International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 06 Issue: 06 | June - 2022 Impact Factor: 7.185 ISSN: 2582-3930

REAL TIME FACE MASK DETECTION USING MOBILENET WITH HARR CASCADED TECHNIQUE

NITHYAA SHRI N S[1], PADMAJA T[2], UMARANI C[3]

[1]Department of CSE-UG, Kingston Engineering College, Vellore-59
[2]Department of CSE-UG, Kingston Engineering College, Vellore-59
[3]Assistant professor, Department of CSE, Kingston Engineering College, Vellore-59

Abstract - Wearing a mask in most of the non-pharmaceutical measures that may be used to reduce the main source of COVID droplets expelled via way of means of an infected individual. To make contributions towards the communal health, this paper pursuits to devise a highly accurate and real-time method that could effectively stumble on non-masks faces in public and thus, implement them to put on mask. Although several researchers have dedicated efforts in designing efficient algorithms for face detection and recognition, there exists an important distinction between 'detection of the face beneath masks' and 'detection of masks over the face. Manual real-time tracking of facemask sporting for a huge group of human beings is turning into a hard task. The aim of this paper is to apply deep learning (DL), which has proven incredible outcomes in lots of real-existence applications, to make sure efficient real-time facemask detection. The proposed method is primarily based on two steps. An off-line step aiming to create a DL model this is capable of stumble on and discover facemasks and whether they are appropriately worn or not. An on-line step that deploys the DL version at edge computing on the way to stumble on mask in real-time. In this study, we advocate to apply MobileNetV2 to detect facemask in real-time. Several experiments are performed and display true performances of the proposed method (99% for training and testing accuracy).

***Key Words*:** Face mask detection, Deep learning.

1.INTRODUCTION

With the reopening of nations from COVID-19 lockdown, Government and Public health departments are recommending face mask as an important measure to make us secure while venturing into public to curtail the unfold of Coronavirus and thereby contributing to public healthcare. Regardless of discourse on clinical sources and diversities in mask, all nations are mandating coverings over the nostril and mouth in public. To mandate using facemasks, it turns into important to plot a few strategies that implement people to use a mask earlier than exposing to public places. This software may be very beneficial in public regions inclusive of airports, railway stations, crowded markets, malls, etc. The proposed approach used over here has two steps. The first step is training the face masks detector with the use of transfer learning technique. The next step is to apply this trained face masks detector on photographs or motion pictures of humans to become aware of if they're wearing a mask.

Although several researchers have dedicated efforts in designing green algorithms for face detection and recognition however there exists an important difference among 'detection of the face below masks' and 'detection of masks over face'. As per the available literature, little or no frame of studies is tried to discover masks over face. Thus, our paper pursuits to a broaden method which could correctly discover masks over the face in public regions (inclusive of airports. railway stations, crowded markets, bus stops, etc.) to curtail the

unfold of Coronavirus and thereby contributing to public healthcare. Further, it isn't smooth to discover faces with/without a mask in public because the dataset to be had for detecting mask on human faces is notably small which leads to the hard training of the model.

So, the idea of transfer learning is used right here to switch the learned kernels from networks trained for the same face detection undertaking on an intensive dataset. The dataset covers numerous face photographs consisting of faces with mask, faces without mask, faces with and without mask in a single picture and difficult photographs without mask. With an intensive dataset containing 45,000 images, our method achieves exquisite accuracy of 98.2%. The most important contribution of the proposed work is given below:

- Develop a novel object detection approach that combines one-level and two -level detectors for accurately detecting the objects in real-time from video streams with transfer learning technique on the back end.
- Improved affine transformation is evolved to crop the facial regions from out-of-control real-time photographs having variations in face size, orientation and background. This step enables in higher localizing the individual that who is violating the facemask norms in public regions/ offices.
- Creation of impartial facemask dataset with imbalance ratio equals to almost one.
- The proposed version requires very less memory, making it without problems deployable for embedded gadgets used for surveillance purposes.

People are pressured via way of means of legal guidelines to put on face mask in public in lots of nations. These policies and legal guidelines had been evolved as a movement to the exponential increase in covid cases and deaths in lots of regions. However, the system of tracking huge groups of humans is turning into greater hard in public regions. So, we can create an automation system for detecting the faces. Here we introduce a facemask detection version that is primarily based totally on computer vision and deep learning. The proposed version may be incorporated with Surveillance Cameras to hinder the COVID-19 transmission via way of means of permitting the detection of folks who are wearing mask and those who are not wearing face mask. The version is integration between deep learning and classical machine learning techniques with Open cv, Tensor flow and Keras.

2. RELATED WORKS

The look is finished with 3 famous baseline models viz. Resnet50, AlexNet and MobileNet. It is found that the proposed approach achieves excessive accuracy (98.2%) at the same time as carried out with ResNet50. The top notch smooth now no longer unusual place ordinary average overall performance of the proposed version is alternatively appropriate for video surveillance devices.[1] COVID-19 has had a brilliant effect at the worldwide. Face masks detection through image processing is one of the excessive-accuracy and green face masks detectors proposed on this paper.[4] The version is informed on a actual worldwide dataset and examined with stay video streaming with an fantastic accuracy. Further the accuracy of the version with precise hyper parameters and more than one people at precise distance and place of the body is done.[3] An online step that deploys the DL version at trouble computing that lets in you to come upon mask in actual time. Several experiments are finished and display accurate performances of the proposed approach (99% for schooling and locating out accuracy). In addition, numerous comparisons with many ultra-current fashions particularly ResNet50, DenseNet, and VGG16 display accurate smooth now no longer unusual place ordinary

average overall performance of the MobileNetV2 in phrases of schooling time and accuracy.[7] As the usual begins off evolved going through numerous ranges of reopening, face mask have come to be a vital detail of our everyday lives to live safe. Wearing face mask may be required that lets in you to socialize or behavior business. So, this software program software program application makes use of a digital virtual digital digicam to come upon if someone is carrying a mask or now no longer.[5] This system is an actual-time software program software program application to come upon people if they are carrying a mask or are without a masks. It has been informed with the dataset that includes round 4000 photographs the use of 224×224 as width and top of the image and function carried out an accuracy rate of 98%. In this research, this version has been informed and compiled with 2 CNN for differentiating accuracy to choose out the brilliant for this form of version. It may be positioned into motion in public regions which embody airports, railways, schools, offices, etc. to test if COVID-19 hints are being adhered to or now no longer.

3.PROPOSED SYSTEM:

In proposed work,

- Train Data getting to know version.
- Apply facemask detector overstay video circulate.

Data Visualization or accumulating records formation Visualization or accumulating information. There are 1800 pictures withinside the "yes" class and 1800 pictures withinside the "no" class detections and by skip the magnificence withinside the speech synthesizer with a purpose to generate the audio voice alert to help the visually impaired.

Providing greater, amplify the records set to consist of a bigger quantity of pictures for schooling. In this step of records expansion, we rotate and turn every picture withinside the dataset.

Training the CNN version, positioned pictures in to schooling set to apply the series version construct via way of means of the keras library. We will educate the version for 10 epochs. By schooling the version with greater epoch to get better accuracy.

Taking enter from camera, take picture enter from video series with the assist of built-in computer camera. Face detection with masks detection set of rules with the assist of Haar Cascades to discover items from the picture and CNN (Convolutional Neural Network) face masks detection.

The version proposed right here is designed and modeled the usage of python libraries particularly TensorFlow, Keras and OpenCV. The version we used is the MobileNetV2 of convolutional neural network. The technique of the usage of MobileNetV2 is known as the usage of Transfer Learning. Transfer getting to know is the usage of a few pre skilled version to train model and get the prediction which saves time and makes the using training extraordinary different models easy. We track the version with the hyper parameters: getting to know rate, quantity of epochs and batch size. The version is skilled with a dataset of pictures with magnificence, with masks and without masks. The dataset has 993 images of with mask class and 1918 images of without mask class.

i. Training the model with the taken dataset
ii. Deploying the model

First feed the dataset withinside the version, run the training program, which trains the models at the given dataset. Then run the detection program, which activates the video stream, captures the frames constantly from the video stream with an anchor field the usage of object detection process. This is handed through the

MobileNetV2 model layers which classifies the picture as without or with masks. If the man or woman is wearing a mask, a green anchor field is displayed and red if no longer wearing a mask with the accuracy for the equal tagged at the anchor field.

3.1 ADVANTAGES

- Complexity is less compared to previous process
- Ability to learn and extract complex features.
- Accuracy is good
- With its simplicity and fast processing time, the proposed algorithm gives better execution time.

4.MODULE DESCRIPTION

- **VIDEO ACQUISITION**
- **DIVIDING INTO FRAMES**
- **FACE DETECTION**
- **FEATURE DETECTION**
- **MASK DETECTION**
- **CLASSIFICATION**

4.1 Dataset and Preprocessing

This section provides an overview of the approach to prepare a deep learning model that will be used later to decide if the person is wearing mask or not. The proposed method will group frames in videos, based on special facial features obtained harrcascaded.

Fig 1. Dataset and Pre-processing

4.2 Video acquisition:

Video acquisition especially entails acquiring the stay video feed of the car driver. Video acquisition is achieved, with the aid of using a digital digicam. Video Acquisition OpenCV offers considerable aid for obtaining and processing stay videos. It is likewise viable to pick whether the video must be captured from the in-constructed webcam or an outside digital digicam with the aid of using putting the proper parameters. As cited earlier, OpenCV does now no longer specify any minimal requirement at the digital digicam, but OpenCV with the aid of using default expects a selected decision of the video this is being recorded, if the resolutions do now no longer match, then an mistake is thrown. This mistake may be countered, with the aid of using over using the default value, which may be achieved, with the aid of using manually specifying the decision of the video being recorded.

4.3 Dividing into frames:

This module is used to take live video as its input and convert it into a series of frames/ images, which are then processed.

Once the video has been acquired, the next step is to divide it into a series of frames/images. This was initially done as a 2-step process. The first step is to grab a frame from the camera or a video file, in our case since the video is not stored, the frame is grabbed from the camera and once this is achieved, the next step is to retrieve the grabbed frame. While retrieving, the image/frame is first decompressed and then retrieved. However, the two-step process took a lot of processing time as the grabbed frame had to be stored temporarily. To overcome this problem, we came up with a single step process, where a single function grabs a frame and returns it by decompressing.

4.4 Face detection:

The face detection characteristic takes one body at a time from the frames supplied through the body grabber, and in every and each body it attempts to stumble on the face of the auto driving force. This is done through utilizing a hard and fast of pre-described samples. Once the frames are efficaciously extracted the subsequent step is to stumble on the face in every of those frames. This is done through utilizing the Haarcascade document for face detection. The Haarcascade document includes some of functions of the face, inclusive of height , width and thresholds of face colors., it's far built through the usage of some of high quality and poor samples. For face detection, we first load the cascade document. Then by skip the obtained body to an facet detection characteristic, which detects all of the feasible items of various sizes withinside the body. To lessen the quantity of processing, rather than detecting items of all feasible sizes, for the reason that face of the auto driving force occupies a huge a part of the image, we are able to specify the brink detector to stumble on handiest items of a selected length, this length is determined primarily based totally at the Haarcascade document, in which every Haarcascade document might be designed for a selected Now, the output the brink detector is saved in an array. Now, the output of the brink detector is then as compared with the cascade document to discover the face withinside the body. Since the cascade includes each high quality and poor samples, it's far required to specify the range of screw ups on which an item detected need to be categorized as a poor sample. In our system, we set this price to 3, which helped in reaching each accuracy in addition to much less processing time. The output of this module is a body with face detected in it.

