



Class Attendance System Based on Face Recognition

Omar Alniemi*^{ORCID}, Hanaa F. Mahmood^{ORCID}

Department of Computer Science, College of Education for Pure Science, University of Mosul, Mosul 41001, Iraq

Corresponding Author Email: omaralniemi@uomosul.edu.iq

<https://doi.org/10.18280/ria.370517>

Received: 22 July 2023

Revised: 25 August 2023

Accepted: 1 September 2023

Available online: 31 October 2023

Keywords:

face recognition, attendance system, FaceNet, Haar Cascaded, Manhattan distance

ABSTRACT

The conventional methodology for recording student attendance, which heavily relies on manual data transcription, is prone to inefficiencies and errors. Consequently, the development of an automated attendance management system has emerged as a critical need for efficient and accurate maintenance of attendance records. This study presents the design and implementation of an automated attendance management system, exploiting face recognition technology for identifying students within a class setting. A unique dataset was curated, consisting of 3900 facial images, captured in five varying positions and under diverse lighting conditions. In the initial phase of the system's operation, images of students are captured via a mobile camera. Subsequently, the Haar Cascaded classifier is utilized for the detection of faces within these captured images, and the FaceNet network is employed to recognize the detected faces. In the subsequent phase, the system registers attendance by cross-referencing the recognized faces with the primary student record. An attendance sheet copy is then dispatched to the teacher. Upon evaluating the system's effectiveness, it was ascertained that the system successfully identifies students and registers their attendance with an impressive accuracy of 97.5%. It outperforms traditional systems in terms of workload reduction, error avoidance, speed, and accuracy. The proposed system holds potential for widespread deployment in institutes and schools for recording attendance and could be extended for employee attendance recording. By reducing human errors and the time required for attendance registration, and by swiftly generating electronic attendance lists, this system signifies a substantial improvement over conventional systems.

1. INTRODUCTION

In contemporary organizations, the procurement of an Attendance Management System (AMS) is transitioning from a luxury to a necessity, serving as a vital tool for maintaining and managing the attendance records of employees or students. Attendance Management Systems are broadly bifurcated into two categories: Manual Attendance Management Systems (MAMS), which utilize traditional pen-and-paper methods, and Automated Attendance Management Systems (AAMS), which leverage software for modern, automated attendance tracking. The latter category further encapsulates a variety of systems such as card-based systems, mobile applications, and biometric-based systems. The traditional or manual methods have been criticized for their time-consuming nature and susceptibility to human error. In contrast, automated systems, particularly those driven by biometrics, offer an efficient and accurate alternative. These systems automate the process of attendance tracking through biometric data, enhancing speed, accuracy, and reducing the burden on administrative staff [1].

However, despite their advantages, AAMS face several challenges. For instance, independent attendance systems may prove ineffective in scenarios where an employee has commitments in multiple locations or works across various offices.

Card-based systems, whose problems lie in using the card incorrectly, losing the card, forgetting the card, or giving the card to a colleague and cheating on attendance.

Heavy reliance on input information manually, which may lead to errors or problems during input process.

Difficulty monitoring and reviewing information in electronic systems Especially if obtaining the results requires direct intervention by the responsible employee, or there is a shortcoming in the preparation of the program [2].

To overcome the aforementioned problems, the development of biometric attendance systems, which are based on reducing human interference in obtaining results [3]. These biometric systems rely on biometrics such as fingerprints, voice, palm prints, iris, and also facial recognition [4]. In systems based on the use of iris or the use of fingerprints, the iris or fingerprints are first stored in a database. Second, finding matches based on the database. However, these systems take a long time, in addition to having some difficulty during their implementation by people [5]. In addition, these systems require direct interaction with the devices used; also, the target area in the survey, whether it is the fingerprint or the iris of the eye, is relatively small to capture by affordable cameras. Besides, direct contact with the devices of these systems may cause infections of some types of diseases.

Consequently, many researchers have searched for better alternative systems in terms of speed, safety, and efficiency. Instead of relying on fingerprints, iris, and others, facial recognition is used. In principle, the systems based on face recognition use the same idea, in terms of storing the face image in the database and then finding the match using the database. Many researchers considered facial recognition systems to be among the best biometric systems [6]. The

researchers concluded that facial recognition systems have high accuracy and provide higher security in terms of health and fraud prevention. In addition, these systems can also be used in schools, universities, companies, hospitals, etc.

2. RELATED WORK

Many research papers have suggested automated attendance management systems that avoid the defects and errors that traditional systems suffer from by relying on attendance recording through face recognition. Some papers suggested the use of deep neural networks, such as:

Kumar et al. [7] proposed an attendance registration system for people who wear a medical mask. The proposed system will alert people who are not wearing masks. The alert informs people that they must wear a medical mask to register attendance. The proposed system includes two basic units: the first unit is the camera unit and the second unit is the software unit. The purpose of the camera unit is to take a frontal picture of people. The purpose of the software unit is to detect the masked face and determine the position of (eyes, eyebrows, and eyelashes). Image processing techniques, as well as the MobileNet deep neural network, are used to detect and identify disguised faces and determine attendance.

In addition, some papers suggested the use of (FaceNet) in their systems, such as: Suguna et al. [8], proposed a system whose priority was to overcome the problems that occur with the use of raspberry pi. Where after saving the images are stolen and used badly. The system was developed by adding a NodeMC board with a camera for taking pictures of the class students. Asymmetric encryption, RSA, was used for protecting the images that were stored in the attendance system. The property used to detect faces is the (dlib) property which is according to the CNN approach. Using FaceNet a vector with 128 dimensions is obtained, for face extraction. The data obtained is considered (input) to train the SVM Classifier. Two Android applications are developed. The first application is for teachers and the second application is for students. Where the teachers use the first application to record students' attendance, students use the second application to record their attendance themselves, by sending a video of their faces to the system server. This system achieved an accuracy of 94%.

