Real-Time Face Mask Detection to Ensure COVID-19 Precautionary Measures

Dr.S.China Venkateswarlu[1], Dr. P. Manikandan[2], Rama Satish K V[3], P.Deepa[4], Dr. B. Swapna[5], Dr Abhijit Janardan Patankar[6]

Department of ECE, Institute of Aeronautical Engineering Dundigal Hyderabad -500043[1]

Department of CSE, Jain University (Deemed to the University), India[2]
Department of AIML, RNS Institute of technology, Bengaluru[3]

Department of CSE, Panimalar Engineering College, Chennai[4]

Department of ECE, Dr MGR Educational and Research Institute, Chennai- 95[5]

Department of Information Technology, D Y Patil college of Engineering Akurdi, Pune 411044[6]

cvenkateswarlus@gmail.com[1], mani.p.mk@gmail.com[2], ramasatish.k.v@rnsit.ac.in[3], pdeepa.3018@gmail.com[4], swapna.eee@drmgrdu.ac.in[5], ajpatankar@dypcoeakurdi.ac.in[6]

Abstract:

The rapid spread of COVID-19 (Coronavirus 2019) around the globe has just brought about a serious public health emergency. The World Health Organization (WHO) has distributed a number of different guidelines in an effort to limit the spread of COVID-19. If a someone is concerned about getting COVID-19, the World Health Organization advises that they wear a mask whenever they are in a public or crowded place. This recommendation applies to both adults and children. Simply looking at someone makes it hard to tell if they are concealing their identity with a mask. In this study, we conduct a comprehensive analysis of the data that was collected and the performance is being measured with deep learning architectures.In this research work, each and every one of the prerequisites for such a model was investigated. The suggested approach uses a deep learning technique - CNN in order to differentiate between labels in an image that have masks and labels that do not have masks. The results of the experiments show that the proposed system achieves 99.77% accuracy on the benchmark datasets, exceeding previous systems and datasets that are considered state-of-the-art in a real-time setting.

Keywords: *Covid -19,Face Mask Detection, Deep Learning Technique - CNN, Accuracy.*

DOI Number: 10.14704/nq.2022.20.9.NQ44568 **Neuro Quantology 2022; 20(9):4900-4906**

I.INTRODUCTION

Coronaviruses (CoV) are a large family of viruses that are responsible for a wide variety of illnesses, ranging from the common cold to potentially lethal respiratory infections such as MERS and SARS [1]. Coronaviruses cause the common cold. According to the World Health Organization (WHO), as of January 2021, there were a total of 90,054,813 confirmed cases of COVID-19 across the globe, with 1,945,610 fatalities. The COVID-19 virus can spread either directly from one person to another or indirectly by contact with contaminated settings. Both modes of transmission are possible. This takes place when the size of pulmonary droplets is anywhere between 5 and 10 micrometers (mm). When a susceptible person comes into close proximity (one meter) with an infected person who is coughing, sneezing, or has other respiratory symptoms and is therefore at risk of being exposed to potentially infectious respiratory droplets through his or her own mouth, nose, or eyes, droplet transmission typically occurs [3]. Droplet transmission can also occur when an infected person is vomiting or diarrhoea [4]. Because an increasing number of people are falling ill with the virus and dying as a result of it, it is very essential that we do everything in our ability to limit it.

In addition to wearing a mask, avoiding direct contact with infected individuals, and often washing one's hands with soap and water for at least 20 seconds at a time, social isolation is one of the most effective methods for preventing the spread of the COVID-19 virus. Other methods include: "Social distancing" refers to the practice of maintaining a healthy distance from one's immediate surroundings [4]. The minimal distance that is recommended to maintain in order to prevent the transmission of the virus by direct physical contact is two meters (six feet) [5,6]. Children under the age of 20 saw an 85% drop in contact as a result of adhering to the social distance rules, but adults over the age of 60 experienced a reduction of 95% in contact [7]. This displays the possibility of "flattening the curve" if the appropriate spacing standards were followed.

