Available online at http://www.advancedscientificjournal.com
http://www.krishmapublication.com
IJMASRI, Vol. 1, issue 9, pp. 161-168, Nov. -2021
https://doi.org/10.53633/ijmasri.2021.1.9.01

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY ADVANCED SCIENTIFIC RESEARCH AND INNOVATION (IJMASRI)

ISSN: 2582-9130 IBI IMPACT FACTOR 1.5 DOI: 10.53633/IJMASRI

RESEARCH ARTICLE

DETECTION OF FACE MASK USING DEEP LEARNING

[1]Manak, [1]Manaswi, [1]Pankaj Kumar and [1]Jyoti Kaushik

[1]*CSE Department, MAIT, Rohini, Delhi*

Abstract

As we know, the COVID-19 pandemic additionally referred to as the coronavirus pandemic, is an ongoing pandemic of coronavirus sickness since 2019. This infectious disease was first detected in Wuhan, China in late 2019. Symptoms of COVID-19 are highly variable, ranging from none to severe health problem. The virus spreads chiefly through the air when individuals are close to one another. It transmits from an infected individual through the droplets as they breathe, cough, sneeze or speak and these droplets then enters another individual via their mouth, nose, or eyes. It might also spread via contaminated surfaces. Individuals stay infectious for up to 2 weeks and can spread the virus albeit they do not have symptoms. As of 1st November 2021, more than 200 million cases are confirmed, with more than 500 million deaths due to COVID-19. The pandemic has been the reason for global social and economic disruption, including the largest global recession since the Great Depression. The recommended preventive measures include social distancing, carrying a mask publicly, ventilation and air-filtering, hand washing, covering one's mouth when sneezing or coughing and self-isolation for individuals exposed. In present endeavour therefore, the author has attempted to make one thing associated with it, that's deciding whether a dividual is carrying a mask or not. The complete investigations area unit distributed in various chapters that embody the current thesis. The performance of our model will be evaluated in precision, accuracy, recall, specificity, and sensitivity that demonstrate the practical application of this model. The system performs with an accuracy of 99.88%, precision of 99.49%, sensitivity of 99.77%, and specificity of 99.6. Thus, this model tracks if people are using masks or not in real-time using a device camera. This model can be used with the current camera infrastructure to enable this tool which can be used in various public places such as markets or railway stations or offices etc.

Keywords: Covid-19, Face Mask Detection, Convolutional Neural Network, MobileNetV2, Precaution.

Introduction

In the late 2019, an extremely infectious disease named as coronavirus disease (COVID-19) has come under reporting initially in Wuhan, China. It has become an extreme health issue all over the world. This pandemic has devastating effects on societies and economies worldwide inflicting a worldwide health crisis. Its associate degree rising metabolic process communicable disease caused by Severe Acute metabolic process Syndrome Coronavirus a pair of (SARS-CoV-2). everywhere the globe, COVID-19 has been an extreme challenge. Several shutdowns in numerous industries because of this deadly gloabal pandemic. Additionally, several sectors like maintena nce works and infrastructure construction and development haven't been suspended because of their important impact on people's routine life.

Earlier, some people used masks to shield themselves from pollution, whereas others used face masks to cover their faces and their emotions from others. Using masks as a protective agent against coronavirus is a necessary, consistent with the World Health Organization (WHO) guidelines. Indeed, using a mask is an efficient methodology of blocking majority of air-borne respiratory infections. Also, the World Health Organization recommends physical distancing to stop this virus from spreading. Everywhere on the globe, nations are fighting against this sort of infectious disease. Several organizations have implemented strict mask rules for the protection against COVID-19. Checking manually if people entering in an area or present in the area are using masks is cumbersome and probably conflicting.

