Real-Time Face Mask Detection to Prevent COVID-19 in Confined Spaces

Aviva Munshi, Anoushka Mehra, Ashna Choudhury, Yogeswari Sahu, Asmit Gupta

Abstract— **In this pandemic, it is getting more and more challenging to keep track of people wearing masks regularly. We cannot solely depend on human efforts to take care of this task. Therefore we need to develop software that can automatically detect whether any given person is wearing a mask or not. In this time of Covid-19 pandemic, it is of utmost importance that we all use masks regularly. But there are always exceptions, and only manual human efforts are meaningless to track who are not wearing masks. Therefore there is a need to develop software that can detect if a person is wearing a mask or not automatically. Currently, face recognition field has seen many challenges and has evolved much. New and more accurate algorithms are being derived by using convolutional architectures. Convolutional architecture also has the benefit of its ability to extract pixel details. Fully convolutional architecture is used for training so that faces are segmented out of the image semantically. We can find whether a person is using a mask not or adequately by feature detection and feature extraction. For accurate model creation, a dataset of morphed images with veneers can be used by the face mask detector. Thus, the created model will be computationally efficient. This will also be easily deployable because it will use MobileNetV2 architecture (Google coral, Raspberry Pi, etc.). Due to Covid-19 pandemic, face mask detection is required, and this framework can be used for that in real-time applications. This project can be used at public places like airports, schools, train stations, workplaces and bus stops by merging with embedded applications. This will lead to a safer environment in pandemic times. This identification process will help us classify individuals by reducing the number of manpower required.**

Index Terms— **Neural network, Image Data generator, Face mask detection, MobileNetV2.**

I. INTRODUCTION

Due to the Covid-19 pandemic, face mask detection plays a vital role in limiting the spread of the corona virus. This pandemic requires the people to wear their face masks continuously , maintain social distancing and wash their hand with sanitizers hourly. Wearing their mask during moving outside or in the home is a very crucial preventive measure and is essential in times that they start exhibiting symptoms related to COVID-19. Therefore, the government and the people need to lower the possibility of transmission of this

Aviva Munshi, Computer Science Student, Vellore Institute of Technology, Vellore, India
Anoushka Mehra, Computer Science Student, Vellore Institute of Technology, Vellore, India
Ashna Choudhury, Computer Science Engineering, Vellore Institute of Technology, Vellore, India
Yogeswari Sahu, Computer Science Engineering, Vellore Institute of Technology, Vellore, India
Asmit Gupta, Computer Science Engineering, Vellore Institute of Technology, Vellore, India

deadly virus and the prevalence of the infections.

Training models through CNN and incorporating convolution operators yields a generic output. With the help of this project, we are trying to create a layer of efficiency to produce more accurate results and a better output. To achieve this efficiency, we are incorporating MobileNetV2, here we will also use an image data generator to augment our dataset (IDG).

Image Data Generator will ensure that more precise outputs will be generated because of the presence of enough images present in the training set. In this, we will also use the "Relu" Activation function because of the presence of non-linear test cases present in our database images. The "Relu" activation function is best suited for the image theoretically. Additionally, we include a double layer in the output layer to identify whether people are wearing masks correctly or not.

II. METHODOLOGY

About the primary technologies:

A. Convolutional Neural Networks

Convolutional Neural Network (ConvNet/CNN) is a Deep Learning system that can accept an image as input, assign relevance (trainable weights and biases) to numerous perspectives in the image, and distinguish between them. When compared to other classification methods, the amount of pre-processing required by a ConvNet is significantly less. While basic approaches require hand-engineering of filters, ConvNets can learn these filters/characteristics with just enough training.