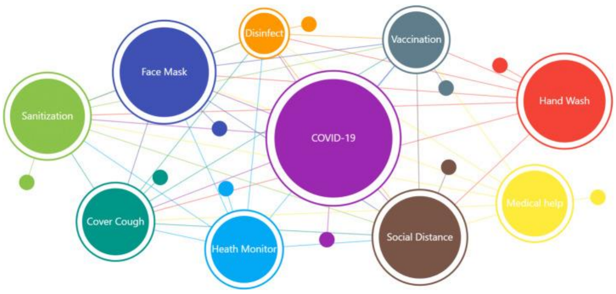

Figure 1. Covid – 19 Safety Measures shows importance of Face – Mask.

In addition to this, the scientific community all around the world has to collaborate in order to stop the spread of the coronavirus pandemic. utilising the most cutting-edge technology available today in terms of both machinery and equipment. There is a wide variety of possible courses of action that can be taken to control the infection. The authors of [8] provide a unique architecture for mask detection that is able to attain an impressively high level of accuracy of 97%. In [5], the face mask detection functionality of PyTorch achieved an accuracy of 97%. In addition, [8-9] presented a cutting-edge technique for discovering the several sorts of masks that are available. We were able to ascertain the conclusion by putting the model through its paces in real time. CNN-based detectors have been trained using face mask datasets that were expressly acquired for the purpose of that paper [10]. Additional study was carried out in order to develop a mobile app that would detect individuals who were out in public masking their identities behind masks [11]. Additionally, in [12], in order to make the already existing dataset more useful, fresh photographs were added to it. The job was finished successfully by utilising the Faster R-CNN model, which produced results that were accurate 99.8 percent of the time. [13] presents a method for checking the correct location of a mask that can be used. In [12], the authors describe the various technical avenues that might be pursued in order to fight the virus.

As a result of the ongoing development of technology, a number of approaches [14 -15] have the potential to be of use to society if they are utilised in an effective manner. A system that operates in real time and is able to place a person who has been watched into one of two categories [17]:

- An individual who conceals their true identity by wearing a mask.
- It's possible that showing your face in this day and age will be to your advantage.

The implementation of such systems could be beneficial to a wide range of public establishments, including hospitals, airports, shopping centers, and other similar locations amongst others. The development of the detector requires a vital phase that involves face recognition in real time. After they have been recognized, the faces that appear in the live feed of a camera are saved as still images and then fed into a classifier. In the following paragraph, we will talk about the several classification algorithms that can be used when putting things into categories. Along the same lines, one may make use of an object detection model. The following findings are based on the most recent advancements in the relevant field:

- The most effective method for preventing the spread of covid-19 is to stay away from other individuals and to always keep your face covered. Because of this, the academic community required a significant amount of knowledge regarding masking strategies.
- A review paper will compare and discuss the performance parameters of a number of

different procedures in order to enhance the usefulness of the document.

II. LITERATURE WORK

Mohamed Loey and his colleagues [10] built a unique model by combining deep transfer learning (ResNet-50) with more traditional machine learning approaches. We eliminated the topmost layer of ResNet-50 and replaced it with three traditional machine learning classifiers (Support vector machine [SVM]), a decision tree, and an ensemble in order to improve the overall performance of the model. One of the four different datasets that were evaluated contained the biggest amount of pictures, including those of persons with and without face masks. As a result, training on this dataset took the longest. Even works that are conceptually comparable to this one fail to remark on the dependability of this kind of data set. [Citation needed] The accuracy of the decision trees classifier was only 68% when applied to the task of categorising false face masks against the training dataset, which consisted of photographs of real faces.

Resnet serves as the backbone of a detection network, while FPN (feature pyramid network) serves as the neck, and a variety of classifiers, predictors, estimators, and other such components serve as the heads [11]. However, because the face mask dataset is so small, learning algorithms have a difficult time understanding how to construct better features. This is because of the limitations of the dataset. There has not been a sufficient amount of effort put into enhancing the reliability of face mask detection.

