International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 08 Issue: 08 | Aug 2021 www.irjet.net p-ISSN: 2395-0072

Deep Learning Technique to Detect Face Mask in the COVID-19 Pandemic Period

Monica M[1], K R Sumana[2]

[1]*PG Student, The National Institute of Engineering, Mysuru, Karnataka, India*
[2]*Assistant Professor, The National Institute of Engineering, Mysuru, Karnataka, India*

Abstract - *The COVID-19 epidemic has had a fast impact on our daily lives, affecting global trade and mobility. Wearing a protective face mask is critical for lowering the chance of infection from a noxious individual during the "pre-symptomatic" stage and preventing viral transmission. Face mask detection has therefore become a critical job in today's global society. Face masking has become the new normal. Many public service providers will need consumers to wear masks appropriately in the near future in order to use their services. As a result, detecting face masks has become a critical duty in aiding worldwide civilization. This project demonstrates a simplified method for accomplishing this goal by combining a Deep Learning methodology known as Convolutional Neural Networks with some fundamental Machine Learning tools such as TensorFlow, Keras, and OpenCV. The suggested technique accurately recognizes the face in the picture and then determines whether or not it is covered by a mask. A cascade classifier and a pre-trained CNN with two 2D convolution layers linked to layers of dense neurons are used in the suggested approach. Here, we look at how to use the Sequential CNN model to find the best parameter values for accurately detecting the existence of masks without overfitting. The model is trained and validated before being deployed, and as a result, the approach achieves an accuracy of up to 95.77 %.*

Key Words: **Convolutional Neural Network (CNN), Open Source Computer Vision (OpenCV) library.**

1.INTRODUCTION

As per the World Health Organization's (WHO) official Situation Report, the COVID illness 2019 (COVID-19) has contaminated over 1,000,000 people. countless people all through the world, bringing about an enormous number of passing Users of COVID19 have announced an expansive scope of incidental effects, going from smooth skin to undeniable ailment. Respiratory issues, for example, shortness of breath or inconvenience breathing, are one of them. Since patients with lung sickness are at a higher danger, COVID-19 illness can make critical challenges. Individuals with respiratory troubles might recognize any individuals (in closeness to them) who has respiratory issues. Irresistible swabs, to be exact. Contact transmission can happen because of a filthy individual's environmental elements, just as different variables, since transporter particles may create on close by surfaces. Wearing a face cover is vital for controlling certain respiratory viral contaminations, like COVID19. The overall population ought to know about whether COVID19 source control ought to be covered. Veiling's potential accentuation regions incorporate diminishing the risk of a pathogenic individual during the "pre-cleanliness" stage and censuring irregular individuals who utilize facial covers to limit the transmission of irresistible diseases. For medical services accomplices, WHO is focusing on clinical covers and respirators. Face distinguishing proof has now become a basic obligation in the present culture across the world. Facial acknowledgment involves recognizing the locale of the face and afterward deciding if it is covered by a mask.

2. LITERATURE REVIEW

Face mask detection includes in distinguishing the area of the face and afterward deciding if it has a cover on it or not. The issue is generally related to general article identification to identify the classes of items. Face recognizable proof completely manages recognizing a particular gathering of elements for example Face. It tends to be refined utilizing the profound learning method called CNN. It has various applications, like self-ruling driving, schooling, observation, etc [1]. CNN is a kind of significant learning model for dealing with data that has an organization configuration, similar to pictures, which is stirred by the relationship of animal visual cortex and planned to thusly and adaptively learn spatial leadership hierarchies of features, from low-to unquestionable level models. CNN is a mathematical foster that is usually made out of three sorts of layers (or building blocks): convolution, pooling, and totally related layers. The underlying two, convolution and pooling layers, perform feature extraction, however the third, a totally related layer, maps the isolated features into indisputable yield, similar to portrayal [2]. Keras is a significant learning API written in Python, running on top of the AI stage TensorFlow. It was made with a consideration on engaging fast experimentation. Having the alternative to go from thought to result anyway fast

