# Efficient Masked Face Recognition Methods using Deep Learning

Abrar Ahmed Mohammed
*School of Computer Science*
*Georgia Institute of Technology*
Atlanta, United States
ORCID: 0000-0003-1435-5841

*Keerthi Uppalapati*
*School of Electronics Engineering*
*Vellore Institute of Technology*
*Vellore, India*
*Email: keerthi.uppalapati2020@vitstudent.ac.in*

*Vankireddy Dimpul Reddy*
*School of Electronics Engineering*
*Vellore Institute of Technology*
*Vellore, India*
*Email: vankireddy.dimpul2020@vitstudent.ac.in*

line 1: 4th Given Name Surname line 2: *dept. name of organization (of Affiliation)*
line 3: *name of organization (of Affiliation)*
line 4: City, Country
line 5: email address or ORCID

***Abstract*—Covid-19 questioned not only people's health but also conventional scientific systems. Due to face masks that were made mandatory to wear, the existing cognitive systems failed to perform in real-time scenarios. The demand to develop face recognition systems that detect people even when they wore masks was naturally high. Deep learning techniques help to solve this problem, working efficiently in detecting user face features and comparing them with a known image database. This paper discusses Deep learning-based high-performing face detection algorithms and evaluates state-of-the-art face recognition algorithms like AlexNet, FaceNet, VGG16, and ResNet-50. The efficiency in detecting non-occluded faces derived from prominent datasets such as LFW, COMASK20, and S-LFW is elaborated. Then the paper's focus shifts toward Masked Face Recognition (MFR) where the authors propose an end-end pipeline to detect faces occluded with a mask. The paper also reviews the existing progress made in object detection, face detection and recognition, occluded face recognition, and masked face recognition to date. Furthermore, the paper addresses the challenges and scope of improvement of MFR methods.***

***Keywords*— masked face recognition, occluded face recognition, cropping-based approach, retina face, facenet, resnet-50, alexnet, deep face, mtcnn, yolov3, histogram of oriented gradients, voila-jones haar like features.***

## I. Introduction

The past decade proved to be an important timeline where the world adapted and moved towards many technological advancements. The massive increase in internet users across the globe justified the scale of disruption caused by the technology. According to [1], in the year 2021, there are 6.2 billion smartphone users in the world out of the 7.7 billion world population. Face detection and Face recognition evolved from novice technologies to a necessity for security, surveillance, and identification purposes. Face detection, as the name suggests detects faces in digital images and video frames. The detection algorithms are also optimized to detect faces from a live feed, for example, traffic surveillance systems use face detection technology to monitor drivers, and the inflow of traffic, and analyze patterns for better traffic management[2]. Furthermore, many derived applications of this technology were used widely such as Driver's fatigue/drowsiness detection[3][4], and pedestrian detection image processing algorithm for automatic traffic light transitions [5]. Similarly, face recognition mimics humans recognizing people, eye-capturing features of the person, and communicating with the brain's memory to identify the person.

Face detection and Face recognition gained popularity due to the numerous applications they are capable of performing and have gained significant research interest from the Machine learning and Computer Vision community. The Covid-19 virus outbreak was witnessed in 2019-2020 and since then the usage of face masks to protect against the SARS-CoV-2 virus has been widely suggested by medical professionals and world organizations. The world was intimidated by the quick spread of the infectious disease and was adapted to wearing a mask. The surveys around usage statistics and trends of face mask usage in [6-9] give insights on adaptiveness and general statistical impressions on an approximate percent of people who wore a mask country-wise. This proved problematic to the face recognition and face detection systems which were not trained to detect and recognize the subject with a mask. Formally referred to as Occluded Face Recognition (OFR) is applying face recognition with the subject's face partially covered with an object like a mask and this area gained the interest of researchers to develop robust and error-free OFR systems to solve the MFR problem.