To summarise AI, it is the flexibleness of an informatics system or computer-controlled mechanism to perform tasks typically associated with intelligent beings. The term is sometimes applied to the project of developing systems endowed with the intellectual processes characteristic of humans, just like the flexibility to reason, discover which implies, generalize, or learn from experience. Since the event of the data process system among the 40s, it has been incontestable that computers area unit usually programmed tocarry out sophisticated tas ks as an example, driving cars, flying aircrafts (autopilot), or even detecting faces in a mobile phone camera. Still, despite continued advances in computer method speed and memory capability, there are areas where no programs which can match human flexibility over wider domains or in tasks requiring plentiful everyday data. On the other hand, some programs have earned the performance levels of human specialists and professionals in liberal arts certain specific tasks, so as that AI throughout this restricted sense is found in various applications such as facial recognition, reverse image search engines, and voice or handwriting recognition. And recognising face mask users using artificial intelligence s just that. It saves human effort and time. And we are aiming just for that.

Down below is the comparison of the accuracy rates of some recent existing work on the same field by some very experienced and renowned authors.

Table 1: Recent work in the same field and their accuracy rates

Accuracy	Reference Number	Date
98.6%	I	29/07/2020
95.77%	II	05/02/2021
99.52%	III	30/03/2021
97.1%	IV	05/2021
98.2%	V	24/06/2021

IMAGE PROCESSING

A. Introduction

Image processing is a method for performing various operations on a picture, to create an enhanced image and/or to extract some useful information from input information. It is a kind of signal during which input is a picture and output could be a picture or characteristics/features related to that picture.

B. ***Image***

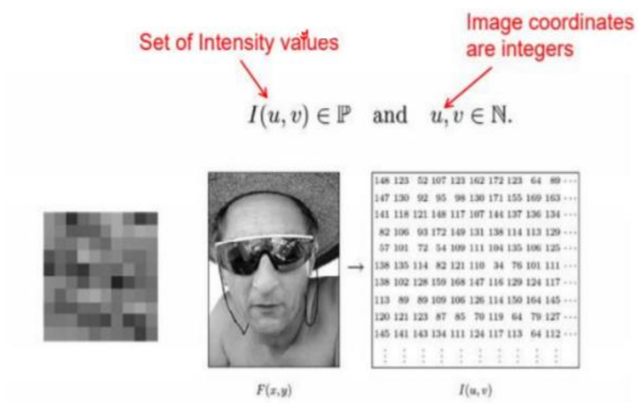

Fig 1. 2-d matrix of intensity values (greyscale)

(a) Digital Image

A digital image is an image made up of tiny picture elements, called pixels. Digital images are finite, discrete quantities of numeric representation for its intensity or grey level that is an output from its two-dimensional functions fed as input by its spatial coordinates denoted with x, y on the x-axis and y-axis, respectively

(b) Image format

- Values per point/pixel (B&W or Grayscale)
- Values per pixel (RGB colours)
- Values per pixel (RGB colours and opacity)

Fig 2. Image format

(c) Image processing

- The algorithm which will change the input image according to the pre-set conditions for creating a new image as an output