Rinkal Keniya was the one who came up with the idea for SocialdistancingNet-19 [12], which has the capability of locating a person's frame and showing labels that can help assess whether or not they are safe if the distance between them and the danger is less than a certain threshold. Webcam detection relies on the fact that people are always going about their business. This outcome is conceivably a consequence of the way users are located on the network. First, we pinpoint the overall location of the scene, and then, based on where the individuals' centres are located, we calculate their distances from one another.

Shashi Yadav presented Single Shot Object Detection (SSD) in OpenCV as a deep learning solution for detecting social distance and masks [13]. SSD is an acronym for "single shot object detection."

The use of this approach had some drawbacks, one of which was that it wrongly labelled as "masked" individuals whose hands or other objects hid their faces. This model does not account for such a range of possibilities. In this particular configuration, however, an SSD can only identify a single human being even if it is capable of detecting many objects inside a single frame.

Retina Facemask is the name of the face mask detector that was developed by Mingijie Jiang and his colleagues [15]. In their strategy to filter out forecasts with low levels of confidence, they used an innovative method for removing items from consideration. They were able to obtain 5.9% and 11.0% higher accuracy than the findings of the most recent research in the field of face detection, as well as 1.5% and 2.3% higher accuracy in the field of face mask detection, respectively. However, they studied the effectiveness of the proposed technique when applied to lightweight neural networks such as Mobile Net. Shashi Yadav et al. [16] proposed using a hybrid method that combines deep learning with a geometrical approach in order to assess social distance and mask use in public places. They utilized a Raspberry pi4 in order to put their theory to the test. Their system uses data from cameras to identify individuals who breach social standards about personal space and the use of garments or masks that conceal one's identity. The system will send a notification to the command center as well as activate an alert whenever a rule is broken in a particular part of the facility. A method known as transfer learning allows for the automatic identification of individuals who are not concealing their faces behind masks. This method is explained in the study [17], which may be found here. This strategy makes use of versions of InceptionV3 that have been trained and tweaked to perfection. Training and testing were done with the help of the simulated masked face dataset (SMFD). In addition, they utilized picture augmentation strategies to increase the total number of training instances, which allowed for a more precise analysis of the results.

III. PROPOSED LEARNING TECHNIQUES

Data Set:

There were numerous face mask detection datasets available on Kaggle. The three groups in one dataset—"with mask," "without mask," and "wearing mask wrongly"—made it stand out. The dataset was highly unclean and skewed, unfortunately. In order to enhance this dataset, it was necessary to add images to ensure that there is an equal number of images in each class and to remove noisy images that could be seen as outliers. My resulting dataset is a hybrid of that plus another dataset that I cleaned up and partitioned equally amongst the various classes.

The data set is divided into three categories, each of which has its own folder. There are three groups: "withmask," "withoutmask," and "maskwearedincorrect." A total of 2,994 images of students in that class are included in each folder.

Length of dataset = 853

Training set = 597

Test set = 256

Convolution Neural Network:

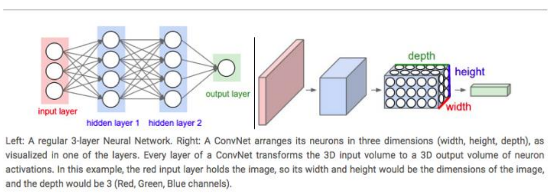

Figure 2.Architecture of CNN for Covid – 19 Face – Mask.

This system is also referred by "CNN," which stands for "convolutional neural network." This type of the neural network is optimized for dealing with matrices of information in two dimensions, such as those observed in images. Examples of such matrices can be found. CNNs are utilised for the most part to identify images and find them in their respective locations.

Data Explorations:

The.jpg files for each dataset are kept in two different folders, each of which is named after the model of the images it contains. Authors use the ***ImageDataBunch.from*** folder() function to load the images and label them based on the name of the folder.