4.5 Eyes detection:

Once the face detection function has detected the face of the automobile driver, the eyes detection function tries to detect the automobile driver's eyes. This is achieved by making use of a set of pre-defined samples.

Once the region of interest is marked, the edge detection technique is applied only on the region of interest, thus reducing the amount of processing significantly. Now, we make use of the same technique as face detection for detecting the eyes by making use of Haarcascade Xml file for eyes detection. But the output obtained was not very efficient, there were more than two objects classified as positive samples, indicating more than two eyes. To overcome this problem, the following steps are taken:

- Out of the detected objects, the object which has the highest surface area is obtained. This is considered as the first positive sample.
- Out of the remaining objects, the object with the highest surface area is determined. This is considered as the second positive sample.
- A check is made to make sure that the two positive samples are not the same.
- Now, we check if the two positive samples have a minimum of 30 pixels from either of the edges.
- Next, we check if the two positive samples have a minimum of 20 pixels apart from each other. After passing the above tests, we conclude that the two objects i.e positive sample 1 and positive sample 2, are the eyes of the automobile driver.

4.6 Mask detection:

After detecting the face of the user, the mask detection function detects if the person is wearing mask or not, by taking into consideration the state of the mask and no

mask state. As the proposed system makes use of OpenCV libraries, there is no necessary minimum resolution requirement on the camera. In the proposed algorithm, first video acquisition is achieved by making use of an external camera placed in front of the automobile driver. The acquired video is then converted into a series of frames/images. The next step is to detect the users face, in each and every frame extracted from the video. As indicated, we start with discussing face detection which has 2 important functions

(a) Identifying the region of interest, and

(b) Detection of face from the above region.

To avoid processing the entire image, we mark the region of interest. By considering the region of interest it is possible to reduce the amount of processing required and speeds up the processing, which is the primary goal of the proposed system.

4.7 CLASSIFICATION

When data has been ready, we apply Deep Learning Technique. We use different classification and ensemble techniques, to predict mental illness. The methods applied on brain MRI dataset. Main objective to apply Deep Learning Techniques to analyse the performance of these methods and find accuracy of them, and also been able to figure out the responsible/important feature which play a major role in prediction.

4.7.1 Classifier Training:

A classifier is a function that takes features as input and generates a class label prediction. Based on the learning function and underlying assumptions, different types of classifiers can be developed. Neuroimaging studies have applied various classifiers for mental illness prediction. The dimensionality issue associated with the relatively large number of features and the small number of samples should be accounted for while applying such classification algorithms.

CNN is a type of Neural Networks widely used for image recognition and image classification. CNN uses supervised learning. CNN consists of filters or neurons that have biases or weights. Every filter takes some inputs and performs convolution on the acquired input. The CNN classifier has four layers; Convolutional, pooling, Rectified Linear Unit (ReLU), and Fully Connected layers.

i. Convolutional layer

This layer extracts the features from the image which is applied as input. The neurons convolve the input image and produce a feature map in the output image, and this output image from this layer is fed as an input to the next convolutional layer.

ii. Pooling layer

This layer is used to decrease the dimensions of the feature map still maintaining all the important features. This layer is usually placed between two convolutional layers.

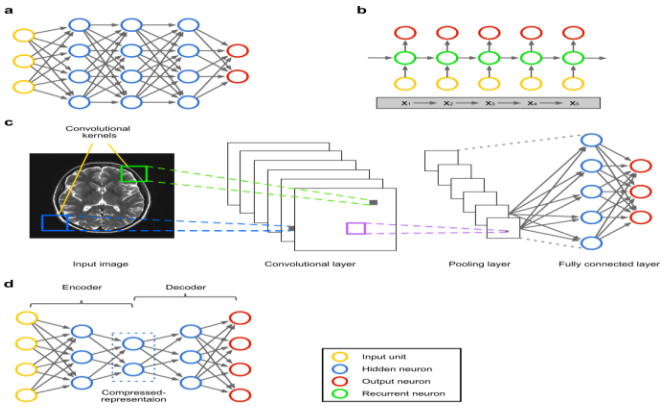

Fig 2. Pooling layer

iii. ReLu layer

ReLu is a non-linear operation which replaces all the negative values in the feature map by zero. It is an element wise operation.

iv. Fully Connected layer

FLC means that each filter in the previous layer is connected to each filter in the next layer. This is used to classify the input image based on the training dataset into various classes.