Input is image, output is image

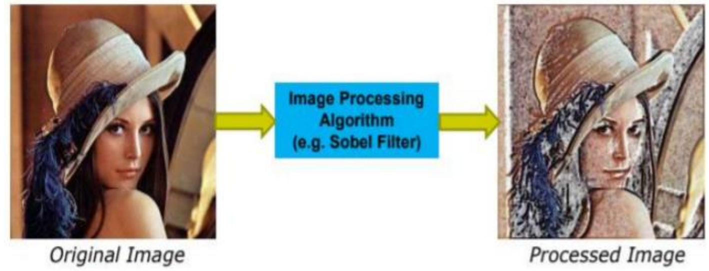

Fig 3. Processed image

(a) Phases of image processing

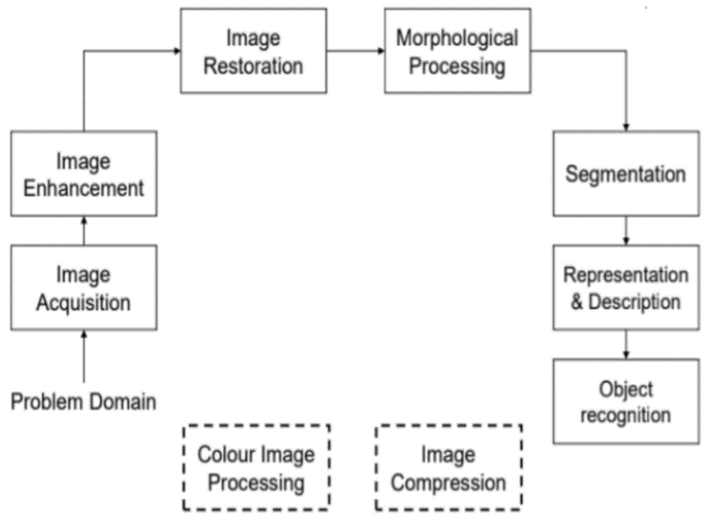

Fig 4. Phases of image processing

FACE MASK DETECTION FRAMEWORK

Facial Recognition requires no physical interaction on behalf of the user. It is accurate and allows for high enrolment and verification rates and it can use the existing hardware infrastructure, existing camaras and image capture devices will work with no problem. In this tough covid-19 times we decided to build something related to it, that is deciding whether a person is wearing a mask or not. For this model, we collected a database of images which consist of about 4095 images with 2165 images containing people wearing face mask and 1930 containing people without face mask. By developing the model and using this image we can predict whether a person is wearing a mask or not using a webcam.

The fig. 5 shows the full proposed twin block framework. The first block consists of the training and testing of the models, while the other block contains the entire framework for testing. For the primary block, our labelled dataset was divided into two parts, with masks and without masks. For every epoch, each model is trained on the training dataset. The training results, in addition because the training accuracy and the training loss, are presented within the kind of curves in figures of "accuracy in terms of epoch" and "loss in terms of epoch," respectively. Each model is validated on the validation dataset, after training. Then, the 2 results are compared with the loss function. a mistake function value tending toward zero means a well-trained model.

