Tech Science Press

Masked Face Recognition Using MobileNet V2 with Transfer Learning

Ratnesh Kumar Shukla[1],* and Arvind Kumar Tiwari[2]

[1]Department of Computer Science & Engineering, Dr. APJ Abdul Kalam Technical University, Lucknow, 226021, India
[2]Department of Computer Science & Engineering, KNIT, Sultanpur, 228118, Uttar Pradesh, India
*Corresponding Author: Ratnesh Kumar Shukla. Email: ratnesh.nitttr@gmail.com

Abstract: Corona virus (COVID-19) is once in a life time calamity that has resulted in thousands of deaths and security concerns. People are using face masks on a regular basis to protect themselves and to help reduce corona virus transmission. During the on-going coronavirus outbreak, one of the major priorities for researchers is to discover effective solution. As important parts of the face are obscured, face identification and verification becomes exceedingly difficult. The suggested method is a transfer learning using MobileNet V2 based technology that uses deep feature such as feature extraction and deep learning model, to identify the problem of face masked identification. In the first stage, we are applying face mask detector to identify the face mask. Then, the proposed approach is applying to the datasets from Canadian Institute for Advanced Research10 (CIFAR10), Modified National Institute of Standards and Technology Database (MNIST), Real World Masked Face Recognition Database (RMFRD), and Stimulated Masked Face Recognition Database (SMFRD). The proposed model is achieving recognition accuracy 99.82% with proposed dataset. This article employs the four pre-programmed models VGG16, VGG19, ResNet50 and ResNet101. To extract the deep features of faces with VGG16 is achieving 99.30% accuracy, VGG19 is achieving 99.54% accuracy, ResNet50 is achieving 78.70% accuracy and ResNet101 is achieving 98.64% accuracy with own dataset. The comparative analysis shows, that our proposed model performs better result in all four previous existing models. The fundamental contribution of this study is to monitor with face mask and without face mask to decreases the pace of corona virus and to detect persons using wearing face masks.

Keywords: Convolutional Neural Network (CNN); deep learning; face recognition system; COVID-19 dataset and machine learning based models

1 Introduction

The Corona virus disease a pandemic, known as COVID-19, began in December 2019. The virus continued spread caused governments around the world to declare a state of emergency for several months. To decreases the spread of corona virus and restore normalcy to world health agencies recommend social separation, wearing face masks and avoiding touching them. The face mask has been

found to aid in the prevention of the spread of this fatal disease. Face masks have required by the government and the World Health Organization in densely populated regions and home care locations in healthcare settings where COVID-19 cases have been reported [1]. COVID-19 is more likely to harm people, who already have chronic diseases such as cancer or diabetes. COVID-19 is an once-in-a-lifetime catastrophe that has founded in a significant number of casualties and identity issues. The wearing face mask is the best solution to protect and minimize the transmission of COVID-19 infection. As the significant aspects of wearing face mask but in this situation face detection and recognition is very difficult to identify the covered faces [2]. Based on the precautions, computer vision system was developed for the prevention of corona virus transmission. Using a mask to conceal one's face can really prevent the spread of this fatal virus and face masks can be used as a preventive measure for any such illnesses. However, it is a challenging task to determine whether or not a person is wearing a face mask and following the rule of social separation, but a deep learning model can recognized face features and anticipate whether or not a person is wearing a face mask [3]. It is important to improve identification and efficacy of existing face detection and recognition technologies on the covered faces. These are some face related applications, such as facial attendance, community entry and exit, control access, source with face detector gates, face authentication and confidentiality based mobile payments. Face recognition are network based security that are working on the faces if failed, there are being big issue in security concern [4]. Binary neural networks (BNNs) are classifying correct position face mask on the face. Such algorithms can be used to reduce viral propagation at corporate building entrances, airports, shopping malls, and other indoor locations in the context of the current COVID-19 epidemic. Implementation of BNNs solves a variety of challenges, giving up to 98% accuracy for Masked Face-Net dataset wearing poses [5]. In these dispersed environments capabilities have required low cost battery powered devices with limited memory. This processing must take place on the face device, with no access to cloud servers, to maintain the environment's security and privacy. Incredible resilience to changes in light facial expression and occlusion influenced by CNN based methods. During this pandemic, MobileNet V2 with a transfer learning model has used to identify the problem's address, masked face identification and face recognition, as well as an occlusion removal approach. The collected characteristics are put into a self-organizing map (SOM), which produces picture clusters with comparable visual content. SOM generates several clusters with their centers and then query picture attributes are compared with all cluster representatives to locate the nearest cluster. Finally, using the Euclidean distance similarity metric, photos are extracted from the nearest cluster [6]. The experimental results in this article illustrate face categorization utilizing re-scaling at data pre-processing. With re-scaling of the face picture with and without mask, pre-processed data offered greater accuracy of face categorization. The accuracy of transfer learning marginally outperforms that of convolutional neural networks [7]. We also compared the accuracy of the suggested technique with the proposed dataset to that of other pre-trained models with the proposed dataset and existing datasets. This article employs the four pre-programmed models VGG16, VGG19, ResNet50 and ResNet101. These models have extracted the deep features of faces with VGG19 and ResNet101 is achieving accuracy with own dataset. The comparative analysis shows that the proposed model performs better in compare to all four previous existing models. The key contribution of this work is the use of MobileNet V2 with transfer learning to monitor face photos with and without masks, and another contribution is the identification of people wearing face masks.

2 Related Works

Recent advancements in deep learning based images editing algorithm have produced hopeful outcome for eliminating items from photos, and struggled to provide realistic results for removing huge complex objects, particularly in facial photographs [8]. A generative adversarial network based network with two discriminators and learning the overall structure with face features and the other for concentrating on the