## II. Background

The research effort around AI has significantly increased over the past few years. Computer Vision, a sub-area of research that primarily consists of the detection and recognition of objects in a scene and applied in various use cases has proven to attract researchers around the world. There are a few challenges such as face misalignment, over or under illumination, occlusions, and facial expressions which may alter the facial features are studied extensively and research done on these limitations increased the efficiency of the systems three-fold. It started with an object detection framework proposed by Paul Voila and Michael Jones [10] based on Haar-like features and introduced the machine learning approach to object detection. Based on this fundamental concept, the researchers improved face detection techniques from time to time. In [11] the authors introduce face detection in varying light conditions, where they have used haar-cascades as a base concept. A multi-face detection system is implemented by the authors in [12] using Haar cascades and the EigenFaces method. A face recognition-based attendance system is created by the authors in [13] using Haar cascades and local binary patterns (LBP). A challenger to Voila-Jones is the Histogram of Oriented Gradients (HOG) method, a popular object detection algorithm that works by counting the occurrences of gradient orientation in localized portions of an image has been presented by Dalal, N., & Triggs, B. in [14].

While the Voila-Jones and HOG established a platform to research object detection, it was not until Deep Learning-based Convolutional Neural Networks (CNNs) that object detection became more robust and accurate. These came particularly famous for their speed in real-time object detection and one such architecture is YOLOv3, developed by Joseph Redmon and Ali Farhadi [15]. Based on this, in the paper [16], a face detector based on YOLOv3 has been implemented which was trained on the Wider face database and the CelebA database and tested on the Face Detection Dataset and Benchmark (FDDB) database resulting in a fast face detector than other algorithms. While it was computationally expensive to apply standard convolutional networks to do object detection, Regions with CNN features (R-CNN) proposed an approach to segment the image into 2000 regions of interest using a selective search algorithm [17] which is based on an exhaustive search and segmentation. In the surveys [18-22], the authors briefly outline the existing face detection systems and their corresponding accuracy on popular face datasets.

Face recognition is the most important application on top of detection. The works of the authors in the paper [23] discuss the common hybrid methods used in face recognition popularly. In [24] the authors proposed a linear regression model for face recognition which gave the best results even in presence of scarf occlusion. In [25-27] the researchers used the local binary patterns (LBP) and histograms to concatenate the facial features of the subject into a single vector representing a face and matching it with the database of face features. A much more robust method is proposed in [28] which uses extended LBP. A decision tree based supervised learning model-based LBP is proposed for face recognition by the authors of [29] which also uses [30] to normalize the illumination of the face, thus achieving promising results.

The statistical method which reduces the dimensionality of the data by extracting principal components is much used in solving face recognition problems. The papers [31-33] propose and implement several use cases of Principal Component Analysis (PCA)-based face recognition systems and study the performance with known databases.

Since FR systems have a wide range of applications in making biometric authentication easier with a no-contact process, it has inspired tech giants such as Facebook and Google to make open-source APIs to recognize faces such as DeepFace, and FaceNet/OpenFace. The works in the papers [34-36] demonstrate the architecture, recognition rate, and various approaches to recognizing a face in the image. These algorithms are trained extensively and are scalable. Furthermore, Amazon's Rekognition and Microsoft's Face API are also state-of-the-art technologies to recognize the face from a database.

The recognition process with occlusion has been differentiated into three broad approaches occlusion robust feature extraction, occlusion aware face recognition, and occlusion recovery-based face recognition. In the paper [37], Min, R., et al. proposed an approach using Gabor wavelets, SVM, and PCA with interesting results during illumination and expression change on the face. Masked FR, is a research topic derived from OFR to propose a viable, robust, and seamless face recognition despite the face being covered with a mask. The research gained its popularity when people adopted wearing a mask as a precautionary measure to prevent the Covid-19 virus spread. The authors Wang, et al., [38] proposed a masked face detection dataset that consists of the largest real-world masked faces and synthetic masked faces. This dataset serves as a benchmark dataset to train and test algorithms. A method to crop the maskless portion of the face and then extract features from it has been proposed [39] and achieved impressive results. CNN-based algorithms were quite popular in OFR and thus they were again used in MFR. A YOLOv3-based approach has led to get a desirable outcome in the paper [40]. Furthermore, the authors in [41] have proposed to take advantage of the combination of deep learning and Local Binary Pattern (LBP) features to recognize the masked face by using RetinaFace, where they extracted LBP features from the masked face's eye, forehead, and eyebrow regions and combined them with features learned from RetinaFace into a unified framework for recognizing masked faces. Finally, an extensive survey has done by Alzu'bi, et al., in their paper [42] which throws a lot of insight into the past, and present works done or being done to solve the Masked Face Recognition problem.