In the study [9], the system aims to record attendance without human intervention. In this system, images of students' faces are captured using the camera linked to the system, the face is detected in the captured image using Max-Margin Object Detection (MMOD) and Histogram of Orientation Gradient (HOG) methods. FaceNet is used for feature extraction and facial recognition. KNN is used for face classification. For testing, the framework is implemented using NVIDIA's Jetson nano. Thus, the performance of the system is relatively good as it takes less time to process and recognize the faces in the captured images. However, this system provides lower accuracy in recognizing faces with a mask.

In addition, Nyein and Oo [10], proposed an attendance system based on facial recognition using the FaceNet network in addition to the Support vector machine. Initially, the raw data is processed to create the dataset used in the training phase. The images captured by the system are considered input images, where OpenCV is used to detect the face in the images, FaceNet is used for feature extraction, and the support vector machine is used as a classifier. By using the classifier, the

features are matched with the trained dataset. Later, an attendance record sheet is created using Excel based on the recognized faces. The system achieved an accuracy rate of 80%. The system faced some practical limitations related to FaceNet, represented by the difficulty in recognizing the face of people wearing additional accessories, which led to a decrease in the accuracy of the system.

While some papers suggested taking advantage of the Internet of Things, such as: Jeong et al. [11], proposed a system for recording attendance by recognizing faces. It is a system based on the Internet of Things. Students' photos are obtained through a camera located, for example, on a mobile phone. After processing the image data for face recognition, the recognition technology is enhanced. For the identification of faces, MTCNN was used, while for face recognition GoogleNet and VGG16 were used. vggface2 was used to train validation models. After the attendance is determined, the results are sent to the database using e-mail. The final result obtained has a precision of approximately 0.988.

Othman and Aydin [12], developed an application for facial recognition using the Internet of Things. Where a motion sensor and camera are installed on a (raspberry pi) device and using the camera to take pictures of the students, and when the required result is obtained and the student is identified, a message is sent via the Telegram application.

Amin et al. [13], proposed another system for recognizing faces using a (raspberry pi) device. After the photos are taken and stored, photos are sent to the cloud for facial recognition using the (cloudlet) system.

While some papers suggested relying on (Haar Cascade Classifier), such as: In the system proposed by Shah et al. [14], each student who registers in the class must record his face via a video clip that takes 3 to 5 seconds. Then the proposed system extracts 200 images from each video. The Haar Cascade Classifier does the process of detecting and classifying faces. The features of the extracted faces are stored in a special database. The front end of the proposed system is designed based on the tkinter library. Although this system succeeded in recognizing the students' faces by 93.1%, it faced some traditional problems represented in taking a lot of time and also not exceeding the noise problems in the pictures in addition to the need to improve the security of the system.

Furthermore, Singh et al. [15], connected a camera to a (raspberry pi) device. With this camera, students are filmed with a video clip. The captured video of the students is considered (the input) for the (raspberry pi) device. Through the captured video and using the viola Jones algorithm, the facial image data is taken for each student. Then the facial features are captured from the image data, as 160,000 features are used (selecting Haar features, integrating images, and then classifying using AdaBoost and Cascade Classifier). Finally, the input images are compared with the images that are already there.

Salman et al. [16], proposed a three-stage system. A camera is connected to a (raspberry pi) device. The images captured by the camera as well as each student's ID will be considered the (input) for the (raspberry pi). Haar Cascades was used to identify the faces of the students in the class. The LBPH algorithm was used for facial recognition. For training, 10 negative and positive images were used for each student, where the number of participating students was 30 students. A special web application has been prepared for the attendance system, with the ability to access the system application via personal computers or smartphones. When the system

recognizes the students' faces, attendance will be recorded. After each student's attendance is recorded, the data is collected in an Excel file. Then the file is uploaded to the server, and then notifications are sent via e-mail to both professors and students. This system has achieved an accuracy of up to 90%.

Last but not least, Bhattacharya et al. [17], proposed an attendance system based on face recognition. Where a camera was installed inside the classroom to take pictures of students. The students' images captured by the camera are stored in a special database. The HOG features are then used to detect the face. After that, the SVM classifier was used to recognize faces and the Viola-Jones algorithm was used to obtain the results. This system is faulty and facing lighting problems, but it can overcome this problem by using higher-resolution cameras and adopting algorithms that may be less sensitive to light.

When reviewing related works, the significant positive impact of technological progress appears on the possibility of faces recognition in terms of accuracy and speed. As deep learning increased the accuracy of face recognition systems by learning neural networks for the features of faces. While the huge potential in the computational power provided by the graphics processing units contributed to a significant increase in the speed of face recognition systems.

In this paper, an attendance recording system based on recognizing students' faces in class is proposed. In general, the system is divided into two main phases. In the first phase, creating a dataset of images, names and student IDs of all students registered in the class. Images are captured and information is recorded using the proposed system camera by the staff responsible for student attendance using the mobile application. While using the application the tablet/mobile camera automatically live captures the students. The system processes the images and creates the dataset. While in the second stage, images of the actual attendee's students are captured via a live video camera linked to the system. The system processes the images captured from the second phase and compares them with the dataset from the first phase to generate a student attendance report. In the proposed system, the Haar Cascaded classifier was used to detect faces in the captured images, while the FaceNet network was used to recognize the detected faces in the captured images.