It has four phases:

1. Model construction

2. Model training

3. Model testing

4. Model evaluation

Model construction depends on machine learning algorithms. In this projects case, it was Convolution Neural Networks. After model construction it is time for model training. Here, the model is trained using training data and expected output for this data. Once the model has been trained it is possible to carry out model testing. During this phase a second set of data is loaded. This data set has never been seen by the model and therefore it's true accuracy will be verified. After the model training is complete, the saved model can be used in the real world. The name of this phase is model evaluation.

5.ARCHITECTURE DIAGRAM:

Designing of system is the process in which it is used to define the interface, modules and data for a system to specify the demand to satisfy. System design is seen as the application of the system theory. The main thing of the design a system is to develop the system architecture by giving the data and information that is necessary for the implementation of a system.

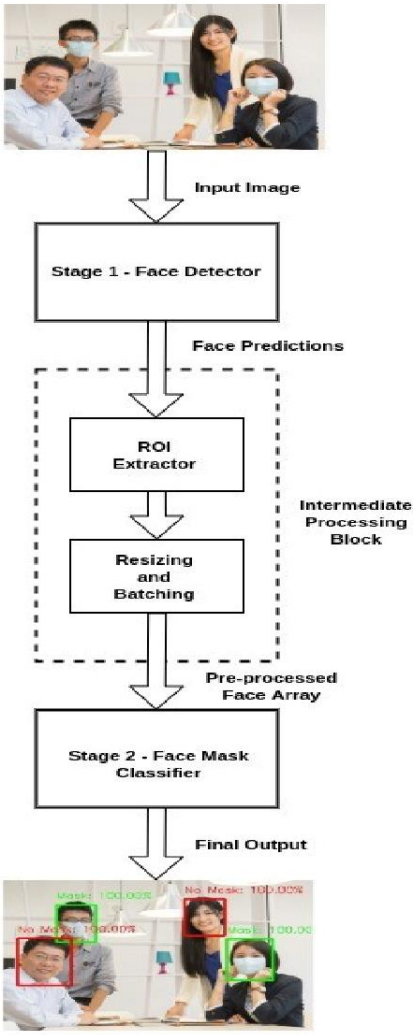

Fig 3: Architecture diagram of face mask detection using CNN

6.TABLES:

Before we visit the operating of CNN's let's cover the fundamentals inclusive of what's a picture is, and the way is it represented. An RGB image is not anything however a matrix of pixel values having 3 planes while a grayscale image is identical however it has a unique plane. let's examine this picture to understand better.

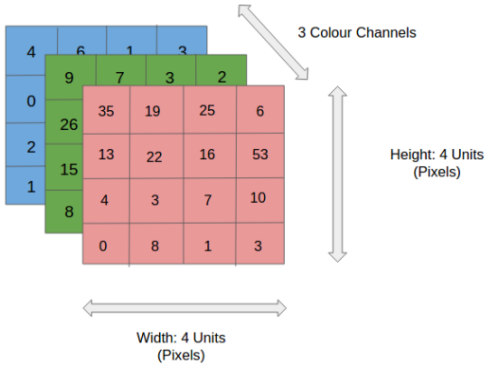

Fig 4: RGB Image

let's stay with grayscale pics as we attempt to recognize how CNNs work.

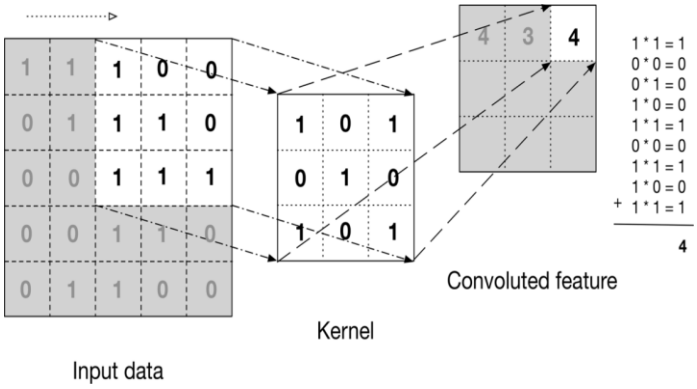

Fig 5: Grayscale Image

The above image suggests what a convolution is. We take a filter/kernel (3×3 matrix) and use it on the input image to get the convolved feature. This convolved feature is handed directly to the subsequent layer. In the case of RGB color, channel takes this animation to recognize its operating

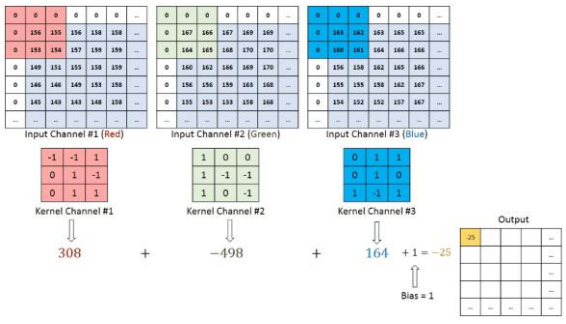

Fig 6: Extraction Image

The first layer typically extracts simple features inclusive of horizontal or diagonal edges. This output is handed directly to the subsequent layer which detects extra complicated capabilities inclusive of corners or combinational edges. As we pass deeper into the network it could discover even extra complicated feature inclusive of objects, faces, etc. Based on the activation map of the very last convolution layer, the classification layer outputs a fixed of self-assurance scores (values among zero and 1) that explain how possibly the picture is to belong to a "class" For instance, when you have a ConvNet that detects cats, dogs, and horses, the output of the very last layer is the opportunity that the input image includes any of these animals.

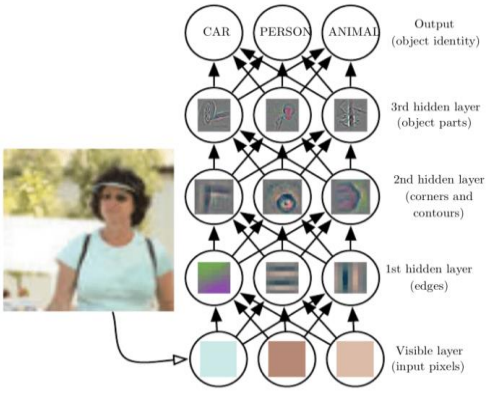

Fig 7: Illustration of a deep gaining knowledge of model

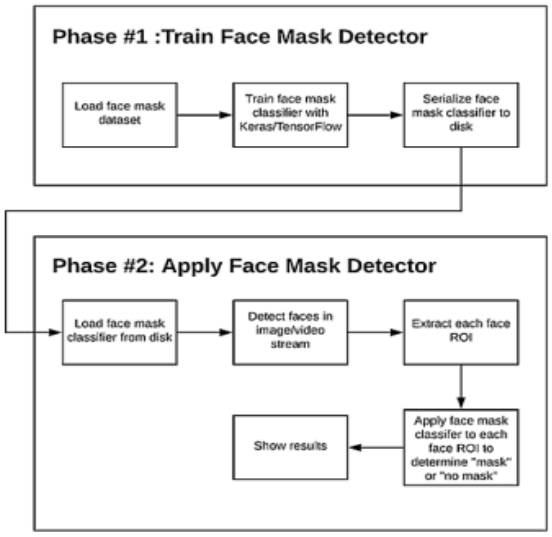

Fig 8: Phases and individual steps for constructing a face masks detector with computer vision and deep learning using Python, OpenCV, and TensorFlow/Keras

7.CONCLUSION:

In this paper, to moderate the spread of the COVID-19 pandemic, measures should be taken. We have demonstrated a facemask detector using Convolutional Neural Network and move learning techniques in neural organizations. To train, validate and test the model, we utilized the dataset that consisted of 993 masked faces pictures and 1918 exposed faces pictures. An efficient CNN model based on MobileNetV2 for Real-time Facemask Detection is presented along with Harr cascaded feature extraction. The proposed approach achieved 99 accuracies in training and testing and can determine whether a mask is appropriately worn or not in real-time video streams. Extensive experiments are conducted to show the good performances of the MobileNetV2 model in detecting facemask in real-time videos.

8.ACKNOWLEDGEMENT

The authors would like to thank Mrs.Umarani C for her suggestions and excellent guidance throughout the project period.

9. REFERENCES:

[1] J. Liu, Y. Pan, M. Li, Z. Chen, L. Tang, C. Lu, J. Wang, "Applications of deep learning to MRI images: A survey," Big Data Mining and Analytics, vol. 1, no. 1, pp. 1-18, 2018.

[2] R. Miotto, F. Wang, S. Wang, X. Jiang, J.T. Dudley, "Deep learning for healthcare: review, opportunities, and challenges," Briefings in bioinformatics, vol. 19, no. 6, pp. 1236-1246, 2018.

[3] W. Boulila, H. Ghandorh, M.A. Khan, F. Ahmed, J. Ahmad, "A novel CNN-LSTM-based approach to predict urban expansion," Ecological Informatics, vol. 64, pp. 101325, 2021.

[4] M. Alkhelaiwi, W. Boulila, J. Ahmad, A. Koubaa, M. Driss, "An Efficient Approach Based on Privacy-Preserving Deep Learning for Satellite Image Classification," Remote Sensing, vol. 13, no. 11, pp. 2221, 2021.

[5] X.W. Chen and X. Lin, "Big data deep learning: challenges and perspectives," IEEE Access, vol. 2, pp. 514-525, 2014.

[6] Q. Zhang, L.T. Yang, Z. Chen, P. Li, "A survey on deep learning for big data," Information Fusion, vol. 42, pp. 146-157, 2018.

[7] W. Boulila, M. Sellami, M. Driss, M. Al-Sarem, M. Safaei, F.A. Ghaleb, "RS-DCNN: A novel distributed convolutional-neuralnetworks based-approach for big remote-sensing image classification," Computers and Electronics in Agriculture, vol. 182, pp. 106014, 2021.

[8] S. Meivel, K.I. Devi, S.U. Maheswari, J.V. Menaka, "Real-time data analysis of face mask detection and social distance measurement using Matlab," Materials Today: Proceedings, 2021.

[9] S. V. Militante and N. V. Dionisio, "Real-Time Facemask Recognition with Alarm System using Deep Learning," 2020 11th IEEE Control and System Graduate Research Colloquium (ICSGRC), pp. 106-110, 2020.

[10] K. Suresh, M.B. Palangappa, S. Bhuvan, "Face Mask Detection by using Optimistic Convolutional Neural Network," 6th International Conference on Inventive Computation Technologies (ICICT), pp. 1084-1089, 2021.

[11] P. Nagrath, R. Jain, A. Madan, R. Arora, P. Kataria, J. Hemanth, "SSDMNV2: A real-time DNN-based face mask detection system using single shot multibox detector and MobileNetV2," Sustainable cities and society, vol. 66, pp. 102692, 2021.

[12] G.J. Chowdary, N.S. Punn, S.K. Sonbhadra, S. Agarwal, "Face mask detection using transfer learning of inceptionv3," International Conference on Big Data Analytics, pp. 81-90, 2020.