as possible is by all accounts essential to doing extraordinary investigation. Keras has a wide assurance of predefined layer types, and besides maintains forming your own layers. Focus layers consolidate Dense (spot thing notwithstanding inclination), Activation (move limit or neuron shape), Dropout (discretionarily set a modest quantity of data units to 0 at every planning update to avoid overfitting), Lambda (wrap a self-emphatic enunciation as a Layer object), and a couple of others [3]. OpenCV is a Python open-source library, which is used for PC vision in Artificial knowledge, Machine Learning, face affirmation, In OpenCV, the CV is a withdrawal sort of a PC vision, which is portrayed as a field of study that helps PCs with understanding the substance of the high level pictures like photographs and accounts. The inspiration driving PC vision is to grasp the substance of the photos. It removes the depiction from the photographs, which may be a thing, a book portrayal, and three-estimation model, and so on [4].

3. PROPOSED SYSTEM

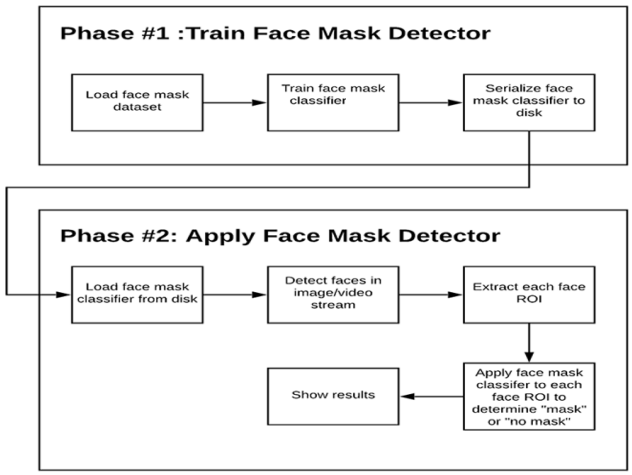

Fig-1: Work Flow of the Proposed System

Our system's work flow is depicted in Fig-1 above. The face mask detector is divided into two phases: training and deployment. In the first stage, we train the face mask classifier model with the dataset we have collected and then save it to a disc.

In the second step, we stack the put away face mask indication from the disc as a face identifier model and use it to differentiate faces from images and video transfers to extract the ROI, i.e. face, and apply a face cover classifier to it to determine if persons are wearing face covers. Finally, it outputs the absolute value as "Mask" or "no Mask.".

4. IMPLEMENTATION

In the proposed technique, a cascade classifier and a pre-trained CNN with two 2D convolution layers connected to layers of dense neurons are employed..

4.1 Data preprocessing

A) Data Visualization

In data preparation, data is converted into a more usable, desired, and understandable format. It can be in any format, such as tables, photographs, videos, graphs, and so forth. These structured data are part of an information model or composition that keeps track of the connections between different things. To deal with picture and video data, the recommended approach employs Numpy and OpenCV. The total number of pictures in our collection has been split into two categories. The 'with mask' class has 1916 photos, whereas the 'without mask' class has 1931 images, as can be seen.

B) Conversion of RGB image to Gray-Scale image

Since it rationalises the approach and minimises computational needs, grayscale is utilised for extracting descriptors instead of working on colour images instantaneously. The method cv2.cvtColor is used to alter the colour space (input picture, flag). Using the flag, you may specify the type of conversion. This scenario uses the flag cv2.COLOR BGR2GRAY for grey conversion. Deep CNNs require a fixed-size input image. A consequence of this is that each image in the collection must be the same size. Using cv2.resize, the grayscale image is resized to 100 by 100 pixels ().

C)Image Reshaping

Most CNNs only accept pictures that have been fine-tuned. Many issues arise as a result, both during the data collecting process and during the actual implementation of the model. Converting photos before adding them to the network might assist to overcome this problem..

