*Grenze International Journal of Engineering and Technology, Jan Issue*

# FMDT: A Face Mask Detection Technology for Pandemic Outbreaks

Muhsina V.P.[1], Prathewsh S Babu[2], Hema P Menon[3] and M.K. Vinayak[4]
[1-4]Department of Computer Science and Engineering, Sreepathy Institute of Management & Technology, Vavanoor, Koottanad, Palakkad, Kerala
Email: {muhsina.vp, hemapmenon}@simat.ac.in, {prathewshsureshbabu, vinayaksunilmk}@gmail.com

***Abstract*—The use of face mask has become a part of the normal life since the rise of Covid-19 pandemic. This period has also seen the difficulties faced by law enforcement officials. This work aims to develop a face mask detection system which would help officials to make people abide by the rules and regulations. The devised system will help the automatic identification of the presence or absence of a mask. Accordingly, the person without mask will be highlighted and alarm raised. Such a system would be helpful for government & private sectors in law enforcement and also to get a consolidated idea on number of people breaking the norms. The Sequential Convolutional Neural Network model has been used for the training and detection of the absence or presence of a face mask.**

***Index Terms*— Convolutional Neural Network, Face Mask Detection, Covid-19.**

## I. INTRODUCTION

The corona virus diseases is started in 2019 (Covid 19) has effected millions of people across the world. The studies founds that the disease has grown to transmitted, come out through sneezing, cough etc. on close contacts with the person who are infected. To curb the spreading of the disease wearing of face mask by all individual when in public places was introduced. Inspite of all rules and regulations, it is fact that many people wherever following them. Hence it become necessary to devise automation based application for identifying defaulters.

A bigger problem rests on the hands of the population who willingly choose to either wear masks ornot. This was made from the perspective of any organization choosing to automate their process of tracking visitors who isn't wearing masks.

Detection of face mask is the complex process because of different types of mask that are available, differences in facial features from person to person and the way the mask has been worn. The development of any such automation process for any real time application requires the use of machine/deep learning based techniques for achieving higher accuracy rate. One such approach is the use of convolutional neural network. This technique

The main objective of the proposed system is to detect if a person has worn the mask or if he/she has worn it properly such that it would prevent the spread infections.

- Security: To stop the spread of infections from others so as to not get affected by viruses or illness.
- Accuracy: To provide more accurate results than thecurrent implemented system.

*Grenze ID: 01.GIJET.9.1.524*

- Robustness: Being robust enough to be used by hospitals and other organizations in order to detect and also increase the security in certain areas for entry or exit.

## II. Literature Survey

It is suggested that there are seven distinct types of information that one derives from seen faces; these are labelled pictorial, structural, visually derived semantic, identity-specific semantic, name, expression, and facial speech codes. The analysis of facial expression, facial speech, and face recognition units can all benefit from the descriptions provided by structural encoding processes, according to a functional model that has been put out. Additionally, it is suggested that the cognitive system actively decides whether the first match is close enough to signify actual recognition or just a "resemblance," with a number of elements thought to influence such decisions. This system consists of a program that can be executed in any computer or device that has a camera. With the help of CNN, two datasets are cross referenced to find if the person in the live video stream is wearing a face mask or not.

Covid-19 Face Mask Detection Using Tensor Flow, Keras and Open CV by Arjya Das, Mohammad Wasif Ansari, Rohini Basak [1]. Here it uses some basic Machine Learning packages like TensorFlow, Keras, OpenCV and Scikit-Learn for face mask detection. The proposed method detects the face from the image correctly and then identifies if it has a mask on it or not and even in motion also. The method attains accuracy up to 95.77 percent and 94.58 percent respectively on two different datasets. Here it finds optimized values of parameters with the help of Sequential Convolutional Neural Network model.

RetinaFaceMask proposed by M. Jiang, X. Fan, H. Yan [2] is a one-stage detector, which consists of a feature pyramid network to fuse high-level semantic information with multiple feature maps, and a novel context attention module to focus on detecting face masks and it also propose a novel cross-class object removal algorithm to reject predictions with low confidences and the high intersection of union. It shows that RetinaFaceMask achieves state-of-the-art results on a public face mask dataset with 2.3 percent and 1.5 percent higher than the baseline result in the face and mask detection precision, respectively, and 11.0 percent and 5.9 percent higher than baseline for recall.