B. MobileNetV2

With regards to classification, object recognition, as well as semantic segmentation for mobile visual recognition, MobileNetV2 is a significant improvement from MobileNetV1. MobileNetV2 is included in the TensorFlow-Slim Image Classification Library, or Colaboratory could be used to get working straight away. It could also download the notebook and analyse it remotely using Jupyter. TF-Hub modules for MobileNetV2 are also available, with pre-trained checkpoints available on Github. The concepts of MobileNetV1 are expanded upon in MobileNetV2, which uses depth - wise separable convolution as a budget building element.V2, on the other hand, includes two additional architectural details:

1) linear bottlenecks between layers
2) quick connections among bottlenecks.

C. Image Data Generator

In Keras, the Image Data Generator class has never used an "additive" technique. It just doesn't put the original data, it randomly transforms it, and then delivers both the actual and the altered data.

The ImageDataGenerator, on the other hand, takes the original information, alters it at random, and thereafter delivers the changed data.

Data augmentation is a set of strategies for generating "new" training data via old ones by incorporating random oscillations and perturbations (while guaranteeing that such data's class labels remain unchanged).

We hope to enhance the model's generalizability while we apply data augmentation.

D. Proposed Approach

The model first uses the abundant dataset of images and analyses them. After analysing, the model passes the acquired input through MobileNetV2 architecture. Feature enhancement is performed on the image as a third step to extract relevant features like: eyes, nose, mask etc. Post that, the image is flattened so that the image can be passed through the multi-layered perceptron. Finally, we can view on the live video feed whether or not a person is wearing a facial mask.

In the initial phase, we lay a heavy emphasis on training the entire model by first augmenting the dataset and then performing the necessary computations.

- Load the augmented face mask dataset.
- Train the face mask classifier to generate accurate outputs.
- Serialize the classifier.

In the second phase we implement our model to generate necessary classification based on input.

- We first load the classifier.
- We then detect faces using our model.
- We then extract necessary features required to detect whether or not a person is wearing a facial mask.

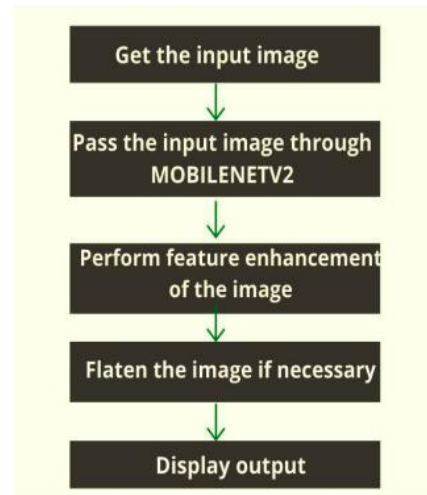Finally, we show that a person is with or without a mask based on the input parameters.
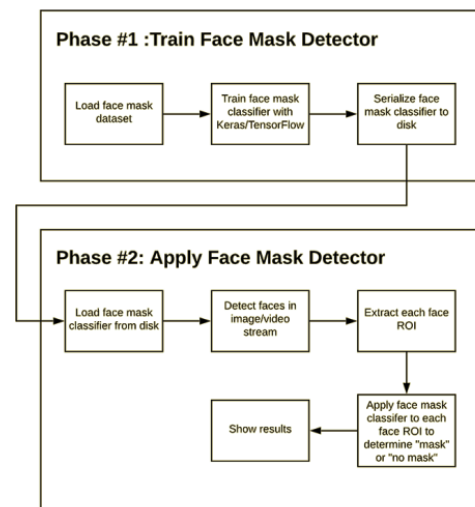

Fig 1: Proposed Approach

Fig 2: Phase Diagram

E. Evaluation Metrics

Accuracy - The most accurate indicator of success is accuracy, and it is solely a matter of the degree of correct number of predicted observations to the observations made in total. One would deduce that the model is the best when they have high accuracy. In order to see the efficiency of the model, you have to check in various parameters. . Accuracy can be defined as TP+TN / TP+FP+FN+TN where TP stands for True Positives, TN stands for True Negatives, FP stands for False Positives and FN stands for False Negatives.