4.2 Training the Model

Many of CNN's computer vision programs are now considered to be among the most advanced in the industry. In the current technique, CNN is utilised sequentially. First Convolution is followed by layers of Rectified Linear Unit (ReLU) and MaxPooling.. When it comes to the Convolution layer, 200 filters are employed. There is a 3x3 section of the image that corresponds to the 2D convolution window. Data regarding the input shape should be given to the

model's primary layer since the model must be informed of the present condition of the information expected. Natural forms can be derived from the following layers. For example, the input shape in this case is data.shape[1:], which returns the information cluster pieces from file 1. This means that the spatial measurements might diminish and the information volume is cushioned by default. Relu is set as the class's initiation boundary. If you're looking for a capacity that's nearly straight and has all the resources of direct models, you'll want to check out the slope plunge methods. This is better contrasted with other enactment capacities when thinking about deep learning exhibition and speculation Max Pooling is used to minimize the spatial components of the yield volume. We set the pool size (number of lines or sections) to 3 x 3 and the state of the yield to (input shape - pool size + 1)/steps, where steps is set to the default value of 1. (1,1). A 100-channel Convolution layer with a 3 x 3 Kernel size is shown in the following example. The layers ReLu and Max-Pooling are then added on top of it.A Flatten layer transforms the highlighted framework into a vector that can be handled by a fully connected neural network classifier in order to integrate the data in CNN. In an attempt to reduce overfitting, a Dropout layer has been used with a haphazard attempt at setting contributions to 0.
This is followed by a Dense Layer of 64 Neurons, as well as ReLu actuation. Last layer (Dense), which has two yields for two classes, uses Softmax actuation work.. The below Fig-2 shows the CNN Architecture.

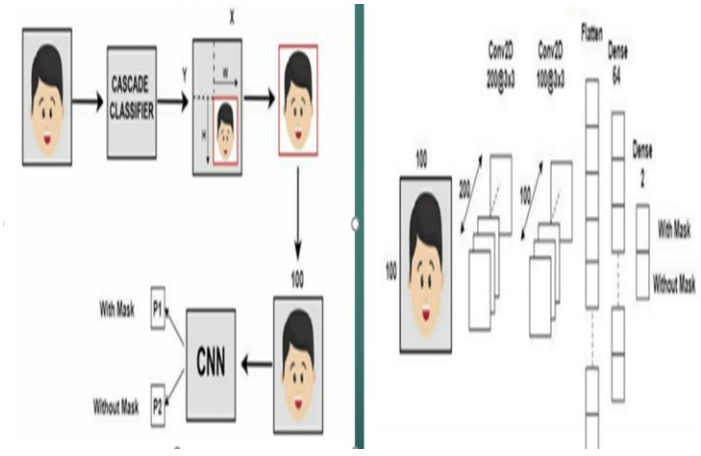

Fig-2: CNN Architecture and Model

4.3 Implementing the Model Using Images

A face detector model is loaded from the disc and used to create a "blob" of data. As the blob travels through the network, we are able to recognise the faces. Mask classifier is applied to the identified face, and it categorizes the output as either "Mask" or "No mask." The below Fig-3 shows the detection of face mask or not from the images.

Fig-3: Detection of Face Mask Through Image.

4.4 Implementing the Model Using Video Stream

When we take the frame dimensions and feed them through the network, we are able to acquire the face detections. This will then be followed by the drawing of a bounding box with a given probability around the identified ROI (face). Once the face mask model has been loaded, the video stream will be initiated to capture the video and then the video will be captured. In the next step, we'll loop through the video frames and use the face mask classifier to recognise and categorize the faces in the video stream. The below Fig-4 shows the face mask detection in video stream.