## III. Methods for Face Detection- an Algorithmic Overview

The motivation for research on FD is to design an algorithm that can match or even outpower the accuracy of humans. A few popular methods are described below

### A. *Yolo v3*

The You Only Look Once version 3 is a successor of the popular YOLO which is a deep learning-based method for best-in-class object detection proposed by Redmon, et al.,[15]. It is particularly used in real-time object detection as it is a fast convolutional neural network. The system uses a single neural network to process the entire image and the network then separates the image into areas, providing bounding boxes and predicting probabilities for each region. The projected probabilities are used to weigh the generated bounding boxes. The non-max suppression technique ensures that the object detection algorithm detects each object just once and discards any false detections before returning the recognized objects and bounding boxes.

### B. MTCNN

The MTCNN is a genius framework for face detection and alignment proposed by Zhang, et al., [43]. The framework has three main steps which compute facial landmarks such as nose, eyes, and mouth which help in facial alignment making it easy for FR algorithms to run recognition. The three stages of the cascaded CNN are described as follows. The first step is to resize the image to various scales to build an image pyramid, which is fed to the three-staged cascaded network named P-net, R-net, and O-net. The P-net is a fully convolutional layer that computes candidate windows and their bounding box regression vectors. Then all candidates are fed into a Refined network(R-net) which is a CNN layer and reduces the count of candidates by using Non-Maximum Suppression (NMS) to combine overlapping candidates. In the final stage, the O-net computes the facial landmarks such as position of eyes, nose, and mouth. The MTCNN is mainly utilized if the application requires both face detection and facial alignment as output.

## IV. Datasets of Masked Faces

The following datasets are the most used in the training and testing of MFR. Each of the datasets can be classified into real-world or synthetic or simulated. Real-world datasets are very limited as they are difficult to collect but using algorithms such as MaskTheFace [44] and Identity aware Mask GAN [45], the authors were able to generate synthetic datasets of Masked faces. Therefore, it is important to understand the metadata and features of a dataset to be used in MFR algorithms.

### A. Simulated masked face recognition dataset (SMFRD)

This is a simulated dataset where the researchers augmented the masks on the faces of public face datasets and created a simulated masked face dataset of 500,000 faces from 10,000 people [46]. The base datasets used were LFW, AgeDB-30, WebFace, and CFP-FP datasets. The sample images of the datasets are presented in Figure 1.

*Figure 1 SMFRD dataset images [46]*

### B. Masked faces in the wild (MFW)-mini

The MFW-mini is made by applying a 3D mask fitting approach called WearMask3D to the existing Faces in the Wild by the authors in [47]. The Masked Faces in the Wild (MFW) micro dataset was created by collecting 3000 photographs of 300 people from the Internet, each including five images of masked faces and five images of unmasked faces.

### C. Comask20

It is a dataset collected from 300 subjects [48]. The dataset has been manually adjusted in instances of blurs and noise. In the end, the dataset obtained was 2754 facial images labeled for 300 different identities.

### D. Synthetic MFW

Mask has been augmented on the faces of the LFW dataset to generate 13,000+ masked images of 5000+ identities. The authors use this dataset of masked and unmasked faces to train the proposed model.

## V. Methods in Face Recognition- an Algorithmic Overview

Face Recognition (FR) systems are used in day-day life in surveillance, offices, and educational institutes for various purposes of authorization and biometric identification. The popular CNN-based robust algorithms which have state-of-the-art recognition accuracy are AlexNet, FaceNet, ResNet-50, and VGG16 which are discussed below

### A. AlexNet

The availability of very large datasets such as ImageNet which had more than 15 million images of 22,000 classes allowed computer vision researchers to train CNN-based models efficiently and in turn reduce error rates in image classification and object recognition. The AlexNet (winner of ILSVRC, 2012) was the first major step in using Deep-CNN to image recognition tasks. The AlexNet architecture had 5 Convolutional layers + 3 Fully Connected layers. The following is the overall architecture as presented by Krizhevsky, A., et al.,[49].

