Face Detection and Recognition for Class Attendance

Chang-Horn Boea, Kok-Why Ngb,\* Su-Cheng Hawc and Palanichamy Naveend

*a,b,c,d Faculty of Computing and Informatics, Multimedia University, 63100, Cyberjaya, Malaysia.*

*\*Corresponding author: kwng@mmu.edu.my*

***Abstract***— **Class attendance is a crucial indicator of students' seriousness towards learning. Many institutions continue to use manual methods, which are usually error-prone and unproductive. By leveraging state-of-the-art computer vision algorithms, the system accurately captures and verifies the identity of individuals attending class. This paper's goal is to investigate and create an automated facial recognition system for classroom attendance in order to increase the precision and effectiveness of attendance tracking. In order to reduce the likelihood of fake attendance and boost efficiency in attendance recording, we propose a system using computer vision technologies, namely Histogram of Oriented Gradients (HOG) with Support Vector Machine (SVM) for face detection and deep Convolutional Neural Networks (CNN) for face identification. The facial recognition system makes it simple and easy to record attendance. All participants have to do is gaze into the camera for the system to automatically record their presence. To evaluate the performance, the system is tested, and its accuracy is compared to the current QR code attendance method used in the authors’ institution. The results of the study demonstrate that the recommended approach is more accurate and efficient than the current procedures. With its real-time face detection and recognition capabilities, the system allows for precise attendance records. This technology ensures accurate and reliable attendance data, empowering organizations to make informed decisions, effectively manage resources, and provide a seamless experience for all students. In addition, the same attendance system can be deployed for any event in an organization.**

***Keywords***—**Face Detection; Face Recognition; HOG; CNN; Class Attendance**

*Manuscript received 15 Oct. 2020; revised 29 Jan. 2021; accepted 2 Feb. 2021. Date of publication 17 Feb. 2021.*

*International Journal on is licensed under a Creative Commons Attribution-Share Alike 4.0 International License.*

**

1. Introduction

Class attendance is an essential metric since it is closely related to students' academic performance [1]. Students that maintain a high attendance rate are more likely to achieve academically [2]. It would be disastrous for education if significant numbers of students were regularly absent because a low attendance rate is directly associated with poor retention. Class attendance is also required for a variety of other reasons, including the fact that it acts as evidence and provides the student with accident protection through personal insurance. Both insurance companies—the student's private insurance and the university's insurance—may have issues if the attendance is not adequately documented or is faulty.

The existing manual class attendance method was inefficient, which was the driving force for this initiative. Both manually inputting attendance and scanning QR codes take time and are susceptible to error. The survey's findings [3] showed that some students cheated their attendance by asking their friends to help them scanning the QR code. Besides, the usual camera-based attendance method also requires that all students line up in front of the camera for it to record each student's attendance. Due to the long time it takes, the class may be interrupted and abbreviated.

This paper therefore presents the notion of an automated class attendance system that relies on the ideas of face detection and face recognition [4], [5]. Every time a student enters the classroom, it will automatically take a photo of their face and register the time to guarantee that the attendance is recorded successfully and efficiently. To be sufficient for validation against the image that is already saved in the database, the captured faces must adhere to a minimal standard [6]. The student's attendance will be recorded if the similarity percentage is higher than a reasonable acceptance rate. To evaluate the effect of the suggested paper, the collected data will be compared to the precision of the current QR Code attendance method [7], [8]. To further improve its functionality, this paper would accept other collaborating developments including Radio-frequency identification (RFID) [9], [10]compatibility.

Arsenovic et al. [11]suggested a solution for facial recognition problems by fusing different contemporary methodologies with cutting-edge deep learning techniques. Their choice for a facial detection algorithm is convolutional neural network (CNN) Cascade. Six CNNs make up the cascade: three for binary classification and three for calibrating bounding boxes. They developed a face detection method using the machine learning framework (A Torch). In contrast to conventional machine learning, CNN can automatically train features to capture graphic variance. The drawback of CNN Cascade face detection is that it needs further methods to improve performance, including parallel processing, and greater computer capacity.

The framework developed a rapid and precise face and object identification system by combining the ideas of the Haar-like feature, Integral Image, AdaBoost Algorithm, and the Cascade Classifier. Ashritha et al. [12] applied the Viola-Jones Algorithm face detection, which uses a cascade classifier to identify human faces. Among all these methods, the framework developed by Viola and Jones has the highest detection rate and is the fastest. The Viola-Jones detection method is efficient for real-time applications due to its speed and dependability. Bhattacharya et al. [13] employed a face-tracking methodology to create face-logs more correctly by first recognizing the face using the Viola & Jones method and then following it frame by frame using the correlation tracker from the dlib package. This approach consumes fewer processing resources since it does not have to recognize the face after switching to a new frame. As a result, creating a face log, or a quick portrayal of the face in a short video has become easier. Sawhney et al. [14] advises recognizing faces by utilizing the 68 identifiers on a person's face. The Viola and Jones algorithm, which is utilized for face landmark recognition, restricted local model-based face tracking, and face bounding area detection, is built based on these facial landmarks.