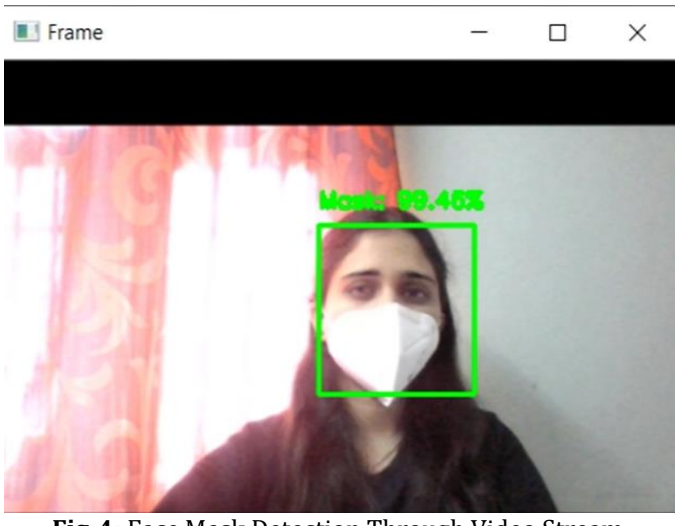

Fig-4: Face Mask Detection Through Video Stream

5. RESULT AND DISCUSSION

The model is trained and validated prior to implementation, and as a consequence, the technique achieves an accuracy of up to 95.77 percent. In order to achieve this degree of accuracy, Max-Pooling is essential. In addition, it reduces the number of parameters that must be learned by the model. The dataset contains a total of 3847 samples, of which 80% are used in the training phase and 20% in the testing stage. Sample-based discretization reduces the input representation's dimensionality. The optimal number of neurons is 64, which is an adequate quantity. Performance may deteriorate when a large number of neurons and filters are employed. The adjusted filter settings and pool size aid in filtering away the primary section (face) of the picture, allowing the presence of a mask to be detected accurately without over-fitting. Because extra training leads to overfitting on the training data, the suggested architecture is trained for 20 epochs. The training loss-accuracy curve is depicted in Fig. 3 below.

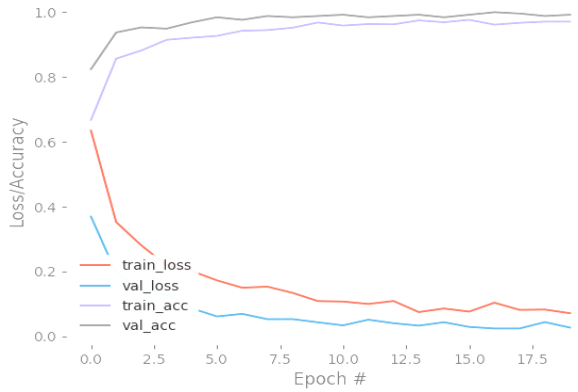

Fig-5: Training Loss and Accuracy Curve on the dataset

Accuracy/loss curves for the COVID-19 face mask detector during training Bends have a high degree of accuracy and no sign of data overfitting. Using this, we'll be able to do face veil identification. As you can see, our test set achieves 99 percent accuracy. The figure demonstrates that there are several overfitting indications, with the approval blunder being less than the preparation blunder. Based on these findings, we are sure that our model will accurately sum up images obtained outside of our preparation and testing set.

6. CONCLUSION

We briefly addressed the work's inspiration right from the start. The model's learning and execution endeavor was then shown at that time. Using basic ML tools and the Deep Learning technique CNN, we were able to attain a reasonably high level of precision. It has the potential to be utilised for a wide range of applications. Given the circumstances surrounding Covid-19, wearing a mask may become essential sooner rather than later. Many public speciality organizations may need consumers to effectively wear masks in order to benefit from their services, and our endeavor may be able to help in this respect..

The established model will continue to make a substantial contribution to the entire medical care system in the future. It may be used to check if someone is wearing the mask appropriately in the future. The model might also be improved to detect whether or not the veil is infection-prone, such as whether or not the cover is N95. The framework can be programmed to recognise people's faces in a live video and identify whether or not they are wearing a face mask. The framework can evaluate if the concerned individual is authorised or not to enter open places such as a market or a medical institution based on the provided data. This company might be utilised in a clinic. a market, or a transportation station, where checking must be done.

REFERENCES

[1] Arjya Das, Mohammad Wasif Ansari, Rohini Basak "Covid-19 Face Mask Detection" 2020 IEEE 17th India Council International Conference (INDICON) | 9781728169163/20/$31.00©2020IEEE|DOI:10.1109/INDICON49873.2020.

[2] "Convolutional neural networks: an overview" Rikiya Yamashita1,2 & Mizuho Nishio1,3 & Richard Kinh Gian Do2 & Kaori Togashi1 Received: 3 March 2018 / Revised: 24 April 2018 /Accepted: 28 May 2018 /Published online: 22 June 2018 # The Author(s) 2018.

[3] "Keras Documentation: About Keras" K. Team, "Keras documentation: About Keras", Keras.io, 2020.

[4] "OpenCV" Opencv.org, 2020. [Online]. Available: https://opencv.org/. 2020