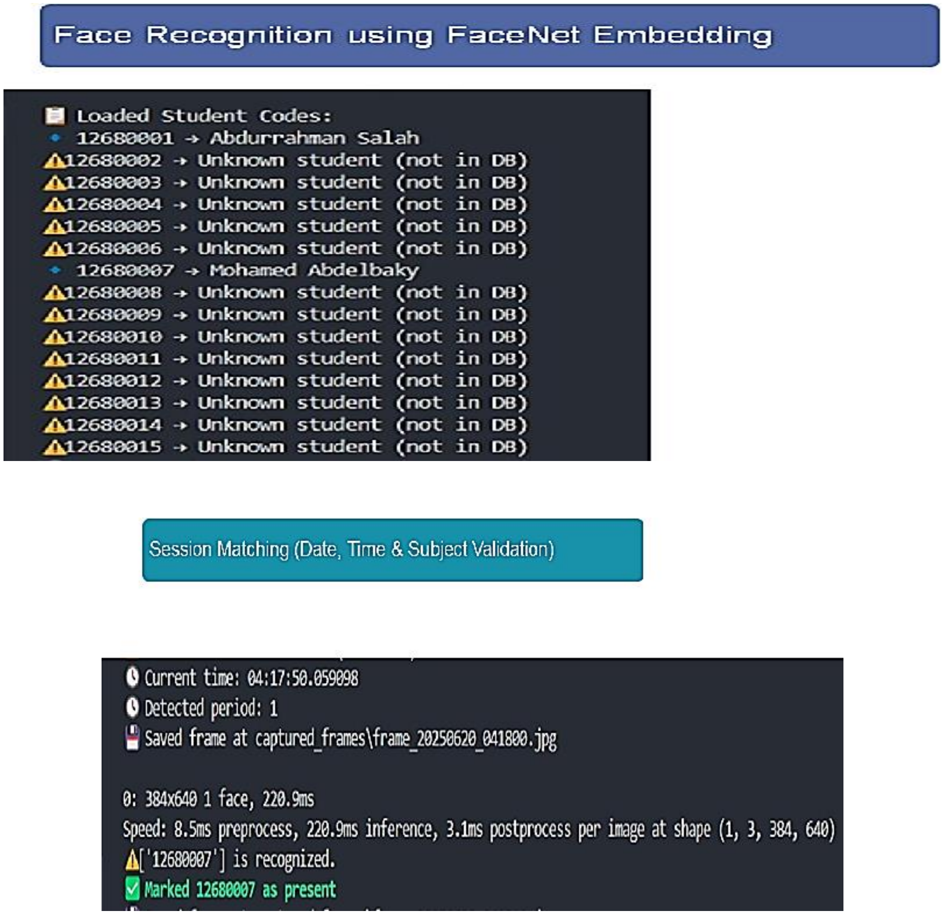


Figure 10. Workflow of attendence system

IV. CONCLUSION

This research presents the design and implementation of a real-time smart attendance system for educational environments, utilizing deep learning-based face detection and recognition techniques. The proposed system integrates the high-speed object detection capabilities of **YOLOv11** with the robust facial representation power of **FaceNet**, enabling accurate identification and automated attendance logging under diverse classroom conditions.

Through a structured workflow involving data acquisition, augmentation, detection, and recognition, the system effectively addresses common challenges such as varying lighting, student pose variations, and partial occlusions. Experimental results across multiple scenarios (easy, medium, and hard conditions) demonstrate that the combined approach achieves high accuracy and responsiveness, confirming its suitability for practical deployment.

Moreover, the system supports seamless interaction among students, instructors, and administrators through its web and mobile interfaces, offering real-time attendance tracking, session management, and data analytics. The use of a pre-trained deep transfer learning model reduces training time and computational cost, while maintaining high recognition precision.