Yolo is a one-stage detector that offers an improvement above conventional single-stage detector (SSD). In R-CNN, SSDs often perform worse than two-stage detectors (TSD). The YOLO method can process up to 45 frames per second (FPS), compared to the average 5 FPS for an R-CNN-based algorithm [15]. Yolo takes an input picture and divides it into grids of squares using its framework. In the picture, each square grid has been subjected to localization and classification. Finally, YOLO is founded to forecast bounding squares with class probabilities in object detection.

The data extraction from tiny portions of an image's edges forms the basis of HOG's feature extraction process [16], [17] . The fundamental objective of HOG feature extraction is to describe the orientation and magnitude values of the pixels in a photo.

The Haar Cascade Method (feature-based), which separates images into groups of those with faces and those without, is used to recognize faces in photos [18]. It seems that there is not much advantage to using the Haar Cascade Method other than having better performance. But issues such as a huge drop in accuracy in weak lighting conditions persist.

1. The Materials And Method
2. *Background and Related Works*

A follow-up to the study [11], in which the author employed CNN cascade for face detection. The machine misrepresents the results due to a fault with the face rotating in a different direction. To address this problem and centre the landmarks as much as possible without distorting the image, the author suggests a facial landmark identification method written in Python. They employ FaceNet, a technique that makes use of deep CNN to create a mapping from face pictures to Euclidean space where distances correspond to the face similarity measures, to produce face embeddings. Finally, classification tasks for the faces are performed using a linear SVM. The proposed method can reach an accuracy of 95.02%. In [13], Spatial convolutions, which slide a kernel across the input feature maps, are a common layer action done on the input picture across layers. Pooling and linear or completely linked layers that take the maximum, average, or Euclidean norm over geographic areas employ weighted sums of all input units. Then, the ReLU layer is used as an activation function for CNN neuron output. Finally, a totally linked layer connects the filters between layers.

An empirical evaluation of a CNN-based face recognition system is provided in [19]–[21]. The recommended method stands out for applying batch normalization for the first and last convolutional layer outputs, which boosts the network's accuracy rates. The faces were categorized using the Softmax Classifier at the fully connected layer stage.

In [11], SVM was employed for the final face classification to determine the identities of each face. However, ultimately, CNN rather than SVM was employed as the primary facial recognition approach in this instance. According to [22], the SVM classifier is now the most widely used technique, and it is utilized to address a wide range of classification problems. SVM has the advantage of being particularly powerful when dealing with large datasets with multiple dimensions. SVM is also useful for dealing with exceedingly high-dimensional data [22]. To improve the accuracy and speed of three-dimensional face identification, the research [23] proposes a three-dimensional face recognition method that combines LBP and SVM. The LBP method is used to extract the feature data from the three-dimensional face-depth picture, which is then classed using the SVM algorithm. The study can show that the algorithm requires less time and has a greater recognition rate.

In [14], a model for the employment of a face recognition approach, Principle Component Analysis (PCA), and Convolutional Neural Networks (CNN) to construct an automated attendance system for students in a class is developed. Following this, a comparison of the identified faces with the database of students' faces should establish a connection between them possible. This approach will be an effective method for keeping track of students' attendance and records. PCA is used in facial recognition to reduce the number of variables. In PCA, each picture in the training set is represented by a linearly weighted eigenvector based on eigenfaces. Eigenfaces, which are faces reduced to a condensed collection of core properties, are the fundamental structural components of learning visuals [14]. Once recognition is complete, an individual is sorted by comparing their current position in eigenface space with the positions of identified people by expecting another picture in the eigenface subspace. The key benefits of PCA for facial recognition are its simplicity, speed, and capacity to adapt findings in response to changes in the human face [14]. Furthermore, PCA beats other algorithms in real-world applications with greater recognition rates and lower false positive rates [12]. In [24], an artificial NN with backpropagation and PCA may be used to build a 3D face recognition system with an accuracy of no more than 95% and a training time of just 9728 seconds. The model recognizes faces more quickly overall. Iterations, overall training time, and face recognition speed are all impacted by the quantity of learning rate levels and hidden nodes.

According to [25], LDA was chosen because it simply takes a little number of processing resources, and the classifier was chosen based on this necessity. As a consequence, the server's load was not too high during classifier training or when several students accessed it at once to process attendance. In order to achieve high face recognition accuracy and short processing time, a classifier was only used in one particular course. The recommended attendance system employed LDA to reach a face recognition performance of 97.29% and only needed 0.000096 seconds for the server to complete the face recognition operation, according to the testing findings [25]. According to [26], LDA may be utilized for face identification as well as extraction of facial characteristics, however, small sample issues are frequently encountered while doing so.

