e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
Volume:03/Issue:05/May-2021 Impact Factor- 5.354 www.irjmets.com

IDENTIFYING COVID-19 FACE MASK VIOLATORS USING CONVOLUTIONAL NEURAL NETWORKS

Harinisankari.K*1, Hemalatha.A.R*2, Ishwariya.B*3, Yuvarekha.M*4

*1,2,3,4Department Of Computer Science And Engineering,

Coimbatore Institute Of Technology, Coimbatore- 641014

ABSTRACT

The Corona virus pandemic is spreading everywhere in the world. Due to the outbreak of Covid-19, it is required that everybody should wear a mask to save themselves and furthermore try not to spread the infection to other people. A few groups adhere to this law however others don't follow it. In certain spots safety officers are delegated to confine the section of individuals without masks. However, in some places where there is no safety officer to guarantee that all are wearing masks, the model proposed in this paper will fill the need. At the entry door, a CCTV camera will monitor everyone. Whenever a person without a mask or wearing a mask improperly is detected, his/her entry is restricted. The model is built using Convolution Neural Network (CNN). A dataset that contains images of people with masks, without masks and improperly wearing masks is collected from various sources to build this model. The built model has achieved an accuracy of 93% during testing. We hope that this model will serve purposeful to lessen the impact of covid-19.

I. INTRODUCTION

Coronavirus or COVID19 (known as Covid) is the most recent scourge infection that has hit human wellbeing. In 2020, the fast spreading of COVID-19 has constrained the World Health Organization (WHO) to announce COVID-19 as a worldwide pandemic. In excess of 5,000,000 cases were contaminated by COVID-19 in under a half year across 188 nations. The World Health Organization (WHO) has proposed various measures to protect every individual from the corona virus like to maintain social distancing, to avoid touching nose, mouth and eyes and to regularly sanitize the hands. Out of that the simplest and effective preventive measure is to wear a mask. The spread of coronavirus can be limited to some extent if and only if people adhere to all the laws proposed by the WHO. But sadly, not all the people are following it. This leads to the spreading of viruses at an exponential rate. The face mask detection system proposed in this system will detect the person at the entry door using a CCTV camera and categorize them into one of the three categories- masked, unmasked, improperly masked. The person will not be allowed to enter if he did not wear a mask or wore a mask improperly. The proposed model is built using a deep learning architecture Convolution Neural Network (CNN). Using CNN, features from the image are extracted and are learned by multiple hidden layers. The model is trained using data collected from several sources. The proposed model achieved training accuracy of 93%. This model can be integrated in real time public places where there are no security officers to ensure that all are masked. The rest of the paper is divided into 4 sections. Section 2 will tell about the past works that were done. Section 3 will tell about the methodology proposed. Section 4 will describe the analysis of the result. Section 5 will be the conclusion.

II. RELATED WORKS

Undoubtedly Convolution Neural Networks is one of the best deep learning algorithms to solve the problems on image processing and classifications [1]. Several CNN architectures are introduced in past years like AlexNet [2], VGGNet [3], GoogLeNet [4], and ResNet [5]. These are some standard CNN architectures. Intensive researches are being undertaken to tune the accuracy of CNN architectures [6].

CNN is a special kind of neural networks. CNN is based on neuroscience. CNN has three main layers like input, hidden and output layer. Each layer is made up of a certain number of nodes. Each node will be having some weights and it will return the activation map. For example if we feed image to CNN, its pixel values will be taken as weight by the neurons. Actually in depth CNN performs two operations 1.convolution(to extract the edges from input image) and 2.pooling(to making pooling very powerful i.e reducing image size and make the image highly intensified).Usually CNN will initially extract the edges from the input images and it will feed it to the

IRJMETS
e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
Volume:03/Issue:05/May-2021 Impact Factor- 5.354 www.irjmets.com

next layer. Then that layer will try to extract corners and colour groups and feed it to next layer. This continues until prediction is being made.Fig.1 depicts this convolution process.