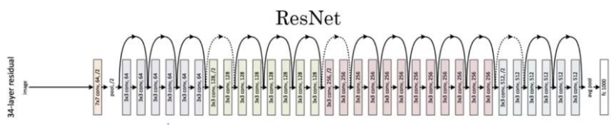

Figure 3.ResNet for Covid – 19 Face – Mask.

- Utilize a model of a ResNet50 Convolutional Neural Net that has already been trained and apply transfer learning to it so that you can train only the weights of the last layer.
- What exactly is the purpose of developing one's ability to transfer? Because in the process of transfer learning, an already-trained neural network that can recognise images is used instead of creating a new network from scratch. After that, you make a few adjustments here and there, some minor and others important, to make the model training fit your requirements. Why exactly do we act in this manner? To train a neural network to an acceptable level, approximately 300,000 photographs are required; however, producing truly spectacular results requires at least one million images.
- Our training set consists of more than 4,000 individual images. Do you believe that this will be sufficient to get a neural network up and operating from the ground up?
- Authorsutilize the construct cnn() method to employ a ResNet18 network that has already been pre-trained. The ImageNet collection was used to teach this network, which resulted in it having access to approximately one million images.

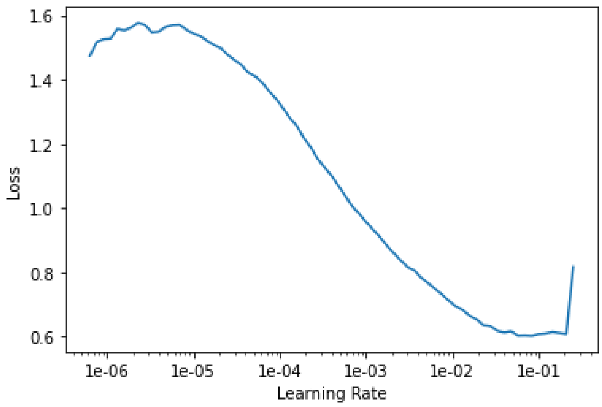

Figure 4. Shows the Learning Rate of CNN

The training model and no. of epochs used in CNN is shown below in table 1.

Table 1. Number. of Epochs used to train in CNN

epoch	train_loss	valid_loss	accuracy	time
0	0.030209	0.001478	0.999338	01:20
1	0.067989	0.232626	0.977483	01:22
2	0.049258	0.01232	0.994702	01:21
3	0.043364	0.010096	0.996689	01:23
4	0.031994	0.014254	0.995364	01:22
5	0.027822	0.000564	1	01:22
6	0.018827	0.451038	0.998013	01:21
7	0.012236	2.847701	0.996027	01:23
8	0.006585	0.002626	0.998676	01:22
9	0.00458	0.005462	0.997351	01:02

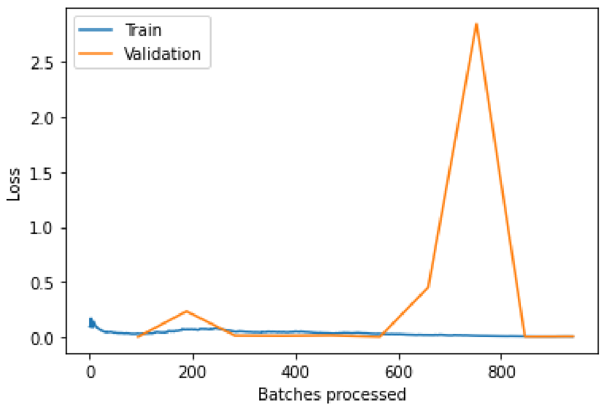

Figure 5. Shows the Train and Validation of CNN.

It is possible to assess the usefulness of a model in a number of different ways. The "confusion matrix" is one approach that might be used in this direction. Along the diagonal of the matrix is where you'll see the total number of correct predictions for each category. The extent to which inaccurate estimates can be made is illustrated by the values in other cells in the table.