Recall - Recall is that the quantitative relation of observations that were correctly predicted to be positive to the observations that were within the right class. Recall is defined as TP / TP+FN where TP stands for True Positives and FN stands for False Negatives.

Precision – Precision is the quantitative relation of the positive instances to the positive instances in total.

F1 Score - F1 Score brings up the weighted average of Precision and Recall into a single metric by calculating their harmonic mean. This score together takes into consideration each of the false positives and the false negatives. it's not as easy to comprehend intuitively as accuracy, but F1 is usually a lot more helpful than accuracy, particularly if the data has got uneven distribution of classes or categories. It is very

https://doi.org/10.31871/IJNTR.7.11.18

**International Journal of New Technology and Research (IJNTR)
ISSN: 2454-4116, Volume-7, Issue-11, November 2021 Pages 17-20**

useful when we have to compare the performance of any two classifiers showing varied precision and recall values. F1 Score is defined as 2*((recall * precision)/ (recall + precision))

III. DATASET ANALYSIS

The dataset we have used for training our model consists of real images of faces wearing masks. We have used the RMFD (Real World Masked Face) dataset and some images were also collected from Kaggle. Overall, our dataset contains 4095 images belonging to two classes:

• with_mask: 2165 images
• without_mask: 1930 images

Fig 3: Dataset

Data augmentation is a set of strategies for generating "new" training images from old ones by adding random vibrations and disturbances, but simultaneously making sure that the class labels are left unaltered. The whole point of applying data augmentation is to focus on improving the model's generalization ability. Hence, an image data generator is utilized to supplement the dataset well before the experiment was carried out, which basically takes the actual data, arbitrarily modifies it, and outputs the altered data. While training the network, ImageDataGenerator guarantees that various versions of each input image are properly considered.

IV. RESULTS

Before conducting this experiment, the original dataset was enhanced using an image data generator. IDG ensures that different variations of a particular picture are factored into the equation when training the network. All inputs were then routed using the MobileNetV2 framework, which makes up the entire model. Whether or not the individual in front of the webcam was wearing a mask, the regular CNN approach was continued, commencing with function extension, reduction, visualization, and clustering. Diagnostics are completed faster in comparison to CNN. Without sacrificing quality, the model used fewer parameters. The "Relu" activation function was used in the hidden layers of CNN attributed to the prevalence of non-linear use cases (pictures). Softmax activation was used in the output nodes to detect if a person is wearing a mask as per the expected likelihood.

The statistics obtained after the research are shown in the table below.

	Precision	Recall	F1-score	Support
User with Face Mask	1	0.83	0.9	383
User without face Mask	0.85	1	0.92	384

Fig 4 : Statistics obtained

The accuracy of the aforementioned model was 96 percent. Precision, recall, f1-score, and support for the with_mask attribute are respectively 1.00, 0.83, 0.90, and 383. For the above-mentioned parameters, the obtained results for the without_mask characteristic are 0.85, 1.00, 0.92, and 384, respectively.

The following photos demonstrate how our proposed system performs. The software utilizes a real – time video to record people's faces and tell them whether or not they're wearing a mask.
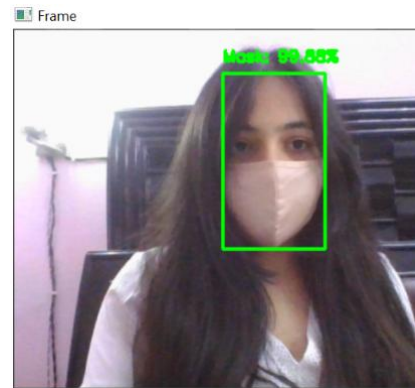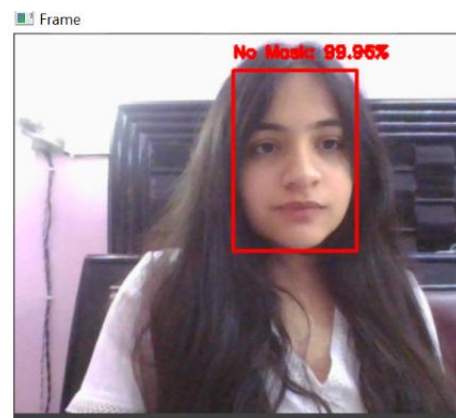