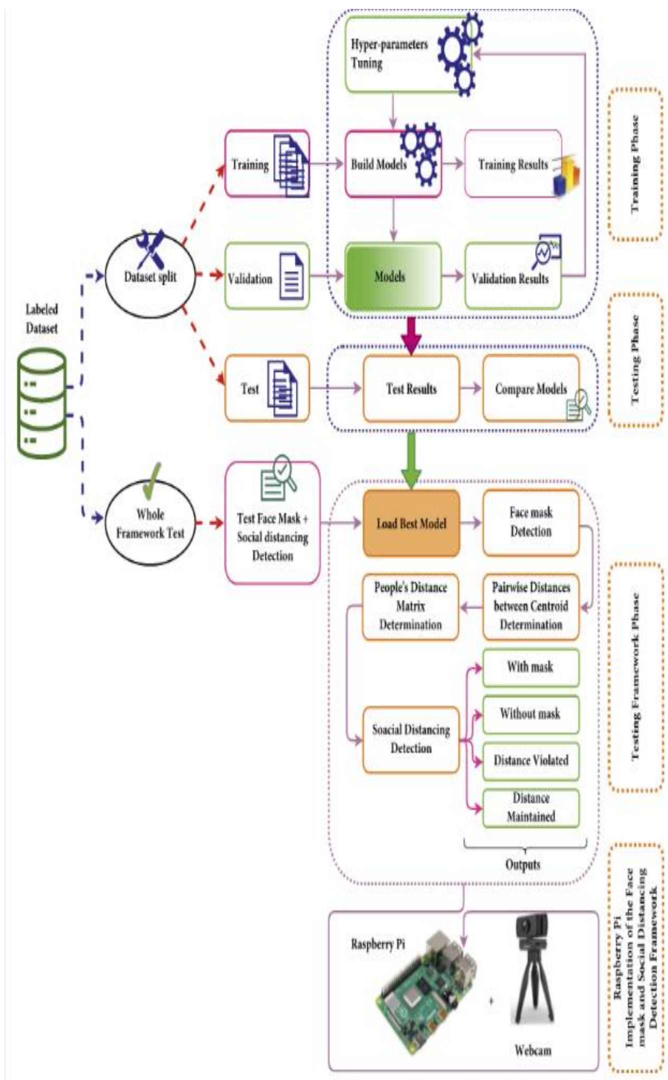

Fig. 5. Face mask detection framework

(a) Overview

- Dataset will be augmented to include a greater number of images for our training. In this step of data augmentation, we will rotate and flip each of the images in our dataset.
- Visualization of total images in the data set will be done.
- We divide our data into the training set which will contain the images on which the CNN model will be trained, and the test set with the images on which our model will be tested. And then we will build our Sequential CNN model with various layers such as Conv2D, MaxPooling2D, Flatten, Dropout and Dense.
- We will fit images in the training set using Kera's library.
- We will be using the Haar Feature-based Cascade Classifiers for detecting the features of the face. This cascade classifier is designed by OpenCV to detect the frontal face by training thousands of images for a better accuracy and good results.
- In the last step, we will use the OpenCV library to run an infinite loop to use our web camera in which we detect the face using the Cascade Classifier. The resulting model will predict the possibility of each of the two classes ([without mask, and with mask]). If the probability of a masked face is higher, the model will highlight it with green box and if probability of unmasked face is detected higher, a red box will be used to highlight the results in the live camera.

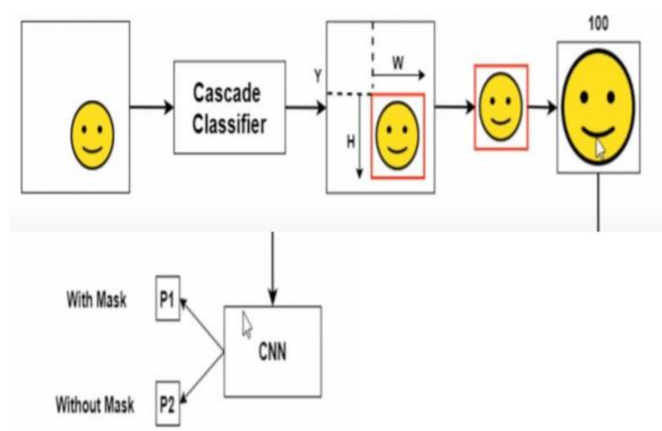

Fig 7. Dataset images example

DATA PREPROCESSING

It is a process in which an image is converted into a grey scale image and then into a specific dimension according to developer 's need. ▪ A grayscale (or grey level) image is simply one in which the only colours are shades of grey. The reason for using the black and white images instead of any other sort of colour image is that less information needs to be provided, hence the processing will be faster for every image and for each pixel.

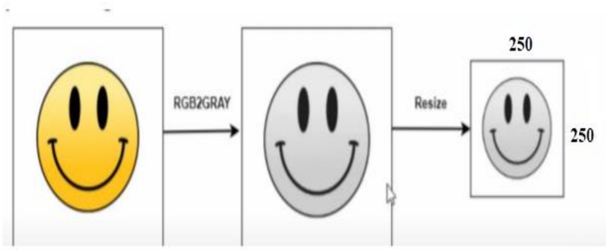

Fig 8. Data pre-processing process

ALGORITHM USED AND CNN

• Convolutional neural network (CNN) is a type of deep learning model for processing data that has a grid pattern, such as images, which is then further used in images recognition, images classifications, objects detections, recognition faces etc.

• During this proposed method, the mask detection model is made using the TensorFlow and Keras libraries. this enables the user to form the new layers for the model step by step. The different layers used for our CNN model is described below.

• The primary layer is that the Conv2D layer with 100 filters and therefore the filter size or the kernel size of 3X3. ReLu function stands for Rectified long measure which may output the input directly if is positive, otherwise, it'll output zero.