deep features missing area, beats prior representative state-of-the art approaches in both qualitative and numerical terms [9]. In facial analysis, masked face recognition was utilized to build a standard and efficient approach for face identification. The Histogram of Gradients (HOG) approach is providing features of the faces as well as the VGG16 and VGG19 deep learning pre training model used to extract important face features from celebrity photographs [10]. The current facial datasets must be enhanced with algorithms that allow face masked to recognise with low false-positive rates and overall high accuracy, without the user dataset, which must be produced by collecting new pictures for authentication [11]. A transfer learning model is used to provide and create fine-tuned state-of-the-art performance accuracy in learning models. InceptionV3 has identified in automated process of recognizing people's faces that are not using face masks [12]. The three datasets are the SMFRD, RMFRD and Labeled Faces in the Wild (LFW). In RMFRD, the support vector machine (SVM) classifier has a testing accuracy of 99.64% [13]. Deep super resolution crack network (SRCNet) achieved 98.70% accuracy and provide outperformed standard in complete deep learning image verification systems in the absence of image super resolution, with kappa decreased by more than 1.5% [14]. ResNet50, AlexNet and MobileNet have been investigated as potential plug-ins with proposed model to get accurate results in less time. When implemented using ResNet50, the technique achieved an accuracy of 98.2% [15]. The wide spread of COVID-19 is wearing face mask in general public areas to potential to lower community spread transmission and the pandemic impact considerably. When acceptance and compliance is nearly universal, wearing face masks is most likely to have the largest community-wide benefits, when they have used in concert with other non-pharmaceutical techniques such as social distancing. In deep learning-based model capable of correctly recognizing persons, those wearing face masks in order to train ResNet50 architecture capable of correctly recognizing masked faces [16]. Wearing face mask are merging to local binary pattern features from the masked face's eye, forehead and eye brow regions with retina face attributes to build a cohesive framework. The COMASK20 dataset had f1-score recognition rate of 87%, whereas the Essex dataset had a f1-score recognition rate of 98% [17]. A novel approach developed by Li et.al. To recognize a face algorithm for the face masked to merging and cropping image based methodology and convolutional block attention module (CBAM). The CBAM module used to target the eyelids and the adjacent regions for each sample's correct cropping is investigated [18]. AlexNet is based on a subset of an artificial neural network (ANN) that is similar to the human visual system. ANNs have grown more important in many computer vision systems in recent years. CNN architecture tailored to detection and identification issues that integrate global and local data at many scales. Using the WIDER Face and Face detection data set and benchmark (FDDB) datasets, the efficiency and practicability of the proposed method for dealing with various scales face identification difficulties have been tested [19]. The autonomous learning rate selection are achieving accuracy of 96.23% related to the COVIDx dataset in just performed 41 epochs [20]. On a public face mask dataset, retina face mask achieves state-of-the-art performance, with precision and recall. Those were 2.3% and 1.5% higher performance than the baseline, respectively, therefore 11% and 5.9% better than the starting point. We also look into the possibility of combining retina face mask with MobileNet, a neural network with low weights for embedded or mobile devices [21]. MobileNetv2 is a recently suggested model to offers both offline and real time performance to provide quick and stable results. The recognition accuracy used as a determinant in ResNet 50 and AlexNet is being improved as well as optimized the CNN algorithm's performance. Extensive study of datasets revealed as improved categorization rates [22]. Set a pixel distance approximation and a threshold to calculate social distance violations between individuals. To determine whether the value of distance exceeds a violation threshold is specified for the minimum social distance barrier. Face categorization is an object detection and pixel band analysis in a variety of face classification algorithms have been used in with mask and without face mask applications [23]. It was discovered that most face classification models were based on static face categorization. There have been many various methods for face classification deployed, however in recent years machine

learning algorithms have been increasingly employed for face classification [24]. It has also been discovered that using CNN for picture classification results in greater accuracy, which is extremely attainable for multi label real-time face classification tasks.

3 Material and Methods

Here, the dataset description and the proposed methodology are discussed.

3.1 Dataset Description

The Real-world Masked Face Recognition Dataset (RMFRD) was created during the COVID-19 outbreak with the objective of improving the facial recognition performance in existing face detection and identification algorithms on masked faces. In this article, an integrated face mask dataset is created with the help of Internet, Kaggle dataset and a different authentic source containing 8169 images. Annotations are used to cropping facial pictures and obtain dimensions labels. The final dataset contains 4066 faces wearing mask images and 4103 faces without wearing mask images. Here, the margins around the face are bounding boxes that are utilized on cropped annotated faces increased the margins around the face cropping. Fig. 1 shows few combinations of faces images without mask and faces wearing mask dataset.

Figure 1: Face image represents without wearing faces mask image (up) and with wearing faces mask image (down)

RMFRD with 5000 images, out of which 525 individuals with masks and 525 subjects without masks in 90000 images, Simulated Masked Face Recognition Dataset (SMFRD) with 500000 simulated masked faces, CIFAR10 with 60000 32 × 32 colour images divided into 10 classes, each with 6000 images and MNIST has a sample size of 60000 training set and a 10000 sample test sets were used for comparative analysis. RMFRD have 5000 images and proposed dataset have 8169 images collected from different sources.

3.2 Model Description

This article presented four pre-trained models such as VGG16, VGG19, ResNet50 and ResNet101 as feature extractors that extracted deep feature from informative regions. Here, the brief description of four pre trained models and the proposed MobileNet V2 based approaches in recognition of faces wearing the mask is given.

3.2.1 VGG16

VGG16 is a convolutional neural network model and it outperforms. AlexNet model is using to build the gradually replacing massive kernel sized filters with 3 × 3 convolutional filter. It was using trained ImageNet dataset, which had 14 million images with 1000 unique classifications. The fact that it had been 16 layers and given it the name VGG16. This model is including convolutional layer, max pooling layer and fully connected layer.

3.2.2 VGG19

A deep learning CNN architecture is used for image categorization. The VGG architecture is combination of 16 convolutional layers and 3 fully linked layer. The VGG19 refers to the fact that there are 19 different strata in all. A VGG19 is a 19 layer CNN, which was trained on millions of image samples that normalizes images using the Zero Center architectural style. The fully linked layer includes ReLU, Dropout, Softmax and classification output. ReLU, Dropout, Softmax and classification output are part of the fully connected layer.

3.2.3 RESNET50

ResNet50 is a form of traditional neural network that enable to build extremely deep neural networks with 150+ layers. It comprises of 50 layers, all of which were trained on the ImageNet dataset. Residual network model are combining with deep architecture parsing in this network. Because of the bottleneck blocks, ResNet50 is more efficient. It is made up of five convolutional neural network blocks those were connected together using shortcuts. Deep Residual Features (DRF) is extracted using the last convolution layer.

3.2.4 RESNET101

ResNet101 is belonging 101 CNN layers. They are bringing a pre-trained model of the network from the ImageNet database, which has trained over a million photos. This network model can be classified photographs into thousands distinct object detection categories including pencils, mouse, keyboards and a wide range of animals. The resultant network has amassed library of deep feature representations in a wide range of pictures. This network model is taking input picture size is 224 × 224 pixels.

3.2.5 MobileNet V2

MobileNetV2 model is a mobile optimized CNN. It has been including inverted residual structure and combination of residual bottleneck levels. The intermediate expansion layer has filtered features using light weight depth wise convolutional layers. MobileNet V2 has overall design a fully connected layer included 32 filters and 19 bottleneck layers.

Above Fig. 2 has included two types of blocks. One is a stride residual block and other strides for shrinking. Each of the blocks has two levels. The depth wise convolution is the first layer, followed by 1 × 1 convolutions with ReLU6. 1 × 1 convolution filter is using in third layer for non-linearity. ReLU has applied in deep networks to classify the output domain and power of a linear classifier on the non-zero part of volume with assertion [25].
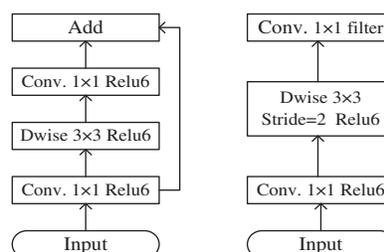
Figure 2: The basic architecture of face detection and recognition in linear bottleneck