In Big Data Concepts and Techniques in Data Processing by B. Suvarnamukhi, M. Seshashayee [3], Big data is a pool of large amounts of data collected from various sources, such as social working sites like facebook and transactional data like online shopping sites. Big data is stored in distributed architecture framework. Hadoop is an open source framework for creating distributed applications that process huge amount of data. Preparing of information can incorporate different activities like highlighting, indexing and searching. It is highly difficult to process in single machine, with huge amount of data.An approach to face detection and recognition proposed by D. Meena and R. Sharan [4]. Face detection is a method to detect a face from an image which have several attributes in that image. In this paper, viola Jones algorithm is used to detect face and principal component analysis for face recognition.

Detecting Masked Faces in the Wild with LLE-CNNs proposed by Shiming Ge, Jia Li, Qiting Ye, Zhao Luo [5] talks about detecting faces with occlusions. This paper first introduces a dataset, denoted as MAFA, with 30, 811 Internet images and 35, 806 masked faces. Faces in the dataset have various orientations and occlusion degrees, while at least one part of each face is occluded by mask. It proposes locally linear embedding (LLE)-CNNs for masked face detection.

Deep Convolutional Neural Network for Image Classification: A Comprehensive Review proposed by W Rawat, Z Wang [6] focuses on the application of CNNs to image classification tasks. It analyzes 1) their early successes, 2) their role in the deep learning renaissance, 3) selected symbolic works that have contributed to their recent popularity, and 4) several improvement attempts by reviewing contributions and challenges of over 300 publications and also it introduces some of their challenges and current trends.

## III. Materials and Methods

This section gives the system architecture and the overall implementation details of the proposed system. Initially there are two datasets. Every images is an array of pixels. Using data visualization, the image meta data is taken. Every image is reshaped. Convolutional layers are developed so as to detect the edge in images. A feature map is created. Max pooling calculates the maximum value for patches of a feature map. When you feed a CNN colored images, those images come in three channels: Red, Green, Blue. Say we have a 32 x 32 image. For each of the 32 x 32 pixels, there is a value for the red channel, the green, and the blue. The number of input channels in the convolution is c, while the number of output channel is c. The filter for

such a convolution is a tensor. A tensor is a generalization of vectors and matrices and is easily understoodas a multidimensional array.The main advantage of using the ReLU(Rectified Linear Unit) function over other activation functions is that it does not activate all the neurons at the same time. In the end with the neural network algorithm, as output many iterations are done and finally accurate prediction and detection is obtained.

Figure1. shows the System Architecture for Face Mask Detection and components are Convolutional layer, Pooling layer (down sampling), Fully connected input layer and Fully connected layer.

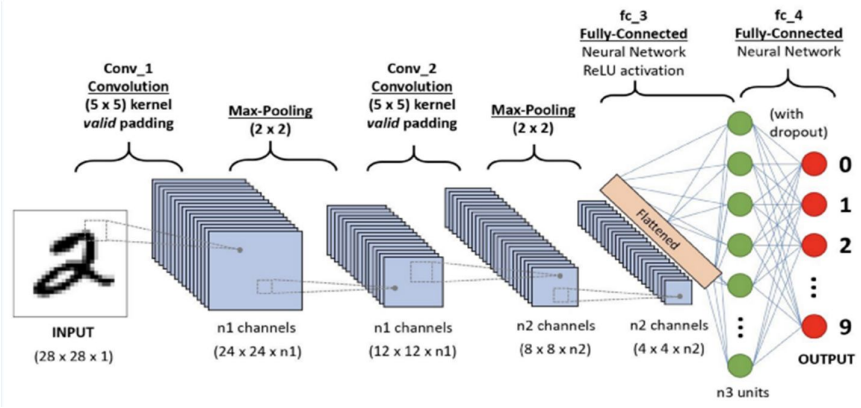



Figure 1: System Architecture for Face Mask Detection PC : Adopted with modification from [7]

The class probability for each feature is predicted by creating a feature map by the convolution layer. This is done by applying a filter that scans the whole image, few pixels at a time. Dimensionality reduction is done using pooling layer to retain the most relevant features. Fully connected input layer "flattens" the outputs generated by previous layers to turn them into a single vector that can be used as an input for the next layer and fully connected layer applies weights over the input generated by the feature analysis to predict an accurate label. The proposed method detects the face from the image correctly even in motion and then identifies if there is a mask or not.

*A. Phase One*

- Data Visualization

The total number of images in the dataset is visualized in both categories – 'with mask' and 'without mask'. The statement categories=os.listdir(datapath) categorizes the list of directories in the specified data path .he variable categories now looks like: ['with mask', 'without mask'].

- Data Pre-Processing

The obtained data needs to be processed for making it ready for further use. This is done by converting the data from the given format to much more user friendly one. This work uses Numpy, OpenCV for this purpose.

