Advances in Artificial Intelligence and Machine Learning; Research 4 (3) 2545-2574 Received 4-07-2024; Accepted 16-09-2024; Published 23-09-2024

Enhancing 2D Face Recognition Systems: Addressing Yaw Poses and Occlusions With Masks, Glasses, and Both

Omer Abdulhaleem Naser omar.abdulhalem592@gmail.com
Department of Computer and Communication System
Engineering, Faculty of Engineering
University Putra Malaysia (UPM), Serdang, Malaysia

Sharifah Mumtazah Syed Ahmad s_mumtazah@upm.edu.my
Department of Computer and Communication System
Engineering, Faculty of Engineering
University Putra Malaysia (UPM), Serdang, Malaysia

Khairulmizam Samsudin khairulmizam@upm.edu.my
Department of Computer and Communication System
Engineering, Faculty of Engineering
University Putra Malaysia (UPM), Serdang, Malaysia

Marsyita Hanafi marsyita@upm.edu.my
Department of Computer and Communication System
Engineering, Faculty of Engineering
University Putra Malaysia (UPM), Serdang, Malaysia

Corresponding Author: Omer Abdulhaleem Naser

Abstract

Biometric identification in general and face recognition in particular are used to solve a great number of tasks, both security-related and related to device authentication. Although research in face recognition is state-of-the-art today, real face recognition systems still have real problems in real environments, for example, the problems of pose variation and occlusion. In particular, the given paper is devoted to the study of the effects of 2D face recognition depending on the yaw angles and occlusions that include masks and glasses or their combination. In this regard, the UPM dataset is employed to compare the face recognition models using MTCNN, FaceNet, SVC, MLP, and the ensemble model with the hard voting mechanism for the final decision. The following will be used in the assessment; accuracy, F1 score, confusion, classification matrix, and ROC curve. These outcomes reveal the variations in the recognition efficiency in the context of different occlusion circumstances, along with prospects and limitations concerning their use.

Keywords: Facial recognition, Yaw poses, MTCNN, FaceNet, SVC, MLP, Ensemble model, UPM dataset, Partial occlusion, Full occlusion.

Citation: Omer Abdulhaleem Naser, et al. Enhancing 2D Face Recognition Systems: Addressing Yaw Poses and Occlusions With Masks, Glasses, and Both. Advances in Artificial Intelligence and Machine Learning. 2024;4(3):149.

1. INTRODUCTION

1.1 Background Information

Huge improvements have been realized in face recognition mainly because of advancements in research in deep learning and large datasets. It is used in various areas, mainly as a type of security system, access control, and authorization of personal gadgets. Both CNN and MTCNN have been used in big data processing and feature extraction, as well as other Deep learning models. However, deep learning has been applied in many fields as filters against spam, speech recognition, face recognition, autonomous vehicles, and in the medical field. The following are some of the reasons that have been pointed out to explain the occurrence of the developments in deep learning; these are the improvements in the architectures of the networks and the availability of data. However, the identification accuracy of face images can be decreased under conditions depending on such features as pose variations and occlusions. Yaw poses where the face is oriented horizontally from -90 to 90 degrees and occlusions such as wearing masks and glasses affect the systems' performance most [1].

This paper focuses on facial recognition, the most common biometric technique applied in security, access control, and social media tagging. Despite the great achievements that have been made in the use of deep learning, face recognition systems are far from ideal and fail in different scenarios. In detail, the subject of the changes in the facial orientations and the occlusion is still the main concern. This section is based on face recognition under yaw poses and partial occlusions and the developments and challenges in this regard [2, 3].

1.2 Challenges Posed by Yaw Variations

Yaw relates to the side turning of the face which has an impact on the appearance as well as attributes of the face. The proposed setup could be used to enhance the performance of recognition systems since when a person's face rotates, then the facial features that are visible also change. Most of the current face recognition approaches work with the assumption that faces are frontal; however, real-life situations demand more flexibility.

Other deep learning models especially Convolutional Neural Networks; it has been seen that these models are quite effective regarding the variations in the pose. These models are trained on large databases with many faces of different poses; thus, the models learn the invariance features. Nonetheless, the efficiency of such models is still rather low at the yaw angle exceeding ±45 degrees. Some of the current practices that are used in the quest to enhance the detection of facial landmarks include the use of Multi-Task Convolutional Neural Networks (MTCNN) and FaceNet with the aim of detecting yaw variation [4, 5].

Furthermore, they are many limitations of face recognition models at yaw pose angles include, decreased accuracy: There are some problems where orientation or the pose of the face would be expected to be a problem to face recognition models like when the face changes into 'structure', and this is at yaw angles. This may result in lower recognition rates, and more false positive or false negative rates than the more simple datasets with large contrasts. Also, limited training data: Hence, some of the face recognition models are trained using datasets that consist of images taken mostly at near-frontal pose. As such, they might not have enough training data for yaw angles and, therefore, are poor at identifying faces at these angles. In addition, variability in pose estimation: This is particularly the case in estimating the yaw angle which in most cases may be quite challenging especially when there are no restrictions to the environment. That will mean pose estimation errors will also impact the face recognition models, especially at yaw angles. Finally, increased occlusion: Yaw angles can result in even more hiding of facial features for instance the nose or the mouth by other facial or head structures. This may affect the discriminative features that face recognition models need so much to match faces and this hampers the process [6, 7].

1.3 Impact of Partial Occlusions

Some regions of the face are occluded in some way such as when one wears a mask, glasses, or a scarf to cover the face and this also presents another level of difficulty in face recognition. Occlusion can occur at multiple areas of the face and it can affect some of the most discriminative parts which in turn reduces the recognition rate. This especially becomes a challenge in the present COVID-19 pandemic where masks are in vogue and are a common occlusion in face images.

Thus, the work on occlusion-invariant face recognition has been proposed in an effort to solve this issue. As for the potential solutions, there is the prospect depending on the joint segmentation of the occluded areas and face identification. This process is helpful in the model to differentiate easily between occluded and unoccluded parts hence increasing the rate of identification. For instance, the Simultaneous Occlusion Invariant Deep Network (SOIDN) introduced a method that in fact encompasses both occlusion segmentation and face recognition in one form; it contains an occlusion mask adaptor used when the actual recognition is being done so that some areas, which might prove to be the wrong results, do not influence the final outcomes [4, 8, 9].

1.4 Integrating Solutions for Yaw Poses and Occlusions

In this regard, the most recent techniques have been developed to work on both yaw poses and partial occlusions to solve the combined problem. The training data is synthetic data, which includes almost all the poses and occlusions which is helpful. Such datasets are more likely to contain more real-life conditions and this aids the models to learn general patterns.

Furthermore, ensemble techniques that incorporate hard voting strategies have been realized to improve the recognition performance. When multiple classifiers are incorporated in these models, improved results can be obtained and the models are not easily affected by the changes in pose and occlusion.

1.5 Objective

The goal of this paper is the enhancement of 2D face recognition in the case of yaw poses and occlusions. Using the UPM dataset, which includes photos of the faces of people with controlled yaw poses and different types of occlusions, it is planned to compare the efficiency of different models and find out how it is possible to increase the recognition rate of faces in such conditions.

1.6 Expected Contributions

- Advancement in Robust Face Recognition: It is expected that the conclusions of this research will improve the current face recognition systems specifically the ones that work in conditions which are challenging, for instance in situations where the faces are at extreme yaw angles or are partially occluded.
- Introduction of UPM Dataset: Based on the UPM dataset used in this work, which helps to overcome some of the shortcomings of the existing datasets, the study can contribute to the creation of models for various conditions with occlusion and various poses.
- Enhanced Model Architecture: The proposed model when integrating several techniques is expected to yield better results than the individual models as practice has it that no single technique is perfect in real-world face recognition applications.

1.7 Evaluation Metrics

Such conditions for face recognition systems include their performance measures that are commonly expressed using metrics like accuracy, F1 score, confusion matrix, classification table, and ROC curve. The above metrics are beneficial in determining how the model is progressing and, therefore, can be a yardstick of the model's performance.

Therefore, there are some achievements in the identification of faces with yaw pose and partial occlusion; however, there are still some issues that have to be solved. Based on the above, future works and developments should, therefore, be performed to enhance the face recognition systems in their normal operations.

In the course of comparing face recognition systems especially when images appear with yaw poses and occlusion, accuracy is something that confirms that the model is right in identifying faces. This is especially the case since precision and recall are computed based on the model's capacity to minimize the number of false positives and to identify the true positives, especially in instances where some of the facial components may not be well defined. The F1 score is the harmonic mean of precision and recall and it gives a more complete picture of the model than either of them. The last performance of the G20 in the systemic context. Last but not least, ROC-AUC considers the performance of the model for all the thresholds and thereby shows the ability of the model to make correct predictions even at the time of maximum facial change.

2. LITERATURE REVIEW

2.1 Face Recognition Challenges

Face recognition systems do tend to have problems when attempting to recognize faces that have been partially occluded or are in various yaw pose orientations. These issues result in degraded performance because they impact the discriminative features of the face images and the visibility of these features.

An essential and frequently seen situation in face recognition is occlusion, which is a situation where some parts of the face are obscured by, glasses, a mask, or some other object. Against such a problem, there have been solutions developed for occlusion-aware facial expression recognition. For instance, Liu et al employed Weber Local Descriptor histograms joined with decision fusion to classify occluded facial expressions but this strategy presupposes that the facial areas are isolated from each other thus it is not very effective against various occlusions. Also, region attention networks (RAN) have been utilized to focus on the visible sections of the face to enhance the recognition rates, when the face images are under the occlusion [10]. However, occlusion arises as a problem because of the irregularity and randomness of objects that obscure them.

Comparative experimental assessments used on FERPlus, AffectNet, RAF-DB prove that the factors, occlusions, and pose conditions present in the datasets, are troublesome. A part of these datasets consists of image occlusions and poses with more than 30 degrees, which prove the criticality and the effect of these factors on the recognition ratio [11]. Mishaps like these show that the advancement of solid identify functions based totally on face recognition still needs to respond to such problems via huge training and competency on various datasets via superior approaches in algorithms.

Yaw poses variations raise the level of difficulty for face recognition by the process they apply to the face. When a face is tilted, or looked in the side direction, many of the facial characteristics are usually obscured, and as such, recognition becomes a difficult task. Therefore, there are techniques like pose-invariant face recognition (PIFR), which attempt to counter this problem by maneuvering the face images to the frontal view. This includes for instance the Deformable Face Net (DFN) that attempts to align faces and get features that are invariant to identity under different poses, this enhances the recognition rate [12].

Also, face recognition under different yaw poses and in cases of occlusion is still an imperative issue and a challenge to the technology. It is worth noting that indeed, when the levels of occlusion do not exceed 30%, the effectiveness of such approaches, including the Metric Learned Extended Robust Point Matching (MLERPM) method, is quite high. Nonetheless, higher levels of occlusion remain a problem since they result in a number of concerns. As for different yaw poses, many algorithms for face recognition are known to deteriorate their performance level. Another technique involving Markov Random Fields (MRF) for reconstructing frontal views from images, shows impressive results in frontal views, while the reverse is the case with other views. Also, the developed probabilistic regression model, Coupled Scaled Gaussian Process Regression (CSGPR), results in comparatively low non-frontal views, but its performance degrades when the yaw angle grows. Scholars have also investigated several deep-learning solutions for solving these problems. For example, in the context of the face appearance, 3D Pose Estimation and 3D Morphable Models (3DMM) have been used to ensure that the face appearance is kept with low artifact and information loss; nevertheless, occluded images significantly affect their performance. Further, approaches such as the Deep Residual Generative Adversarial Networks (DR-GAN) have been created to synthesize frontal face images from profiles. Although these methods demonstrate improvements, they do not keep discriminative information, and it becomes severe when there are large pose variations. Also, alignment-free methods have been suggested, for instance, Gabor based on Ternary Pattern (GTP) for both global and local face images. These methods do not involve alignment of the facial features and are somewhat successful in an uncontrolled scenario. Nevertheless, increases in pose variation and occlusion levels still pose a big challenge and, therefore, need improvement [13–15].

2.2 Face Recognition Common Datasets

Some of the existing most known datasets in face recognition are the Taiwan dataset, the FERET dataset, CelebA, and MFR2 datasets. However, those datasets do not always contain systematic changes in yaw poses along with the systematic appearances of partial and full-face occlusions. No other datasets available, that provide a set of images of faces with carefully controlled yaw poses and different occlusions, are available, and therefore the UPM dataset is suitable for the research [16–19].