3.2.6 MobileNet V2 with Transfer Learning

MobileNet V2 is used to detect and recognised faces in occlusion. On the other hand, the facemask fully removed the major discriminative features of the nose and mouths, making most existing techniques are ineffective for mean facial recognition in practice. Face masks have been used to protect against the intake of dangerous viruses and particles in many aspects of daily life. They are also an excellent choice in healthcare for offering bi-directional protection against airborne illnesses. It is critical to wear and position the mask appropriately for it to work properly. CNNs are a great option for face identification and recognition, as well as categorization of proper face mask wearing and placement. The inference equipment needs to be low-cost, compact and energy efficient, while having sufficient storage and processing capacity to perform correct latency CNNs (See Fig. 3). Deep neural network models may take days or weeks to train for very large data sets. Reusing the weights of pretrained patterns, such as image recognition tasks on ImageNet, for the standard computer vision benchmark data sets, is one approach to speed up for this process. You may download and utilize top level models for your own computer view difficulties or you may directly combine top level models into a new model. There are two steps to implementing transfer learning. First, we are using convolutional layers in original model for extracting the features. We trained and classified the last level of images to classify the problem. Convolutional layers have fixed in first layer and classification layers are trained. During the second stage, all layers have unlocked and the system is retrained. The feature extraction of characteristics in this approach is appropriate for proposed task.
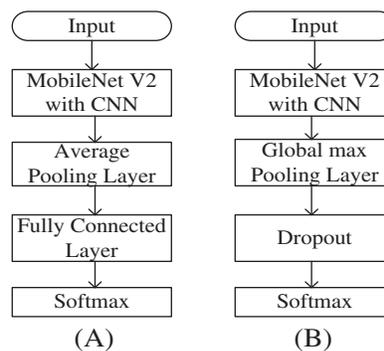
Figure 3: The Architecture of MobileNet V2 and MobileNet V2 transfer learning: (A) Architecture of MobileNet V2 and (B) Architecture of MobileNet V2 model with transfer learning

1. Transfer learning is the process of using models trained on one problem as a springboard for solving another.
2. Pre-trained models can be utilized simply as preprocessing feature extraction or incorporated into completely new models because of transfer learning adaptability.
3. Keras makes it easy to access several of the best ImageNet image recognition models, such as VGG16, VGG19, ResNet50 and RESNET101.

As shown in Fig. 3, the parameters of each convolutional layer in the pre-trained with MobileNet V2 model and delete the last two layers (average pooling layer and fully connected layer). In face classification MobileNet V2 with transfer learning model are added the global max pooling layer and dropout layer. The dropout layer and global max pooling layer for previous model parameters and prevent overfitting. The core concept of transfer learning is to apply a model learned on a big dataset to a smaller dataset. The object recognition freezes the network's early convolutional layers and trains just the last levels to predict them. The aim is to extract large and low-level characteristics from convolutionary

levels, such as edges, patterns and gradients, which are essential across the pi, while the next layers identify particular components inside an image, eyes or wheels.

3.3 Proposed Model of the Wearing Face Mask and Without Face Mask Images

The quality of face detection and recognition in mask recognition model are finding actual photos rather than digitally rendered images, the people are wearing masks and also collect photos of people's faces that will confuse the model as to whether or not they are wearing a mask. Machine learning is using particularly facial mask detector algorithms to identify abuse and limit of corona virus spread, basically indoors. MobileNet V2 is using transfer learning to identify not just whether or not a mask is employed, but also other mistakes that are typically overlooked yet may contribute to viral propagation [26]. This kind of deep quantization has numerous advantages over other methods of quantization. As a result, the masked face recognition in the actual world is achievable. Aside from that, masked areas of the face vary from one face to another, resulting in pictures of varying proportions. To solve this question, the deep quantization suggested allows pictures of varying sizes to be classified. In addition, the proposed technique does not need the mission region be trained once the mask has been removed. Face recognition is improved when a mask is worn during the coronavirus pandemic because it is more generalized. There are following mechanism involved in suggested model:

1. Create a facial dataset with mask and without mask.
2. The facial landmarks and multi task convolutional neural network (MTCNN) is using to detect the face.
3. A facial algorithm for mask detection is used to establish if a person is wearing a face mask or not.
4. If a face is found with wearing mask then apply the feature extraction on eyes using facial landmark detection algorithms and see the correct face of the persons then go to step 4.
5. Otherwise go to next step 4.
6. Face recognition using proposed MobileNet V2 with Transfer learning based approach.
7. Find the image of the correct person's face.

Fig. 4 shows the face detection and recognition with wearing mask and without wearing mask. This model used MTCNN for face detection and MobileNet V2 with transfer learning for face recognition. Here, gathering data from with wearing mask and without wearing mask datasets. After collecting this information and trained the model with the help of a library in python using with keras and tensorflow. The attributes of the face features are listed in this library. Then the uses MTCNN for face detection of faces. After detecting the faces, a cascade classifier is used to remove each one. Then apply the facial landmark detection on eye for detecting eye features. Then the proposed model of MobileNet V2 with transfer learning have used to recognizing wearing face mask in face images. The Inception V3 Model has working brilliantly and providing high accuracy results, such as testing accuracy of the model has 97.11%. Rescaling has been proven to improving the two class mask detection models accuracy. The reason of unsatisfactory accuracy is low-quality of images and lack of the data.

In this article a 2D rotation of the eye locations to make them horizontal in wearing face mask and without wearing face mask faces is used (See Fig. 5). Here, the right and left eyeballs of the Region of Interest (RoI) are collecting and used to extract ocular feature points. These RoI windows have now implied to obtain corner point features from the detection windows obtained for eye brows and lips. The keras and tensorflow classifiers to detect facial features in images and videos and provide an accurate result in face detection and recognition. Here, python libraries are used to extract each characteristic of faces such as the eyes, lips, nose, width and height of the faces. This model used MTCNN for face detecting and MobileNet V2 with Transfer learning for face recognition.
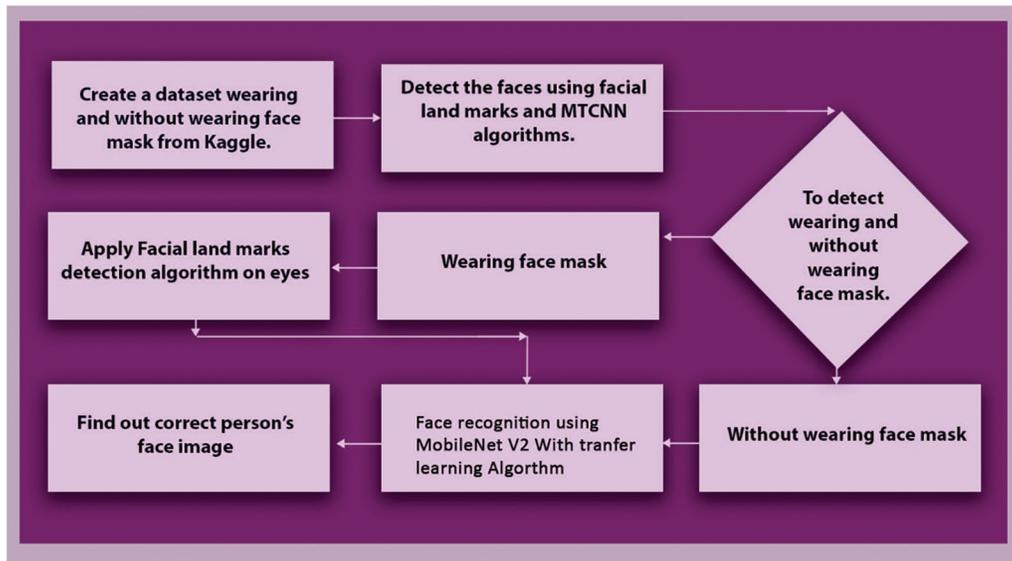
Figure 4: Proposed architecture for face detection and recognition from input image with mask and mask