*B. Phase Two*

This work makes use of a sequential based CNN model. The main indicator of the training progression isthe loss function. The loss function tells us how bad the model prediction on training and evaluation data is. Herewe have to make a clear distinction between the training and evaluation dataset. The training dataset is used to train the network meaning that training data are used to modify the model variables

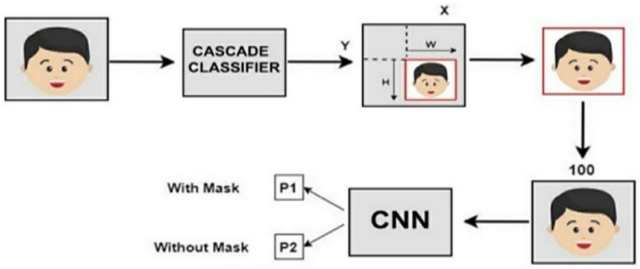



Figure 2. Overview Model

TABLE I: DETAILS OF SELECTED

| Name of Dataset | Size(in MB) |
|---|---|
| With Mask | 25.3 |
| Without Mask | 25.4 |

*C. Dataset Selection and Creation*

Dataset was created from images using kaggle and google images the dataset consist of two types of images. In total there are 1376 images of faces in which 690 images have people wearing masks and 686 wearing no masks. Defective faces were fixed and artificial masks had to be added in some images. The details of dataset is given in TABLE 1 and sample image data with mask and without mask also given Figure 3 and Figure 4 respectively.

Figure 3. Dataset with mask

Figure 4: Dataset without mask

❖ Fine-Tuning of Pre-Trained Model

In the proposed work, facemask detection is achieved through deep neural networks because of their better performance than other classification algorithms.. To train the network faster and cost-effective, deep learning-based transfer learning is applied here. Transfer learning allows to transferring of the trained knowledge of the neural network in terms of parametric weights to the new model. It boosts the performance of the new model even when it is trained on a small dataset. There are several pre-trained models like AlexNet, MobileNet, ResNet50 etc. that had been trained with 14 million images from the ImageNet dataset. In the proposed model, MobileNetV2 is chosen as a pre-trained model for facemask classification.

❖ Image Complexity Prediction

It is observed that dataset, we consider primarily, contains two major classes that is, mask and non-mask class but the mask class further, contains an inherent variety of occlusions other than surgical/cloth facemask, for example, occlusion of ROI by other objects like a person, hand, hair or some food items. These occlusions are found to impact the performance of face and mask detection. Thus, obtaining an optimal trade-off between accuracy and computational time for face detection is not a trivial task. So, an image complexity predictor is being proposed here. Its purpose is to split data into soft versus hard images at the initial level followed by a mask and non-mask classification at a later level through a facemask classifier.

❖ Loss Function And Optimization

The stochastic gradient descent optimization algorithm is applied to update the model parameters with a learning rate of 0.03. Further, there exist numerous loss functions in PyTorch but one which is most suitable with balance data is cross-entropy loss. Activation functions that are not interpretable as probabilities (i.e. negative or greater than 1 or sum of output not equals to 1) should not be selected. Since Softmax guarantees to generate well-behaved probabilities distribution over categorical variable so it is chosen in our proposed model

$$Loss = \sum_{\forall_x} t\,(x) \log\,(e(x))$$

Further, the loss function over N images (also known as cost function over complete system) in binary classification can be formulated as given in equation

$$Loss = \frac{1}{N} \sum_{x} \sum_{n=1}^{N} t_n(x) \log\,(e_n(x))$$

*D. Training and Testing*

❖*Supervised Pre-Training*

Here, it is discriminatively pre-trained the CNN on the original biased dataset. The pre-training was performed using the open-source Caffe python library. Obtaining a top-1 error rate 1.8% higher on validation set. This discrepancy may cause due to simplified training approach.

❖ Supervised Pre-Training With Domain-Specific Fine-Tuning

The other approach is to first remove the inherent bias present in the available dataset and then execute supervised learning over a domain-specific balanced dataset. The bias is alleviated by applying random over-sampling (ROS) with data augmentation. The formula used for computing the imbalance ratio is given by equation

$$\rho = \frac{Count(majority(D_i))}{Count(minority(D_i))}$$

Here, D refers to image Dataset, majority (Di) and minority (Di) return the majority and minority class of D. Count(X) returns the number of images in any arbitrary class x. After data balancing, stochastic gradient descent (SGD) training of CNN parameters with a learning rate of 0.003 is set over wrapped region proposals. The low learning rate allows fine-tuning of the model without clobbering the initialization. We added 2025 negative windows with 50 background windows to increase non-mask dataset 22 KB. The balancing leads to a reduction in the top-1 error rate of 3.7

## IV. Results and Discussions

To measure the effectiveness of the proposed system, it is compared with the existing system. But the existing system is implemented for images. So first of all, it is extended into video. Then compare it with the modified system.