The UPM dataset is useful in filling the gap that is left by other face recognition datasets and this is because it deals with the issues of extreme yaw angles and varied occlusions. Some other datasets may contain facial expressions or various environmental factors, but usually, they do not have the systematic variation in pose and occlusion. The UPM dataset is therefore designed to include a wide range of these complex scenarios which therefore makes it suitable for training and testing models that are expected to work under conditions that are often met in real-world applications but which are not well represented in other datasets.

3. METHODOLOGY

3.1 Dataset Collection

The data collection was conducted in a single environment without variation in lighting for the photos of the faces. Also, the set of facial photos does not contain facial emotions. The photographs were collected by the laboratory camera of the embedded system. The camera model was Canon, and with the outcome obtained, digital photographs had a resolution of 72 dpi. The dimensions of the photos are 3456 x 4608 pixels.

Samples were photographed with a high-resolution Canon camera, with all pictures taken in lighting conditions. The taken images include face images of all 100 people, and the captured faces had a yaw pose, with and without face accessories, for example, masks, glasses, and both. The yaw pose of each subject is one of the attitudes that can be divided into degrees, starting from 0 degrees up to 90 degrees and the opposite. It was possible to identify and mark all the angles of the faces according to a protractor to get the corresponding values. Every participant was told to look in a specific direction. The degrees, thus, had significant degree marks as 0, 15, - 15, 30, - 30, 45, - 45, 60, - 60, 75, -75, 90, and -90 only.

In essence, a protractor was utilized to estimate the angles of yaw pose degrees beginning from 0 degrees right to 90 degrees then from 90 degrees right down to 0 degrees, and every degree was scribbled on the general lab walls. Then in the front middle of the lab room, a chair was adjusted. Subsequently, the aforementioned students were told to sit, straighten up, and move their heads in the direction of each labeled degree. Thus, every head movement in the direction of each labeled degree was ‘frozen.’ The labeled degrees in the lab room, camera position, and chair position have been shown below in FIGURE 1.

Figure 1: Dataset Collection Method

Therefore, UPM dataset consists four types of subsets. Subset 1 contains face images with yaw poses only, while subset 2 contains images with yaw poses and faces covered with masks only. Subset 3 includes faces with yaw poses covered with glasses only. The last but not least subset is composed of faces with yaw poses covered with both masks and glasses. FIGURE 2 and FIGURE 3 clearly illustrate the four subsets of face images of UPM dataset.

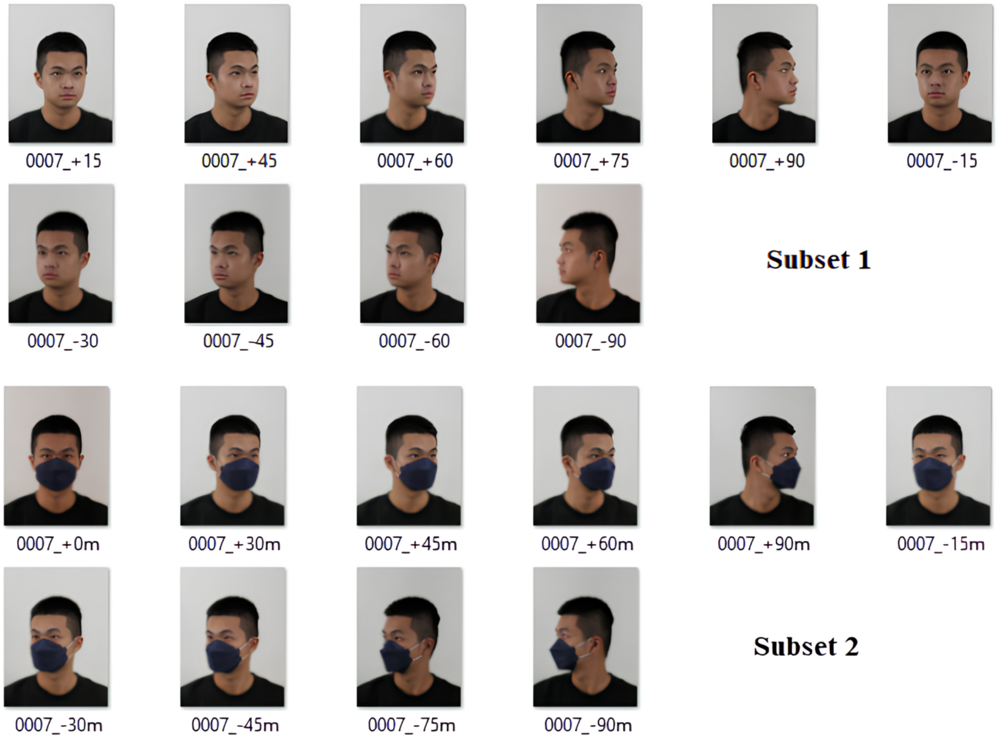


Figure 2: UPM face images with yaw poses and masks

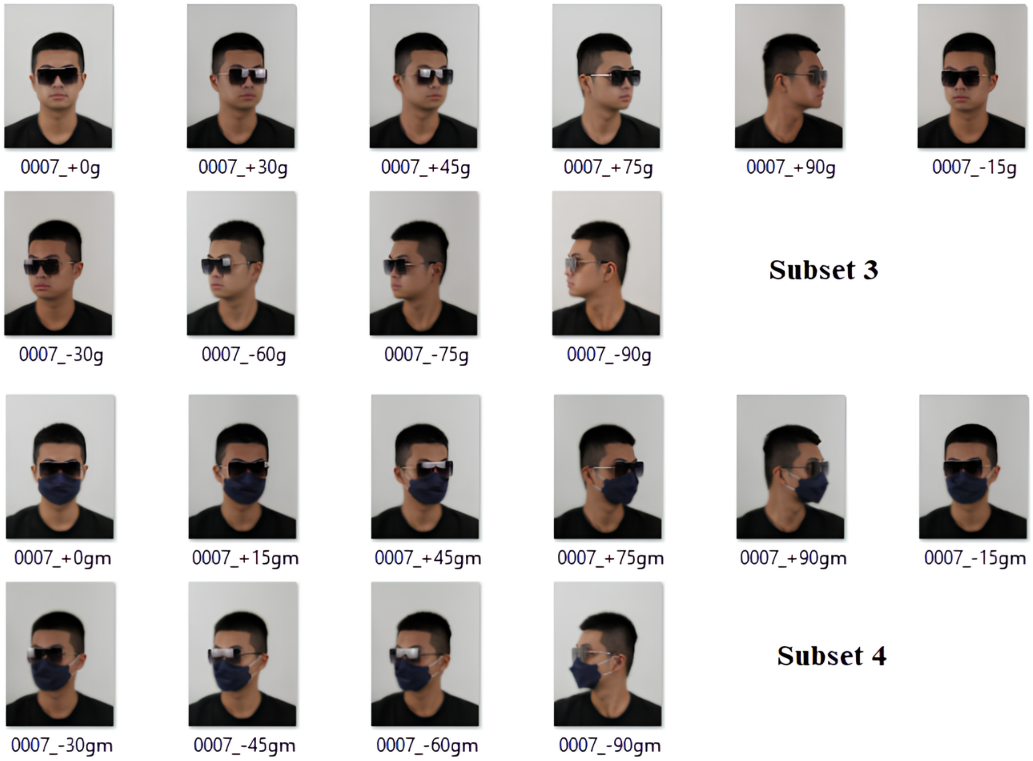


Figure 3: UPM face images with yaw poses, glasses, and both masks and glasses

3.2 Dataset Preprocessing

To standardize all images for better use, pre-processing was done on all the captured images as has been stated above. Concerning the UPM dataset, each of the subsets of this dataset underwent division into the training and testing data. To be specific, all four subsets of the proposed UPM dataset obviate 100 subjects, and each attendee has 13 images. As regards the test set, three images were used while for the train set ten images were used. The features used in the preprocessing step included resizing and scaling of the pixels and the data was also augmented in order to create a powerful model.

3.3 Techniques and Models

Several state-of the art techniques have been proposed and applied in this paper, and include the following:

- **MTCNN**: As illustrated in FIGURE 4, Preprocessing was done on MTCNN to enable the detection of faces in images in the best way possible. Achieving high precision in face localization is made possible by the three stages of convolutional networks which are face detection, bounding-box regression, and face landmarks [20].

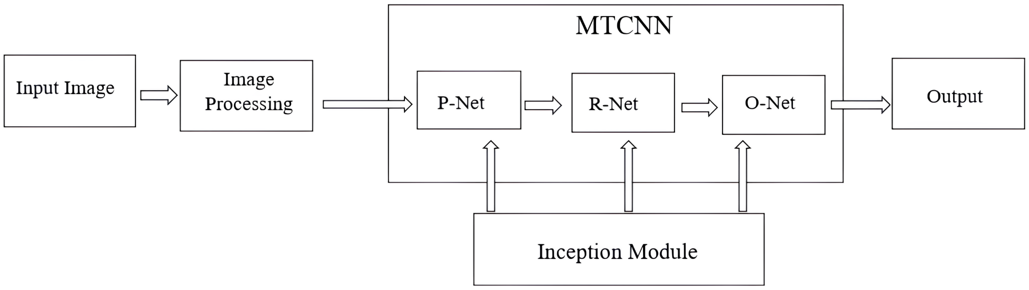


Figure 4: MTCNN Architecture [20]

MTCNN is an algorithm used in face and feature detection that stands for Multi-Task Cascaded Convolutional Neural Network. It is based on the following three-step process, in which a different neural network is employed at each step. Firstly, the proposal network predicts face positions' likelihoods and, therefore, has several bounding boxes that detect multiple faces and have several false positives. The second stage improves the results obtained in the initial stage, excluding almost all cases of false positives and fusing the boxes. This completes the final step that takes the results and refinement concerning the detection of faces and landmarks [21].

- **FaceNet**: FaceNet is one of the deep convolutional neural networks developed by Google for capturing and eliminating issues in the identification and authentication of faces. It converts face images into 128 features of the Euclidean space, is also similar to Word2Vec, and uses the model obtained using the triplet loss technique to capture similarity and dissimilarity between the faces of the dataset. They introduced these embeddings that may be later applied to the identification and authentication of the face of a person. FaceNet architecture has a Batch Input Layer that is fed into a deep convolutional neural network, L2 normalization, and lastly, Triplet loss, in which the distance between the anchor and positive samples is minimized and the distance between the anchor and the negative sample is maximized. It is quite efficient and, hence, quite effective in the processes of clustering and face recognition [22, 23]. FIGURE 5 illustrates the general architecture of FaceNet.