The Eigenface Algorithm is better at recognizing faces because it matches the characteristics rather than the full face [27]. Additionally, it provides the opportunity to find the perfect match despite certain facial changes that do not alter the proportion of the features as they truly are. Additionally, by overcoming the system's incapacity to differentiate between identical twins and unfamiliar faces, the override option made accessible to the lecturer enhances its capabilities. Additionally, [22] employed the Eigenface algorithm that is based on PCA. The algorithm will encode the image and compare the results to the previously acquired decoded image. Eigen's face accuracy is a disadvantage, and it is only suited for photos with frontal faces [22]. The PCA technique and the EigenFace database are used to offer a solution [28]. However, there are various restrictions on how facial recognition systems may be used, including issues with image quality, size, the angle of the face, and illumination. Techniques like histogram equalization, illumination invariant, and PCA are utilized to get around these problems.

The Local Binary Pattern (LBP), a simple yet effective texture operator, adds a threshold to each pixel's neighbours in order to identify it in an image. By merging the Local Binary Pattern (LBP) algorithm with advanced image processing techniques including Contrast Adjustment, Bilateral Filter, Histogram Equalization, and Image Blending, [29] provided a novel way for improving face recognition accuracy. The facial recognition system's overall accuracy will increase when the LBP codes are updated. The experimental results of the study imply that our technique is highly reliable for facial recognition systems that might be applied in real-world situations as an automatic attendance system [29]. Patterns of Local Binary In contrast to PCA based on [30], which requires rebuilding both the Eigenfaces and mean each time a new student is introduced to a class, the LBPs technique may dynamically refresh the classifier model with fresh face images. The computing and storage requirements for correlation algorithms are very high. As a result, the face recognition system requires feature reduction and face representation. Accordingly, LBPH is often the favoured approach in computer vision, image processing, and pattern recognition, and it is suitable for feature extraction since it depicts the texture and structure of an image [31], [32]. Radius, Grid X, Grid Y, and Neighbours are the four parameters used by LBPH. The author will next do the LBP operation, which thresholds each pixel's 3-by-3 neighbours with the value of the focus pixel and uses the result as a parallel number to label the pixels in an image, after which the approach will be practiced.

In [33], a template of the acquired faces is constructed using the Fisherfaces Algorithm and the Fisher Linear Discriminant (FLD). When the light changed but the face did not, the technique was designed and used with a 54.17% accuracy rate verification. When comparing diverse facial expressions during the verification and subsequent changes in light during registration, the system's accuracy was 70.83%. Despite the fact that these algorithms have significantly quicker performance rates, 70% accuracy at peak performance is still not very good. Using more advanced face recognition and classification algorithms like the Fisher face algorithm can maximize the separation between classes throughout the training data phase [18].

1. *Process for Start Class Attendance*

In Fig. 1, the Start Class Attendance comprises three main parts: video processing, face detection, and face recognition. This section will be separated into three parts for simplicity of understanding.

|  |
| --- |
| A diagram of a process flow  Description automatically generated |

Fig. 1 Flowchart for Start Class Attendance

1. Video Processing Techniques: There are several different types of video processing techniques, such as video compression and image enhancement, that help to provide a clearer image for face identification and recognition. On the contrary, a simpler and more reliable approach would involve processing every other frame, as the detection and recognition model does not require processing every single frame to operate effectively. By reducing the number of processed frames by half, the model's performance in detecting faces can be significantly improved. In addition to the aforementioned approach, another common strategy to be implemented in this paper is down-sampling the frames. This technique effectively reduces the computational demands for both face detection and face recognition. It's important to note that only the system itself will process the lower-resolution video stream, while users accessing the web application will continue to see the full-resolution video stream, ensuring an optimal user experience. By resolving the computational challenge without compromising the video render resolution, both computational efficiency and user satisfaction can be achieved simultaneously.
2. Face Detection Techniques: At present, a combination of Histogram of Oriented Gradients (HOG) technique and a SVM classifier is employed for face detection in video frames. While the SVM classifier is used to categorize the presence of the face in a particular video frame, the HOG feature descriptor is utilized to identify facial characteristics. For the time being, face detection is performed using a pre-trained model. Because HOG and SVM complement each other so well, they are frequently used jointly for the bulk of object identification tasks, including face detection, which is a well-known task in the field. The combination of HOG and SVM exhibits exceptional robustness against various factors, including illumination variations and face orientations in the video frames. In the next sub-section, we present an alternative face detection strategy called Multi-task CNN (MTCNN), allowing for a comparative analysis of these two approaches. The objective is to identify the face recognition attendance system that best suits the situation, considering factors such as model accuracy and efficiency. By comparing and evaluating both methods, an informed decision can be made to select the most suitable system for the paper.
3. Face Recognition Techniques: Prior to initiating the face recognition process, the program will initially retrieve the face encodings from the database. The system utilizes a deep learning model, specifically a deep Convolutional Neural Network (CNN) that has been trained on diverse datasets, for facial recognition. After a face has been detected and forwarded from the aforementioned face detection method, the face recognition algorithm proceeds to encode its facial features in order to generate face encodings. The library utilizes a pre-trained model to generate a 128-dimensional face embedding by encoding the facial features. This face embedding is then matched with the face encodings of known faces stored in a database. If the distance between the two face encodings falls above a specific threshold, the detected face is considered a match to a known face. The approach relies on the pre-trained model's capacity to generalize from a comprehensive labeled face dataset, making it highly effective and reliable. The integration of techniques such as HOG for face detection and Support Vector Machines (SVM) serves to enhance the overall performance of the entire face recognition system pipeline.
4. *Process for Generate Encoding*