(A) Origional Image   (B) Landmark Detection   (C) Original Image   (D) Landmark Detection
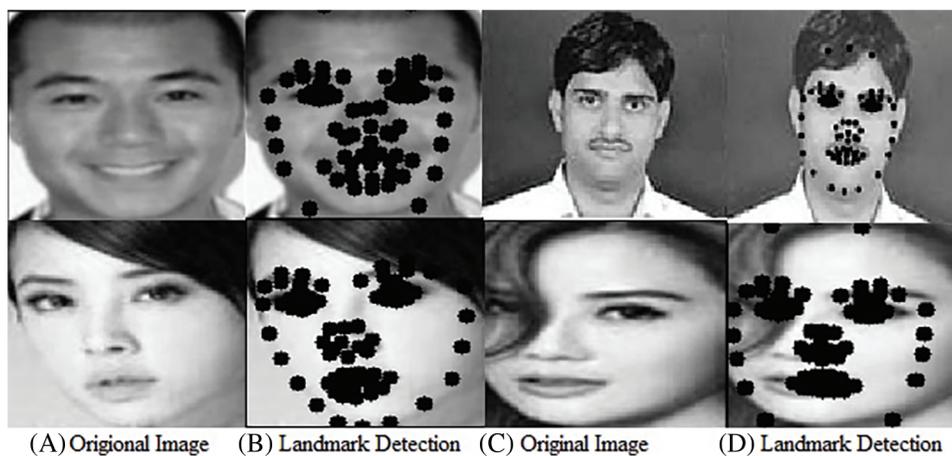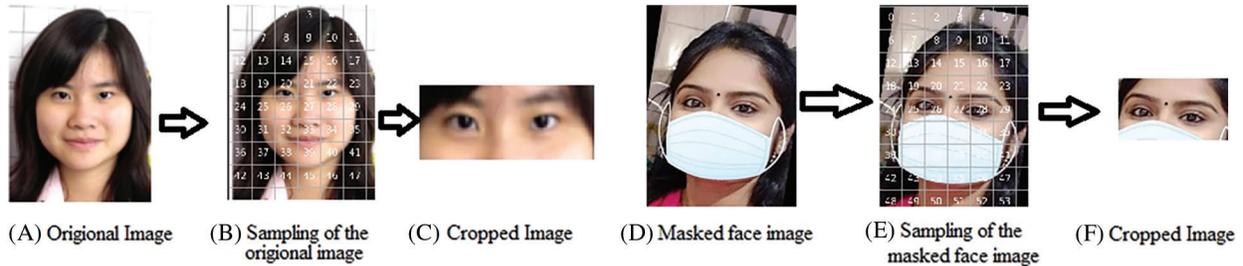
Figure 5: (A) Original and input image is showing the faces and (B) is showing the facial land marks of the faces. They are including 77 facial land marks of the faces. (C) and (D) are showing original and input image and facial land marks features on the face images

Fig. 5 depicts the output result of facial land marks detection; single camera face landmark detection may be represents in screen coordinate space, with X, Y, and Z normalized, coordinates respectively. Despite the suitability for some application, this format is unable to directly allow augmented reality (AR) capabilities such as the alignment of a virtual 3D object with a recognised face. By aligning faces can be altered, facial landmarks may be used to build in pictures [27]. As a result of transfer learning, a network is used to train the model for 3D landmarks using synthetic rendered data. Real-world data shows that the network functioned successfully. An uncropped video frame supports the 3dimensional face landmark networks. The model generates 3D point locations by referring in given face data.

In Fig. 6 shows the grid value of the input image in large scale and then we take a screen shot of the eyes. Eye is an important and effective feature of the face identification in hide face with face mask it is very

difficult to identify the face. Therefore reducing this problem, we are applying grid algorithm to detect the feature of eyes [28].

(A) Origional Image   (B) Sampling of the origional image   (C) Cropped Image   (D) Masked face image   (E) Sampling of the masked face image   (F) Cropped Image

Figure 6: (A) Original and input image of the person, (B) this is showing the grid image of the original face image, (C) cropped image is showing the property of the eyes, (D) masked face image is showing the wearing mask, (E) this image is showing the sampling face with wearing masked face image and (F) cropped image is showing the eyes for identification
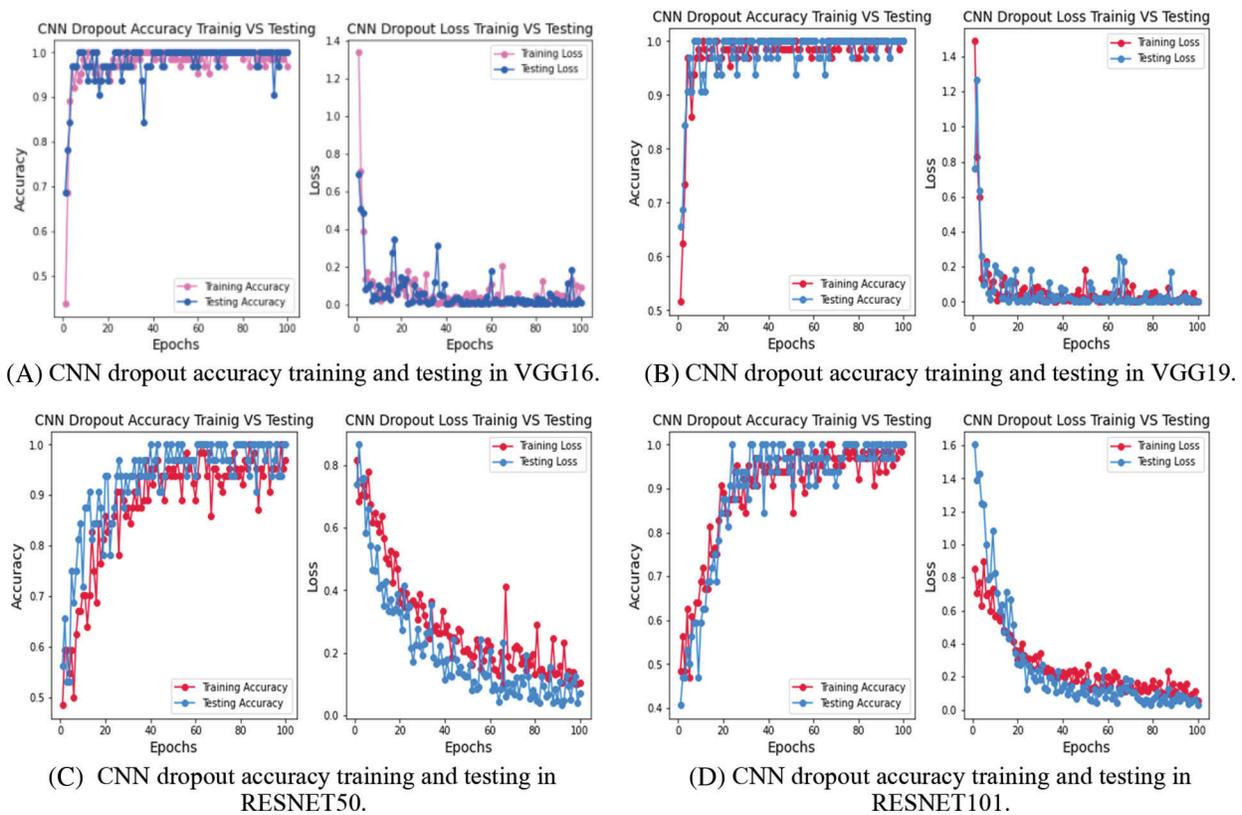
4 Result & Analysis