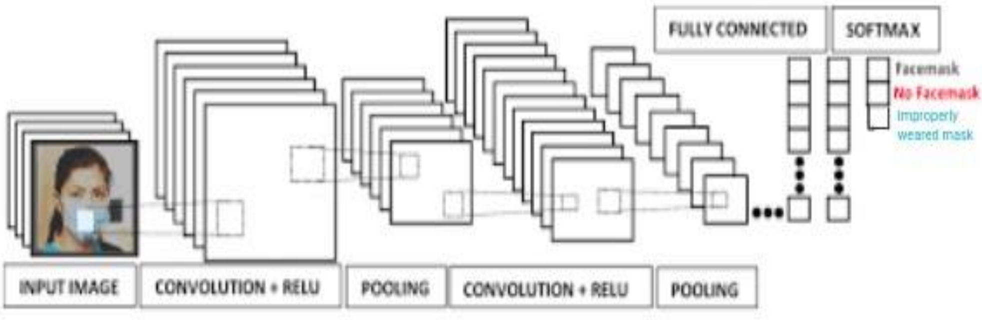


Fig.1

III. PROPOSED METHODOLOGY

The system focuses on building and training a CNN model that will be able to predict the faces and then classify it into one of the following three categories- masked, unmasked or improperly masked. The Fig.2 depicts the overall block diagram of the proposed model. It is split into two phases. The first phase focuses on building and training the model. The second phase focuses on applying the model to test the data.

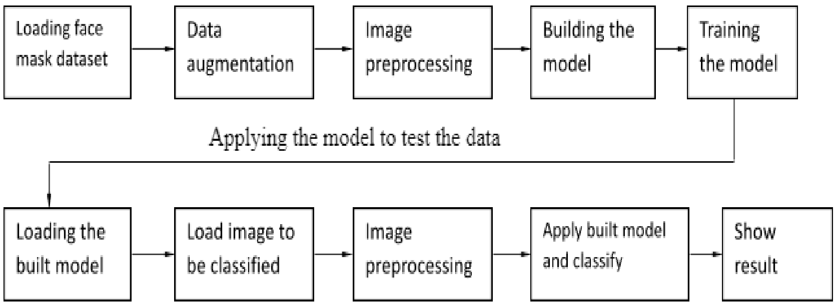
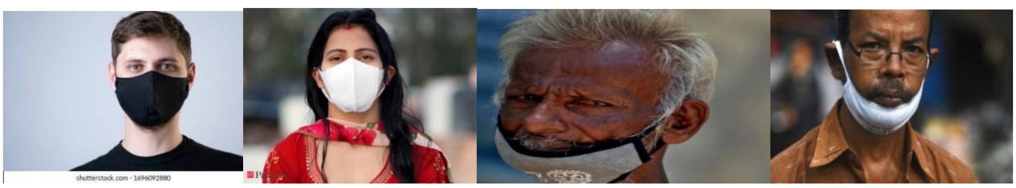


Fig.2

Several images are collected from various sources for training and testing the model. The images are labelled into three categories- masked, unmasked, improperly masked. Fig.3 shows some sample images from respective categories. To increase the size of the dataset, the data collected are augmented. Image augmentation is a technique used to increase the size of the training dataset by artificially modifying images in the dataset. Image augmentation will be very helpful if we want to make our dataset size larger. This is done because limitations of the dataset will make the CNN model to learn at a poor rate. The augmentation techniques that are used are rescaling, rotation range of 20, zooming range of 0.15 pixels ,shifting by width of range 0.2, shifting by height of range 0.2, flipping the image horizontally etc. The data taken from various sources will have different resolutions and sizes. Hence they are pre-processed to clean versions which could be fed into the CNN model. All the images of dataset are resized to 70x70 pixels. Each image will have three channels (RGB). So the dimension of each image is 70x70x3. Once the images are resized, then they will be converted from list to numpy arrays. Then the numpy arrays are normalized (i.e dividing each value in numpy array by 255). Because when we consider an image, its pixel values ranges from 0 to 255. This is a wide range. So if we normalise the image numpy array, then each value will be in a same distribution and also it will help in training the model faster.