### B. FaceNet

It was proposed by Schroff, F., et al., in the paper "FaceNet: A Unified Embedding for Face Recognition and Clustering" [50]. It employs deep convolutional networks in conjunction with triplet loss. The architecture of the CNN is described as follows. It is a 22-layer Deep Neural Network that takes the person's facial image as input and outputs a face embedding of a 128-dimensional vector that represents the most important features of a face.

### C. ResNet-50

It is a model based on Deep Residual learning and one variant from the family of Resnet proposed in [51]. It has 48 Convolutional layers and 2 other layers such as the SoftMax function at the output. This network combines Residual network integrations with Deep architecture parsing. Because of the bottleneck blocks, training with ResNet-50 is faster and it is made up of five convolutional blocks having shortcuts placed between them. Deep Residual Features are extracted using the last convolution layer.

### D. VGG16

It was proposed by Simonyan, K., & Zisserman, A., [52]. VGG-16 has 16 layers and is a popular deep-learning image categorization algorithm due to its simplicity of implementation. The input to the model is 224×224×3. Spatial pooling is achieved by using five max-pooling layers that follow parts of the convolutional layers. Max-pooling is done with stride 2 over a 22-pixel window. Out of 16 layers, three are Fully Connected layers that precede 13 Convolutional layers, and the final layer is a soft-max layer.

## VI. Performance of Face Recognition Algorithms

Table I summarizes and compares the discussed algorithms when performing MFR feature extraction in the subsequent sections. Based on the table the authors think that using FaceNet for the task of MFR is the ideal approach as it is faster than AlexNet, has fewer layers than ResNet-50, and has fewer trainable parameters which make it a computationally cheap algorithm.

TABLE I. Summary of CNN-based FR algorithms

| Model | Training Dataset | Trainable parameters | Total layers (Convolutional + other) |
|---|---|---|---|
| AlexNet | ImageNet | 62M | 5+3 |
| VGG16 | ImageNet | 138M | 13+3 |
| ResNet-50 | ImageNet | 25M | 48+2 |
| FaceNet | Google | 140M or 7.5M | 22 |

The authors have experimented with the above-discussed pre-trained Face Recognition models on the benchmark LFW dataset [53] which contains non-occluded faces. OpenCV and Keras have been used to do this experiment. Transfer learning which is a machine learning technique where a model trained on one task is re-purposed on a second related task is used to train and test on the LFW dataset. Table II presents the results of the experiment which used SVM as a classification approach

TABLE II. Performance metrics of FR algorithms on Occlusion free faces

| Model | Accuracy (%) |
|---|---|
| AlexNet | 98.1 |
| VGG16 | 98.9 |
| ResNet-50 | 99.2 |
| FaceNet | 99.6 |

## VII. Proposed Pipeline for Masked Face Recognition

### A. Occlusion Handling Approach

The occlusion due to wearing the mask which covers some of the primary landmarks of the face such as mouth, chin, and nose must be handled by one of three methods i.e., Matching, Face Restoration, or Occlusion disregard. The occlusion disregard approach is used to prevent a bad restoration of the unmasked face using the reconstruction. A general idea is to detect the facial landmarks and develop a baseline of rules to crop the image such that the mask region is discarded or given less weightage in the classifier and only the unmasked face is used for the feature extraction process. The authors used this approach in this paper's proposed MFR pipeline to deal with mask occlusion.

### B. Flowchart of the proposed pipeline

Figure 2 illustrates the approach that was used to experiment with the Masked Face Recognition task.

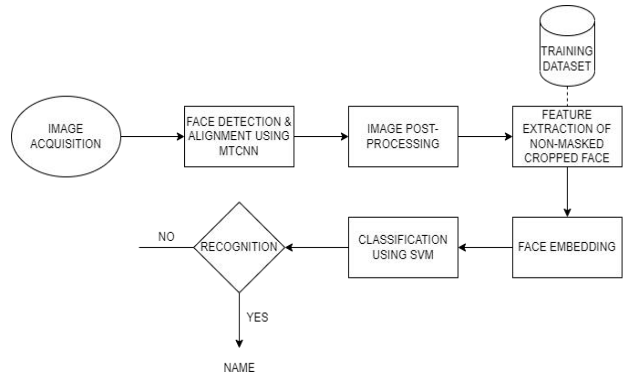


*Figure 2 Flowchart of the proposed MFR*

The COMASK20 dataset and MFW-mini datasets which consist of simulated masked faces are used as they are openly available for research purposes. Then the Face Detection method of MTCNN is applied to detect the face and draw a bounding box around the face detected. The MTCNN algorithm is also used to align the face by doing a 2-D rotation of the image w.r.t landmark location of the eye which is the output of the MTCNN detector.