This article is used different model and found sufficient result of face recognition. This article uses four pre-trained model such as VGG16, VGG19, ResNet50 and ResNet101. These algorithms have provided a classical and accurate result with faces wearing mask and faces without wearing mask. Here, shows the result of the different models with running 100 epochs. Dropout simulates sparse activation from a particular layer, which, as an unexpected side effect, pushes the network to learn a sparse representation. As a result, it might be implied instead of activity regularization in autoencoders models to encourage sparse representations. By eliminating half of the feature detectors on each training example at random, this "overfitting" is considerably decreased. This avoids complicated co-adaptations in which a feature detector is only useful for conjunction with a number of other feature detectors. The combinatorial enormous range of internal circumstances, it must work instead of each neuron learns to identify a trait that is typically beneficial in giving the correct answer. For position prediction, tiny item recognition is performed utilizing feature maps from a shallower layer containing more efficient information.

The Fig. 7 is showing dropout accuracy and testing of VGG16, VGG19, ResNet50 and ResNet101. It has been regularizing technique of mimicking the parallel training in a large number of neural networks with different typologies. Convolutional layer outputs have ignored or drop out at random during training. The resultant convolutional layer appears and behaves simultaneously different than the preceding layer, with a different nodes and connections. Each modification of a layer during training is done with a fresh "view" of the present layer. Dropout adds noise into the training process by requesting that nodes within a convolutional layer and take care of responsibility in there inputs on a database from probabilistic basis. They are suppressing the suggested datasets overfitting value. Our suggested dataset has a high level of accuracy. They are exhibited various models in the above picture, but they were all operating with the same dataset and producing good results. These methods have a technique that ensures that the model and the datasets were in sync. Dropout has a regularization technique to improve accuracy and training in large number of convolutional neural networks of different typologies at the same time. By required nodes in convolutional layer take less input and provide better performance on a probabilistic basis and also dropout introduces noise into the training process.

The advantage of transfer learning is a conditional identification framework with wearing a face mask. Face size, parameter count, depth, and accuracy are all evaluated with and without the face mask. The feature extraction layers from each of the aforementioned models have been employed. The classification layers are

only trained, and the transfer model locked. Fine-tuning is performed on epochs 100, enabling all layers to learn. Other configuration information may be found in the training section of Fig. 8 that displays the training progression in each model. In COVID-19 pandemic situation occurs death ratio caused by the new virus known as severe acute respiratory syndrome coronavirus 2 (SARSCoV2) for the precaution use face masks. Several authors have presented various applications to detect the face mask and their proper usage. They have proposed a slew of concepts, all with the objective of creating applications that can recognise with face masks. While the precision of our top findings appears to be lower than that of previous comparable research, this is due to the fact that our method analyse a more intricate problem, which results in lower accuracy.

(A) CNN dropout accuracy training and testing in VGG16.

(B) CNN dropout accuracy training and testing in VGG19.

(C) CNN dropout accuracy training and testing in RESNET50.

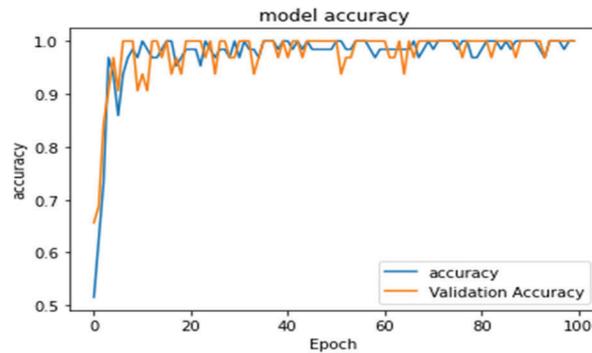(D) CNN dropout accuracy training and testing in RESNET101.

Figure 7: (A) represents the dropout accuracy and training loss using CNN in VGG16, (B) represents the dropout accuracy and training loss using CNN in VGG19, (C) represents the dropout accuracy and training loss using CNN in RESNET50 and (D) Represents the dropout accuracy and training loss using CNN in RESNET101

The above Fig. 9 shows the model accuracy and model loss function in VGG16, VGG19, ResNet50 and ResNet10 model as with using different proposed dataset. They performed good result with using proposed dataset. After running 100 epochs, they were performing a regular and continuously linear output.

This article used MobileNet V2 with Transfer learning for face recognition of wearing and without wearing mask. The proposed model is achieving a loss of 0.0041, an accuracy of 99.81%, a validation loss of 0.0169, and a validation accuracy of 99.47%. This model achieves a strong balance between loss function and accuracy benefit (See Tab. 1).
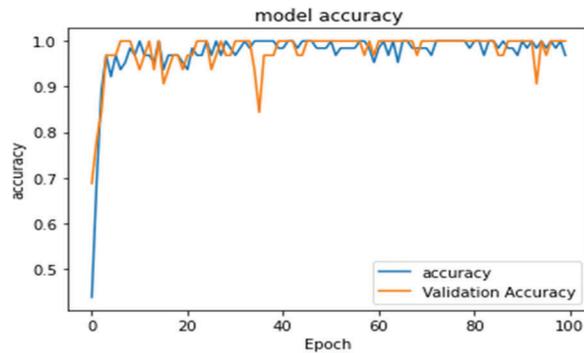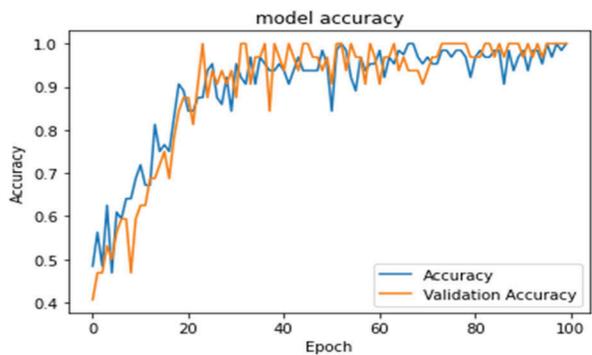
(A) Model accuracy training and testing in VGG16.

(B) Model accuracy training and testing in VGG19.
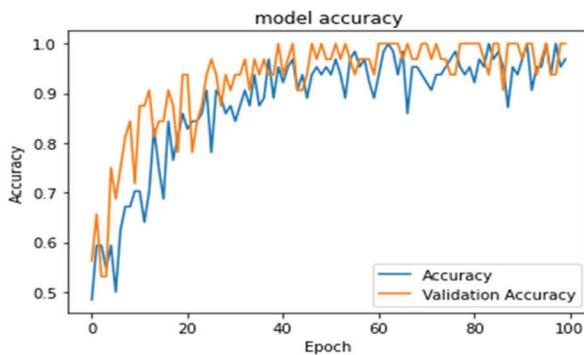
(C) Model accuracy training and testing in
RESNET50.