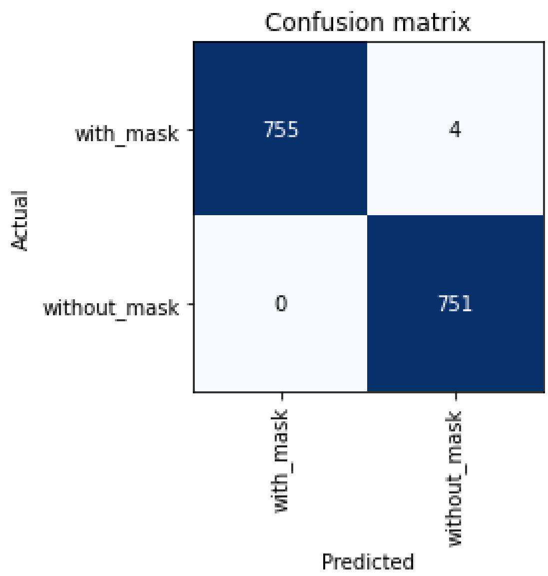

Figure 6. Shows the Confusion Matrix of CNN.

Figure 7. Shows the Prediction of Without and With Mask of CNN.

Once the model is trained, it is time to put it to use. The architecture of the model as well as the training parameters need to be stored before it can be deployed in production. This is performed through the use of the export method. The exported model is kept in a PKL file, which is a specific kind of file that pickle generates (a Python module).

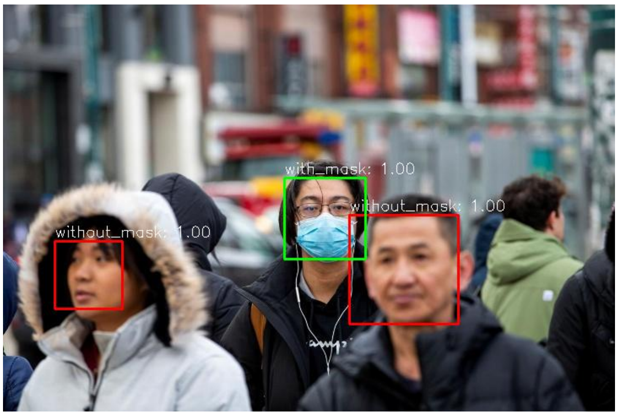

Figure 8. Shows the Result of Without and With Mask of CNN.

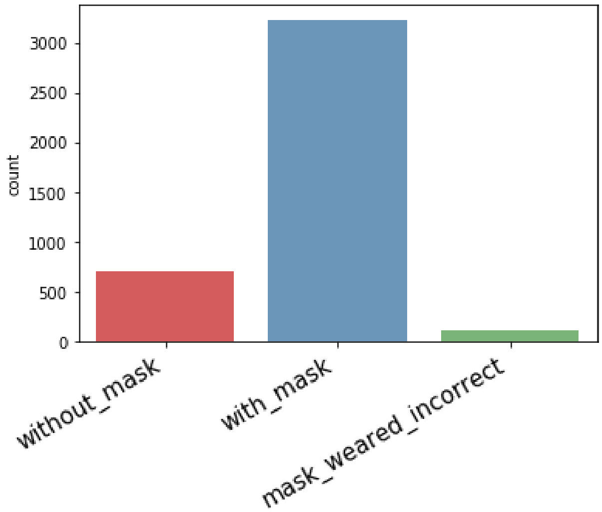

Figure 9. Shows the Result of Without and With Mask of CNN in Graphical Method.

IV. CONCLUSION

This research examines the use of deep learning strategies to predict a person wears mask or not. Because of COVID-19, the current situation of health all around the world is a disaster. Several countries are cooperating closely in an attempt to halt or at least significantly delay the spread of COVID-19. According to both World meter [8] and Google News COVID-19 [7], places with a high population density are more likely to see a rapid spread of the virus. Multiple studies have shown that wearing face masks in public places is an effective way to stop the spread of the COVID-19 virus. Because of this, the majority of countries require people to wear masks in public places. In this kind of environment, it is nearly hard to keep an eye on other people. For the purpose of this investigation, we devised a simplified CNN model to recognize persons without requiring them to conceal their identities with disguises. During testing, the proposed model attained an accuracy level that was an average of 99.77%. Utilizing a substantial quantity of data in future iterations of the proposed study is one way in which the study might be made more effective. It is possible to further develop it to identify people hiding behind masks by identifying the sorts of masks that they wear and building a facial recognition system that can be used in a range of scenarios. This may be accomplished by classifying the masks that the individuals wear.