3. HAAR CASCADE CLASSIFIER

It is considered one of the most famous algorithms based on machine learning for detecting objects in real time, whether in video or digital images. The Haar Cascade Classifier algorithm is characterized by its high speed performance because it does not depend on complex mathematical operations, due to its dependence on integral images, this also reflected in its ability to perform very efficiently on devices with limited capabilities. The algorithm is also characterized by the possibility of being satisfied with weak training data, because it uses the method of training weak classifiers and extracting a strong classifier [18]. The term Haar refers to the mathematical operation of a rectangle, while Haar features include edge, line, and rectangle. With regard to face detection, the (Haar feature) is the main core of the Haar Cascade Classifier, which is used to identify the feature in the input image, where each feature has one value.

This value is obtained by subtracting the sum of the pixel

values under the white rectangle from the sum of the pixel values under the black rectangle [19, 20]. This algorithm uses statistics to identify the face in the image. It does not depend on the values of all the pixels in the image, but only on the values of the features. Positive images and negative images are used for training, where the images are distinguished based on learning to recognize the lips, nose, and eyes [21]. Positive images required for training must contain a face, while negative images do not contain a face. Through the Features, only positive images will be classified into black and white rectangles [22, 23]. (Haar Features) is used to search for features in the input image. Simply the algorithm scans the input image from the upper left corner of the image and ends at the lower right corner. The scanning process is done through a square window of a certain size first, looking for features. When features of a certain value are found as a face, a square is drawn directly around the face. After that, the square window is resized several times, and the scanning process is restarted again. The purpose of resizing the window is to select faces of different sizes. The algorithm consists of four basic concepts:

3.1 Features

The classification in the algorithm is based on simple features, rather than pixels, for two reasons, the first reason is that the features encode based on little data. The second reason is that the system runs faster when it is based on features rather than pixels. The Haar-like features can be divided into five types, left-right, top-bottom, diagonal, horizontal-middle, and vertical-middle, as shown in Figure 1. Top-bottom is obtained by rotating left-right counterclockwise, and horizontal-middle is obtained by rotating vertical-middle counterclockwise. These features are used to detect the existence of the face in the input image, by scanning the image using the multi-scale feature window. However, due to the huge number of features that can be obtained, which may exceed 160,000 features, and which need a lot of time to calculate, the integral image was proposed [24].

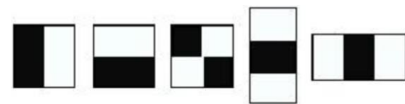


Figure 1. Haar-like features [24]

3.2 Integral image

It can be defined as the intermediate representation of the image. It is used to avoid the many arithmetic operations needed to calculate the (Haar feature), and this helps to quickly detect the features in the image. Instead of arithmetic operations to calculate each pixel, sub-rectangles are created, then matrices of these rectangles are created and finally used to calculate the values of Haar features [25]. However, not all of the Haar features are useful in the object identification process. For this, and to identify useful features, (AdaBoost) is relied upon.

3.3 AdaBoost

Used to select only useful features for detecting the object among the many features in the image. The working principle of AdaBoost is to combine weak classifiers to produce strong

classifiers. As long as a single integral image is weak, a group of integral images can produce a stronger classifier [26].

The AdaBoost algorithm iteratively trains weak classifiers on the set of samples that were misclassified in the previous training. Where in each re-training the algorithm increases the values of the wrongly classified samples, which means an increase in interest in them for the purpose of improving the performance of the classifiers. The learning process of the classifiers is repeated a certain number of times and then these classifiers are aligned to make one powerful classifier [27]. The steps of the algorithm can be summarized as follows:

- 1-Give equal values (weights) to all training samples.
- 2-Training the classifier according to equal values for all samples.
- 3-Calculate the value of the classifier. A lower value of error means a higher value of the classifier.
- 4-Depending on the values of the classifier, the values of the misclassified samples are increased, in order to focus more on them in the next training. Fit the new values so that their sum becomes one.
- 5-Repeating the training process a certain number of times, and at each training iteration, the value of the classifier is calculated, and the values of the misclassified samples are increased and adjusted.
- 6-Obtaining a strong final classifier by merging the previous classifiers according to their values and finding their average [28].

3.4 Cascade classifier

It is a classifier consisting of several levels. Creates a simple classifier with high speed and correct results, by excluding negative images and keeping positive images. This classifier consists of at least three levels of filters. Where the images entered in the first filter level of the classifiers are calculated, if the result of the feature in the filter is negative, the input is rejected, while if the result is positive, it moves to the next filter level, and so on. Until the number of windows decreases and the face detection approaches [25, 26]. Figure 2 shows the sequential classifier.

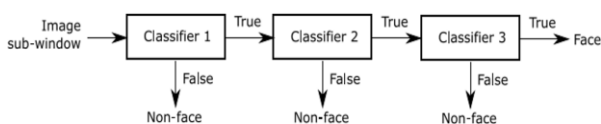


Figure 2. Cascade classifier [26]