The Generate Encoding feature is accessible only on the admin side (Fig. 2). It is responsible for producing face encodings for the faces stored in the database. This function retrieves images from the database and creates a face encoding for each face using the face detection and recognition techniques discussed earlier. The system also checks for duplicate face encodings, indicating that the face images are highly similar. In such cases, the duplicate image is removed. Finally, the database is updated with the newly created face encodings for the images. These face encodings are then utilized for comparison with real-time captured faces using the aforementioned face recognition methods.

|  |
| --- |
| A diagram of a database  Description automatically generated |

Fig. 2. Flowchart for Generate Encoding

1. Results And Discussion
2. *Evaluation of Face Detection Methods*

The testing dataset employed in this study comprises a collection of manually gathered short videos that depict various scenarios involving students entering a classroom. The purpose of including multiple scenarios is to evaluate the resilience and adaptability of the face detection and recognition models when confronted with real-life situations. These scenarios encompass instances where students intentionally evade the camera, students wearing items that partially obstruct their faces, such as spectacles, masks, or hats, multiple students appearing simultaneously, and even situations with low lighting conditions. By incorporating these diverse scenarios, the experiment aims to assess the effectiveness and versatility of the face detection and recognition models in practical, real-world contexts.

The purpose of this experiment is to assess the performance of the face detection methods across various real-world scenarios. The evaluation focuses on the models' effectiveness in accurately detecting and isolating facial segments within the images. This effectiveness is determined by measuring the accuracy of the models, which is calculated as the ratio of correctly detected actual faces to the total number of potential faces detected by the models. Additionally, the time taken by the models to detect faces in each video frame is also considered. The experimental input consists of raw videos capturing different scenarios of students entering a classroom, recorded by a camera. The output of the models includes a list of cropped face images extracted from each frame of the raw video file, a video compilation with rectangles encompassing the potential faces detected by the models, and statistical information such as accuracy and processing time.

Unfortunately, this experiment has certain limitations. One major drawback is the absence of ground truth in using real-world scenarios as the testing set for the models. Consequently, the experiment is unable to capture important metrics commonly utilized in face detection evaluation, such as recall, precision, and F1 score. The decision to opt for real-world scenarios without a ground truth over well-known online datasets for face detection evaluation stems from various reasons. Firstly, online datasets are often not reflective of practical real-life situations, as they are typically pre-processed. These datasets are commonly used for training and evaluating models, resulting in face images of the highest quality in terms of lighting conditions, facial angles, and absence of obstructing objects. In order to thoroughly evaluate the models' effectiveness in the context of the paper, which entails a face recognition attendance system, it is crucial that the models exhibit high robustness and minimize errors. Erroneous data in an attendance system can have severe consequences. Therefore, the experiment's ultimate decision is to manually collect real-world scenarios as the testing datasets to evaluate the efficacy of the detection methods as a deployable face recognition attendance system.

Fig. 3: Overall accuracy of detection methods in different scenario

Referring to Fig. 3, the multiple bar plot demonstrates the overall accuracy of different face detection methods. Notably, the HOG method exhibits the highest overall accuracy compared to other methods. This finding indicates that the HOG method demonstrates robustness and effectiveness in accurately detecting faces across various conditions and scenarios. The proposed HOG approach utilizes a sliding window technique to extract HOG features from an image, which are then input into an SVM classifier for face detection.

Based on the plot, the MTCNN method [34] achieves the second-highest accuracy, slightly surpassing HOG in the "test\_nolook" testing set. This set comprises a video of a student entering the classroom while intentionally avoiding eye contact with the camera. However, the accuracy of the MTCNN method significantly drops in the "test\_hat\_specs" testing set. This set consists of a video where a student walks into the classroom wearing a hat and spectacles. The decreased accuracy reveals the method's limitation in detecting faces when they are partially obstructed by items. This observation suggests that the MTCNN method may not be the most suitable choice for a face detection method in an attendance system, particularly when considering that hats and spectacles are common accessories worn by students. Table 1 contains image frames that are extracted from the "test\_nolook" and "test\_hat\_specs" testing videos, illustrating examples of these scenarios.