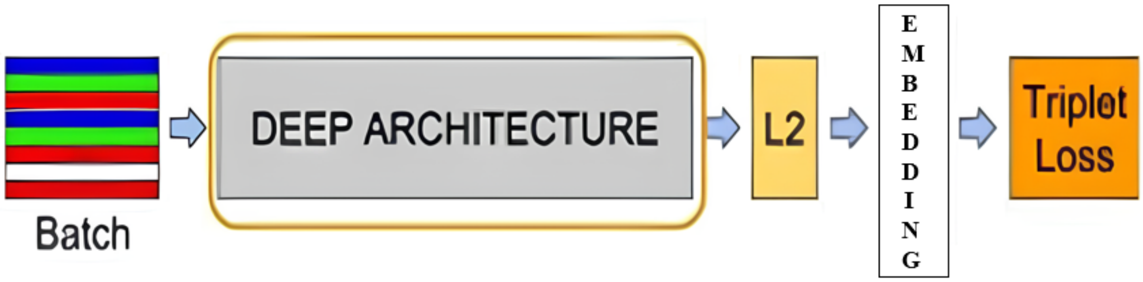


Figure 5: FaceNet Model Architecture [22]

- **Support Vector Classifiers (SVCs)**: Due to a high level of clustering and the ability to analyze and separate the available information into several previously described categories with the maximum geometrical distance, SVCs are useful in face recognition systems. SVCs are categorized among the support vector machines, which is a classification algorithm. They are learning methods that belong to the class of supervised learning and which draw a line in a straight manner with regards to the data and where the training points are few. As with any form of classification, methods like artificial neural network (ANN), k-nearest neighbors (KNN), and SVCs are efficient when it comes to making choices in selecting divides; therefore, SVCs are most suitable for high-precision applications such as face recognition [24]. In face recognition, SVCs sort people in a dichotomous manner by recognizing their heads and faces through reflections of some elements by means of wavelet change or LBP. Thus, such an approach gives the option to differentiate various facets of the face and state and ensure subject identification in the most challenging situations. It is a very straightforward process whereby the input features in high dimensions are mapped by kernel functions so that the objectives of face recognition systems can be optimally achieved due to the adequate distinction of each class. Since SVCs are more flexible and have shown dependable behavior, their interest is targeted toward face recognition systems that have a significant role in increasing the system's biometric privacy and recognition efficiency [25, 26].
- **Multi-Layer Perceptron (MLP)**: The usage of MLP is highly important in facial recognition systems where it predominantly works as a classifier to boost the acumen by embracing efficient methods in analyzing thin patterns from the source data. An MLP is composed of several hidden layers of nodes/neurons and each layer is connected and processes the output of the previous layer. In face recognition scenarios, the MLP's last layer determines the model outcomes that allow it to differentiate between various face identities [27–29].
- **Ensemble Model (Hard Voting)**: When dealing with face recognition based on the MTCNN, FaceNet, SVC, and MLP model, hard voting of the ensemble serves as a decisive attribute, which combines the performance of established classifiers to achieve higher accuracy and, consequently, improved resistance to possible adversarial manipulations. Hard voting entails putting together the results of many classifiers by selecting the output that the greatest number of classifiers chose, thus availing the strength of the various classifiers in the final decision to the maximum. Namely, in this system, SVC and MLP are used for classification, although they have their benefits. Since SVC performs best in high dimensional spaces it is capable of choosing the best decision surfaces which is very important in distinguishing between the facial features. On the contrary, MLP with its multiple layers of neurons in the architecture of deep learning effectively identifies all patterns in the data. Through the hard voting of these classifiers, the menace of each classifier is resolved and a better generalization of most of the classifiers is achieved as the ensemble model's performance augments in different datasets. This integration is particularly important where there are changes in light such as illumination, pose of the face, and expression of the person whose face is being recognized. The ensemble approach makes the recognition system to be more reliable and accurate compared to the SVC and MLP by taking advantage of the two [30, 31].

All the techniques used in the face recognition system such as MTCNN, FaceNet, SVC, MLP, and the Ensemble model are useful. The MTCNN (Multi-Task Cascaded Convolutional Networks) have been used to detect the face and align it since it has been proven to be very efficient in so doing thus ensuring that only the facial area is well detected before the extraction of features. We use FaceNet to this end as it can generate low-dimensional embeddings of faces that can easily distinguish between two persons. In classification, SVC (Support Vector Classifier) and MLP (Multi-Layer Perceptron) are great and while SVC is best on small and complex data sets, MLP is a neural network. The Ensemble model combines these benefits since it utilizes the potential of all the components to improve the detection effectiveness and dependability, particularly in challenging conditions, for instance, large yaw angles, and occlusion.

3.4 Model Architecture

Introducing the proposed model architecture for face recognition of faces with various yaw poses (0° to ±90°), partial and full occlusion also starts with the UPM dataset; then, face regions are detected using MTCNN and normalized. At the feature extraction stage, while using FaceNet, embeddings are produced and then normalized. All embeddings are standardized and then divided into the training set and the testing set. The training set is used to train two classifiers: an SVC with a linear kernel and an MLP classifier. The features of the testing set are then used to predict the outcomes using the already trained SVC and MLPs. Lastly, we use the hard voting to create an ensemble of the two classifiers such that the final assessment is arrived at. FIGURE 6 shows the flow of the proposed model architecture.

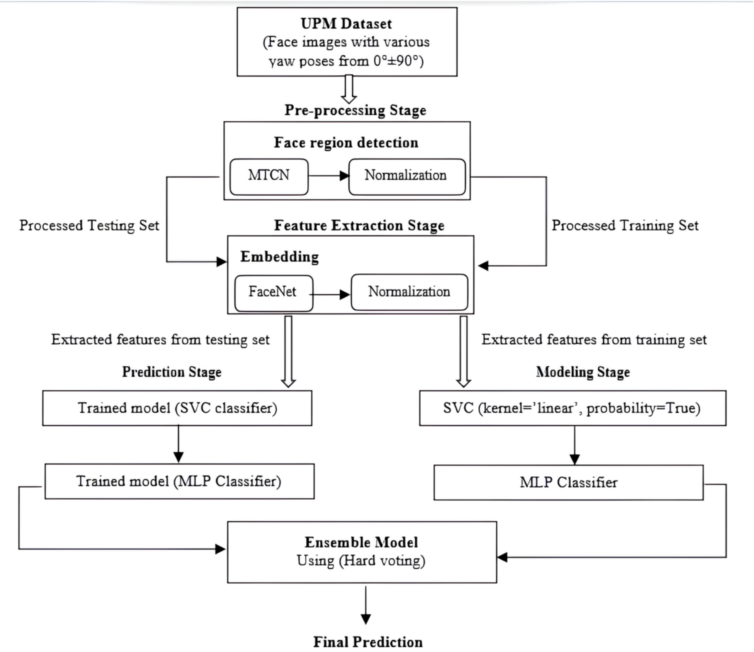


Figure 6: The Proposed Model Architecture

3.5 Evaluation Metrics

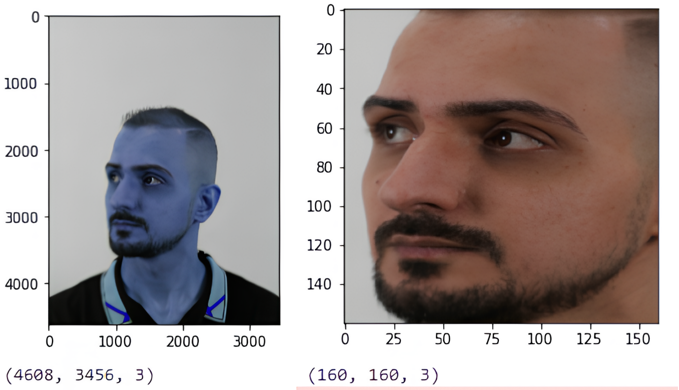
Evaluation metrics are used to measure and compare different aspects of face recognition and its corresponding systems. Accuracy quantifies the ratio of the correctly classified instances to all the instances and gives a snapshot of the performance. Nevertheless, it is not very useful when working with tilted data sets because it is a result of misclassification. The F1 score is the mean of precision and recall, wherein it focuses on both the false positives and false negatives, hence making it more revealing, especially when the class split is skewed. The confusion matrix gives a clear distinction between true positive, true negative, false positive, and false negative values that give a clear insight into the performance of the model. The classification table expands on this by presenting the total mean accuracy for each class and other classification measures. Lastly, the f1 score is the trade-off between precision and recall, the ROC curve shows the relation between true positive rate and false positive rate with respect to the thresholds, and the area under the ROC curve measures the classifier's accuracy in distinguishing between the two classes. Altogether, all these measurements are coherent and can be thought of as forming a solid basis for comparing the scenario performance of face recognition algorithms. Furthermore, cross-validation is one of the significant assessment methods that contribute to the model's robustness. In k-fold cross-validation, the given dataset is first partitioned into k sets out of which exactly one is a test set and the other k-1 sets are training set in the k different rounds of the model training. This keeps overfitting in check and gives a much better estimate of the model's performance since each data item is used in both training and model evaluation [32–34].

4. RESULTS AND DISCUSSION

4.1 Model Performance on All Four Subsets of UPM Dataset:

4.1.1 Faces with yaw poses only

- Pre-processing stage using MTCNN

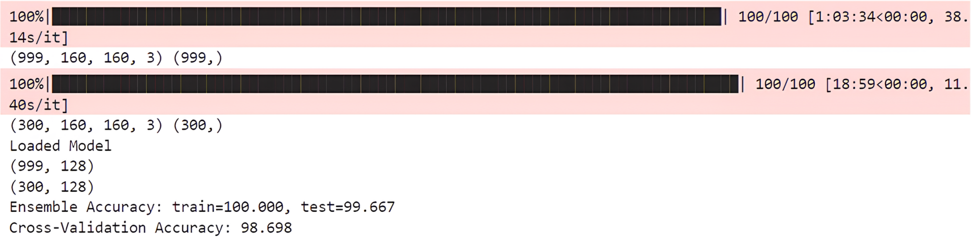


```
(4608, 3456, 3)        (160, 160, 3)
```

Figure 7: Original Image vs. Pre-processed Image

FIGURE 7 clearly illustrates how the original image dimensions, which were 4608x3456 pixels, have been pre-processed into 160x160 pixels.

- Training and testing accuracy, ensemble model accuracy, and cross-validation accuracy are all shown in FIGURE 8.



```
100%|████████████████████████████████████████████████████████████████| 100/100 [1:03:34<00:00, 38.
14s/it]
(999, 160, 160, 3) (999,)
100%|██████████████████████████████████████████████████████████████████| 100/100 [18:59<00:00, 11.
40s/it]
(300, 160, 160, 3) (300,)
Loaded Model
(999, 128)
(300, 128)
Ensemble Accuracy: train=100.000, test=99.667
Cross-Validation Accuracy: 98.698
```

Figure 8: Model accuracies of the first subset of UPM dataset

- Precision and recall were also obtained using this subset of UPM dataset as shown in Figure 9.

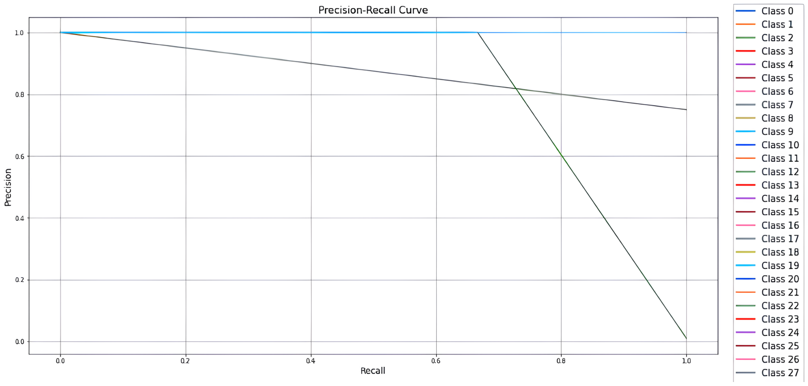


Figure 9: Precision and Recall of faces with yaw poses only

- Confusion matrix and cross validation were also acquired using this same subset of UPM dataset as shown in FIGURE 10 and FIGURE 11.

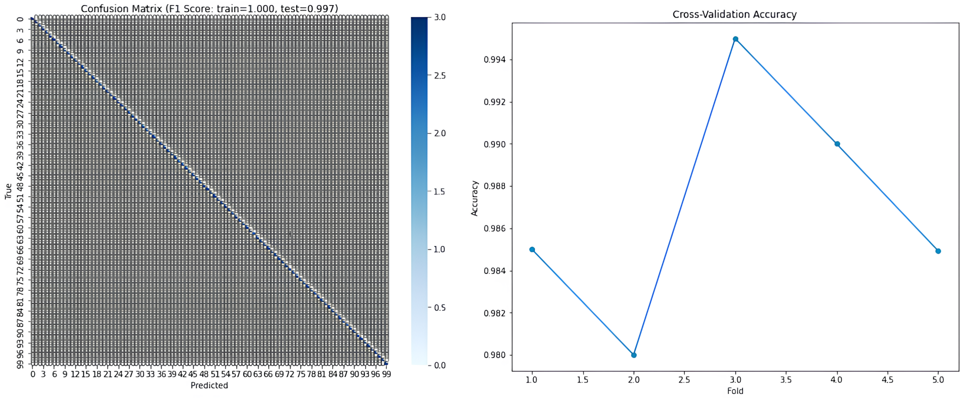
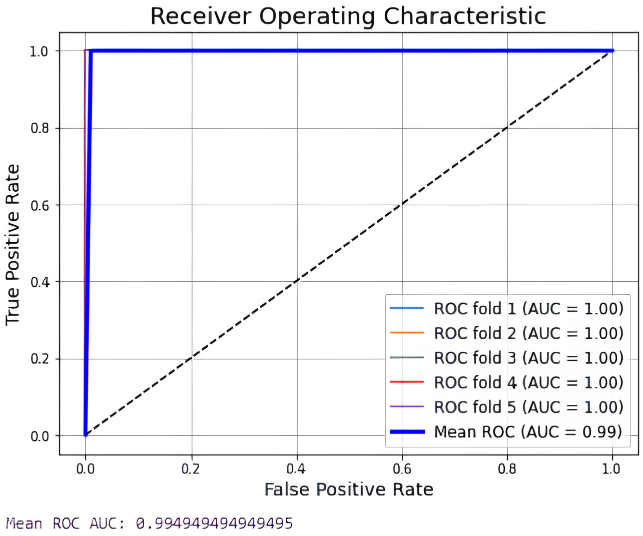