4. FACE RECOGNITION USING FACENET

It is a deep learning-based neural network that was pre-trained to efficiently recognize faces based on Convolutional Neural Networks (CNN) by three researchers at Google in 2015 [29, 30]. When tested to verify and identify people, it had an accuracy of more than 95% [31]. The FaceNet algorithm provides impressive results in the field of face recognition compared to other algorithms, in terms of illumination conditions, speed, and accuracy. As the SVM algorithm takes longer time than the FaceNet algorithm in computational processors, while the accuracy of the PCV algorithm is affected by illumination conditions compared to the FaceNet algorithm. On the other hand, FaceNet algorithm achieved higher accuracy value in face recognition compared to ArcFace algorithm using the same dataset and under the same

conditions [32, 33]. FaceNet aims to extract facial features from an image, where the face image is embedded in the form of a digital vector that accurately represents the features of the faces, so that the digital representation is close in value for similar faces and divergent in value for non-similar faces. In addition, the digital embedding used in FaceNet was so accurate that it could handle different lighting contrasts and different facial expressions of the same person [34]. The input to the neural network is an image of a person's face, while the output from the neural network is a 128-digit vector. The vector represents the features of the face. This vector can be called (embedding). This embed is stored to be proactively processed and categorized, to obtain high-accuracy results [35]. In machine learning, the process of re-representing (high) dimensional data such as images into (low) dimensional data such as embeddings, has become a common practice these days. The vectors, which can be defined as specific points in the Cartesian system, represent the inclusions; Figure 3 shows the structure of FaceNet.



Figure 3. FaceNet structure [29]

By calculating the inclusions, a person's face can be identified. The vectors for the same person are very close together, while the vectors for several people are different. If the inclusion of a specific image (x) is close to the inclusion of a specific person's image (y), then it can be said that the image (x) is for the person (y). The neural network uses a learning method based on positive and negative examples, called triple loss as shown in Figure 4.

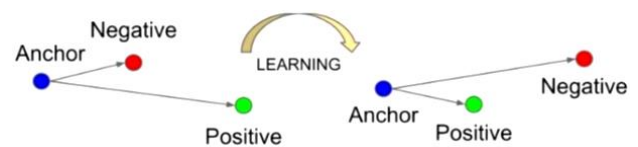


Figure 4. Triple loss [30]

The input format for FaceNet training is (160×160×3), so the face images in the dataset are resized to this format. FaceNet learns to embed the image of a single face in a vector of 128 digits. In general, FaceNet is built based on 22 layers as shown in Table 1 [36], and triplet loss is used to optimize the weight of the network. Therefore, when the images are inserted into the neural network, the images will be converted into vectors. After that, an image called (an anchor image) is selected, then an image of the specific person is identified, called (a positive image) that is similar to the anchor image, and an image of another person is identified, called (the negative image) that is not identical to the anchor image [37]. The benefit of triplet loss is to reduce the difference between the input image and images that have similar features. As triplet loss brings similar faces closer together and diverges different faces from each other, where it updates the weights of the samples during training to help the network classify the samples by calculating the difference between the positive, negative and anchor faces and then decreasing the difference between the positive and anchor faces and increasing the difference between the negative and anchor faces [38].

Table 1. FaceNet layers [34]

layer	Size in	Size out	Kernel	Param	FLPS
conv1	220 × 220 × 3	110 × 110 × 64		9 K	115 M
pool1	110 × 110 × 64	55 × 55 × 64	7 × 7 × 3, 2	0	
rnorm1	55 × 55 × 64	55 × 55 × 64	3 × 3 × 64, 2	0	
conv2a	55 × 55 × 64	55 × 55 × 64		4 K	13 M
conv2	55 × 55 × 64	55 × 55 × 192	1 × 1 × 64, 1	111 K	335 M
rnorm2	55 × 55 × 192	55 × 55 × 192	3 × 3 × 64, 1	0	
pool2	55 × 55 × 192	28 × 28 × 192		0	
conv3a	28 × 28 × 192	28 × 28 × 192	3 × 3 × 192, 2	37 K	29 M
conv3	28 × 28 × 192	28 × 28 × 384	1 × 1 × 192, 1	664 K	521 M
pool3	28 × 28 × 384	14 × 14 × 384	3 × 3 × 192, 1	0	
conv4a	14 × 14 × 384	14 × 14 × 384	3 × 3 × 384, 2	148 K	29 M
conv4	14 × 14 × 384	14 × 14 × 256	1 × 1 × 384, 1	885 K	173 M
conv5a	14 × 14 × 256	14 × 14 × 256	3 × 3 × 384, 1	66 K	13 M
conv5	14 × 14 × 256	14 × 14 × 256	1 × 1 × 256, 1	590 K	116 M
conv6a	14 × 14 × 256	14 × 14 × 256	3 × 3 × 256, 1	66 K	13 M
conv6	14 × 14 × 256	14 × 14 × 256	1 × 1 × 256, 1	590 K	116 M
pool4	14 × 14 × 256	7 × 7 × 256	3 × 3 × 256, 1	0	
concat	7 × 7 × 256	7 × 7 × 256	3 × 3 × 256, 2	0	
fc1	7 × 7 × 256	1 × 32 × 128		103 M	103 M
fc2	1 × 32 × 128	1 × 32 × 128	maxout p = 2	34 M	34 M
fc7128	1 × 32 × 128	1 × 1 × 128	maxout p = 2	524 K	0.5 M
	1 × 1 × 128	1 × 1 × 128			

5. PROPOSED MODEL

In the proposed system shown in Figure 5, the dataset is created for all students registered in the class, The students were females and males between the ages of 18 to 28. The system administrators input the students' data (name and student ID) and also upload the students' typical images to the system, using a mobile video camera with 12 mp wide sensor f/1.6 aperture 26 mm focal length linked to the system. The single video for each student is divided into thirty frames, which include images of the student's face from five different positions which are from the front, 45° to the right side, 90° to the right side, 45° to the left side, 90° to the left side in real classroom environment under multiple lighting conditions where three lighting conditions used (high, medium, low). The dataset of the proposed model includes 3900 faces images belong to 130 students, 30 faces images for each student. This dataset divided in to 80% samples for training and 20% samples for testing. The dataset consists of 64% males and 36% females. Consent is obtained from the students and university after explain to then what kind of data will be collected and how it will be used.