Based on the findings presented in Fig. 3, it is evident that the Haar Cascade method [18] performs poorly compared to the other two detection methods. Notably, the method shows a complete failure in identifying any faces in both the "test\_mask" and "test\_hat\_specs" testing sets. The "test\_mask" video involves a student entering the classroom while wearing a mask throughout the entire duration. This experiment provides a comprehensive demonstration of the significant limitations of the Haar Cascade method in real-world scenarios when compared to the other detection methods. The inability of the Haar Cascade method to detect faces accurately, particularly when masks or hats are present, highlights its inadequacy for applications requiring robust and reliable face detection, such as an attendance system.

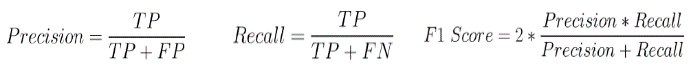
1. *Evaluation of Face Detection Methods*

The overall experiment for the face recognition methods is closely linked to the experiment described for the face detection methods. In this experiment, the testing dataset will consist of cropped images obtained from the output of the HOG detection method, which has been identified as producing the highest quality output among the mentioned detection models. The rationale behind using these cropped images from real-world scenario videos aligns with the earlier explanation, as the objective is to assess the robustness and capabilities of the recognition models in recognizing faces in real-world scenarios rather than relying solely on pre-existing datasets of face images.

Numerous experiments have already been conducted using well-known online datasets, and the results of these experiments are readily available online for the two pretrained face recognition models that will be employed in the testing phase. The Dlib face recognition model achieves an accuracy of 99.38% on the Labeled Faces in the Wild benchmark dataset, while the Facenet Pytorch model [34] claims a 99.65% accuracy after being trained on the VGGFace2 dataset. Therefore, it is more appropriate to evaluate these recognition models using manually collected real-world datasets to gauge their performance in recognizing faces within a classroom environment.

By utilizing the manually collected real-world datasets, the experiment aims to provide a more realistic and representative evaluation of the face recognition models' effectiveness in a practical setting, as opposed to relying solely on well-curated online datasets. This approach ensures that the models are tested against real-world scenarios and can accurately recognize faces within the context of a classroom environment. In order to ensure comprehensive coverage of the student's face encodings, each model will analyze the facial features captured in seven images taken from various angles. Subsequently, the minimal distance between the face encodings and the detected face images created by the face detection model will be calculated.

The output of the recognition models will undergo evaluation based on four commonly used metrics: accuracy, recall, precision, and F1-score. These metrics are widely employed for validating face recognition models. Additionally, the time taken by the recognition models to perform face recognition on the same set of segmented images will also be taken into consideration. The results of these evaluations will be presented in the form of plots in the subsequent results section. Equation 1 below shows the formula for precision, recall and F1-score.

 (1)

Notably, the Facenet model [34] performs well, particularly in the test\_base and test\_nolook datasets, outperforming the Dlib model [7] in terms of accuracy, recall, and F1-score. However, for the test\_nolook dataset, although the Facenet model surpasses the Dlib model by a considerable margin in accuracy, recall, and F1-score, its precision is not perfect (less than 1), indicating the possibility of errors in the final attendance results produced by the Facenet model. In contrast, despite performing comparatively worse than the Facenet model in other metrics, the Dlib model achieves a precision of 1, implying that all the predicted faces are correct, ensuring accurate attendance output.

Turning to the test\_specs dataset, it is evident that the Facenet model fails entirely, as it does not recognize any faces from the dataset, despite the dataset containing decent quality face-segmented images generated by the detection model, as shown in Fig. 5. On the other hand, the Dlib model remains relatively consistent in performance, exhibiting similar rates as observed in the aforementioned datasets, while maintaining a precision of 1.

In the case of the test\_mask dataset, as depicted in Fig. 6, it is anticipated that both recognition models would fail entirely. This is because the student in the dataset is wearing a mask, which obscures a significant portion of the facial landmarks that are crucial for recognition. Consequently, the limited availability of recognizable facial features would hinder both the Facenet and Dlib models from accurately identifying and recognizing the faces in the test\_mask dataset.

A collage of a person's face

Description automatically generatedFig. 4 Overall time spent for each detection model in various datasets

A collage of a person wearing a mask

Description automatically generatedFig. 5 Example face images of the test\_specs dataset

Fig. 6. Example face images of the test\_mask dataset

For the remaining datasets, namely test\_hat\_specs, test\_3ppl, and test\_dim, the results for all four metrics between the Facenet model and Dlib model exhibit similar patterns. The Dlib model consistently outperforms the Facenet model by a small margin. It is worth noting that the Dlib model maintains a perfect precision score across all datasets except the test\_mask dataset, which is the most crucial metric among the four. This signifies that the attendance data produced by the Dlib model will always be 100% accurate, ensuring the reliability of the system.