Fig 5- with mask demo

Fig 6- without mask demo

V. CONCLUSION

In this project, the training and the implementation of the proposed real-time face mask detection that was proposed due to the rise of the epidemic worldwide was made by training the deep learning model built up by utilizing image dataset that was classified into two major categories namely with facial mask and without facial mask. The widely used OpenCV library used for computer vision applications and deep neural network technique gave out positive results. The

proposed method is different and novel because MobilenetV2 image classifier was utilized to classify images, consequently resulting in better and more efficient performance. Many previous studies have had problems with their results and findings, although some of them have been able to improve the accuracy of the quality of their image dataset. As the dataset was collected from various sources and the images used in the dataset are processed to improve the accuracy of the results, the problem of many inaccurate predictions will be successfully removed from the model. In our bid to make this project a novel and innovative initiative, we have added our own personal touch to a salient part, namely the dataset. Along with the images we have obtained from various data sources that augment our results, we have also included our own pictures. Not only that but we also lay a heavy emphasis on training the entire model by first augmenting the dataset and then performing the necessary computations. Image Data Generator takes the unique data, alters it randomly, and then delivers the changed data. The application of this in the real world would be the biggest challenge out in the future. The proposed solution in the project would assist the pertinent authorities in dealing with this global epidemic appropriately.

VI. FUTURE WORK

Many previous studies had issues with their conclusions, while others were able to improve their accuracy with their dataset. Because the dataset was compiled from a number of distinct reassessments and the images used within the dataset were physically wiped clean to improve the accuracy of the conclusions, the issue of several incorrect predictions was effectively eliminated from the version. International packages are going to be a lot more difficult to come by in the future. The version should, ideally, assist the responsible government in dealing with this major pandemic problem that has spread throughout most of the globe.

We would also recommend a similar system for the future that is able to differentiate between different kinds of masks and is able to assign a percentage of "safety" accordingly. For instance, double masking is preferred over a single cloth mask as it reduces the risk of transmission. Similarly, N95 masks and KN95 masks are also shown to have better effectiveness. By putting such standards in place, we are making enclosed spaces even safer by reducing the spread of the virus - something we need dearly in order to return back to normal.

Acknowledgment

REFERENCES

[1] Suresh, K., Palangappa, M. B., &Bhuvan, S. (2021, January). Face Mask Detection by using Optimistic Convolutional Neural Network. In 2021 6th International Conference on Inventive Computation Technologies (ICICT) (pp. 1084-1089). IEEE.

[2] Sachdeva, R. (2020). Face Mask Detection System. Available at SSRN 3755508.

[3] Nagrath, P., Jain, R., Madan, A., Arora, R., Kataria, P., & Hemanth, J. (2020). SSDMNV2: A real-time DNN-based face mask detection system using single shot multibox detector and MobileNetV2. Sustainable cities and society, 102692.

[4] Kavin, M. (2021). Advanced Face Mask Detection System. International Research Journal on Advanced Science Hub, 3, 112-115.

[5] Curtis, J., Mospens, L., Nelson, D., Wilson, S., & Kilo, A. (2021). A Low-Cost Face Mask Detection System.

[6] Hussain, S., Yu, Y., Ayoub, M., Khan, A., Rehman, R., Wahid, J. A., & Hou, W. (2021). IoT and Deep Learning-Based Approach for Rapid Screening and Face Mask Detection for Infection Spread Control of COVID-19. Applied Sciences, 11(8), 3495.

[7] Mata, B. U. (2021). FACE MASK DETECTION USING CONVOLUTIONAL NEURAL NETWORK. Journal of Natural Remedies, 21(12 (1)), 14-19