Haar Cascaded Classifier is used to detect the location of the face in the image. The value of 1.2 is specified for the zoom factor of the Haar Cascade search box, which is used while searching for (features) of the face in the image. Where the search begins with a small box and then grows by 1.2 and continues to increase by this value each time. While the next search box is set two pixels away from the first box, and so on. There are too many Haar Cascade features (160.000 features) that identify the face in the image and by searching for these (features), the location of the face is detected in the image.

FaceNet network has been proposed to create an (embedding database) of face images of people to be identified. The input to the FaceNet network is the face image, while the outputs are the facial features in digital format, as each output face consists of its own set of digits (128 digits), different from the set of digits of the other faces. The unique set of digits for a particular person's face is called an (embedding). If a

different person's face is input, a different embedding will be output. Based on this idea, a person is defined by his unique embedding. Then, after obtaining the unique embedding responsible for identifying the person, the embedding is saved in a special database for embeddings we named (embeddings database). The database includes the names of people with their embeddings. After input the face into the FaceNet network and have an embedding as the output for the first image, the same process will be repeated with all images in the (images database), and each time an embedding is produced, it will be saved in (embedding database).

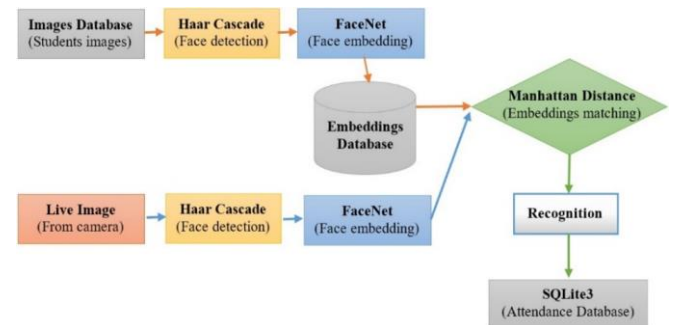


Figure 5. Proposed model

In the second phase of the proposed system, the input is from the live camera that is used to capture the students who are in the classroom, the Haar Cascade classifier is used with the captured video frames to identify the location of the face in the image, by the distinctive features of the face. The Haar Cascade classifier helps discard the unhelpful parts of an image and focus on the parts of a face that are useful in image processing and face detection. the output of Haar Cascade will be fed into FaceNet. FaceNet produces an embedding that is directly compared with all items of (Embeddings Database), or in other words, the embeddings of images that are input by the live camera will be compared with the embeddings of all images in (Image Database).

Manhattan distance is proposed to find the amount of difference between the two embedding vectors. Where Manhattan distance is commonly used to find the distance between two points in a two-dimensional space, and an n-dimensional space. The lower the distance value, the more similar the faces. There are two embedding vectors with 128 digits, the first one is obtained from (Embeddings Database) and is indicated by the symbol (ED), and assumes that $ED = (y_1, y_2, \dots, y_{128})$, the second one is the embedding of the face to be recognized, input from the live camera is indicated by the symbol (CD), and assume that $CD = (x_1, x_2, \dots, x_{128})$, the required equation will be as follows:

$$D_{(EDCD)} = \sum_1^{128} |x - y| \quad (1)$$

By finding the smallest distance (D) between (CD) and (ED), can specify which embedding belong to which person. The process of comparing the embedding of the image from the camera will be repeated with the number of embeddings saved in the (embeddings database), until it is found the smallest distance (D).

SQLite3 is used in the proposed model for database management. The list of names of attendees and absent students is saved using SQLite3. Where, each face coming from the camera and recognized (with the lowest distance value), will take the name attached to the embedding, which matched this face. The names are saved in the list of known faces (names of attendees). A comparison is made between the list of all students' names and the list of those present (known faces) to identify the names of the absent students and save them in the database. The proposed system was tested using a real time video, where a group of students and non-registered persons participated in the test.

6. RESULTS AND DISCUSSION

First, the teacher uses their mobile device to capture real time video for the students in the classroom. The proposed system application allows the mobile device to be linked to the system computer, to use the mobile device camera. The application of the proposed system also allows teachers to access the system computer, where attendance data is collected and stored. Directly the Haar Cascade algorithm is implemented on captured video, to get individual faces by using line and edge features. The Haar Cascade algorithm focuses on the parts of the face that are useful for detection (region of interest). By cropping, the algorithm discards the other parts of the image that are not useful in detection and matching operations. Once the faces are detected, they are extracted and stored. The detected face will be the input of the FaceNet network, but before feeding the detected face to the FaceNet network, and in order to train the model, the face is cropped, normalized, and changed the size to $(160 \times 160 \times 3)$, where the FaceNet network has been pre-trained to receive input with specifications of this size and normalization.

The Resnet architecture is adopted in the construction of the FaceNet network, which consists of 22 layers. When training, the parameters were equipped with a learning rate of 0.1, while the epoch value was 1000. The triple loss function was used to optimize the network weight. Triplet Loss reduces the difference between the positive input and the anchor if they

have the same identity, and increases the difference between the negative input and the anchor if they do not have the same identity. Each face coming from the camera and recognized (with the lowest distance value), will take the name attached to the embedding, which matched this face. The names are saved in the list of known faces (names of attendees). A comparison is made between the list of all students' names and the list of those present (known faces) to identify the names of the absent students and save them in the database.