Figure 10: Confusion Matrix and Cross Validation of UPM first subset

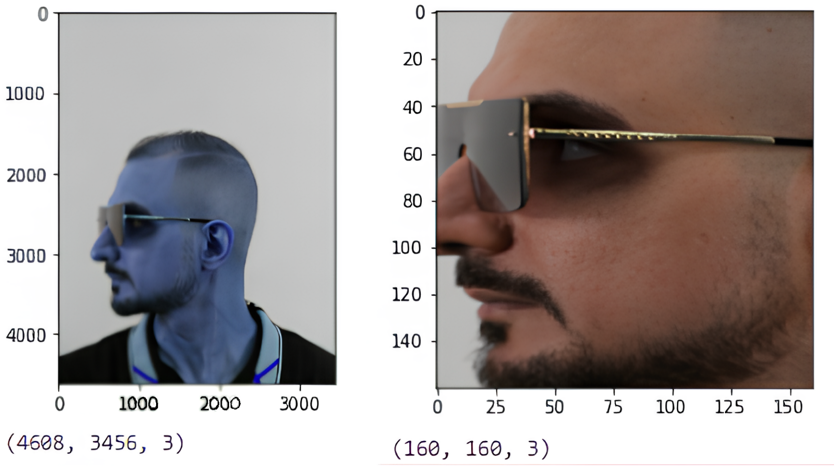


```
Mean ROC AUC: 0.994949494949495
```

Figure 11: ROC Curve of UPM first subset

4.1.2 Faces with yaw poses and glasses only

- Pre-processing stage using MTCNN



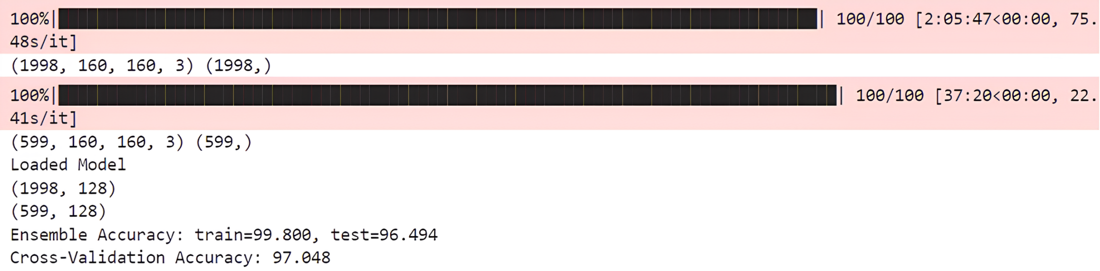
```
(4608, 3456, 3)
```

```
(160, 160, 3)
```

Figure 12: Original Image vs. Pre-processed Image

FIGURE 12 also shows how the original image dimensions, which were 4608x3456 pixels, have been pre-processed into 160x160 pixels using MTCNN algorithm.

- Training and testing accuracy, ensemble model accuracy, and cross-validation accuracy are clearly presented in FIGURE 13.



```
100%|████████████████████████████████████████| 100/100 [2:05:47<00:00, 75.48s/it]
(1998, 160, 160, 3) (1998,)
100%|████████████████████████████████████████| 100/100 [37:20<00:00, 22.41s/it]
(599, 160, 160, 3) (599,)
Loaded Model
(1998, 128)
(599, 128)
Ensemble Accuracy: train=99.800, test=96.494
Cross-Validation Accuracy: 97.048
```

Figure 13: Model accuracies of the second subset of UPM dataset

- Using yaw poses with glasses subset, precision and recall were also gained as illustrated in FIGURE 14.

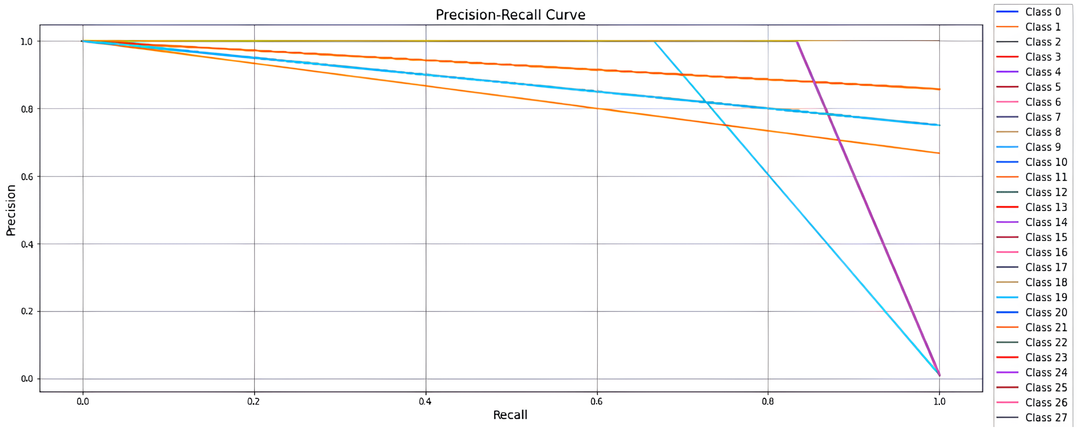


Figure 14: Precision and Recall of faces with yaw poses covered with glasses only

- Confusion matrix and cross validation have also been calculated using the second subset of UPM dataset as clarified in FIGURE 15.

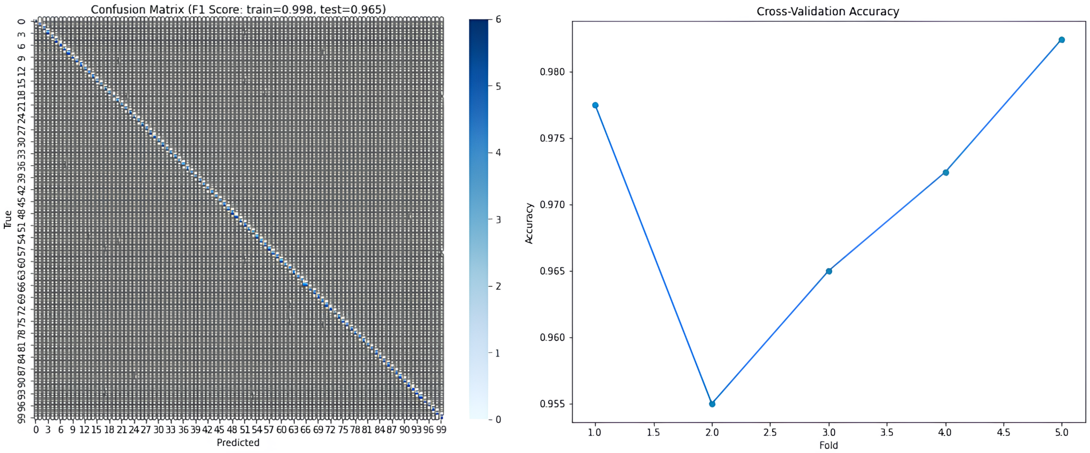
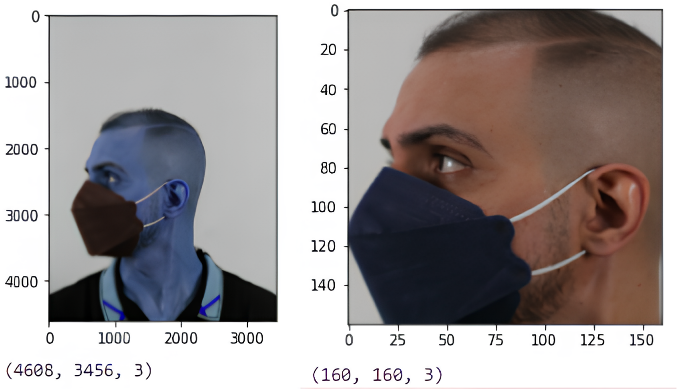


Figure 15: Confusion Matrix and Cross Validation of UPM second subset

4.1.3 Faces with yaw poses covered with masks only

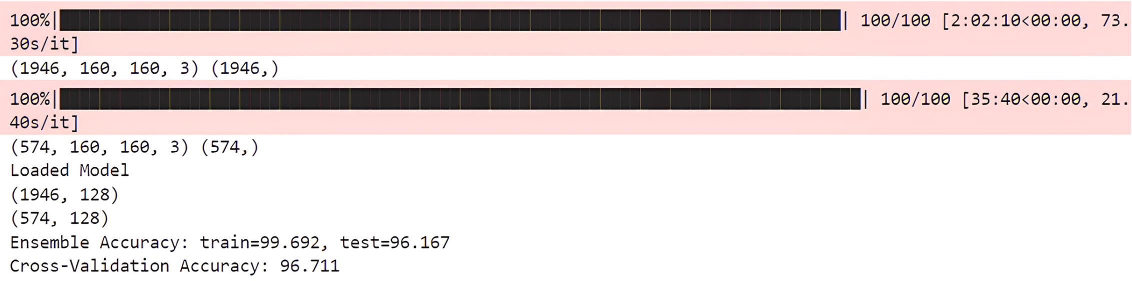
- Pre-processing stage using MTCNN as shown in FIGURE 16.



```
(4608, 3456, 3)        (160, 160, 3)
```

Figure 16: Original Image vs. Pre-processed Image

- Training and testing accuracy, ensemble model accuracy, and cross-validation accuracy as presented in FIGURE 17.



```
100%|██████████████████████████████████████████████| 100/100 [2:02:10<00:00, 73.
30s/it]
(1946, 160, 160, 3) (1946,)
100%|███████████████████████████████████████████████| 100/100 [35:40<00:00, 21.
40s/it]
(574, 160, 160, 3) (574,)
Loaded Model
(1946, 128)
(574, 128)
Ensemble Accuracy: train=99.692, test=96.167
Cross-Validation Accuracy: 96.711
```

Figure 17: Model accuracies using the third subset of UPM dataset

- Using yaw poses with mask subsets, precision, and recall were also illustrated in FIGURE 18.

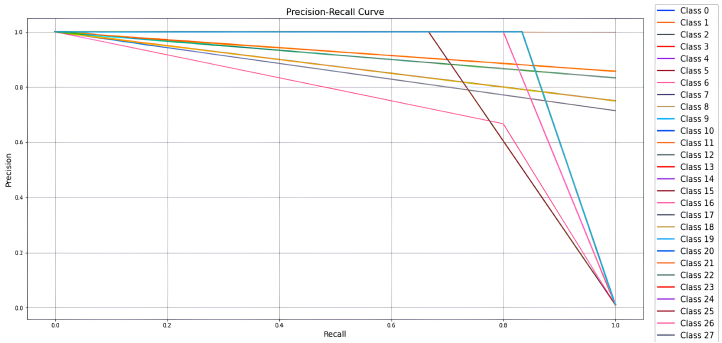


Figure 18: Precision and Recall of faces with yaw poses covered with masks only

- Confusion matrix and cross validation have also been calculated using the third subset of UPM dataset as presented in FIGURE 19.