• Subsequent layer is again a Conv2D layer with another 100 filters of an equivalent filter size 3X3.

• And in the next step, we use the Flatten () layer to flatten all the layers into one 1D layer.

• After the Flatten layer, the Dropout (0.5) layer is used to stop the model from overfitting.

• Finally, towards the top, we use the Dense layer with 50 units and therefore the activation function as 'ReLu'

• The last layer of our model is going to be another Dense Layer, with only two units and therefore the activation function used are going to be the 'SoftMax' function.

• The SoftMax function outputs a vector which can represent the probability distributions of every of the input units.

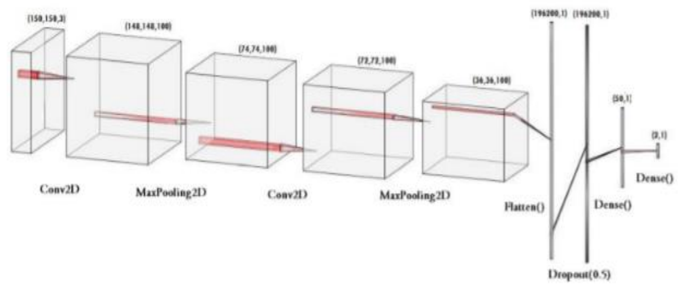

Fig. 9. CNN working diagram

(a) Training the CNN model

Training the convolution neural network is one of the most important step in the construction of the whole model where we fit our images in the training set and the test set to our Sequential model, we built using keras library. The model has been trained for 75 epochs (iterations). However, we can train for a greater number of epochs to attain higher accuracy lest there occurs over-fitting. The batch size was 25, ensuring we had a greater number of batches to test and initial learning rate was set at 1e-4 to ensure further greater accuracy.

(b) Labelling the information

Once the model is completed, we will obtain two probable outcomes. ['0' as 'without mask' and '1' as 'with mask']. The masked and unmasked individuals will be identified by a rectangle over their face of different colours using the RGB values. 'RED' will be for 'without mask' and 'GREEN' will be for 'with mask.

(c) Importing the face detection program

First, we need to implement face detection. In this, we are using the Haar Feature-based Cascade Classifiers for detecting the features of the face. This cascade classifier is designed by OpenCV to detect the frontal face by training thousands of images.

(d) Detecting the faces without the masks

We use the OpenCV library to run an infinite loop to use our web camera in which we detect the face using the Cascade Classifier. The piece of code that specifies the camera which is to be used for the application will be written as *webcam =cv2.VideoCapture(0*) The model will predict the possibility of each of the two classes ([without mask, with mask]). Based on which probability is higher, the model will compare the outcomes and the it will display an appropriate label around the faces.

Result

This model aims to demonstrate the future uses of the system, so tests are administered using the recall metrics, Precision, and therefore the corresponding macro average and weighted average the target of using these metrics is to gauge the model under various conditions. Recall and precision indicate the power of the model to properly detect true positives. Recall also considers the false negatives detected, and therefore the precision of the false positives detected by the model. False positives, during this case of face detection with masks, occur when an object is labelled as a face. for instance, the system frames a plant as a face with a mask, because it is fake that a plant may be a positive face. the explanations this will occur in our system are numerous, which is why diligence is required to gather an outsized database in order that the model being trained can better distinguish the specified classes. False negatives occur when a face isn't detected within the first stage, because the face has covered areas that make classification difficult; during this proposal, this first classifier may be a pre-developed tool. On the opposite hand, the F1-score provides a worldwide measure of the system's performance, it's a mixture of precision and recall (in one value), with 0 being low performance and 1 being the simplest possible performance (all cases detected correctly). By considering the macro average metric, user can get a general idea of the typical of all the experiments, while the weighted average establishes a mean measure of all the experiments but considering the number of observations of every class. Thus, within the event that a category features a higher score, the ultimate weighted average score won't be suffering from it but will provides a value of importance to every score counting on the quantity of observation. When considering these metrics, what's sought is to verify the robustness of the tactic by classifying both classes.