The main reason why precision is prioritized over other metrics for an attendance system is because a higher precision value indicates a lower chance of false positives. This is crucial in preventing accidental attendance for the wrong individuals and reducing incorrect identifications. In an attendance system, the accuracy of the attendance record is solely dependent on precision, rather than overall accuracy. Even if only one image out of all the images is successfully recognized, which indicates a low accuracy but a high precision, it is considered an accurate attendance record. Moreover, as the number of faces increases, there is a higher chance of the recognition model making mistakes and falsely identifying a particular face as belonging to another individual. This can lead to incorrect attendance records and undermine the reliability and accuracy of the system. Therefore, striking a balance between processing efficiency and maintaining a high level of precision is crucial in designing an effective and reliable attendance system.

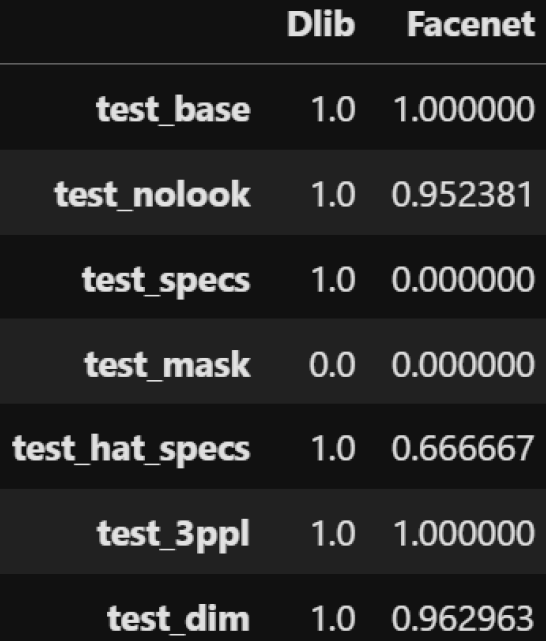


Fig. 7 Overall precision of the recognition methods in different scenarios

Fig. 8: Overall precision of the recognition methods in different scenarios

The ultimate goal is to ensure that the attendance system is accurate and avoids marking the wrong person for attendance. If the recognition model is too sensitive, it may produce incorrect attendance data by associating a face with the wrong individual. This is reflected in a low precision score. Given the significance of precision, Fig. 7 and Fig. 8 displays the overall precision values of the recognition methods across different dataset scenarios. Based on the figure, despite both recognition models demonstrating high precision, the Dlib model outperforms the Facenet model with a perfect precision score. This implies that the attendance produced by the Dlib model achieves 100% accurate in real-world scenarios, which is suitable for a face recognition attendance system.

While the Facenet model is more optimized and efficient in recognizing faces due to its architecture, it is important to prioritize the integrity and accuracy of the attendance record over processing speed. A higher precision is crucial to ensure that the output attendance record is fully accurate, even if it means sacrificing time efficiency. The entire attendance system would be rendered useless if the final output of the attendance record contains erroneous data. Thus, prioritizing precision over efficiency is justified to maintain a reliable and accurate face recognition system. Minimizing the risk of errors or false positives is paramount in a face recognition attendance system, and therefore, the trade-off of longer processing time for increased accuracy is a reasonable decision.

The current procedural framework delineates the task into distinct phases, specifically, detection and recognition. This segmentation is grounded in the principle that, should the initial detection phase fail to identify a face, the system will seamlessly transition to the subsequent frame for further face detection endeavors. In instances where facial detection proves unsuccessful, the recognition process is deliberately omitted in order to mitigate computational resource expenditure and enhance overall system performance. This bifurcated approach ensures an optimized utilization of resources by circumventing redundant recognition attempts on frames where facial presence remains undetected, thereby streamlining the overall operational efficiency of the system.

1. Conclusion

This paper has successfully researched and developed an efficient and accurate attendance system. The HOG with SVM is used for the face detection model and deep CNN is chosen for the face recognition model. Nevertheless, there are rooms for improvement in this research work. The face recognition attendance system may be improved by contrasting various face detection and recognition models. The selected model is suggested to have the capacity to process a large amount of data, work within the specified financial restrictions, and reliably identify and track individuals.

References

[1] P. Secreto and E. Tabo, “Impact of Synchronous Class Attendance on the Academic Performance of Undergraduate Students,” *International Journal of Educational Management and Development Studies*, vol. 4, no. 1, 2023, doi: 10.53378/352968.