(D) Model accuracy training and testing in
RESNET101.

Figure 8: (A) is representing model accuracy value using VGG16 in 100 epochs. (B) is representing model accuracy value using VGG19 in 100 epochs. (C) is representing model accuracy value using RESNET50 in 100 epochs. (D) is representing model accuracy value using RESNET101 in 100 epochs

Fig. 10 depicts the output of the training loss and validation of accuracy in wearing face mask and without wearing face mask. Using MobileNet V2 with transfer learning is performed after 100 epochs and after 500 epochs. It is found that the proposed model achieved 99.81% accuracy with proposed dataset at 100 epochs and found 99.82% accuracy of 500 epochs observed. Face recognition with and without a mask are achieved using MobileNet V2 with transfer learning in this work. The suggested model has a precision of 99.81%, a loss of 0.0041%, and a validation loss of 0.0169 percent. The loss function and the accuracy advantage are well-balanced in this model. The suggested dataset is imbalanced, with 4066 masked and 4103 non-masked faces.

As a result, this article employed cropping the oversample specific photographs without wearing masks in order to get an equal number of modified images and full-faced photos for comparative analysis. This study has already covered the four pre-trained models like VGG16, VGG19, ResNet50 and ResNet101 by normalized 2D faces. Individual deep features from the recommended dataset and final convolutional layers CIFAR, MNIST RMFRD, SMFRD are used in the proposed model MobileNet V2 with transfer learning for comparison analysis. Meanwhile in terms of accuracy and losses are discovered.
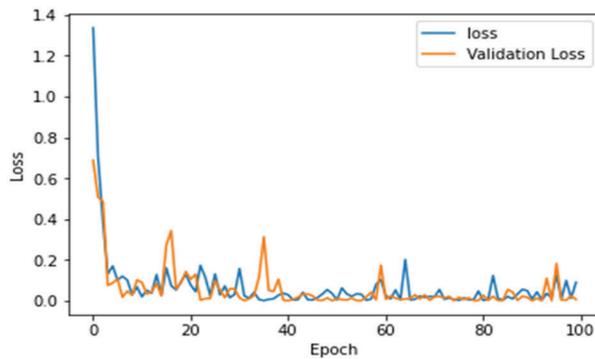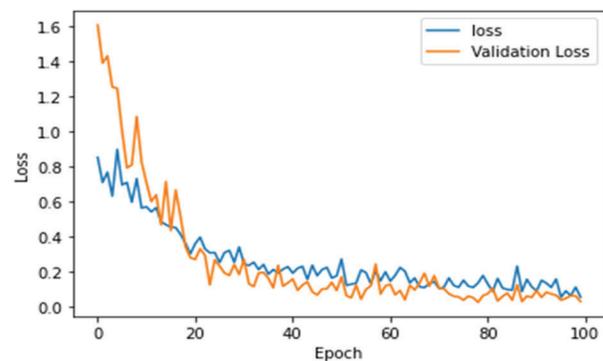
(A) Loss function in VGG16 using 100 epochs
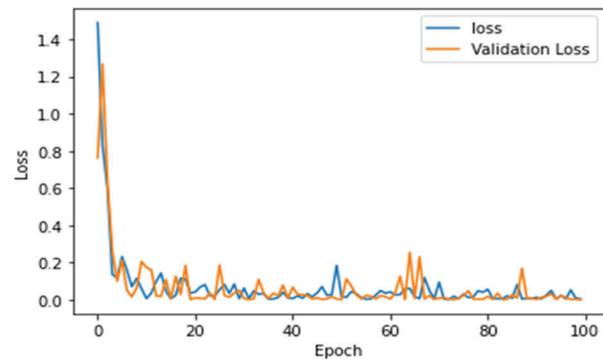
(B) Loss function in VGG16 using 100 epochs
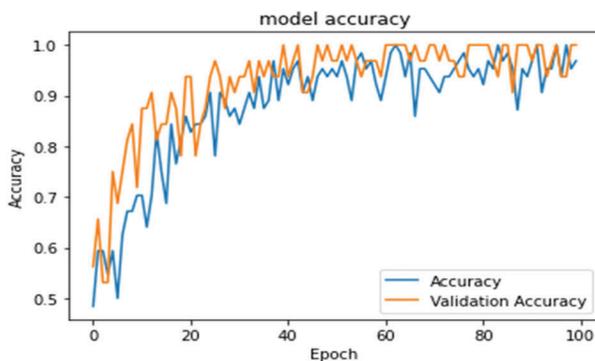
(C) Loss function in RESNET50 using 100 epochs

(D) Loss function in RESNET101 using 100 epochs

Figure 9: (A) is representing loss function in VGG16 using 100 epochs. (B) is representing loss function in VGG19 using 100 epochs. (C) is representing loss function in RESNET50 using 100 epochs. (D) is representing loss function in RESNET101 using 100 epochs

Table 1: Represents the value of loss, validation-loss, recognition accuracy and validation accuracy with different-different number of epochs

Number of epochs	Loss	Validation loss	Recognition accuracy	Validation accuracy
10	0.0428	0.0329	0.9893	0.9908
20	0.0267	0.0274	0.9923	0.9934
50	0.0108	0.0190	0.9974	0.9934
100	0.0041	0.0169	0.9981	0.9947
500	0.0031	0.0158	0.9982	0.9948
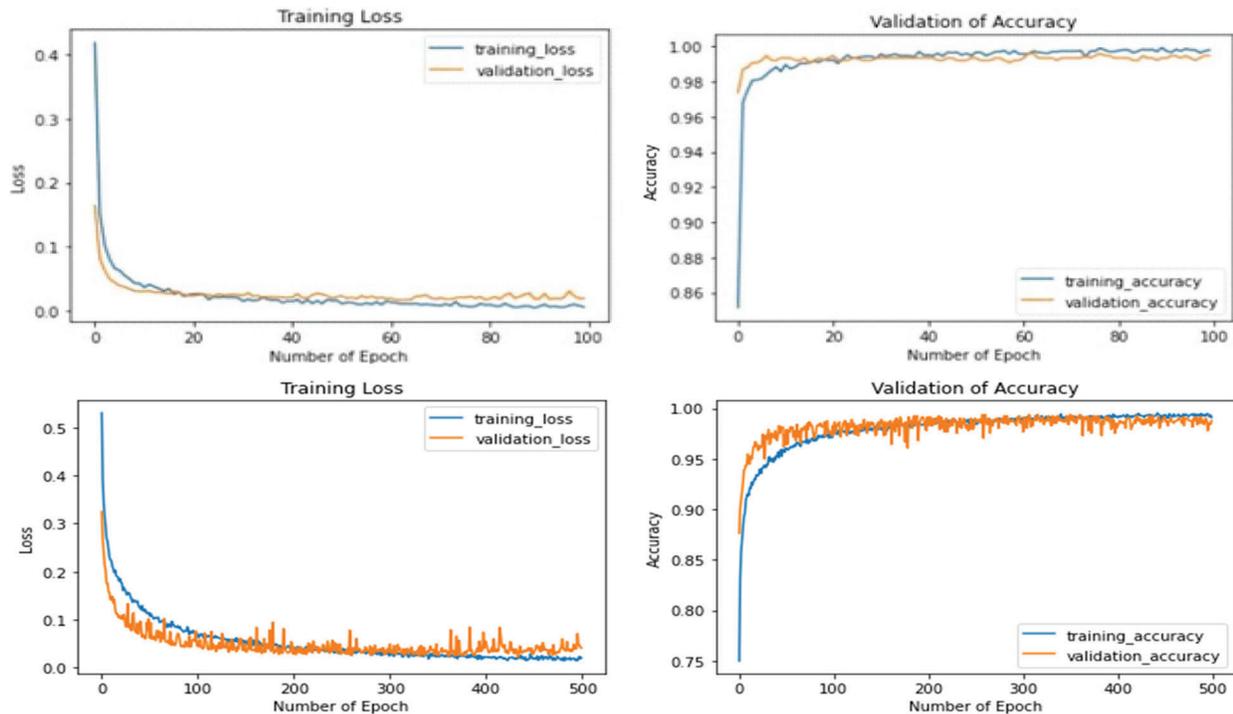
Figure 10: Above Fig. 10 is representing the training loss and validation of accuracy in wearing face mask and without wearing face mask to using proposed algorithm. It is presenting result of training loss and accuracy after running 100 epochs and 500 epochs

5 Comparative Analyses of Different Models Using Our Proposed Dataset

This article presents the facial photographs pre-processed as indicated in Section 3.3. Here, the proposed dataset has unbalanced that contains 4066 masked faces and 4103 non masked faces. So comparative analysis of this article used cropping to oversample certain without wearing mask to obtain an equal dimension of modified images and face images. Using the normalization technique for two dimensional faces, this article already covered the four pre trained models VGG16, VGG19, ResNet50 and ResNet101. Proposed model MobileNet V2 with transfer learning for comparative analysis individually to use deep features extracted from their final convolutional layers CIFAR, MNIST RMFRD, SMFRD and modified proposed datasets. It has discovered that the proposed model based on MobileNet V2 with transfer learning outperformed other models in terms of accuracy (See Tab. 2). The face photographs with and without masks are used, despite the fact that wearing a mask is critical for this situation. Some people were engaging in unethical behaviour during COVID-19. As a result, it's difficult to tell who is real. This article attempted to address this issue by creating a model for detecting faces with and without masks, as well as matching the faces in the dataset. The proposed models have been compared for the VGG16, VGG19, ResNet50 and ResNet101. Different classifiers with different optimizers must be assessed in order to create a system that can be applied on a big scale. This article compares classifiers such as MobileNetV2, RESNET101, RESNET50, VGG16, and VGG19 with Optimizer ADAM. The computation are used to support the use of contactless facial authentication while wearing a face mask and provided 99.82% accuracy. However, due to lack of facial feature information, masked face detection is a difficult process. By leveraging RetinaFace, we can extract local binary pattern characteristics from the masked face's eye, forehead, and eyebow regions and integrate them with features learned from RetinaFace to create a unified framework for detecting masked faces.

**Table 2:** Result analysis of face recognition accuracy using different model with different datasets

Models	Datasets	Accuracy (%)
VGG16	CIFAR10 [23]	93.56
	MNIST [15]	92.7
	RMFRD [2]	91.3
	SMFRD [2]	85.6
	Proposed Dataset	99.302
VGG19	CIFAR10 [23]	92.9
	MNIST [15]	93.3
	RMFRD [2]	87.3
	SMFRD [2]	87.5
	Proposed dataset	99.548
RESNET 50	CIFAR10 [23]	93.57
	MNIST [15]	98.97
	RMFRD [2]	87.9
	SMFRD [2]	88.9
	Proposed dataset	78.700
RESNET 101	CIFAR10 [23]	94.68
	MNIST [15]	97.89
	RMFRD [2]	88.91
	SMFRD [2]	89.82
	Proposed dataset	98.641
MobileNet V2 with transfer learning	CIFAR10	92.4
	MNIST	93.9
	RMFRD	70
	SMFRD	69
	Proposed dataset	99.82

The above Fig. 11 is showing the output result of different model with different datasets. Proposed models have been worked well with different datasets. Our proposed model is working well in classification of the object in given datasets and doing well in face classification. CIFAR, MNIST, RMFRD, SMFRD, and suggested datasets are used to construct a state-of-the-art method.

6 Discussions

This article, represents a comprehensive review of the most current publications on viral infection prevention, including wearing face mask in community region, that is protected the prevention of virus transmission, as well as concerns of security acceptability based on the user's community and surroundings. Wearing face mask are a technique to prevent the spread of corona virus and other respiratory illnesses in affected areas. However, wearing a mask alone may not provide adequate

protection against viruses; additional, equally critical measures must be followed. If masks are used, they must be used in combination with hand hygiene and other measures to prevent COVID-19 transmission. As everyone in COVID-19 situation wearing mask, identifying the theft facial recognition was incredibly difficult. At the time of consequence, determining the theft's facial recognition will be tough. To overcome the challenge, in this article first collected data on whether or not people use face masks. Then proposed model is used to identify the faces in the database. MTCNN are used to detect faces, whereas MobileNet V2 with transfer learning has used to identify them. Data collection from datasets with and without masks is shown previous figure. The model trained using the Tensor Flow and Keras libraries once the data was collected. A list of the qualities of the faces may be found in this library. The MTCNN algorithm used to detect faces. After discovering each face, this article used a cascade classifier to exclude it. Recent breakthrough in deep learning and artificial intelligence has helped to improve the efficacy of face image-based machine learning and/or synthesizing by better comprehending facial expressions. Here, to extract a variety of little characteristics from a huge shot of a human face. This ensures that the camera can detect even the tiniest changes, otherwise go unnoticed through the naked eye. In the last stage the faces with and without masks were distinguished and classified. This article also used various datasets and combines them with proposed model and other existing models. When we are including millions of images they will provide accurate and simultaneous accuracy in the result shows in the comparative analysis. In comparison to other models, the proposed model provides sutaible accuracy and loss function.
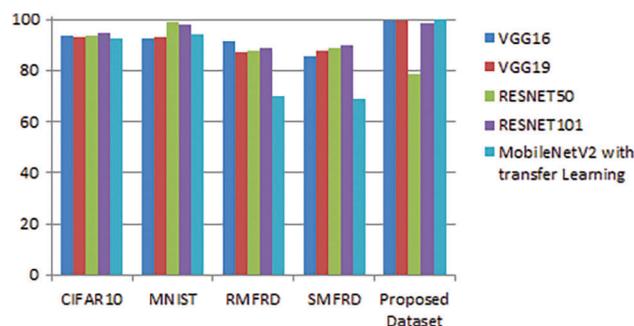
Figure 11: Graphical representation of the different model with different datasets

7 Conclusions

The amid of COVID-19 is an once-in-a-lifetime calamity that has caused a massive number of deaths and security issues. People were wearing face masks on a daily basis to protect themselves and to help prevent the transmission of the deadly virus. This article, proposed a MobileNet V2 with transfer learning based technique for tackle the difficulty of masked face identification using occlusion reduction and deep learning-based features. The first step removed the mask from the masked part of the face. Then the proposed approach applied on CIFAR, MNIST RMFRD and SMFRD datasets. The proposed model has been achieved 99.82% recognition accuracy. This article also examines the assistance of four pre trained models such as VGG16, VGG19, ResNet50 and ResNet101. This proposed dataset has been used VGG16, VGG19, ResNet50, ResNet101 and VGG16 for achieving an accuracy 99.30%. VGG19, RESNEt50, and RESNET101has achieved accuracy 99.54%, 78.70% and 98.64%, respectively. The comparative analysis showed that the proposed model performs better in comparison to all pre-trained models, which already existed. The use of MobileNet V2 with transfer learning to monitor with and without face masks to limit the speed of the corona virus as well as to detect persons using face masks is the work's main contribution.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] C. Jagadeeswari and M. U. Theja, "Performance evaluation of intelligent face mask detection system with various deep learning classifiers," *International Journal of Advanced Science and Technology*, vol. 29, no. 11, pp. 3083–3087, 2020.

[2] W. Hariri, "Efficient masked face recognition method during the covid-19 pandemic," 2021. [Online]. Available: https://arxiv.org/abs/2105.03026.

[3] K. Teke, A. Manjare and S. Jamdar, "Survey on face mask detection using deep learning," *International Journal on Data Science and Machine Learning with Applications*, vol. 1, no. 1, pp. 1–9, 2021.

[4] Z. Wang, G. Wang, B. Huang, Z. Xiong, W. H. Hong *et al.,* "Masked face recognition dataset and application," 2020. [Online]. Available: https://arxiv.org/abs/2003.09093.

[5] N. Fasfous, M. R. Vemparala, A. Frickenstein, L. Frickenstein, M. Badawy *et al.,* "BinaryCoP: Binary neural network-based COVID-19 face-mask wear and positioning predictor on edge devices," in *IEEE Int. Parallel and Distributed Processing Symp. Workshops (IPDPSW)*, Portland, OR, USA, IEEE, pp. 108–115, 2021.

[6] V. P. Singh and R. Srivastava, "Automated and effective content-based mammogram retrieval using wavelet based CS-LBP feature and self-organizing map," *Biocybernetics and Biomedical Engineering*, vol. 38, no. 1, pp. 90–105, 2018.

[7] J. Tomás, A. Rego, S. Viciano-Tudela and J. Lloret, "Incorrect facemask-wearing detection using convolutional neural networks with transfer learning," *Healthcare*, vol. 9, no. 8, pp. 1–17, 2021.

[8] A. Alzu'bi, F. Albalas, T. Al-Hadhrami, L. B. Younis and A. Bashayreh, "Masked face recognition using deep learning: A review," *Electronics*, vol. 10, no. 21, pp. 1–35, 2021.

[9] N. U. Din, K. Javed, S. Bae and J. Yi, "A novel GAN-based network for unmasking of masked face," *IEEE Access*, vol. 8, pp. 44276–44287, 2020.

[10] M. Y. Saib and S. Pudaruth, "Is face recognition with masks possible?," *International Journal of Advanced Computer Science and Applications (IJACSA)*, vol. 12, no. 7, pp. 43–50, 2021.

[11] S. Kumar, A. Jain, A. Kumar Agarwal, S. Rani and A. Ghimire, "Object-based image retrieval using the U-net-based neural network," *Computational Intelligence and Neuroscience*, vol. 2021, no. 2, pp. 1–14, 2021.

[12] A. Anwar and A. Raychowdhury, "Masked face recognition for secure authentication," 2020. [Online]. Available: https://arxiv.org/abs/2008.11104.

[13] G. J. Chowdary, N. S. Punn, S. K. Sonbhadra and S. Agarwal, "Face mask detection using transfer learning of inceptionv3," in *Int. Conf. on Big Data Analytics*, Springer, Cham, vol. 12581, pp. 81–90, 2020.

[14] M. Loey, G. Manogaran, M. H. N. Taha and N. E. M. Khalifa, "A hybrid deep transfer learning model with machine learning methods for face mask detection in the era of the COVID-19 pandemic," *Measurement Science Direct Elsevier*, vol. 167, no. 5, pp. 1–14, 2021.

[15] B. Qin and D. Li, "Identifying facemask-wearing condition using image super-resolution with classification network to prevent COVID-19," *Sensors*, vol. 20, no. 18, pp. 1–23, 2020.

[16] S. Sethi, M. Kathuria and T. Kaushik, "Face mask detection using deep learning: An approach to reduce risk of Coronavirus spread," *Journal of Biomedical Informatics*, vol. 120, no. 5, pp. 1–12, 2021.

[17] S. E. Eikenberry, M. Mancuso, E. Iboi, T. Phan, K. Eikenberry *et al.,* "To mask or not to mask: Modeling the potential for face mask use by the general public to curtail the COVID-19 pandemic," *Infectious Disease Modelling*, vol. 5, no. 8, pp. 293–308, 2020.

[18] H. N. Vu, M. H. Nguyen and C. Pham, "Masked face recognition with convolutional neural networks and local binary patterns," *Applied Intelligence*, vol. 4374, pp. 1–16, 2021.

[19] Y. Li, K. Guo, Y. Lu and L. Liu, "Cropping and attention based approach for masked face recognition," *Applied Intelligence*, vol. 51, no. 5, pp. 3012–3025, 2021.

[20] H. Mliki, S. Dammak and E. Fendri, "An improved multi-scale face detection using convolutional neural network," *Signal Image and Video Processing*, vol. 14, no. 7, pp. 1345–1353, 2020.

[21] M. Farooq and A. Hafeez, "Covid-resnet: A deep learning framework for screening of covid19 from radiographs," 2020. [Online]. Available: https://arxiv.org/abs/2003.14395.

[22] X. Su, M. Gao, J. Ren, Y. Li, M. Dong *et al.,* "Face mask detection and classification via deep transfer learning," *Multimedia Tools and Applications*, vol. 81, pp. 1–20, 2021.

[23] A. Kumar and A. Jain, "Image smog restoration using oblique gradient profile prior and energy minimization," *Frontiers of Computer Science*, vol. 15, no. 6, pp. 1–7, 2021.

[24] S. Almabdy and L. Elrefaei, "Deep convolutional neural network-based approaches for face recognition," *Applied Sciences*, vol. 9, no. 20, pp. 1–21, 2019.

[25] R. C. Çalik and M. F. Demirci, "Cifar-10 image classification with convolutional neural networks for embedded systems," in *IEEE/ACS 15th Int. Conf. on Computer Systems and Applications (AICCSA)*, Aqaba, Jordan, IEEE, pp. 1–6, 2018.

[26] Y. Li, "Facemask detection using inception V3 model and effect on accuracy of data preprocessing methods," *Journal of Physics*, vol. 2010, no. 1, pp. 1–7, 2021.

[27] W. Sun, G. Dai, X. Zhang, X. He and X. Chen, "TBE-Net: A three-branch embedding network with part-aware ability and feature complementary learning for vehicle re-identification," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–13, 2021. https://dx.doi.org/10.1109/TITS.2021.3130403.

[28] W. Sun, L. Dai, X. Zhang, P. Chang and X. He, "RSOD: Real-time small object detection algorithm in UAV-based traffic monitoring," *Applied Intelligence*, vol. 52, pp. 1–16, 2021.