During a successful execution of the program, an accuracy of 99.88% was obtained using Deep Supervised Learning CNN. The model calculated the accuracy percentage and plotted the data in the accuracy vs epoch plot, and results were tested and found accurate as well. During the successful run, expected class matched with the correct class, showing that the model worked. Individuals using a mask were set to class 0 and those who weren't were set to class 1. Fig. 9 shows a picture of a person who is not wearing a mask, therefore the model correctly predicted that the input image matches class 1. Fig 10 is an example of a person who is using a mask, thus, the model predicted that the input image matches class 0. And finally, fig. 11 shows an example of person wearing mask partially, that is, on their chin. The model correctly showed the results as given. All the results shown are predicted with a high accuracy of 100%. The lighting in the surroundings need to be adequate for better results. In dim lights, the model will work, though without a great accuracy.

We have used one-fifth of the total dataset for testing, and the remaining dataset was used for the training of face classification; Fig. 13 shows the graph plotting the training results. The accuracy of this model when is run at 75 epochs, comes out to be 99.88%, which is one of the highest among recent work done in the existing models. Our model has the following numbers as the results, it has an accuracy: 99.88%, specificity: 99.6%, sensitivity: 99.77%, and precision: 99.49%.

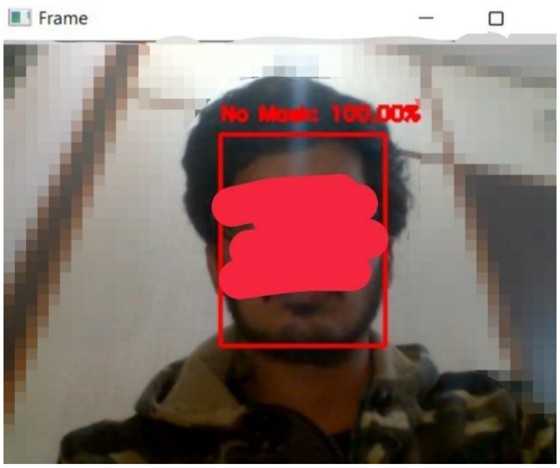

Fig. 10. No mask

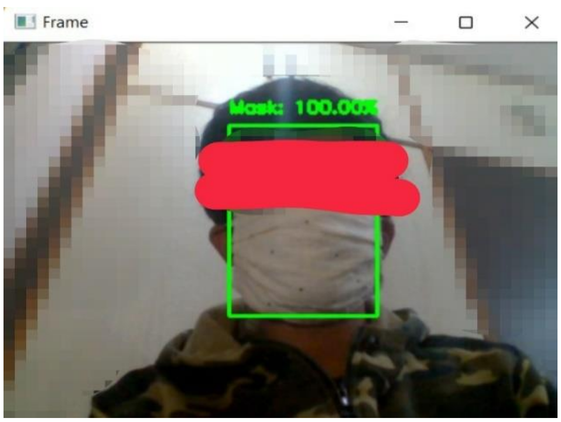

Fig. 11. Mask wearing

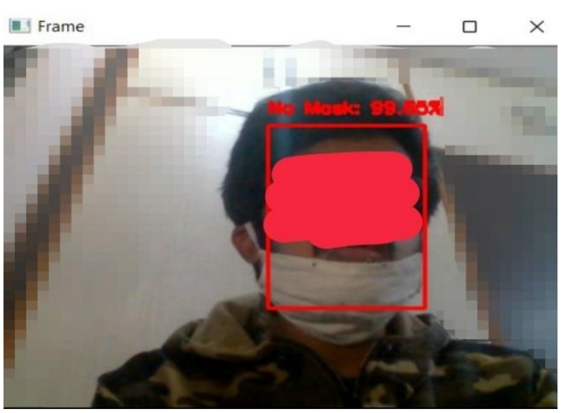

Fig. 12. Partial mask

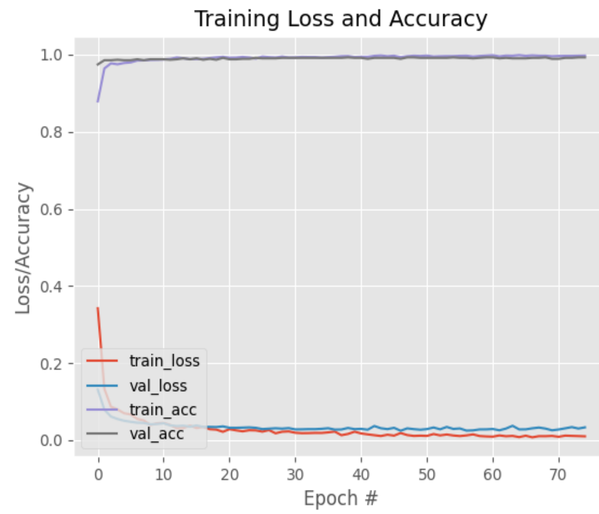

Fig. 13. Analysis data plot

Conclusion

There is an urgency in controlling COVID-19 all over the globe and because of that, real-time mask and social distancing detection applications valuable and important. The authors have tried to demonstrate a working model of real time face detection application using Convolutional Neural Network. After studying various other published work about the current outbreak, the authors reviewed the existing work and technology in the Face Mask Detection. Then the model was constructed and trained using a dataset built from various online sources like Google Images and Kaggle. The results were analysed on the factors like sensitivity, precision, accuracy, and specificity and the model exhibiting very accurate results as compared to the previous work done.