In conclusion, the integration of YOLOv11 and FaceNet provides a scalable, secure, and efficient solution to automate classroom attendance. Future work may focus on enhancing privacy through edge computing, supporting large-scale deployment across institutions, and integrating emotion or behavior analysis for improved student engagement monitoring.

REFERENCES

[1] Omar Abdul Rahman Salim et al., 2018, "Class Attendance Management System Using Face Recognition" in 7th International Conference on Computer and Communication Engineering (ICCCE), 978-1-5386-6992- 1/18/$31.00 ©2018 IEEE.

[2] Dr. Palanivel. N et al., 2019, "Automated Attendance Systems Using Face Recognition by K-Means Algorithm", in Proceeding of International Conference on Systems Computation Automation and Networking, @IEEE 978-1-7281-1524-5.

[3] Shreyak Sawhney et al., 2019, "Real-Time Smart Attendance System using Face Recognition Techniques " in 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence), 978-1-5386- 5933-5/19/$31.00 c 2019 IEEE.

[4] Jaehoon Paul Jeong et al., 2020, "IAAS: IoT-Based Automatic Attendance System with Photo Face Recognition in Smart Campus", in Univ of Calif Santa Barbara, 978-1-7281-6758-9/20/$31.00 ©2020 IEEE.

[5] Asri Nuhi et al., 2020, "Smart Attendance System using QR Code", in 9th Mediterranean Conference on Embedded Computing, 978-1-7281-6949-1/20/$31.00 ©2020 IEEE.

[6] Soumitra Chowdhury et al., 2020, "Development of an Automatic Class Attendance System using CNNbased Face Recognition "in Emerging Technology in Computing, Communication and Electronics (ETCCE), 978-1-6654-1962-8/20/$31.00 ©2020 IEEE.

[7] Shizhen Huang et al., 2020, "Attendance System Based On Dynamic Face Recognition", in International Conference On Communications, Information System and Computer Engineering (CISCE), 978-1-7281-9761- 6/20//$31.00 ©2021 IEEE.

[8] M. Arsenovic, S. a. A. A. Sladojevic and D. Stefanovic, "FaceTime—Deep learning based face recognition attendance system," 2017 IEEE 15th International Symposium on Intelligent Systems and Informatics (SISY), pp. 000053--000058, 2017.

[9] R. Fu, D. Wang, D. Li and Z. Luo, "University classroom attendance based on deep learning," 2017 10th International Conference on Intelligent Computation Technology and Automation (ICICTA), pp. 128--131, 2017.

[10] M. Zulfiqar, F. Syed, M. J. Khan and K. Khurshid, "Deep Face Recognition for Biometric Authentication," in 2019 International Conference on Electrical, Communication, and Computer Engineering (ICECCE), 2019.

[11] Y.-Q. Wang, "An analysis of the Viola-Jones faces detection algorithm," Image Processing On Line, vol. 4, pp. 128--148, 2014

[12] Deng, J., Guo, J., Niannan, X., & Zafeiriou, S. (2019). ArcFace: Additive angular margin loss for deep face recognition. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 4690–4699.

[13] Rawat, S., & Jain, H. (2021). Real-time Face Mask Detection using YOLOv3 and OpenCV. In 2021 5th International Conference on Computing Methodologies and Communication(ICCMC),pp.563–567.IEEE.

[14] Li, Y., & Deng, W. (2019). Reliable Crowdsourcing and Deep Locality-Preserving Learning for Unconstrained Face Recognition. IEEE Transactions on Image Processing, 28(1), 354-36

[15] Bochkovskiy, A., Wang, C. Y., & Liao, H. Y. M. (2020). YOLOv4: Optimal Speed and Accuracy of Object Detection. arXiv preprint arXiv:2004.10934.

[16] https://docs.ultralytics.com/datasets/detect/coco/

[17] Al-Qershi, A., & Manasrah, A. (2021). Real-Time Face Recognition-Based Attendance System for Classroom Environment. Indonesian Journal of Electrical Engineering and Computer Science, 23(3), 1546–1553.