The masked face with no or minimal background is then fed to the Image post-processing block where the occlusion disregard approach is used to remove the occlusion or mask region. This is a fundamentally challenging task when considering different shapes of the face and diverse features. The method of calculating the approximate un-masked region is done using the method presented in [39] by Li, Y., et al., Since the MTCNN generates the landmark coordinates, let us assume the coordinate of the left eye is M1(a, b) and the right eye is M2(c, d). The mid position of the two eyes is given as E((a + b)/2, (c + d)/2). Thereafter, the Euclidean distance between the key points of the eyes is calculated as the reference distance L = sqrt[(a − b)^2 + (c − d)^2]. The x-axis coordinate of the cropping point is (a + c)/2, the y-axis is (b + d)/2 + l, indicated by C((a + c)/2, (b + d)/2+l), where $l \subseteq [0.4L, 1.2L]$. Lastly, crop the images through the cropping points parallel to the x-axis.

When testing this theory on the Synthetic Masked LFW dataset with l=0.6L the following Figure 3 shows the cropped images which were earlier detected and aligned by MTCNN in the Face Detection block.

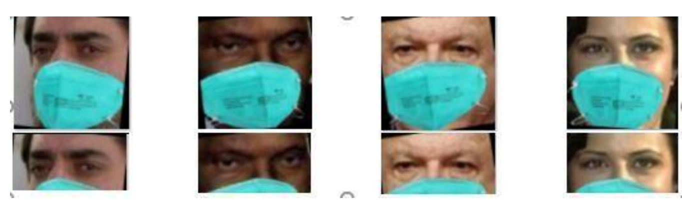

*Figure 3 Original vs Cropped faces using the method in [39]*

As represented in the 2nd row of Figure 3, the mask has been discarded mostly and only the useful landmarks are left for further blocks. The FaceNet model is used in the next step to extract features of the upper face like from the areas of eyebrows, nose bridge, eyes, and forehead. The algorithm works perfectly well in fewer features than expected scenarios and generates a 128-dimensional vector termed as FaceNet embeddings as feature vectors. Then an ML model is trained using the training dataset and a linear SVM is used for classification. The clustering methods such as K-nearest neighbors can also use instead of SVM and achieve the same results. A hyperplane is constructed for each of the features differentiating n-dimensional features of all the inputs of the dataset. The FaceNet embedding is then classified by calculating the Euclidean distance to the nearest features' labels. Thus, top-n feature vectors whose distance is nearest to the features of FaceNet embedding are returned. If the features match with a confidence score of over 90% the face is recognized successfully.

## VIII. EXPERIMENTAL RESULTS

The Face Detection using MTCNN detected masked faces with 99.8% accuracy and the cropping filter did solve the problem of occlusion due to the face mask. Accuracy is the ratio of the number of right predictions to the total number of samples, which may be described as Accuracy = (TP + TN)/(TP + TN + FP + FN)

The proposed model's performance in terms of Accuracy of recognizing masked faces is described in Table III and Table IV.

TABLE III. PERFORMANCE OF PROPOSED MFR MODEL ON S-LFW DATASET

| Model | Accuracy (%) |
|---|---|
| AlexNet +SVM | 81.4 |
| VGG16 +SVM | 91.0 |
| ResNet-50 +SVM | 87.6 |
| FaceNet+SVM | 98.2 |
| **ArcFace +SVM** | **99.5** |

TABLE IV. PERFORMANCE OF PROPOSED MFR MODEL ON COMASK20 DATASET

| Model | Accuracy (%) |
|---|---|
| AlexNet +SVM | 62.9 |
| VGG16 +SVM | 83.3 |
| ResNet-50 +SVM | 79.6 |
| FaceNet+SVM | 85.0 |
| **ArcFace +SVM** | **89.6** |

The ArcFace +SVM model outperforms others on both datasets of S-LFW and COMASK20. Due to the few numbers of masked images in COMASK20 and face deviations in the dataset like the person looking to the left without a mask but looking straight with a mask, the model is only giving optimistic results, but it recognizes better than the existing systems using this dataset.