Images with mask **Images with improper mask**

Images without mask

Fig.3

The numpy array (image) will be shuffled. This will avoid data of same distribution sitting together. Then the entire dataset is splitted into training set (80%) and validation set (20%). Then the model will be trained using the training data set. The CNN architecture will be having convolution layer followed by pooling layer (max pooling/avg pooling). In this case max pooling is adopted. Each layer is associated with RELU activation function which is one of the most commonly used. Then finally the output layer will be having SOFTMAX activation function which is used for classification problems. Once the training is over we have to check the accuracy and loss attained by the model. After analysing it, if there is any over fitting or under fitting or low accuracy, then we have to tune the hyper parameters like learning rate, number of epochs, adopting regularization method etc.

IV. RESULT ANALYSIS

The performance of the model is measured using the accuracy and loss rate. A good model must have reasonable accuracy and minimal loss rate. This model face mask detection system is created and tested successfully. Initially we had a dataset of size 300 in each category . But when we trained the model using this small dataset , the trained model resulted in over fitting. i.e the model performed well on training samples but performed poorly on validation set . Then we performed some data augmentation steps like rotation , flipping vertically, flipping horizontally, zooming, changing the contrast of the image, blurring , etc manually , then the dataset size becomes 700 in each category. The final trained model had an accuracy of 93%. Each convolution layer is followed by a max pooling layer . Then we trained the model for different epochs , learning rate and batch size . Some of them are as follows .

• When learning rate = 0.001, epochs=10, we got a training accuracy of
61% and validation accuracy of 64% .

• When learning rate = 0.001, epochs=30, there was some increase in
accuracy and decrease in loss for both validation and training set . The
training accuracy was 82% and the validation accuracy was 85% .

• When learning rate=0.0001, epochs=100 , the training accuracy was 93%
and the validation accuracy was 95% .

V. OUTPUT SCREENSHOTS

Figure4

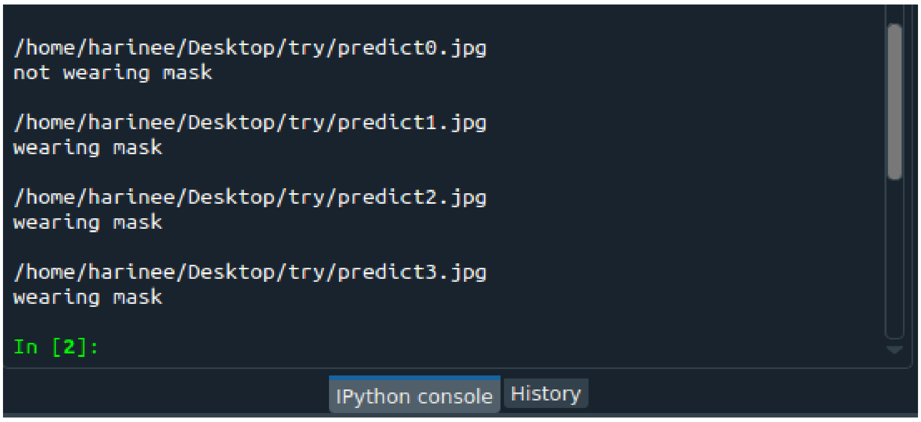


Figure5

VI. CONCLUSION

This face mask detection project is the need of the hour. We hope that a face mask detector will definitely contribute to the public health care. Convolution Neural Network is the foundation of this system. The model proposed in this paper predicts people in real time images and categorizes them accurately. We are working on improving the accuracy of the model by tuning the hyper parameters. The system proposed in this paper will reduce the manual power to impose the law at the same time, it ensures that everyone adheres to the law.

VII. REFERENCES

[1] An Introduction to Convolutional Neural Networks, Keiron O' Sheaand Ryan Nash,arXiv:1511.08458v2 [cs.NE] 2 Dec 2015

[2] Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet classification with deep convolutional neural networks.Commun. ACM 2017, 60, 84–90.

[3] Simonyan, K.; Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv,2014, arXiv:1409.1556.

[4] Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A.Going Deeper with Convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 1–9.

[5] He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the 2016,IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016;pp. 770–778.

[6] Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin.