Real Time Face Mask Identification Using Deep Learning

M Sujaritha¹, S Kabilan¹, M Manikandan¹ and S Nanda Kisore¹
¹Department of Computer Science and Engineering, Sri Krishna College of Engineering and Technology, Coimbatore, India.
sujaritham@skcet.ac.in, 17eucs072@skcet.ac.in, 17eucs090@skcet.ac.in , 17eucs101@skcet.ac.in

Abstract. Corona virus disease (COVID-19), one of the most life-threatening diseases which is highly harmful to the people in all over the world. The government of all countries in the world are in the situation to make effective steps to prevent the spread of COVID-19 in the society. If this situation continuous means the extinction of human beings may possible. So, the foremost action towards the virus is to stop spreading from the infected to the non-infected persons. To prevent spreading of the virus, one has to maintain social distancing from other people and personal hygiene. The main mode of transmission of the COVID-19 is by the droplets of saliva and the discharge from the nose. The governments laid a strict rule to wear a mask in the public places to prevent the spreading of the virus. Various technological techniques are well applied in pandemic studies. In this case we applied a deep learning algorithm to determine whether or not an individual is wearing a face mask. This research resulted in identification of face mask in the public places which results in the lowering of spread of the virus. This model is developed using the Convolutional Neural Network (CNN). It extracts the facial landmarks of the person and finds the face mask region. It is trained by datasets and predicts the result.

Keywords: COVID-19, Face Mask, Segmentation, Deep Learning algorithm, Prediction.

1. Introduction
The COVID-19 case was first discovered in Wuhan, Hubei Province of China, and spread rapidly in many International Countries. The virus broadens around the world in a minuscule period of time and became a pandemic and collapsed a health system and the life of the people in all countries. The most common symptoms of the disease are fever, dry cough, sore throat, headache, weakness, muscle pain, diarrhea, and shortness of breath, fatigue, nausea, loss of smell and taste which leads to severe attack of pneumonia, respiratory failure, heart and liver problems and finally leads to death of the person. The knowledge and the effect of the COVID-19 should be well informed to the public. As of now, the statistics shows that 115M people were affected. In these, 2.55M people were died due to the attack of COVID-19. The virus is transmitted primarily by close contact with an infected individual. Some preventive steps are washing hand regularly, physical distancing, containment, oxygenating of indoor spaces, covering coughs and sneezes and finally the use of face masks in the civil areas are highly recommended to prevent the spreading of the deadliest disease in the society. The entire human beings are facing the crisis, in the meantime the nature took its own healing procedure by closing of the ozone hole, the birds and other creatures were seen everywhere peacefully. So, this pandemic situation has taught us that the world is not for humans alone and to stay united and should help each other. Many issues benefit from the implementation of artificial intelligence techniques. Many experiments using
artificial intelligence techniques have been performed in the COVID-19 pandemic. Since the outbreak, the pandemic has had an effect on the entire planet, not only in terms of health and hygiene, but also on the consumer confidence. There are currently no known therapies for this virus, but maintaining personal disinfection and social distance will help to minimize the spread of infection. Although the number of infected persons is high, the number of deaths is low. So, here the control of spreading the virus in the society is to follow the wearing of the face masks.

In this paper, analysis and prediction of the face mask identification is done using Support Vector Machine (SVM), Convolutional Neural Network (CNN), Input Layer, Pooling Layer, Fully Connected Layer, SoftMax or Logistic Layer.

2. Literature Survey
Coronavirus Disease 2019 (COVID-19) spread rapidly across the world and emerged as a comprehensive pandemic. Facemask use is predominant for pestilential disease control, but the activeness of facemasks has been dwindled, owing to inappropriate use. There have been no studies on the automatic detection of facemask-wearing conditions that have been released. By integrating image super resolution and classification networks, the built face mask recognition method quantifies three category classification based on 2D facial images [1]. The proposed algorithm consists of four major steps: Face detection and cropping, image super-resolution, and facemask-wearing condition recognition are all examples of image pre-processing [2].