## IX. FUTURE RESEARCH DIRECTIONS

MFR has varying use cases such as face identification in smartphones and attendance systems. The problem to recognize faces with different racial features than the dataset the algorithm was pre-trained on must be addressed by proposing a novel approach. A publicly available high-resolution Indian simulated masked face dataset is needed for training the models to recognize people of Indian origin.

## X. CONCLUSION

This paper presented a Masked Face Recognition pipeline to recognize people even though they are wearing a mask to prevent the spread of covid-19. The existing statistical & Deep learning-based Face Detection algorithms are studied and Multi-Task Cascaded Convolutional Neural Network (MTCNN) algorithm is chosen to detect faces in the proposed pipeline. A cropping-based approach is used to remove the occluded face region. Then the feature extraction is done using a comprehensive list of popular CNN-based algorithms suchas AlexNet, VGG-16, ResNet-50, FaceNet, and ArcFace. Their performance in recognizing the face is presented as result. The ArcFace+SVM model achieved 99.6% on the simulated LFW dataset and 89.6% on the COMASK20 dataset.

## REFERENCES


[1] S. O'Dea. (2022). *Smartphone subscriptions worldwide 2016-2027.* https://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide/

[2] Luo, D., Lu, X., Hu, C., Peng, X., & Wang, W. (2016, December). A robust traffic driver face detection method. In *2016 International Conference on Progress in Informatics and Computing (PIC)* (pp. 285-288). IEEE.

[3] Ahmed, M., Masood, S., Ahmad, M., & Abd El-Latif, A. A. (2021). Intelligent Driver Drowsiness Detection for Traffic Safety Based on Multi CNN Deep Model and Facial Subsampling. *IEEE Transactions on Intelligent Transportation Systems*.

[4] Alioua, N., Amine, A., Rziza, M., & Aboutajdine, D. (2011, August). Driver's fatigue and drowsiness detection to reduce traffic accidents on road. In *International Conference on Computer Analysis of Images and Patterns* (pp. 397-404). Springer, Berlin, Heidelberg.

[5] Manlises, C. O., Martinez, J. M., Belenzo, J. L., Perez, C. K., & Postrero, M. K. T. A. (2015, December). Real-time integrated CCTV using face and pedestrian detection image processing algorithm for automatic traffic light transitions. In *2015 International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM)* (pp. 1-4). IEEE.

[6] Worby, C. J., & Chang, H. H. (2020). Face mask use in the general population and optimal resource allocation during the COVID-19 pandemic. *Nature communications*, *11*(1), 1-9.

[7] Badillo-Goicoechea, E., Chang, T. H., Kim, E., LaRocca, S., Morris, K., Deng, X., ... & Stuart, E. A. (2021). Global trends and predictors of face mask usage during the COVID-19 pandemic. *BMC public health*, *21*(1), 1-12.

[8] Kwon, S., Joshi, A. D., Lo, C. H., Drew, D. A., Nguyen, L. H., Guo, C. G., ... & Chan, A. T. (2021). Association of social distancing and face mask use with risk of COVID-19. *Nature Communications*, *12*(1), 1-10.

[9] Katharina Buchholz. (2020). *Asians Still Most Likely to Wear Face Masks Due to COVID-19.* https://www.statista.com/chart/21452/share-of-people-wearing-face-masks-per-country-covid-19/

[10] Viola, P., & Jones, M. (2001, December). Rapid object detection using a boosted cascade of simple features. In *Proceedings of the 2001 IEEE computer society conference on computer vision and pattern recognition. CVPR 2001* (Vol. 1, pp. I-I). Ieee.

[11] Hapsari, D. T. P., Berliana, C. G., Winda, P., & Soeleman, M. A. (2018, September). Face detection using Haar cascade in difference illumination. In *2018 International Seminar on Application for Technology of Information and Communication* (pp. 555-559). IEEE.