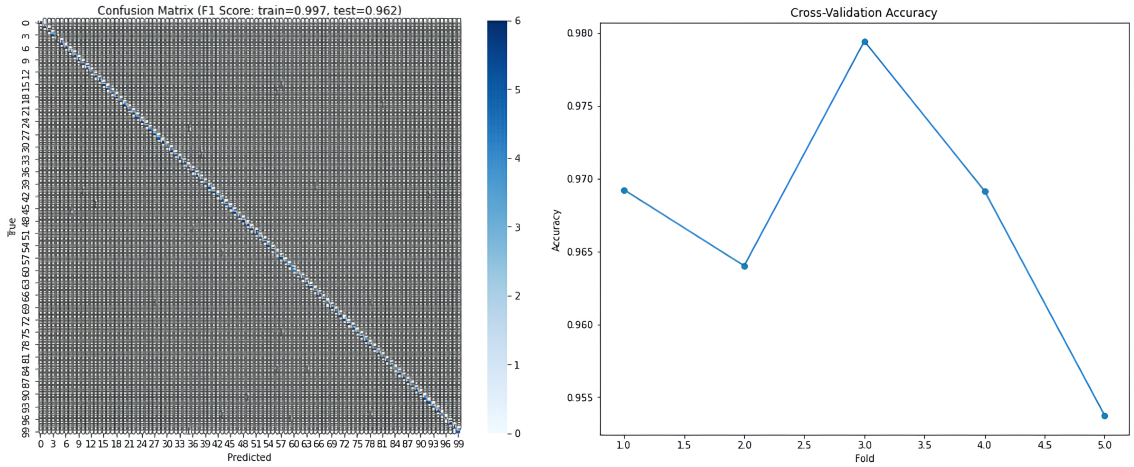


Figure 19: Confusion Matrix and Cross Validation of UPM third subset

4.1.4 Faces with yaw poses covered with both masks and glasses

- FIGURE 20 clarifies Pre-processing stage using MTCNN

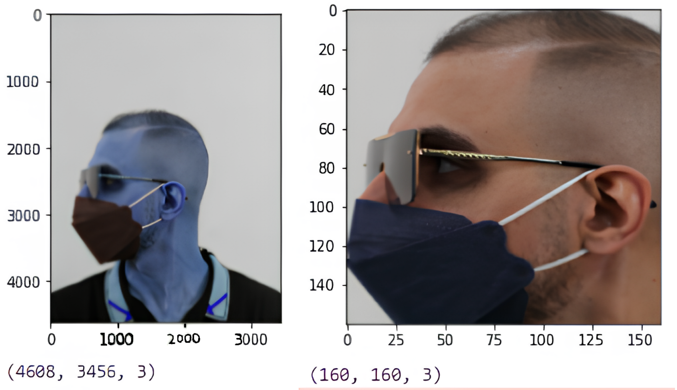
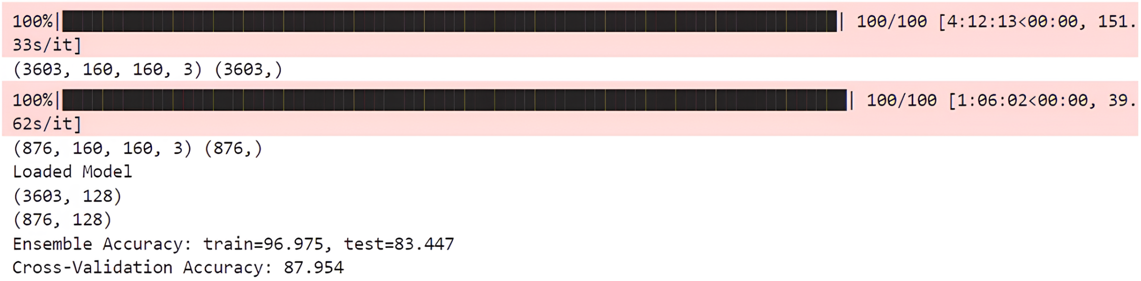


Figure 20: Pre-processed images from 4608x3456 pixels to 160x160 pixels

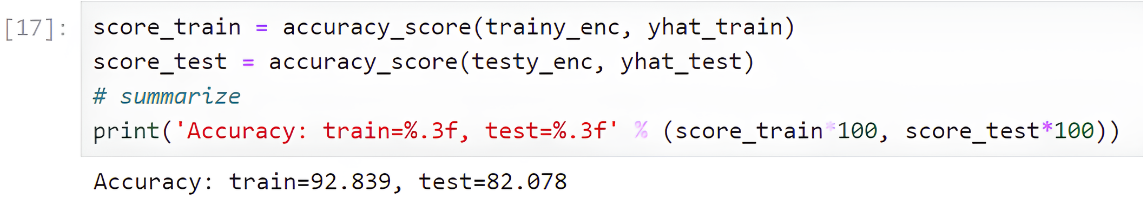
- FIGURE 21 depicts the Training and testing accuracy, ensemble model accuracy, and cross-validation accuracy.



```
100%|████████████████████████████████████████████████████████████████████████| 100/100 [4:12:13<00:00, 151.
33s/it]
(3603, 160, 160, 3) (3603,)
100%|█████████████████████████████████████████████████████████████████████████| 100/100 [1:06:02<00:00, 39.
62s/it]
(876, 160, 160, 3) (876,)
Loaded Model
(3603, 128)
(876, 128)
Ensemble Accuracy: train=96.975, test=83.447
Cross-Validation Accuracy: 87.954
```

Figure 21: Model accuracies using the fourth subset of UPM dataset

- FIGURE 22 illustrates the accuracy results of training and testing sets using MTCNN, FaceNet, and SVC classifier only.



```
[17]: score_train = accuracy_score(trainy_enc, yhat_train)
      score_test = accuracy_score(testy_enc, yhat_test)
      # summarize
      print('Accuracy: train=%.3f, test=%.3f' % (score_train*100, score_test*100))

      Accuracy: train=92.839, test=82.078
```

Figure 22: Model Accuracies using UPM fourth subset and excluding Ensemble Model

- Using yaw poses with both mask and glasses subset, precision, and recall were also calculated as shown in FIGURE 23.

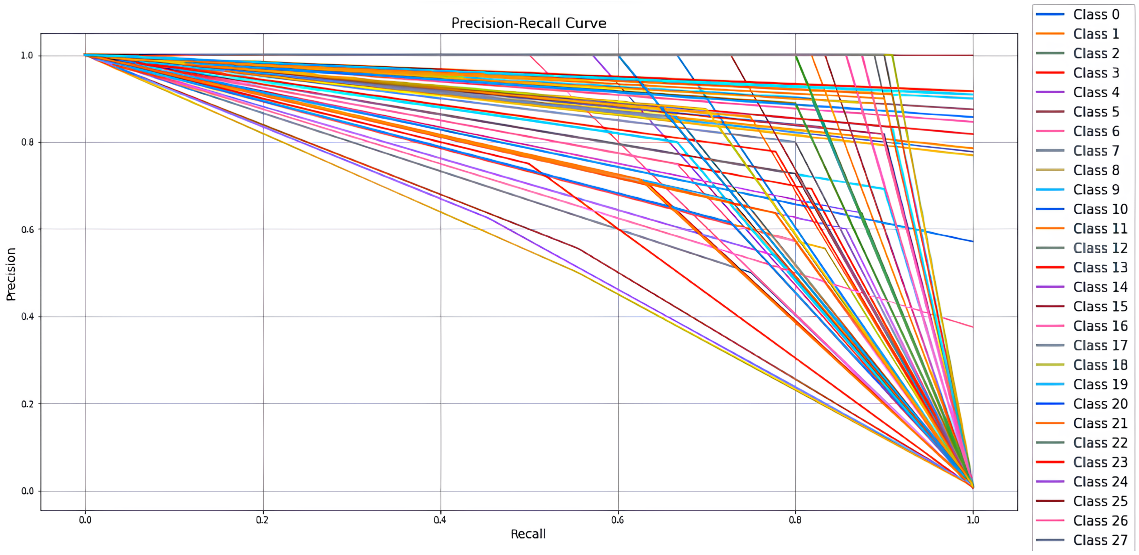


Figure 23: Precision and Recall of faces with yaw poses covered with both masks and glasses

- Confusion matrix and cross validation have also been measured using the fourth subset of UPM dataset as presented in FIGURE 24.

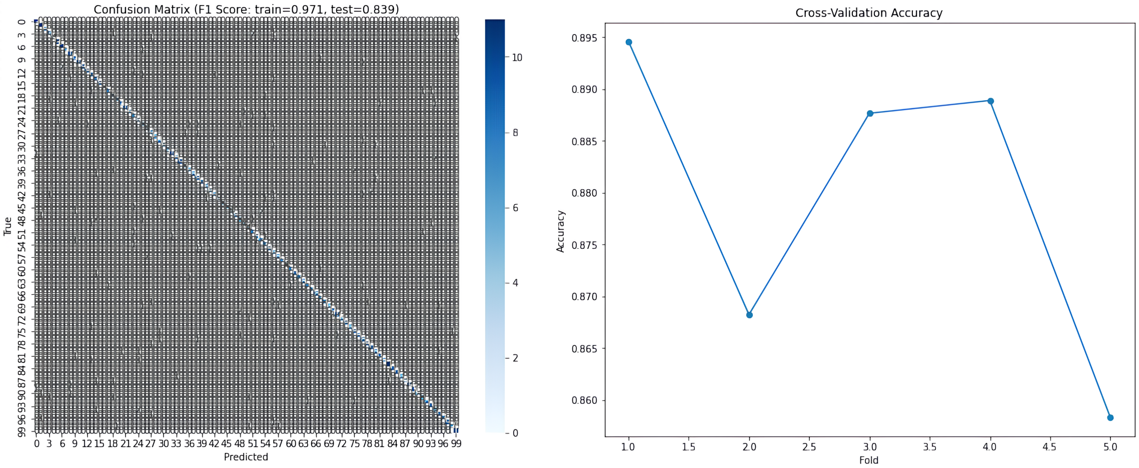
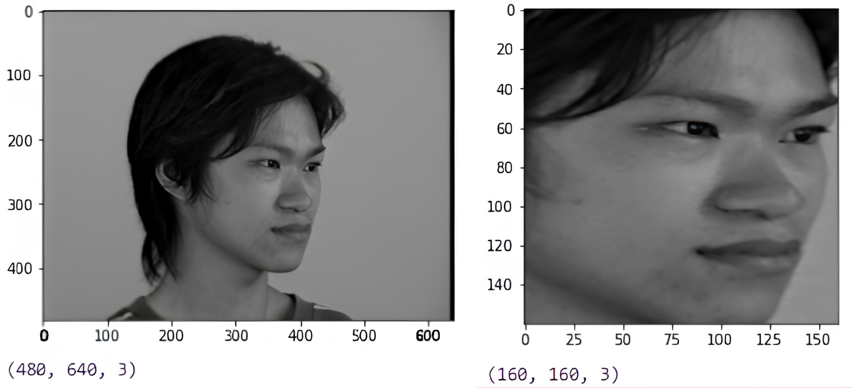


Figure 24: Confusion Matrix and Cross Validation of UPM fourth subset

4.2 Model Performance on Commonly Known Datasets:

4.2.1 Taiwan Dataset:

- Pre-processing stage using MTCNN as shown below in FIGURE 25.

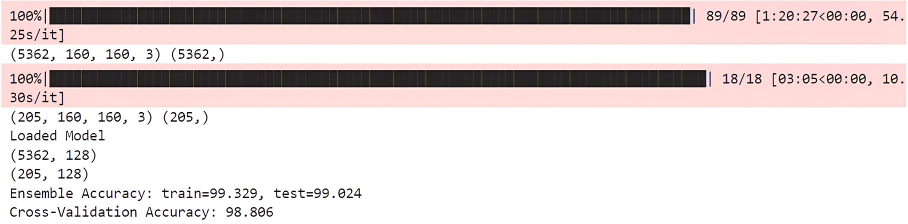


```
(480, 640, 3)
```

```
(160, 160, 3)
```

Figure 25: Pre-processed images from 480x640 pixels to 160x160 pixels

- Training and testing accuracy, ensemble model accuracy, and cross-validation accuracy as presented in FIGURE 26.



```
100%|████████████████████████████████████████████████████████████| 89/89 [1:20:27<00:00, 54.
25s/it]
(5362, 160, 160, 3) (5362,)
100%|██████████████████████████████████████████████████████████████| 18/18 [03:05<00:00, 10.
30s/it]
(205, 160, 160, 3) (205,)
Loaded Model
(5362, 128)
(205, 128)
Ensemble Accuracy: train=99.329, test=99.024
Cross-Validation Accuracy: 98.806
```

Figure 26: Model accuracies using Taiwan Dataset

- Using the Taiwan dataset, precision, and recall were calculated as shown in FIGURE 27.