### *A. Detection Analysis*

The proposed system was able to detect a maximum of 3-4 faces with or without masks in a single frame compared to previous existing system. The faces with mask are displayed inside a green rectangle and faces without masks in a red rectangle. The accuracy is in terms of 100% and the model's training and detection loss are plotted on a graph for further analysis. The results are shown in following figures (Figure 5 to Figure 12).

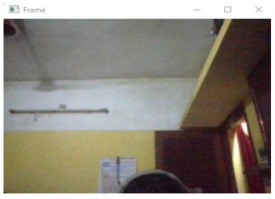
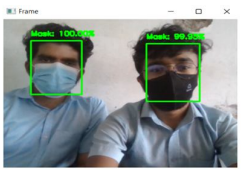


Figure 5: Frame

Figure6: Detection with mask

Figure 7: Detection with mask



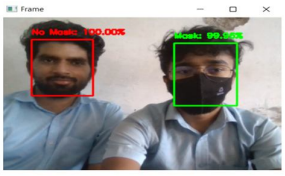


Figure 8. Both with and without mask ( one light and one dark )

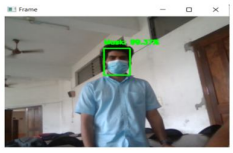


Figure 9. Increased detection range

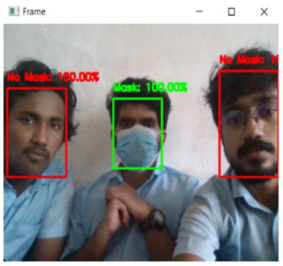


Figure 10. Detection on multiple faces


Figure 11. Training - Accuracy loss graph (Before optimization)

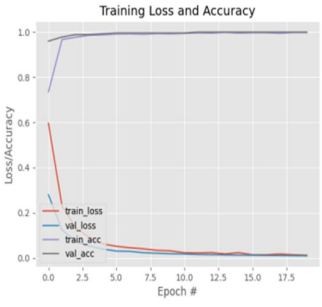



Figure 12. Training - Accuracy loss graph (After optimization)

## V. Conclusion

With advent of Covid-19 Pandemic wearing a mask has become mandatory in public places. This paper discuss the development of a face mask detection technology using sequential CNN. This work implemented using Keras/Tensorflow and OpenCV . The training accuracy loss graph before and after optimization shows that the model accuracy increases in the different epoch and stabilizes after some time. 3800 face samples including both with mask and without mash were used for the training the model. The developed system was tested at real time taking in to consideration the various possibilities and was founds to give an accuracy of 99%. The system thus developed would be very useful for integration into any surveillance device for detecting the persons who are not wearing the mask.

## Acknowledgment

The authors would like to acknowledge Ms. Divya Visakh, Ms. P Athira and Mr. Jithin for their support in development of this work.

## References

[1] Das, A., Wasif Ansari, M., Basak, R. (2020). Covid-19 Face Mask Detection Using TensorFlow, Keras and OpenCV. 2020 IEEE 17th India Council International Conference (INDICON)

[2] M. Jiang, X. Fan and H. Yan "RetinaMask: A Face Mask detector", arXiv.org, 2020. [Online]. Available: https://arxiv.org/abs/2005.03950. 2020

[3] Suvarnamukhi and M. Seshashayee "Big Data Concepts and Techniques in Data Processing", International Journal of Computer Sciences and Engineering, vol. 6, no. 10, pp. 712-714, 2018. Available: 10.26438/ijcse/v6i10.712714.

[4] Meena and R. Sharan "An approach to face detection and recognition," 2016 International Conference on Recent Advances and Innovations in Engineering (ICRAIE), Jaipur, 2016, pp. 1-6, doi: 10.1109/ICRAIE.2016.7939462.

[5] S. Ge, J. Li, Q. Ye and Z. Luo "Detecting Masked Faces in the Wild with LLE-CNNs," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, 2017, pp. 426-434, doi: 10.1109/CVPR.2017.53.

[6] W Rawat, Z Wang "Deep Convolutional Neural Network for Image Classification: A Comprehensive Review", Neural Computation (Volume: 29, Issue: 9, Sept. 2017), ieeexplore.ieee.org, doi: 10.1162/necoa00990

[7] Souradip Chakraborty & Rajesh Shreedhar Bhat "Reducing the Carbon Foot Prints of CNNs at the cost of interactions- Depthwise & Pointwise Convolution", article published in Towards Data Science - Mar 20, 2020