It is more difficult to train deep neural networks. A residual learning system is presented to make the training of networks that are significantly deeper than previously used networks simpler. Instead of learning not used functions, we obviously rebuild the layers to learn residual functions with regard to the layer inputs [3-5]. We present extensive experimental affirmation demonstrating that these residual networks are serene to revitalize and can obtain exactness from significantly expand depth. We appraise residual nets with up to 152 layers of depth on the dataset, which is eight layers deeper than VGG nets but with less complications.

3. Model implementation and methods
In order to solve the problem, this research follows a set of steps. Starting with the dataset preparation, which must be included in the dataset preprocessing. The test data must be set before the classifier can be implemented. The classifier can be used after the test data has been set. The Deep Learning Algorithms are compared in the final stage. The steps followed for the implementation of the process is shown in Figure 1.

![Figure 1. General Process](image_url)

The following Figure 2 is a more detailed explanation of the process. Each of the steps will be explained in sub-chapters.
3.1 Dataset

The training of CNN requires thousands of images as their training dataset. It is used 2 label images which are the person is wearing face mask and another one in which the person is not wearing the face mask to train our dataset. These are the images which we have collected from various sources including Google images and research the dataset in Kaggle website which is the largely utilized database for technicians etc.,[6-9]. Here 678 masked images and 664 unmasked images are the dataset. The following Figure 3 represents the dataset images for this process.

![Figure 3. Dataset Images](image)

3.2 Preprocessing

The dataset contains images, based on the people wearing face mask and the people not wearing face mask. In this paper we use a label encoder preprocessing technique to convert and the Image quality increases. This proper dataset is cleave into training and testing. Training data used to train the machine. Training process contains both feature and label. The testing data should be used to predict the label class.

3.3 Model Selection

Deep Learning is a subfield of Artificial Intelligence that deals with the development of techniques and methods that enable computers to learn. In this study, to analyze the supervised machine learning algorithms. Here, used supervised learning classification algorithm to compare the algorithm working process [10-12].

3.4 Support Vector Machine

Support vector machines (SVM) are supervised learning process that can be used to classify and predict data. The data points that are adjacent to the decision surface are known as support vectors (or hyperplane).

Define the hyperplanes $H$ such that:
\[ w \cdot x + b \geq +1 \text{ when } y_i = +1 \]
\[ w \cdot x + b \leq -1 \text{ when } y_i = -1 \]

H1 and H2 are the planes:
H1: \[ w \cdot x + b = +1 \]
H2: \[ w \cdot x + b = -1 \]

The points H1 and H2 are tips of the Support.
Figure 4 represents the boundary for the decision.

**Figure 4.** Maximum margin hyper lane

### 3.5 Convolutional Neural Network (CNN)

Day by day, computer vision progresses at a breakneck pace. It’s one of the reasons is using a deep learning methodology now a days in various process. It is commonly applied to analyzing visual imaginary. A CNN is made up of neurons, each of which binds to all of the neurons in the next layer. When we speak of “vision, the phrase convolutional neural network (CNN) comes to mind since it is often used. Face recognition, image classification, and other applications of CNN in computer vision are examples. It behaves in the same way as a simple neural network. CNN even has learnable parameters, such as weights and biases, similar to a neural network [13]. Figure 5 depicts the feature extraction process.

**Figure 5.** CNN process
3.6 Classification of Images using MobileNetV2
This model has been used for the classification for Deep Neural Network. The adding of newly training layers is done and trained on the dataset which is used to determine the features of masked image and unmasked image [14-15]. Figure 6 is the classification process of this model.

3.7 Convolutional Layer
This layer is the CNN’s basic building block. The main process of this layer is to combine the two functions and the resultant third function is the output of this process. The mechanism which works on this process is called sliding window mechanism. The convolution of the input image matrix and the kernel gives the output with the help of the Keras [16]. Figure 7 represents the general convolution process of this model.

\[ C(T) = (A \ast B)(x) = \int_{-\infty}^{\infty} A(T) \times B(T - x) dT \]

3.8 Input Layer
This layer consists of the neurons which is the initial input for the whole process and in the neural network it represents the image pixel matrix. It is also consisting of the visual layers as well as the hidden layers.

3.9 Pooling Layer
Pooling layers is used to reduce the dimensions of the data. Activation maps are supplied by convolutional layers. On activation maps, the pooling layer uses non-linear down sampling. Pooling is a combative strategy. Smaller filter sizes are being used more often, and pooling is being abandoned. Figure 8 represents the two types of pooling process [17].
a. Max Pooling
This pooling gives the output of the maximum value of the selected region.

b. Average Pooling
This pooling gives the output of the average values in the current region.

3.10 Fully Connected Layers
This layer is considered as the most successful in classifying and recognizing images [18]. The model having the layers appended which have connections to the activation layers. This layer helps in classification of the images with multi or binary classification. The input of this layer is the resultant output of the antecedent layer which is pooling layer.

3.11 SoftMax or Logistic Layer
The SoftMax layer is CNN's final layer. It is placed at the base of the Fully Connected layer. The input of this layer is transformed into the binary values which can be represented as the probabilities. It is the last activation function which helps to normalize the output to probability distribution.

3.12 Output Layer
Output layer contains the result of the input and is often referred as the at most last layer of the convolutional neural network process. The output is shown in the real time using standard camera [19-22]. Figure 9 is the final output of our process.

4. Performance Evaluation Metrics
Here, the performance metrics namely accuracy, precision, recall and f1 score is to be discussed below. Figure 10 depicts the various performance metrics values.

\[
\text{Accuracy} = \frac{(T_p+T_n)}{(T_p+F_p+F_n+T_n)}
\]
Precision = \( \frac{Tp}{(Tp+Fn)} \)

Recall = \( \frac{Tp}{(Tp+Fn)} \)

F1 score = \( \frac{2 \times (Recall \times Precision)}{(Recall + Precision)} \)

Where,

Tp = True positive
Tn = True negative
Fp = False positive
Fn = False negative

![Performance Metrics](image)

**Figure 10.** Performance result

5. Experimental Setup

This learning algorithms are executed Programming language in windows operating system with core i5 and with 8GB of RAM and 2.8GHz of processor speed.

6. Result and Discussion

The work of face masks reduces transmission of COVID-19 from infected one to non-infected. This study has proved to give high accuracy values. In data set, features are analyzed and preprocessed, choosing best fit input that predicts whether the mask is wore or not. Since the features are linear and simple, it is greatly fits for the proposed model. The performance of other algorithms compared to CNN is low. The model developed using CNN algorithm shows the best among all with the accuracy of 98%. The accuracy of four classification algorithms will be measured in this report. Following the classification procedure, a review of the accuracy results reveals that CNN Architecture achieved a high accuracy of 98 percent. The comparison of the accuracy of the different algorithm used in the process of mask identification and their performance is given in the below Figure 11.
Besides, the results view displayed. To investigate the accuracy of estimation in this report, confusion matrix and ROC curve was used. In figure 12, the ROC curve of the performance measurement of the various classification problems is shown.

![Figure 11. Accuracy of different algorithms](image)

**Figure 11. Accuracy of different algorithms**

Figure 13 represents the confusion matrix for this method. The output of a classification model on a collection of test data for which the true values are known is represented by a confusion matrix. The uncertainty matrix is easy to grasp.

![Figure 12. ROC curve](image)

**Figure 12. ROC curve**

![Figure 13. Confusion matrix](image)

**Figure 13. Confusion matrix**

7. Conclusion and Future Scope

As of today, the Covid-19 wave has infected over 200 nations. Healthcare sectors have crashed due to a dramatic rise in the number of cases. The study was to identify the mask is wore or not. The use of the Deep Learning algorithms in the field of medical and health is very useful and it takes to the advanced
level. In today’s world everything is about technical. So, the identification of face mask is very essential in today’s world. In this paper we use CNN algorithm to attain at most accuracy with 98%. It shows that the use CNN is best comparing to other algorithms.

This study helps us to identify the person is in higher risk of getting affected by the COVID-19. So, as a result of this paper, the scope for the future works is many. Train and test the greater number of images in the datasets to achieve more speed and accuracy in identifying. It can be applied in the entrance of the companies, hospitals, public places, airports, railway stations to detect whether the employees or publics wore face mask or not. It can also be installed in the Public CCTV cameras.

References

[1] Preethas R and Suresh G R 2019 Identifying Facemask-Wearing Condition Using Image Super-Resolution with Classification Network to Prevent COVID-19.
[2] Jobin Christ M C and Parvathi R M S 2018 Deep Residual Learning for Image Recognition.
[3] Zhongyuan Wang, Guangcheng Wang 2017 Masked Face Recognition Dataset and Application.
[4] Mamata and Kalas 2018 Additive Angular Margin Loss for Deep Face Recognition.
[5] Hassner T, Masi I, and Medioni G 2016 Regressing Robust and Discriminative 3D Morphable Models with a very Deep Neural Network.
[6] Viola P and Jones M 2001 Rapid object detection using a boosted cascade of simple features IEEE computer society conference on computer vision and pattern recognition, vol. 1, pp. I–I.
[7] Girshick R, Donahue J, Darrell T, and Malik J, Rich feature hierarchies for accurate object detection and semantic segmentation, IEEE conf. computer vision and pattern recognition, pp. 580–587.
[8] D. Devikanniga, A. Ramu, and A. Haldorai, Efficient Diagnosis of Liver Disease using Support Vector Machine Optimized with Crows Search Algorithm, EAI Endorsed Transactions on Energy Web, p. 164177, Jul. 2018. doi:10.4108/eai.13-7-2018.164177
[9] H. Anandakumar and K. Umamaheswari, Supervised machine learning techniques in cognitive radio networks during cooperative spectrum handovers, Cluster Computing, vol. 20, no. 2, pp. 1505–1515, Mar. 2017.
[10] Liu W, Anguelov D, Erhan D, Szegedy C, Reed S, Fu C Y, and Berg A C 2016 Single shot multibox detector, European conf. computer vision, Springer, pp. 21–37.
[11] Redmon J, Divvala S, Girshick R, and Farhadi A 2016 You only look once: Unified, real-time object detection, IEEE conf. computer vision and pattern recognition, pp. 779–788.
[12] Lin T Y, Goyal P, Girshick K, He K, and Dollar P 2017 Focal loss for dense object detection.
[13] Rosebrock A 2020. COVID-19: Face Mask Detector With Opencv, Keras/Tensorflow, And Deep Learning.
[14] Francois Chollet 2017 Deep Learning with Depthwise Separable Convolutions IEEE conf. computer vision and pattern recognition, pp. 1251-1258.
[15] Sandler M, Andrew Howard, Menglong Zhu, Andrey Zhmoginov, and Liang-Chieh Chen, 2018 MobileNetV2: Inverted Residuals and Linear Bottlenecks, IEEE conf. computer vision and pattern recognition.
[16] Hinton and Salakhutdinov G E 2006 Reducing the dimensionality of data with neural networks. Science pp. 313, 504–507.
[17] Chen J, He F, Zhang Y, Sun G and Deng M 2020 Combining Superpixel Pooling and Multi-Scale Feature Fusion.
[18] Yang S, Luo P, Loy C C and Tang X 2015 Facial Parts Responses to Face Detection: A Deep Learning Approach. IEEE Conf. Computer Vision, pp. 3676–3684.
[19] Yi D, Lei Z, Liao S and Li S Z, 2014 Learning Face Representation from Scratch.
[20] Kingma D P, Ba J and Adam 2014 A Method for Stochastic Optimization.
[21] Wang Z, Bovik A C, Sheikh H R and Simoncelli E P 2004 Image quality assessment: From error visibility to structural similarity IEEE Trans. Image Process, vol 13, pp. 600–612.
[22] Zota P A, Oberste M S, Monroe S S, Nix W A, Campagnoli R, Icenogle J P, Penaranda, Bankamp B, Maher K and Chen et al M H 2003 Characterization of a novel coronavirus associated with severe acute respiratory syndrome, vol. 300, no. 5624, pp. 1394–1399.