The proposed system was tested using a real time video, where a group of students and non-registered persons participated in the test. The student attendance database is created after the face recognition process is completed, and it is dealt with using a simple web application. In addition, the attendance management system can be accessed either via a mobile phone or a computer, as SQLite3 files are collected and uploaded to the cloud server. The application of the proposed system uses the SMTP protocol to send e-mail and is characterized by a clear and easy-to-use user interface, Figure 6 shows the main user interface of the attendance management administrator. The proposed system records student attendance instead of manual registration by teachers. To ensure information security, the administrator logs into the system by entering a username and password. The attendance management administrator can add a student to the database, by inputting his required data and taking the appropriate video. The application also allows the attendance management administrator to easily check students' attendance and send the attendance status to the teacher and student via e-mail.

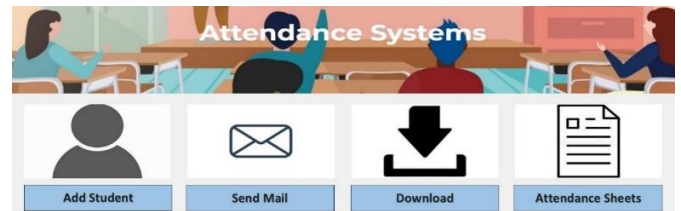


Figure 6. The main user interface of the attendance management administrator



(a) Detection and recognition of multiple faces



(b) Detection and recognition of faces in multiple angles

Figure 7. The proposed system results

When testing the proposed system to determine its effectiveness, it was exposed to several conditions represented in different lighting, facial movements in several directions the system succeeded in recognizing multi face with an accuracy of 97.5% Therefore, it can be said that the proposed system is highly effective for managing student attendance successfully. To test the quality of the system, as shown in Figure 7, and the distance between the student and the camera.

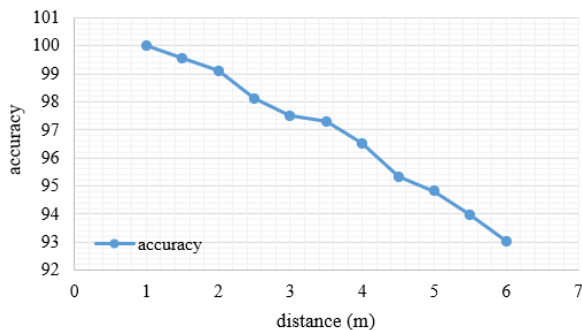


Figure 8. Inversely proportional to the accuracy of the system with distance

The distance between the student and the camera is not fixed, due to the use of mobile cameras. In addition, the teacher's use of the camera is not the same every time, and the student may choose to sit on a different seat each time. Figure 8 shows the inverse proportion to the accuracy of the proposed system. The greater distance between the camera and the student, the less accurate the system is in recognizing the student.

The system achieved 99.5% for the detection of a single face in the image and 98.9% for recognition, also the system achieved 98.8% for the detection of multi faces and 97.5% for recognition. After multiple observations were used for train the proposed system. Using the camera several times in one lecture increases the accuracy of the system in recognition, reduces limitations related to camera blind spots and eliminates the effect of poor capturing angles. If some students' faces are not captured the first time, they will be captured in the next time. however, if an image of a student's face is captured more than once, the proposed system will delete the redundant faces and keep one face for one student, thus ensuring that the student's attendance is not recorded more than once.

Creating special dataset considering multiple perspectives, under multiple lighting brightness and multidistance, in the same classrooms environment increases the robust of the proposed system

7. CONCLUSIONS

The aim of the proposed system attendance system is to obtain correct and accurate attendance records in a quick automatic manner and without the intervention of the human factor as much as possible, to avoid the practical mistakes that face the previous traditional systems. The proposed system relied on recognizing students' faces in recording attendance, using a live video camera. The Haar Cascade classifier was used for face detection, and the pre-trained FaceNet network was used for face recognition. The proposed system achieved an accuracy of 97.5% in identifying students.

The operation of the proposed system, requires relying on only one camera, which reduces the required cost compared to

other systems. It also succeeded in reducing the limitations related to camera blind spots as well as removing the effect of poor capturing angles.

The proposed system can send the attendance status to the teacher and the student via e-mail, and it also can store and retrieve attendance records to benefit from them at the required time. Teachers and students can get the benefits of this systems where attendance register automatically without annoying them, furthermore, both can get a copy of the attendance information. When testing the proposed system in multiple conditions in terms of different lighting, different facial positions, and the different distances between the student and the camera, the system was characterized by high accuracy and speed in recording attendance without the need for manual work or the use of expensive and complex devices.

The deployment cost of the proposed system is much lower than other systems, the teachers and students can install system application on personal mobile or computer to deal with proposed attendance system, furthermore teachers can use mobile camera or computer camera to register the attendance without need to install equipment, which cost money and time.

REFERENCES

- [1] Raj, A.A., Shoheb, M., Arvind, K., Chethan, K.S. (2020). Face recognition based smart attendance system. In 2020 International Conference on Intelligent Engineering and Management (ICIEM), London, UK, pp. 354-357. <https://doi.org/10.1109/ICIEM48762.2020.9160184>
- [2] Kariapper, R. (2021). Attendance system using RFID, IOT and machine learning: A two-factor verification approach. *Journal of Advanced Research in Dynamical and Control Systems*, 12(6): 3285-3297. <https://doi.org/10.5373/jardcs/v12i6/20202653>
- [3] Abuzar, M., bin Ahmad, A., bin Ahmad, A.A. (2020). A survey on student attendance system using face recognition. In 2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), Noida, India, pp. 1252-1257. <https://doi.org/10.1109/ICRITO48877.2020.9197815>
- [4] Sunaryono, D., Siswantoro, J., Anggoro, R. (2021). An android based course attendance system using face recognition. *Journal of King Saud University - Computer and Information Sciences*, 33(3): 304-312. <https://doi.org/10.1016/j.jksuci.2019.01.006>
- [5] Ruhitha, V., Prudhvi Raj, V.N., Geetha, G. (2019). Implementation of IOT based attendance management system on raspberry pi. In 2019 International Conference on Intelligent Sustainable Systems (ICISS), Palladam, India, pp. 584-587. <https://doi.org/10.1109/ISS1.2019.8908092>
- [6] Elias, S.J., Hatim, S.M., Hassan, N.A., Abd Latif, L.M., Ahmad, R.B., Darus, M.Y., Shahuddin, A.Z. (2019). Face recognition attendance system using Local Binary Pattern (LBP). *Bulletin of Electrical Engineering and Informatics*, 8(1): 239-245. <https://doi.org/10.11591/eei.v8i1.1439>
- [7] Kumar, R.S., Rajendran, A., Amrutha, V., Raghu, G.T. (2021). Deep learning model for face mask based attendance system in the era of the COVID-19 pandemic. In 22021 7th International Conference on Advanced Computing and Communication Systems (ICACCS),

- Coimbatore, India, pp. 1741-1746. <https://doi.org/10.1109/ICACCS51430.2021.9441735>
- [8] Suguna, G.C., Kavitha, H.S., Sunita, S. (2021). Face recognition system for realtime applications using SVM combined with FaceNet and MTCNN. *International Journal of Electrical Engineering and Technology (IJEET)*, 12(6): 328-335. <https://doi.org/10.34218/IJEET.12.6.2021.031>
- [9] Shanmuhappriya, M. (2021). Automatic attendance monitoring system using deep learning. In *Proceedings of the First International Conference on Combinatorial and Optimization*, Chennai, India. <https://doi.org/10.4108/eai.7-12-2021.2314601>
- [10] Nyein, T., Oo, A.N. (2019). University classroom attendance system using FaceNet and support vector machine. In *2019 International Conference on Advanced Information Technologies (ICAIT)*, Yangon, Myanmar, pp. 171-176. <https://doi.org/10.1109/AITC.2019.8921316>
- [11] Jeong, J.P., Kim, M., Lee, Y., Lingga, P. (2020). IAAS: IoT-based automatic attendance system with photo face recognition in smart campus. In *2020 International Conference on Information and Communication Technology Convergence (ICTC)*, Jeju, Korea (South), pp. 363-366. <https://doi.org/10.1109/ICTC49870.2020.9289276>
- [12] Othman, N.A., Aydin, I. (2018). A face recognition method in the Internet of Things for security applications in smart homes and cities. In *2018 6th International Istanbul Smart Grids and Cities Congress and Fair (ICSG)*, Istanbul, Turkey, pp. 20-24. <https://doi.org/10.1109/SGCF.2018.8408934>
- [13] Amin, A.H.M., Ahmad, N.M., Ali, A.M.M. (2016). Decentralized face recognition scheme for distributed video surveillance in IoT-cloud infrastructure. In *2016 IEEE Region 10 Symposium (TENSYP)*, Bali, Indonesia, pp. 119-124. <https://doi.org/10.1109/TENCONSpring.2016.7519389>
- [14] Shah, K., Bhandare, D., Bhirud, S. (2020). Face recognition-based automated attendance system. In *International Conference on Innovative Computing and Communications: Proceedings of ICICC 2020*, pp. 945-952. https://doi.org/10.1007/978-981-15-5113-0_79
- [15] Singh, D.N., Sri, M.K., Mounika, K. (2019). IOT based automated attendance with face recognition system. *International Journal of Innovative Technology and Exploring Engineering*, 8(6S4): 450-456. <https://doi.org/10.35940/ijitee.F1093.0486S419>
- [16] Salman, H., Uddin, M.N., Acheampong, S., Xu, H. (2019). Design and implementation of IoT based class attendance monitoring system using computer vision and embedded Linux platform. In *Web, Artificial Intelligence and Network Applications: Proceedings of the Workshops of the 33rd International Conference on Advanced Information Networking and Applications (WAINA-2019)*, Matsue, Japan, pp. 25-34. https://doi.org/10.1007/978-3-030-15035-8_3
- [17] Bhattacharya, S., Nainala, G.S., Das, P., Routray, A. (2018). Smart attendance monitoring system (SAMS): A face recognition based attendance system for classroom environment. In *2018 IEEE 18th International Conference on Advanced Learning Technologies (ICALT)*, Mumbai, India, pp. 358-360. <https://doi.org/10.1109/ICALT.2018.00090>
- [18] Riyantoko, P.A., Sugiarto, Hindrayani, K.M. (2021). Facial emotion detection using Haar-cascade classifier and convolutional neural networks. *Journal of Physics: Conference Series*, 1844(1): 012004. <https://doi.org/10.1088/1742-6596/1844/1/012004>
- [19] Indira, D., Sumalatha, L., Markapudi, B.R. (2021). Multi facial expression recognition (MFER) for identifying customer satisfaction on products using deep CNN and Haar cascade classifier. *IOP Conference Series: Materials Science and Engineering*, 1074(1): 012033. <https://doi.org/10.1088/1757-899x/1074/1/012033>
- [20] Ahmad, A.H., Saon, S., Abd Kadir Mahamad, C.D., Wiwoho, S., Mudjanarko, S.M.S.N., Hariadi, M. (2021). Real time face recognition of video surveillance system using Haar cascade classifier. *Indonesian Journal of Electrical Engineering and Computer Science*, 21(3): 1389-1399. <https://doi.org/10.11591/ijeecs.v21.i3.pp1389-1399>
- [21] Kulkarni, P., T M, R. (2021). Video based sub-categorized facial emotion detection using LBP and edge computing. *Revue d'Intelligence Artificielle*, 35(1): 55-61. <https://doi.org/10.18280/ria.350106>
- [22] Malhotra, S., Aggarwal, V., Mangal, H., Nagrath, P., Jain, R. (2021). Comparison between attendance system implemented through Haar cascade classifier and face recognition library. *IOP Conference Series: Materials Science and Engineering*, 1022(1): 012045. <https://doi.org/10.1088/1757-899X/1022/1/012045>
- [23] Labib, R.P.M.D., Hadi, S., Widayaka, P.D. (2021). Low cost system for face mask detection based haar cascade classifier method. *MATRIK: Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer*, 21(1): 21-30. <https://doi.org/10.30812/matrik.v21i1.1187>
- [24] Priadana, A., Habibi, M. (2019). Face detection using Haar cascades to filter selfie face image on Instagram. In *2019 International Conference of Artificial Intelligence and Information Technology (ICAIIIT)*, Yogyakarta, Indonesia, pp. 6-9. <https://doi.org/10.1109/ICAIIIT.2019.8834526>
- [25] PNithish Srیمان, K., Raj Kumar, P., Naveen, A., Saravana Kumar, R. (2021). Comparison of Paul Viola - Michael Jones algorithm and hog algorithm for face detection. *IOP Conference Series: Materials Science and Engineering*, 1084(1): 012014. <https://doi.org/10.1088/1757-899x/1084/1/012014>
- [26] Rahmad, C., Asmara, R.A., Putra, D.R.H., Dharma, I., Darmono, H., Muhiqqin, I. (2020). Comparison of Viola-Jones Haar Cascade Classifier and Histogram of Oriented Gradients (HOG) for face detection. *IOP Conference Series: Materials Science and Engineering*, 732(1): 012038. <https://doi.org/10.1088/1757-899X/732/1/012038>
- [27] Namah, A.A., Mirza, N.M., Al-Zuky, A.A. (2022). Target detection in video images using HOG-based cascade classifier. *Revue d'Intelligence Artificielle*, 36, (5): 709-715. <https://doi.org/10.18280/ria.360507>
- [28] Sevinç, E. (2022). An empowered AdaBoost algorithm implementation: A COVID-19 dataset study. *Computers & Industrial Engineering*, 165: 107912. <https://doi.org/10.1016/j.cie.2021.107912>
- [29] Raj, A., Raj, A., Ahmad, I. (2021). Smart attendance monitoring system with computer vision using IoT. *Journal of Mobile Multimedia*, 17(1-3): 115-125. <https://doi.org/10.13052/jmm1550-4646.17135>

- [30] Tiwari, H., Goyal, S., Agrawal, R., Pawar, M. (2018). Live attendance system via face recognition. *International Journal for Research in Applied Science & Engineering Technology (IJRASET)*, 6(IV): 3891-3898. <https://doi.org/10.22214/ijraset.2018.4639>
- [31] Pavez, V., Hermosilla, G., Pizarro, F., Fingerhuth, S., Yunge, D. (2022). Thermal image generation for robust face recognition. *Applied Sciences*, 12(1): 497. <https://doi.org/10.3390/app12010497>
- [32] Xiang, W. (2022). The research and analysis of different face recognition algorithms. *Journal of Physics: Conference Series*, 2386(1): 012036. <https://doi.org/10.1088/1742-6596/2386/1/012036>
- [33] Firmansyah, A., Kusumasari, T.F., Alam, E.N. (2023). Comparison of face recognition accuracy of ArcFace, FaceNet and Facenet512 models on deepface framework. In *2023 International Conference on Computer Science, Information Technology and Engineering (ICCoSITE)*, Jakarta, Indonesia, pp. 535-539. <https://doi.org/10.1109/ICCoSITE57641.2023.10127799>
- [34] Dmello, R., Yerremreddy, S., Basu, S., Bhitle, T., Kokate, Y., Gharpure, P. (2019). Automated facial recognition attendance system leveraging IoT cameras. In *2019 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, Noida, India, pp. 556-561. <https://doi.org/10.1109/CONFLUENCE.2019.8776924>
- [35] Dzhargarov, A.I., Suleymanova, M.A., Zolkin, A.L. (2020). Face recognition methods. *IOP Conference Series: Materials Science and Engineering*, 862(4): 042046. <https://doi.org/10.1088/1757-899X/862/4/042046>
- [36] Schroff, F., Kalenichenko, D., Philbin, J. (2015). FaceNet: A unified embedding for face recognition and clustering. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, MA, USA, pp. 815-823. <https://doi.org/10.1109/CVPR.2015.7298682>
- [37] Praneesh, M., Napoleon, D. (2022). Face recognition for secure online payment with proxy detection using face net classifier. *International Journal of Research Publication and Reviews*, 3(1): 159-161. <https://doi.org/10.2139/ssrn.4140233>
- [38] Adhinata, F.D., Rakhmadani, D.P., Wijayanto, D. (2021). Fatigue detection on face image using FaceNet algorithm and K-nearest neighbor classifier. *Journal of Information Systems Engineering and Business Intelligence*, 7(1): 22-30. <https://doi.org/10.20473/jisebi.7.1.22-30>