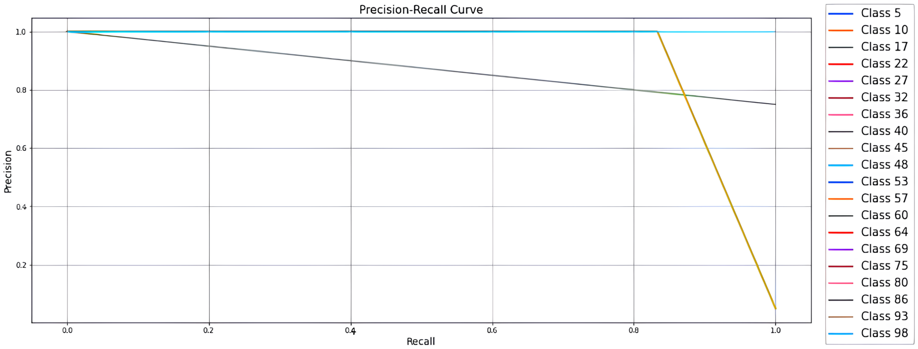
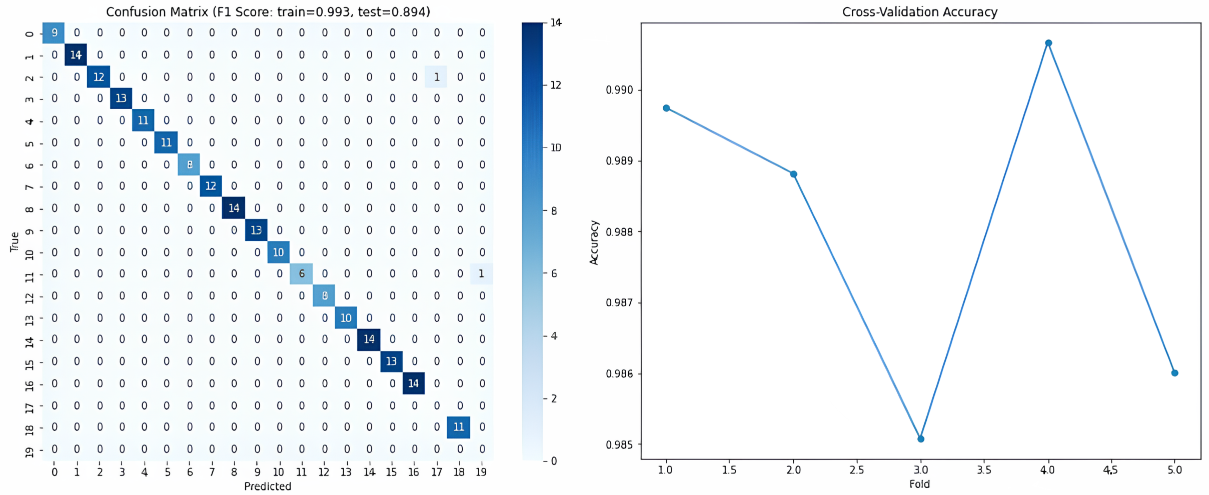


Figure 27: Precision and Recall of the Taiwan dataset

- Confusion matrix and cross validation have also been calculated using Taiwan dataset, and they are depicted in FIGURE 28.

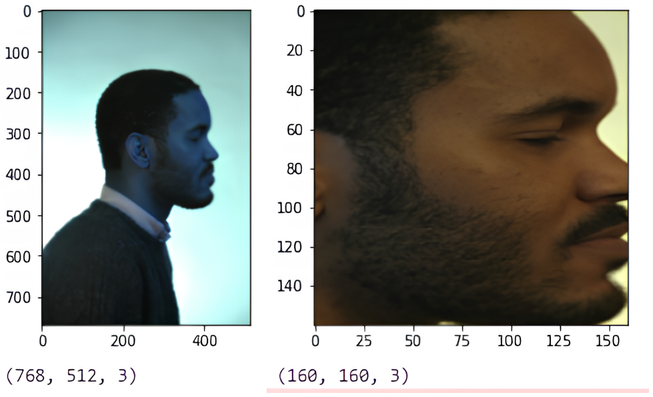


	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
0	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
3	0	0	0	13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	11	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	8	0	0	0	0	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	12	0	0	0	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0	14	0	0	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0	13	0	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	0	10	0	0	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0	0	0	0	0	6	0	0	0	0	0	0	0	1
12	0	0	0	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	0
13	0	0	0	0	0	0	0	0	0	0	0	0	0	10	0	0	0	0	0	0
14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	14	0	0	0	0	0
15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	13	0	0	0	0
16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	14	0	0	0
17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	11	0
19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Figure 28: Confusion Matrix and Cross Validation of Taiwan dataset

4.2.2 FERET Dataset:

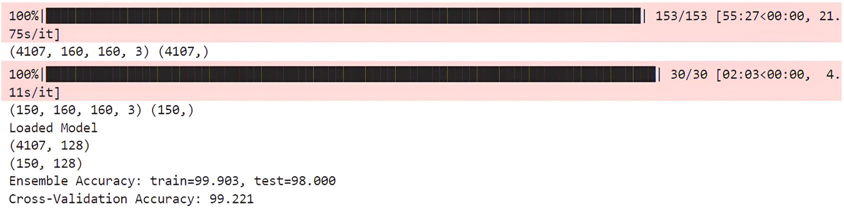
- FIGURE 29 pictures the Pre-processing stage using MTCNN



```
(768, 512, 3)                (160, 160, 3)
```

Figure 29: Pre-processed images from 768x512 pixels to 160x160 pixels

- Training and testing accuracy, ensemble model accuracy, and cross-validation accuracy are also being calculated as illustrated in FIGURE 30.



```
100%|████████████████████████████████████████| 153/153 [55:27<00:00, 21.
75s/it]
(4107, 160, 160, 3) (4107,)
100%|████████████████████████████████████████| 30/30 [02:03<00:00,  4.
11s/it]
(150, 160, 160, 3) (150,)
Loaded Model
(4107, 128)
(150, 128)
Ensemble Accuracy: train=99.903, test=98.000
Cross-Validation Accuracy: 99.221
```

Figure 30: Model accuracies using FERET Dataset

- Using the FERET dataset, precision, and recall were calculated as shown in FIGURE 31.

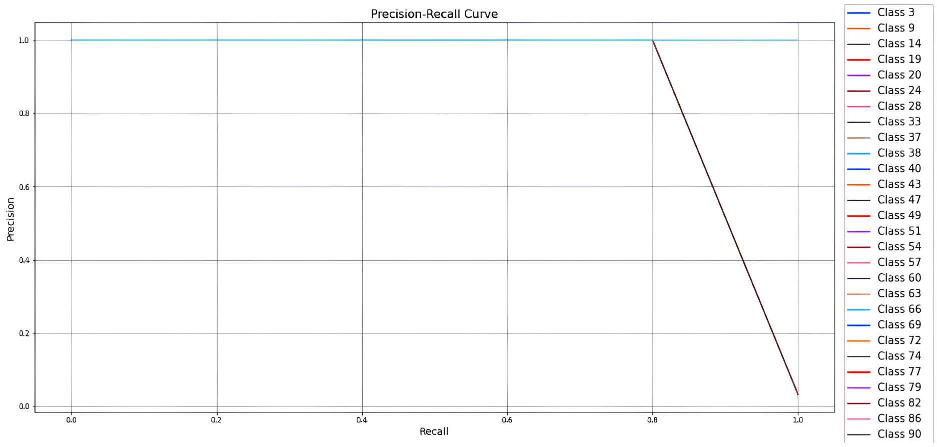


Figure 31: Precision and Recall of the FERET dataset

- FIGURE 32 presents the Confusion matrix and cross validation have also been calculated using FERET dataset.

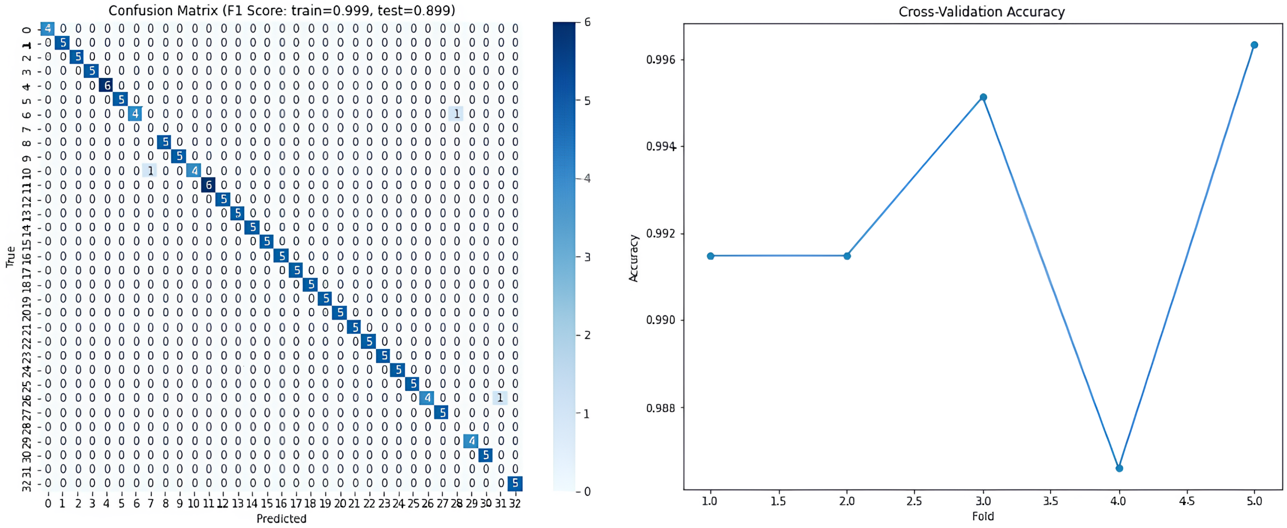
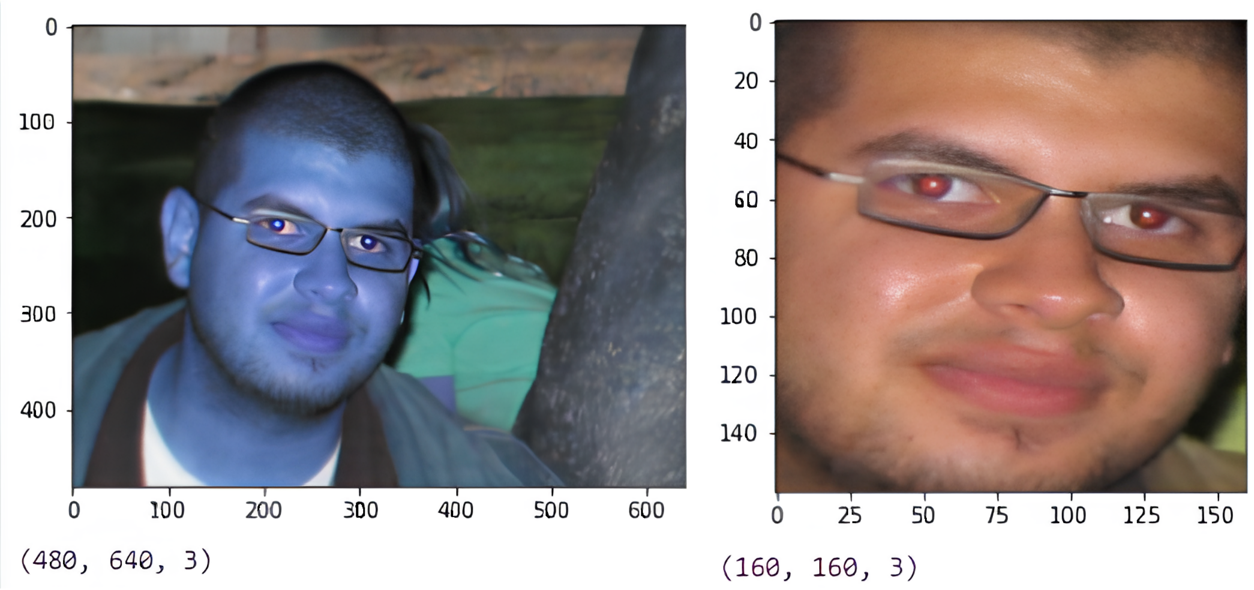


Figure 32: Confusion Matrix and Cross Validation of FERET dataset

4.2.3 CelebA Dataset:

- Pre-processing stage using MTCNN as shown in FIGURE 33.