[2] D. Qutishat, R. Obeidallah, and Y. Qawasmeh, “An Overview of Attendance and Participation in Online Class During the COVID Pandemic: A Case Study,” *International Journal of Interactive Mobile Technologies*, vol. 16, no. 4, 2022, doi: 10.3991/ijim.v16i04.27103.

[3] S. Tachmammedov, Y. K. Hooi, and K. S. Kalid, “Automated multi-factor analytics for cheat-proofing attendance-taking,” in *ACM International Conference Proceeding Series*, 2018. doi: 10.1145/3185089.3185093.

[4] M. CLARO, J. K. LAURE, B. OGDOC, and D. GUPIT, “Class Attendance Using Face Recognition,” *SMCC Higher Education Research Journal*, vol. 2, no. 1, 2019, doi: 10.18868/ccs.02.060120.01.

[5] P. Netinant, N. Akkharasup-Anan, and M. Rakhiran, “Class Attendance System using Unimodal Face Recognition System based on Internet of Educational Things,” in *Proceedings of the 6th IEEE Eurasian Conference on Educational Innovation 2023: Educational Innovations and Emerging Technologies, ECEI 2023*, 2023. doi: 10.1109/ECEI57668.2023.10105374.

[6] R. Karthikeyan, G. Nalinashini, R. Chithrakkannan, K. Mohanraj, and B. Uma Maheswari, “Implementation of GUI for class attendance using face detection and recognition system,” *International Journal of Innovative Technology and Exploring Engineering*, vol. 9, no. 1, 2019, doi: 10.35940/ijitee.A9122.119119.

[7] Pamith Madusanka Kumara, Mehrdad Tahmasebi, and Devika Sethu, “Automatic Attendance Recording System Using Facial Recognition,” *Malaysian Journal of Science and Advanced Technology*, 2021, doi: 10.56532/mjsat.v1i2.12.

[8] C. Annubaha, A. P. Widodo, and K. Adi, “Implementation of eigenface method and support vector machine for face recognition absence information system,” *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 26, no. 3, 2022, doi: 10.11591/ijeecs.v26.i3.pp1624-1633.

[9] M. S. Akbar, P. Sarker, A. T. Mansoor, A. M. Al Ashray, and J. Uddin, “Face Recognition and RFID Verified Attendance System,” in *Proceedings - 2018 International Conference on Computing, Electronics and Communications Engineering, iCCECE 2018*, 2019. doi: 10.1109/iCCECOME.2018.8658705.

[10] W. Zeng, Q. Meng, and R. Li, “Design of intelligent classroom attendance system based on face recognition,” in *Proceedings of 2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference, ITNEC 2019*, 2019. doi: 10.1109/ITNEC.2019.8729496.

[11] M. Arsenovic, S. Sladojevic, A. Anderla, and D. Stefanovic, “FaceTime - Deep learning based face recognition attendance system,” in *SISY 2017 - IEEE 15th International Symposium on Intelligent Systems and Informatics, Proceedings*, 2017. doi: 10.1109/SISY.2017.8080587.

[12] K. Ashritha, K. Ashritha, and S. Bhukya, “Automated Attendance System Using Face Recognition,” *Int J Res Appl Sci Eng Technol*, vol. 10, no. 6, pp. 2096–2099, Jun. 2022, doi: 10.22214/ijraset.2022.44212.

[13] S. Bhattacharya, G. S. Nainala, P. Das, and A. Routray, “Smart attendance monitoring system (SAMS): A face recognition based attendance system for classroom environment,” in *Proceedings - IEEE 18th International Conference on Advanced Learning Technologies, ICALT 2018*, 2018. doi: 10.1109/ICALT.2018.00090.

[14] S. Sawhney, K. Kacker, S. Jain, S. N. Singh, and R. Garg, “Real-time smart attendance system using face recognition techniques,” in *Proceedings of the 9th International Conference On Cloud Computing, Data Science and Engineering, Confluence 2019*, 2019. doi: 10.1109/CONFLUENCE.2019.8776934.

[15] S. Khan, A. Akram, and N. Usman, “Real Time Automatic Attendance System for Face Recognition Using Face API and OpenCV,” *Wirel Pers Commun*, vol. 113, no. 1, 2020, doi: 10.1007/s11277-020-07224-2.

[16] A. Raj, A. Raj, and I. Ahmad, “Smart Attendance Monitoring System with Computer Vision Using IOT,” *Journal of Mobile Multimedia*, vol. 17, no. 1–3, 2021, doi: 10.13052/jmm1550-4646.17135.

[17] D. D. Nguyen, X. H. Nguyen, T. T. Than, and M. S. Nguyen, “Automated Attendance System in the Classroom Using Artificial Intelligence and Internet of Things Technology,” in *Proceedings - 2021 8th NAFOSTED Conference on Information and Computer Science, NICS 2021*, 2021. doi: 10.1109/NICS54270.2021.9700991.