For the future upgrades, this technique could be used on smart sensors where Cloud services could be used collect multimedia data, or maybe a drone system, that can gather images through its camera to detected objects and upload the data to the Cloud for analyses. There could also be work done in developing mobile phone-based application which would be very handy and useful in filtering huge crowds based on whether people are wearing masks or not.

Reference

1. M. Inamdar and Mehendale. N. 2020. “Real-Time Face Mask Identification Using Facemasknet Deep Learning Network”. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3663305
2. Das. M., Wasif Ansari and Basak. R. 2020. "Covid-19 Face Mask Detection Using TensorFlow, Keras and OpenCV," 2020 IEEE 17th India Council International Conference (INDICON), 2020, pp. 1-5, doi: 10.1109/INDICON49873.2020.9342585.
3. Talahua, J.S., Buele, J, Calvopina, P, Varela-Aldás, J. 2021, Facial Recognition System for People with and without Face Mask in Times of the COVID-19 Pandemic. Sustainability, 13, 6900. https://doi.org/10.3390/su13126900
4. Boulos., Mira. M. 2021. "Facial Recognition and Face Mask Detection Using Machine Learning Techniques". Theses, Dissertations and Culminating Projects. 728. https://digitalcommons.montclair.edu/etd/728
5. Sethi. S., Kathuria. M and Kaushik ,T. 2021. Face mask detection using deep learning: An approach to reduce risk of Coronavirus spread. J Biomed Inform. 2021;120:103848. doi:10.1016/j.jbi.2021.103848
6. Egger, M. Ley, M and Hanke, S. 2019. Emotion Recognition from Physiological Signal Analysis: A Review. In Proceedings of the Electronic Notes in Theoretical Computer Science, Yangzhou, China, 14–17. Elsevier B.V.: Larnaca, Cyprus,
7. Adjabi, I., Ouahabi, A. Benzaoui, A and Taleb-Ahmed, A. 2020. Past, present, and future of face recognition: A review. Electronics, 9, 1188. [Google Scholar]
8. Voulodimos, A., Doulamis, N. Doulamis, A and Protopapadakis, E. 2018. Deep Learning for Computer Vision: A Brief Review. Comput. Intell. Neurosci. [Google Scholar] [CrossRef]

*****