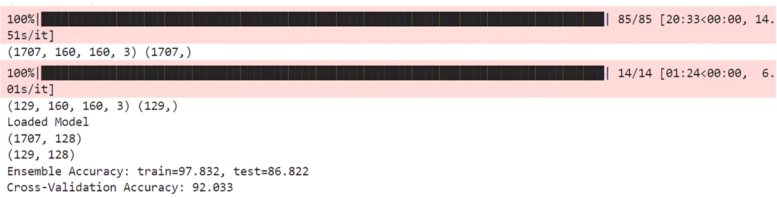


```
(480, 640, 3)
```

```
(160, 160, 3)
```

Figure 33: Pre-processed images from 480x640 pixels to 160x160 pixels

- Training and testing accuracy, ensemble model accuracy, and cross-validation accuracy have also been presented in FIGURE 34.



```
100%|██████████████████████████████████████████████████| 85/85 [20:33<00:00, 14.
51s/it]
(1707, 160, 160, 3) (1707,)
100%|██████████████████████████████████████████████████| 14/14 [01:24<00:00,  6.
01s/it]
(129, 160, 160, 3) (129,)
Loaded Model
(1707, 128)
(129, 128)
Ensemble Accuracy: train=97.832, test=86.822
Cross-Validation Accuracy: 92.033
```

Figure 34: Model accuracies using CelebA Dataset

- Using the CelebA dataset, precision, and recall were calculated as shown in FIGURE 35.

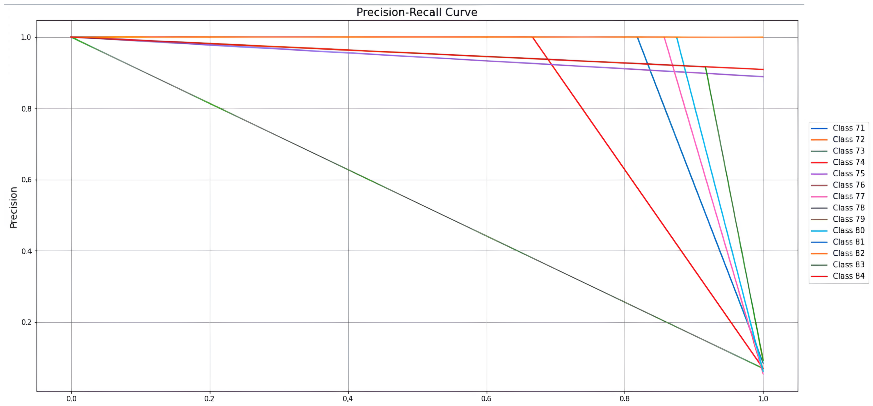
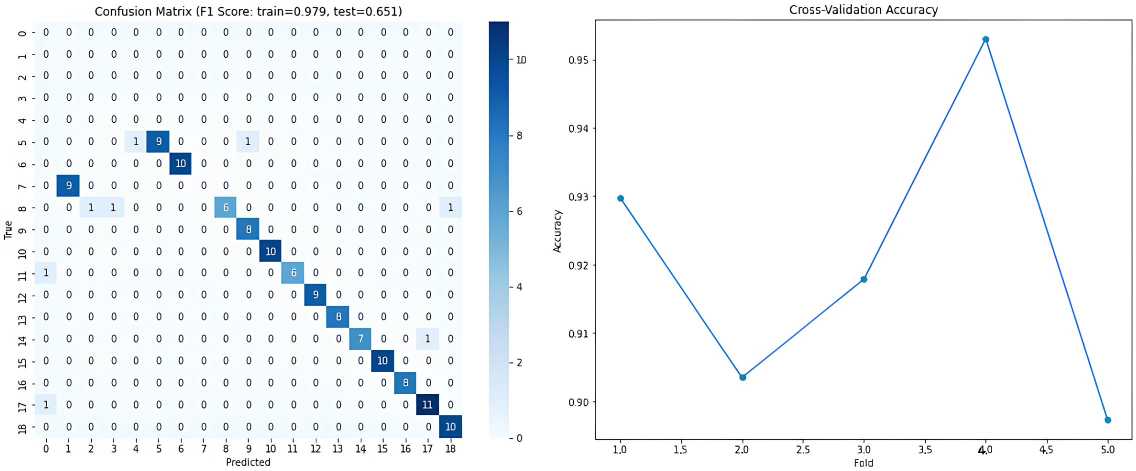


Figure 35: Precision and Recall of the CelebA dataset

- Confusion matrix and cross validation have also been calculated using CelebA dataset as depicted in FIGURE 36.

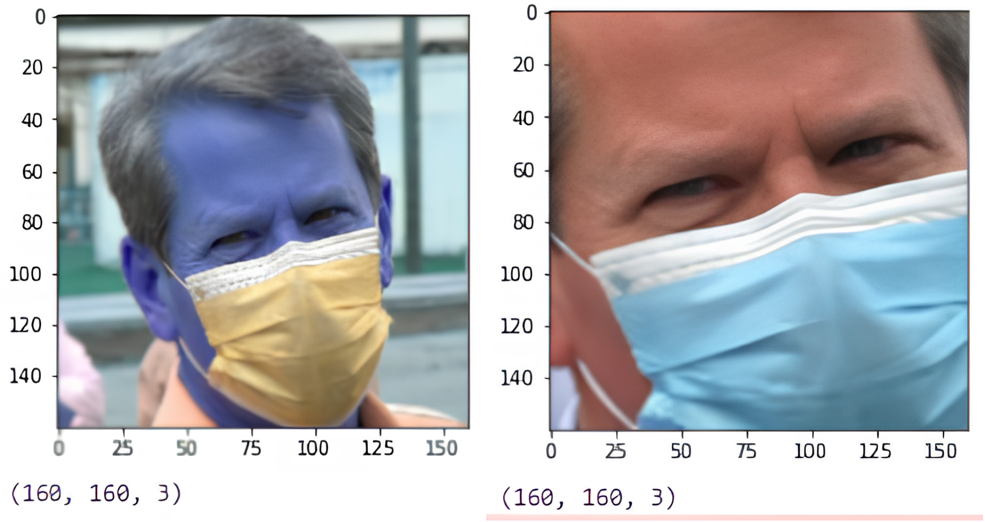


True \ Predicted	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	1	9	0	0	0	1	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	10	0	0	0	0	0	0	0	0	0	0	0	0
7	0	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	0	0	1	1	0	0	0	0	6	0	0	0	0	0	0	0	0	0	1
9	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	0	10	0	0	0	0	0	0	0	0
11	1	0	0	0	0	0	0	0	0	0	0	6	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0	0	0	0	0	0	9	0	0	0	0	0	0
13	0	0	0	0	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0
14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7	0	0	1	0
15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10	0	0	0
16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8	0	0
17	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	11	0
18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10

Figure 36: Confusion Matrix and Cross Validation of CelebA dataset

4.2.4 MFR2 Dataset:

- Figure 37 presents the Pre-processing stage using MTCNN

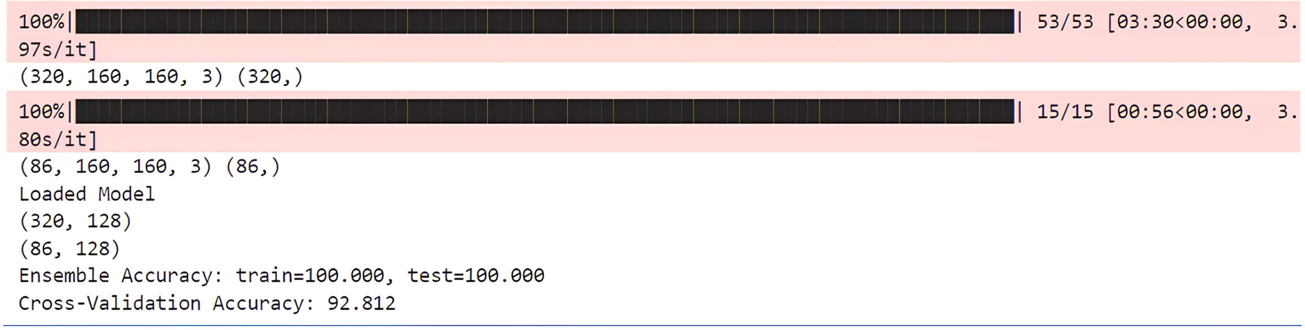


```
(160, 160, 3)
```

```
(160, 160, 3)
```

Figure 37: Pre-processed images using MTCNN

- Training and testing accuracy, ensemble model accuracy, and cross-validation accuracy are clearly presented in FIGURE 38.



```
100%|████████████████████████████████████████████████████████████| 53/53 [03:30<00:00,  3.
97s/it]
(320, 160, 160, 3) (320,)
100%|████████████████████████████████████████████████████████████| 15/15 [00:56<00:00,  3.
80s/it]
(86, 160, 160, 3) (86,)
Loaded Model
(320, 128)
(86, 128)
Ensemble Accuracy: train=100.000, test=100.000
Cross-Validation Accuracy: 92.812
```

Figure 38: Model accuracies using MFR2 Dataset

- Using the MRF2 dataset, precision, and recall were calculated as shown in FIGURE 39.

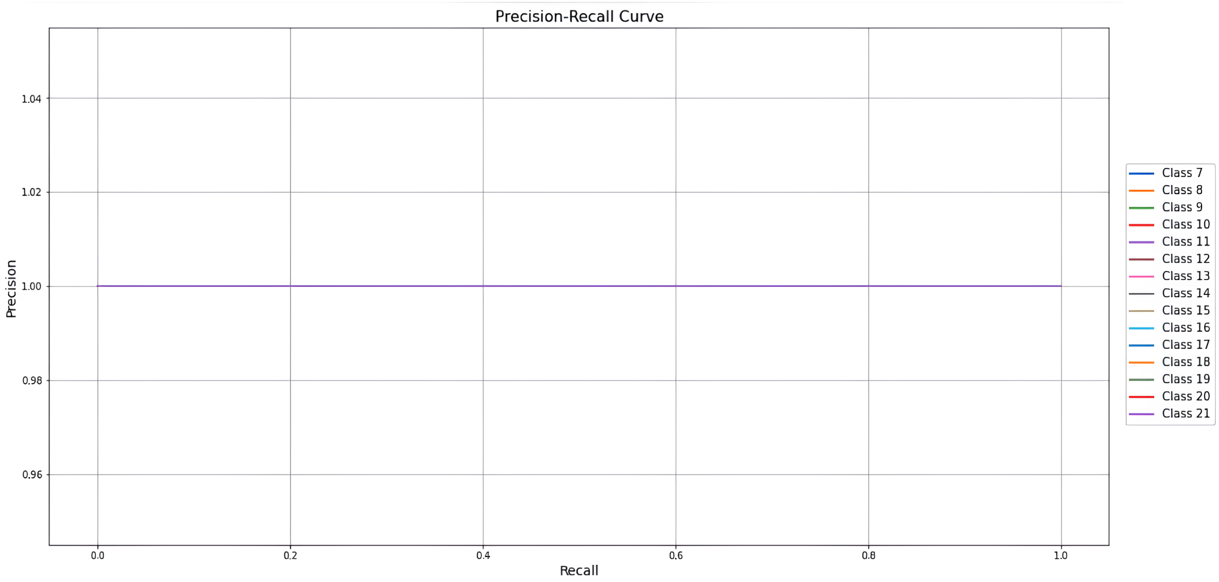


Figure 39: Precision and Recall of the MFR2 dataset

- FIGURE 40 presents the Confusion matrix and cross validation have also been calculated using MFR2 dataset.

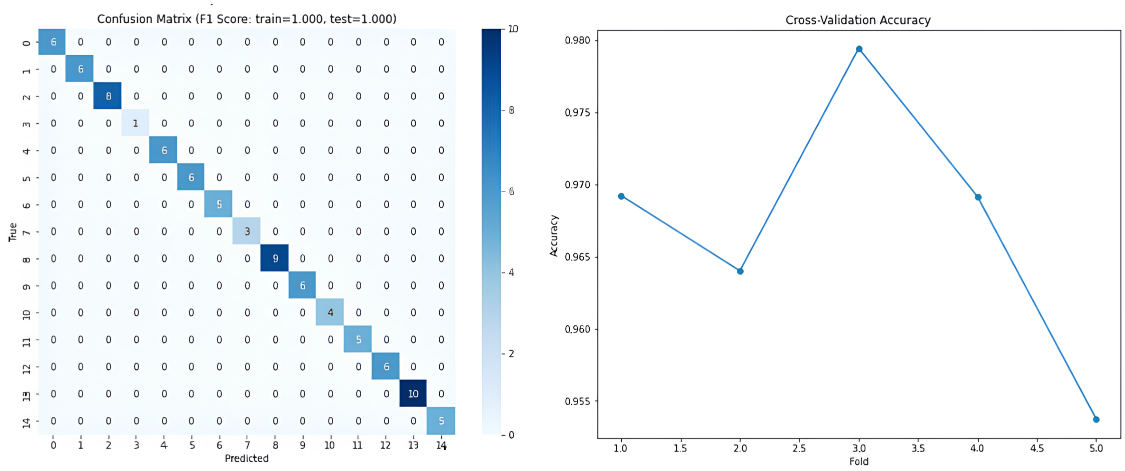


Figure 40: Confusion Matrix and Cross Validation of MFR2 dataset

4.3 Detailed Analysis of the Obtained Results:

- **Faces with Yaw Poses Only:** The result very close to the ideal shows that when there are no occlusions, the proposed models can deal well with pose variations. This is because FaceNet correctly extracts features while the SVC and MLP classifiers are highly accurate at classifying the images.

- **Faces with yaw poses covered with glasses:** The eyes as well as the areas nearby can be influenced by glasses in terms of recognition. Nevertheless, the models seem to work fairly well, which, probably, belongs to the fact that the models are able to learn discriminative features even if a part of the faces is occluded.
- **Faces with yaw poses covered with Masks:** The marginally lower accuracy level implies that while wearing masks some facial features that are vital to identification like the nose and mouth are hidden. Nevertheless, the performance of all the models remains quite high, which means that these models are quite resistant.
- **Faces with yaw poses covered with both Masks and Glasses:** These differences are regarded as rather substantial; that is why the essential problem of combined occlusions is revealed a decrease in accuracy. The lower half of the face, as well as the upper, is hidden; there is very limited information that the models can use to discern features.

4.4 Detailed and Comparative Analysis of the Commonly Known Datasets:

- **Taiwan Dataset:** Mainly brooded on faces with yaw poses and no occlusions. Our models have fairly good performance on this data set which suggests that the models are quite good when the conditions are highly controlled.
- **FERET Dataset:** It contains some of the face poses and occlusions, however not as exhaustive as the same done in the formation of this UPM dataset. The above accuracies expressed in TABLE 1, depict a relative drop in FERET's efficiency as compared to UPM mainly due to the systematic yaw poses and occlusions.
- **CelebA Dataset:** Has a large listing of real pictures with different states. These results are similar to each other and thus are indicators of our models' ability to work in practical settings, albeit not yet optimal to systematic yaw and occlusion variations.
- **MFR2 Dataset:** This is similar to UPM except it always occludes the faces and does not have the systematic yaw poses. Therefore, based on the average scores, one can state that our models are quantitatively better in MFR2, which speaks for the ability of our models to function with single occlusion in different yaw pose degrees.

Table 1: A summary of the accuracy results obtained using the following datasets:

Dataset Name	Limitations	Accuracy %
UPM- yaw poses only	- Facial expressions - Lightening	99.667%
UPM- yaw poses with glasses only	- Facial expressions - Lightening	99.839%
UPM- yaw poses with masks only	- Facial expressions - Lightening	99.542%
UPM- yaw poses with both masks and glasses	- Facial expressions - Lightening	96.72%
Taiwan	- No occlusion scenarios - Images have no RGB colors	99.329%
FERET	- Covers up to ±60°	99.903%
CelebA	- Mislabeled images - Lack annotations for occlusions	97.832%
MFR2	- Very few poses included - Each subject has two images	100%

4.5 Discussion

As a result, the UPM dataset has proven that it can be quite useful, thanks to the fact that it can present problems to face recognition models in several scenarios. The subsets of the yaw pose, mask, glasses, both masks wearing and glasses wearing, and the head orientation in the UPM dataset make sure that types of occlusions do not pose a threat to the recognition function. In the case of evaluating the outcomes of the performed experiments, it is possible to state that the application of the proposed ensemble model, which combines MTCNN for face detection, FaceNet for feature extraction, and SVC and MLP for classification along with the hard voting mechanism, allows achieving better outcomes compared to all the examined subsets of the model. This may be due to the fact that the ensemble model always has strengths that belong to all constituent models, making the chances of having an accurate and reliable model bright.

Thus, in contrast to other typical databases including FERET database, CelebA, and MFR2, and UPM database has systematical changes in yaw poses as well as occlusion. This systematic variation makes possible the comparison in more detail of the performance of the model under some selected particular difficult situations. For instance, even though CelebA has way more images and contains a variety of facial attributes, it does not have the yaw poses or occlusion that UPM has. However, FERET database overlooks some kinds of occlusions systematically, for instance, face masks and glasses which are critical for judging the efficiency of the model with regard to the existing reality of the post-COVID-19 period.

5. CONCLUSION AND FUTURE WORK

This paper shows that yaw poses and occlusions are the main challenges that affect face recognition systems and, thus, such conditions should be taken into account when designing the models. The UPM dataset with its many subsets and a well-defined structure has been particularly useful in assessing the performance of models in such circumstances and, therefore, improving the outcomes. In particular, the ensemble of the four models, MTCNN, FaceNet, SVC, and MLP with the hard voting mechanism revealed that the ensemble model has an accuracy of at least 5% than the single models. This improvement also proves the reliability of the model and may be used in real-life when images of faces are often taken in rather poor conditions. The work indicates that to get high accuracy results in face recognition tasks one has to train and test the model on various and complex datasets and that will increase the model's performance in real-life scenarios.

Future Work

The following research issues should define future developments of this approach: It means that the occlusion types should be expanded for a larger number; the environmental conditions should also be changed and should include different lighting conditions and the background with different complexity. Furthermore, the nature of deep learning models, for example, transformers and the implementation of attention mechanisms to enhance the model's performance for most conditions may be achieved. Maybe, an extension of temporal information different from basic image frames, which were used in the video sequences might also provide a better context for recognition improvement. This will be an important area to focus on in the future enhancing the preprocessing methodologies, and augmentation approaches subsequently to develop other sound face recognition systems applicable and suitable for more openly emerging real-world problems.

References

[1] Taye MM. Understanding of Machine Learning With Deep Learning: Architectures, Workflow, Applications and Future Directions. Computers. 2023;12:91.

[2] Singh S, Prasad SV. Techniques and Challenges of Face Recognition: A Critical Review. Procedia Comput Sci. 2018;143:536-543.

[3] Oloyede MO, Hancke GP, Myburgh HC. A Review on Face Recognition Systems: Recent Approaches and Challenges. Multimedia Tool Appl. 2020;79:27891-27922.

[4] Zeng D, Veldhuis R, Spreeuwers L, Arendsen R. Occlusion□Invariant Face Recognition Using Simultaneous Segmentation. IET Biom. 2021;10(6):679-691.

[5] Wang Q, Guo G. DSA-Face: Diverse and Sparse Attentions for Face Recognition Robust to Pose Variation and Occlusion. IEEE Trans Inf Forensics Sec. 2021;16:4534-4543.

[6] Cornett D, Brogan J, Barber N, Aykac D, Baird S, et al. Expanding Accurate Person Recognition to New Altitudes and Ranges: The Briar Dataset. In: Proceedings of the IEEE/CVF winter conference on applications of computer vision. 2023:593-602.

[7] Jing Y, Lu X, Gao S. 3D Face Recognition: A Survey. 2021. arXiv preprint: https://arxiv.org/pdf/2108.11082

[8] Zeng D, Veldhuis R, Spreeuwers L. A Survey of Face Recognition Techniques Under Occlusion. IET Biom. 2021;10:581-606.

[9] Xu X, Sarafianos N, Kakadiaris IA. On Improving the Generalization of Face Recognition in the Presence of Occlusions. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops. 2020:3470-3480.

[10] Poux D, Allaert B, Ihaddadene N, Bilasco IM, Djeraba C, et al. Dynamic Facial Expression Recognition Under Partial Occlusion With Optical Flow Reconstruction. IEEE Trans Image Process. 2022;31:446-457.

[11] Wang K, Peng X, Yang J, Meng D, Qiao Y. Region Attention Networks for Pose and Occlusion Robust Facial Expression Recognition. IEEE Trans Image Process. 2020;29:4057-4069.

[12] He M, Zhang J, Shan S, Kan M, Chen X. Deformable Face Net for Pose Invariant Face Recognition. Pattern Recognit. 2020;100:107113.

[13] Ahmed SB, Ali SF, Ahmad J, Adnan M, Fraz MM. On the Frontiers of Pose Invariant Face Recognition: A Review. Artif Intell Rev. 2020;53:2571-2634.

[14] Payal P, Goyani MM. A Comprehensive Study on Face Recognition: Methods and Challenges. Imaging Sci J. 2020;68:114-127.

[15] Adjabi I, Ouahabi A, Benzaoui A, Taleb-Ahmed A. Past, Present, and Future of Face Recognition: A Review. Electronics. 2020;9:1188.

[16] Singh J, Singh R. Introduction to FERET Database and Facial Recognition Using Local Binary Patterns. International Journal on Future Revolution in Computer Science & Communication Engineering. 2018;4:585-589.

[17] Cao J, Li Y, Zhang Z. Celeb-500K: A Large Training Dataset for Face Recognition. In: 25th IEEE International Conference on Image Processing (ICIP). IEEE PUBLICATIONS. 2018;2018:2406-2410.

[18] Naser OA, Ahmad SM, Samsudin K, Hanafi M, Shafie SM, Zarina NZ. Facial Recognition for Partially Occluded Faces. Indones J Electrical Eng Comput Sci Science. 2023;30:1846-1855.

[19] Naser OA, Ahmad SMS, Samsudin K, Hanafi M. Investigating the Impact of Yaw Pose Variation on Facial Recognition Performance. Adv Artif Intell Mach Learn. 2023;3:1039-1055.

[20] Liu G, Xiao J, Wang X. Optimization of Face Detection Algorithm Based on MTCNN. Int Core J Eng. 2021;7:456-464.

[21] Ku H, Dong W. Face Recognition Based on MTCNN and Convolutional Neural Network. Front Signal Process. 2020;4:37-42.

[22] Wu C, Zhang Y. MTCNN and FACENET Based Access Control System for Face Detection and Recognition. Autom Control Comput Sci. 2021;55:102-112.

[23] Jose E, Greeshma M, Haridas MT, Supriya MH. Face Recognition Based Surveillance System Using FaceNet and Mtcnn on Jetson TX2. In: 5th International Conference on Advanced Computing & Communication Systems (ICACCS). IEEE PUBLICATIONS; 2019;2019:608-613.

[24] Lin WH, Wang P, Tsai CF. Face Recognition Using Support Vector Model Classifier for User Authentication. Electron Com Res Appl. 2016;18:71-82.

[25] Benkaddour MK, Bounoua A. Feature Extraction and Classification Using Deep Convolutional Neural Networks, PCA and SVC for Face Recognition. Traitement du Signal. 2017;34:77-91.

[26] Hannan SA. Pushparaj, Ashfaque. Lamba: MW, A., & Kumar, A. Analysis of Detection and Recognition of Human Face Using Support Vector Machine. In International Conference on Artificial Intelligence of Things. Cham: Springer Nature Switzerland.2023:86-98.

[27] Ali M, Diwan A, Kumar D. Attendance System Optimization Through Deep Learning Face Recognition. Int J Comput Digit Syst. 2024;15:1527-1540.

[28] Abbas Q, Albalawi TS, Perumal G, Celebi ME. Automatic Face Recognition System Using Deep Convolutional Mixer Architecture and AdaBoost Classifier. Appl Sci. 2023;13:9880.

[29] Aouani H, Ben Ayed Y. Deep Facial Expression Detection Using Viola-Jones Algorithm, CNN-MLP and CNN-SVM. Soc Netw Anal Min. 2024;14:65.

[30] Shah A, Ali B, Habib M, Frnda J, Ullah I, et al. An Ensemble Face Recognition Mechanism Based on Three-Way Decisions. J King Saud Univ Comput Inf Sci. 2023;35:196-208.

[31] Shareef AQ, Kurnaz S. Deep Learning Based COVID-19 Detection via Hard Voting Ensemble Method. Wirel Personal Commun. 2023:1-12.

[32] Ali MA, Meselhy Eltoukhy M, Rajeena P P F, Gaber T. Efficient Thermal Face Recognition Method Using Optimized Curvelet Features for Biometric Authentication. PLOS ONE. 2023;18:e0287349.

[33] Valero-Carreras D, Alcaraz J, Landete M. Comparing Two Svm Models Through Different Metrics Based on the Confusion Matrix. Comput Oper Res. 2023;152:106131.

[34] Rahim A, Zhong Y, Ahmad T, Ahmad S, Pławiak P, et al. Enhancing Smart Home Security: Anomaly Detection and Face Recognition in Smart Home IoT Devices Using Logit-Boosted CNN Models. Sensors. 2023;23:6979.