[18] J. Harikrishnan, A. Sudarsan, A. Sadashiv, and A. S. Remya Ajai, “Vision-Face Recognition Attendance Monitoring System for Surveillance using Deep Learning Technology and Computer Vision,” in *Proceedings - International Conference on Vision Towards Emerging Trends in Communication and Networking, ViTECoN 2019*, 2019. doi: 10.1109/ViTECoN.2019.8899418.

[19] M. Coskun, A. Ucar, O. Yildirim, and Y. Demir, “Face recognition based on convolutional neural network,” in *2017 International Conference on Modern Electrical and Energy Systems (MEES)*, IEEE, Nov. 2017, pp. 376–379. doi: 10.1109/MEES.2017.8248937.

[20] G. Lou and H. Shi, “Face image recognition based on convolutional neural network,” *China Communications*, vol. 17, no. 2, 2020, doi: 10.23919/JCC.2020.02.010.

[21] W. Liu, L. Zhou, and J. Chen, “Face recognition based on lightweight convolutional neural networks,” *Information (Switzerland)*, vol. 12, no. 5, 2021, doi: 10.3390/info12050191.

[22] M. Lal, K. Kumar, R. H. Arain, A. Maitlo, S. A. Ruk, and H. Shaikh, “Study of face recognition techniques: A survey,” *International Journal of Advanced Computer Science and Applications*, vol. 9, no. 6, 2018, doi: 10.14569/IJACSA.2018.090606.

[23] L. Shi, X. Wang, and Y. Shen, “Research on 3D face recognition method based on LBP and SVM,” *Optik (Stuttg)*, vol. 220, 2020, doi: 10.1016/j.ijleo.2020.165157.

[24] Jonathan, A. Kusnadi, and D. Julio, “Security system with 3 dimensional face recognition using PCA method and neural networks algorithm,” in *Proceedings of 2017 4th International Conference on New Media Studies, CONMEDIA 2017*, 2017. doi: 10.1109/CONMEDIA.2017.8266048.

[25] D. Sunaryono, J. Siswantoro, and R. Anggoro, “An android based course attendance system using face recognition,” *Journal of King Saud University - Computer and Information Sciences*, vol. 33, no. 3, 2021, doi: 10.1016/j.jksuci.2019.01.006.

[26] H. Yang and X. Han, “Face recognition attendance system based on real-time video processing,” *IEEE Access*, vol. 8, 2020, doi: 10.1109/ACCESS.2020.3007205.

[27] R. A. Abbas Helmi, S. Salsabil Bin Eddy Yusuf, A. Jamal, and M. I. Bin Abdullah, “Face Recognition Automatic Class Attendance System (FRACAS),” in *2019 IEEE International Conference on Automatic Control and Intelligent Systems, I2CACIS 2019 - Proceedings*, 2019. doi: 10.1109/I2CACIS.2019.8825049.

[28] E. Rekha and P. Ramaprasad, “An efficient automated attendance management system based on Eigen Face recognition,” in *Proceedings of the 7th International Conference Confluence 2017 on Cloud Computing, Data Science and Engineering*, 2017. doi: 10.1109/CONFLUENCE.2017.7943223.

[29] S. M. Bah and F. Ming, “An improved face recognition algorithm and its application in attendance management system,” *Array*, vol. 5, 2020, doi: 10.1016/j.array.2019.100014.

[30] O. A. R. Salim, R. F. Olanrewaju, and W. A. Balogun, “Class Attendance Management System Using Face Recognition,” in *Proceedings of the 2018 7th International Conference on Computer and Communication Engineering, ICCCE 2018*, 2018. doi: 10.1109/ICCCE.2018.8539274.

[31] M. A. Abuzneid and A. Mahmood, “Enhanced human face recognition using LBPH descriptor, multi-KNN, and back-propagation neural network,” *IEEE Access*, vol. 6, 2018, doi: 10.1109/ACCESS.2018.2825310.

[32] I. M. Sayem and M. S. Chowdhury, “Integrating face recognition security system with the internet of things,” in *Proceedings - International Conference on Machine Learning and Data Engineering, iCMLDE 2018*, 2019. doi: 10.1109/iCMLDE.2018.00013.

[33] K. Okokpujie, E. Noma-Osaghae, S. John, K. A. Grace, and I. P. Okokpujie, “A face recognition attendance system with GSM notification,” in *2017 IEEE 3rd International Conference on Electro-Technology for National Development, NIGERCON 2017*, 2017. doi: 10.1109/NIGERCON.2017.8281895.

[34] F. Schroff, D. Kalenichenko, and J. Philbin, “FaceNet: A unified embedding for face recognition and clustering,” in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2015. doi: 10.1109/CVPR.2015.7298682.