[12] Mantoro, T., & Ayu, M. A. (2018, May). Multi-faces recognition process using Haar cascades and eigenface methods. In *2018 6th International Conference on Multimedia Computing and Systems (ICMCS)* (pp. 1-5). IEEE.

[13] Chinimilli, B. T., Anjali, T., Kotturi, A., Kaipu, V. R., & Mandapati, J. V. (2020, June). Face recognition-based attendance system using haar cascade and local binary pattern histogram algorithm. In *2020 4th International Conference on Trends in Electronics and Informatics (ICOEI)(48184)* (pp. 701-704). IEEE.

[14] Dalal, N., & Triggs, B. (2005, June). Histograms of oriented gradients for human detection. In *2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05)* (Vol. 1, pp. 886-893). Ieee.

[15] Redmon, J., & Farhadi, A. (2018). Yolov3: An incremental improvement. *arXiv preprint arXiv:1804.02767*.

[16] Li, C., Wang, R., Li, J., & Fei, L. (2020). Face detection based on YOLOv3. In *Recent Trends in Intelligent Computing, Communication and Devices* (pp. 277-284). Springer, Singapore.

[17] Uijlings, J. R., Van De Sande, K. E., Gevers, T., & Smeulders, A. W. (2013). Selective search for object recognition. *International journal of computer vision*, *104*(2), 154-171.

[18] Zafeiriou, S., Zhang, C., & Zhang, Z. (2015). A survey on face detection in the wild: past, present, and future. *Computer Vision and Image Understanding*, *138*, 1-24.

[19] Kumar, A., Kaur, A., & Kumar, M. (2019). Face detection techniques: a review. *Artificial Intelligence Review*, *52*(2), 927-948.

[20] Hjelmås, E., & Low, B. K. (2001). Face detection: A survey. *Computer vision and image understanding*, *83*(3), 236-274.

[21] Zhang, C., & Zhang, Z. (2010). A survey of recent advances in face detection.

[22] Razzaq, A. N., Ghazali, R., El Abbadi, N. K., & Dosh, M. (2022). A Comprehensive Survey on Face Detection Techniques. *Webology*, *19*(1).

[23] Oloyede, M. O., Hancke, G. P., & Myburgh, H. C. (2020). A review on face recognition systems: recent approaches and challenges. *Multimedia Tools and Applications*, *79*(37), 27891-27922.

[24] Naseem, I., Togneri, R., & Bennamoun, M. (2010). Linear regression for face recognition. *IEEE transactions on pattern analysis and machine intelligence*, *32*(11), 2106-2112.

[25] Rahim, M. A., Azam, M. S., Hossain, N., & Islam, M. R. (2013). Face recognition using local binary patterns (LBP). *Global Journal of Computer Science and Technology*.

[26] Ahonen, T., Hadid, A., & Pietikäinen, M. (2004, May). Face recognition with local binary patterns. In *European conference on computer vision* (pp. 469-481). Springer, Berlin, Heidelberg.

[27] Ahonen, T., Hadid, A., & Pietikainen, M. (2006). Face description with local binary patterns: Application to face recognition. *IEEE transactions on pattern analysis and machine intelligence*, *28*(12), 2037-2041.

[28] Liu, L., Fieguth, P., Zhao, G., Pietikäinen, M., & Hu, D. (2016). Extended local binary patterns for face recognition. *Information Sciences*, *358*, 56-72.

[29] Maturana, D., Mery, D., & Soto, Á. (2010, November). Face recognition with decision tree-based local binary patterns. In *Asian Conference on Computer Vision* (pp. 618-629). Springer, Berlin, Heidelberg.

[30] Tan, X., & Triggs, B. (2010). Enhanced local texture feature sets for face recognition under difficult lighting conditions. *IEEE transactions on image processing*, *19*(6), 1635-1650.

[31] Karim, T. F., Lipu, M. S. H., Rahman, M. L., & Sultana, F. (2010, July). Face recognition using PCA-based method. In *2010 IEEE International Conference on Advanced Management Science (ICAMS 2010)* (Vol. 3, pp. 158-162). IEEE.

[32] Ebied, H. M. (2012, May). Feature extraction using PCA and Kernel-PCA for face recognition. In *2012 8th International Conference on Informatics and Systems (INFOS)* (pp. MM-72). IEEE.

[33] Faruqe, M. O., & Hasan, M. A. M. (2009, August). Face recognition using PCA and SVM. In *2009 3rd international conference on anti-counterfeiting, security, and identification in communication* (pp. 97-101). IEEE.

[34] Schroff, F., Kalenichenko, D., & Philbin, J. (2015). Facenet: A unified embedding for face recognition and clustering. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 815-823).

[35] Jose, E., Greeshma, M., Haridas, M. T., & Supriya, M. H. (2019, March). Face recognition based surveillance system using facenet and mtcnn on jetson tx2. In *2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS)* (pp. 608-613). IEEE.

[36] Taigman, Y., Yang, M., Ranzato, M. A., & Wolf, L. (2014). Deepface: Closing the gap to human-level performance in face verification. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1701-1708).

[37] Min, R., Hadid, A., & Dugelay, J. L. (2011, March). Improving the recognition of faces occluded by facial accessories. In *2011 IEEE International Conference on Automatic Face & Gesture Recognition (FG)* (pp. 442-447). IEEE.

[38] Wang, Z., Wang, G., Huang, B., Xiong, Z., Hong, Q., Wu, H., ... & Liang, J. (2020). Masked face recognition dataset and application. *arXiv preprint arXiv:2003.09093*.

[39] Li, Y., Guo, K., Lu, Y., & Liu, L. (2021). Cropping and attention-based approach for masked face recognition. *Applied Intelligence*, *51*(5), 3012-3025.

[40] Mhadgut, S. (2021, July). Masked Face Detection and Recognition System in Real Time using YOLOv3 to combat COVID-19. In *2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT)* (pp. 1-7). IEEE.

[41] Vu, H. N., Nguyen, M. H., & Pham, C. (2021). Masked face recognition with convolutional neural networks and local binary patterns. *Applied Intelligence*, 1-16.

[42] Alzu'bi, A., Albalas, F., Al-Hadhrami, T., Younis, L. B., & Bashayreh, A. (2021). Masked Face Recognition Using Deep Learning: A Review. *Electronics*, *10*(21), 2666.

[43] Zhang, K., Zhang, Z., Li, Z., & Qiao, Y. (2016). Joint face detection and alignment using multitask cascaded convolutional networks. *IEEE signal processing letters*, *23*(10), 1499-1503.

[44] Anwar, A., & Raychowdhury, A. (2020). Masked face recognition for secure authentication. *arXiv preprint arXiv:2008.11104*.

[45] Geng, M., Peng, P., Huang, Y., & Tian, Y. (2020, October). Masked face recognition with generative data augmentation and domain constrained ranking. In *Proceedings of the 28th ACM International Conference on Multimedia* (pp. 2246-2254).

[46] Wang, Z., Wang, G., Huang, B., Xiong, Z., Hong, Q., Wu, H., ... & Liang, J. (2020). Masked face recognition dataset and application. *arXiv preprint arXiv:2003.09093*.

[47] Hong, J. H., Kim, H., Kim, M., Nam, G. P., Cho, J., Ko, H. S., & Kim, I. J. (2021, September). A 3D model-based approach for fitting masks to faces in the wild. In *2021 IEEE International Conference on Image Processing (ICIP)* (pp. 235-239). IEEE.

[48] Vu, H. N., Nguyen, M. H., & Pham, C. (2022). Masked face recognition with convolutional neural networks and local binary patterns. Applied Intelligence, 52(5), 5497-5512.

[49] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). Imagenet classification with deep convolutional neural networks. Communications of the ACM, 60(6), 84-90.

[50] Schroff, F., Kalenichenko, D., & Philbin, J. (2015). Facenet: A unified embedding for face recognition and clustering. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 815-823).

[51] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770-778).

[52] Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.

*[53]* Huang, G. B., Mattar, M., Berg, T., & Learned-Miller, E. (2008, October). Labeled faces in the wild: A database forstudying face recognition in unconstrained environments. In *Workshop on faces in'Real-Life'Images: detection, alignment, and recognition.*