REFERENCES

[1] Adarsh P, Rathi P, Kumar M. 2020 6th international conference on advanced computing and communication systems (ICACCS) 2020. YOLO v3-tiny: object detection and recognition using one stage improved model.

[2] Albawi S, Mohammed TA, Al-Zawi S. 2017 international conference on engineering and technology (ICET) 2017. Understanding of a convolutional neural network.

[3] Ardabili S, Mosavi A, Várkonyi-Kóczy AR. Advances in machine learning modeling reviewing hybrid and ensemble methods. In: Várkonyi-Kóczy A, editor. Engineering for sustainable future. INTER-ACADEMIA 2019. Lecture notes in networks and systems. Springer; 2020.

[4] Chandan G, Jain A, Jain H, Mohana (2018) Real time object detection and tracking using deep learning and OpenCV. In:2018 international conference on inventive research in computing applications (ICIRCA). 10.1109/ICIRCA.2018.8597266

[5] Ghasemi E, Gholizadeh H, Adoko AC. Evaluation of rockburst occurrence and intensity in underground structures using decision tree approach. *Eng Comput.* 2020;36(2):213–225. doi: 10.1007/s00366-018-00695-9.

[6] Jean-Daniel Fekete, “Infrastructure,” in Mastering The Information Age - Solving Problems with Visual Analytics, Daniel Keim and Jrn Kohlhammer and Geoffrey Ellis and Florian Mansmann, Ed. Eurographics Assoc., 2010, ch. 6, pp. 87–108.

[7] M. Beaudouin-Lafon, S. Huot, M. Nancel, W. Mackay, E. Pietriga, R. Primet, J. Wagner, O. Chapuis, C. Pillias, J. R. Eagan, T. Gjerlufsen, and C. Klokmose, "Multisurface interaction in the wild room," Computer, vol. 45, no. 4, pp. 48 –56, april 2012.

[8] J. Talbot, Z. DeVito, and P. Hanrahan, "Riposte: a trace-driven compiler and parallel vm for vector code in r," in Proceedings of the 21st international conference on Parallel architectures and compilation techniques (PACT '12). ACM, 2012, pp. 43–52.

[9] Shashi Yadav,Goel Institute of Technology and Management, Dr. A.P.J. Abdul Kalam Technical University ,Deep Learning based Safe Social Distancing and Face Mask Detection in Public Areas for COVID-19 Safety Guidelines Adherence (2020).

[10] Rinkal Keniya · Ninad Mehendale, Real-time social distancing detector using Socialdistancing-Net19 deep learning network(2020).

[11] Indhu Jain, Mr. Sudhir Goswami; A Comparative Study of Various Image Restoration techniques with different types of blur, International Journal Of Research In Computer Applications And Robotics (2015).

[12] Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun; Deep Residual Learning for Image Recognition (2015).

[13] Joseph Redmon, Ali Farhadi from University of Washington ;YOLOv3: An Incremental Improvement.(2018).

[14] "MobileNetV2: Inverted Residuals and Linear Bottlenecks", The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018.

[15] Bastanfard A, Takahashi H, Nakajima M (2004) Toward E-appearance of human face and hair by age, expression and rejuvenation. In: 2004 *International* conference *on Cyberworlds*. IEEE, pp 306–311.

[16] Bastanfard A, Bastanfard O, Takahashi H, Nakajima M (2004) Toward anthropometrics simulation of face rejuvenation and skin cosmetic. Comput Animation Virt Worlds 15(3–4):347–352.

[17] Ansari M, Singh DK (2021) Human detection techniques for real time surveillance: A comprehensive survey. Multimed Tools Appl 80(6):8759–8808.