Proceedings of the Sixth International Conference on Inventive Computation Technologies [ICICT 2021]
IEEE Xplore Part Number: CFP21F70-ART; ISBN: 978-1-7281-8501-9

Face Mask Detection by using Optimistic Convolutional Neural Network

Suresh K[1], Palangappa MB[2], Bhuvan S[3]

[1,2,3] Department of Computer Science, Amrita School of Arts and Sciences, Mysore, Amrita Vishwa Vidyapeetham, India
suresh.kalaimani@gmail.com, palangappa456@gmail.com, bvsm9826@gmail.com

***Abstract*—COVID-19 pandemic has rapidly increased health crises globally and is affecting our day-to-day lifestyle. A motive for survival recommendations is to wear a safe facemask, stay protected against the transmission of coronavirus. By wearing a facemask, the most effective preventive care must be taken against COVID-19. Monitoring manually if the individuals are wearing facemask correctly and to notify the victim in public and crowded areas is a difficult task. This paper approaches a simplified way to achieve facemask detection and notifying the individual if not wearing facemask. Using Kaggle datasets, the proposed system/model is trained and examined. The system runs in real-time and detects if an individual face has facemask if not then notify the individual personally through text message. The mask is extracted from real-time faces in public and is fed as an input into convolutional neural network (CNN).**

Keywords— Face Mask Detection, Convolutional Neural Networks (CNNs), Kaggle Datasets, Public Safety, COVID-19, OpenCV.

I. INTRODUCTION

In 2020, COVID-19 being a pandemic disease had rapid spread of virus over the world creating a red alert in global health, humanity and everyday lifestyle of humans and daily lives had deep impact. All around the globe, illness on respiratory such as Severe Acute Respiratory Syndrome (SARS), and Middle East Respiratory Syndrome (MERS) are spreading, on December 2019 in Wuhan, China a new critical respiratory illness has arisen and has infected millions of peoples and lost millions of helpless lives in more than 200 countries according to world health organization (WHO) and declared global pandemic. The spread of the virus is through close contact with the people and in overcrowded public areas [1]. The guidelines listed by the WHO, primary precaution should be taken to prevent the spread of virus is to wear facemask and maintain social distance. An efficient based computer vision approach aims on real-time application which monitors the individuals in public whether wearing facemask or not and safe social distancing by implementing raspberry pi4 model to the monitor and spot violation through public cameras.

Modern deep learning algorithms are mixed with techniques of geometry to build robust models which are able to cover aspects such as, detection, tracking, and validation [2]. This paper addresses building a hybrid deep and classical machine learning model to detect facemask. The model consists of two elements. The first element is designed using Resnet50 for feature extraction, and the second element is built using Support Vector Machine (SVM), decision tree, and ensemble algorithm for facemask classification process [3]. An architecture is trained efficiently using deep learning on a dataset that has images of individual's faces with and without facemask. The system architecture prevents spreading of virus by finding out individuals who are not wearing facemask in smart cities by monitoring with public Closed-Circuit-Television (CCTV) cameras. An individual is spotted if not wearing a facial mask and reported corresponding authority [4].

COVID-19 disease 2019 has large scale negative impact on health across the globe. One major form of defense against viruses is to wear face masks in public places. Retinaace Mask detector, a high accuracy facemask detector. The proposed detector is a one-stage detector that has a pyramid network feature used to fuse multiple feature maps with high-level semantic information, and a novel-context attention module is used for facemask detection. With low anonymity and strong union intersection estimates are refused using the proposed novel cross-class object elimination algorithm [5]. COVID-19 pandemic gave a breakout in global health, occurs badly need of facemask has a basic protection mechanism to break the chain of spreading of coronavirus. Aims to detect facemask on human face on live streaming video and also with human face images. Face detector model is built using deep learning and Single Shot Detector(SSD).

The presence of facemask in image or video stream is done using basic concepts of transfer learning in neural networks [6]. Coronavirus has rapidly affected day-to-day life in turn disturbing the world trade and movements. A simplified method detects the face from the image and identifies if it has facemask on it or not using basic machine learning packages such as TensorFlow, Keras, OpenCV and Scikit-Learn. The model is able to detect facemask even if the face is in motion [7]. It is a difficult task to monitor if the people are wearing masks manually in public and crowded areas. Using the pre-trained state-of-art deep learning architecture, InceptionV3, an automated facemask detection model is developed using fine-tuning methodology. Model is skilled/trained using a Simulated Masked Face Dataset (SMFD). Here, on the public face dataset mask is put and then it's simulated. This is used to better training and testing of the model. The image augmentation technique is used for improved training and testing of the model for restricted data usability [8].

2021 6th International Conference on Inventive Computation Technologies (ICICT) | 978-1-7281-8501-9/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICICT50816.2021.9358653

In this paper, we proposed to build a real-time facemask detection model using Convolution Neural Network (CNN) which is a class of Deep Neural Network (DNN), most commonly used in image classification and recognition. The proposed model can be implanted in surveillance cameras in organizations, schools, universities, shopping malls, multiplex etc. which helps to monitor individuals automatically whether they are wearing facemask, if not, spot them and report to higher authorities as well as notify them personally through text. This model helps to break the chain of spreading of virus when in close contact and reduces the positive cases which are rapidly increasing day-by-day and the rate of losing helpless lives can be controlled.

II. RELATED WORK

The facemask detection model has become a most required and essential model during COVID-19 pandemic. Since manually monitoring whether people are wearing facemask in public and crowded areas, [1] research paper had built real-time automated model integrated with surveillance cameras in public which detects whether people are wearing facemask and maintaining social distance in public areas and report to the respective authorities using computer vision and implementing raspberry pi4. Global health crises have occurred due to coronavirus, facial masks form a basic prevention from the virus, hence, [2] hybrid model is built using classical and deep machine learning consisting of two components. The first component is for feature extraction is by using Resnet50 and the second is for classification processing of mask is using ensemble, support vector machine (SVM) algorithms and decision tree. Uses three datasets after investigation. Simulated Masked Face Dataset (SMFD) is the first dataset , the second dataset is Labeled Faces in the Wild (LFW) and third one is the Real-World Masked Face Dataset (RMFD). SVM learning algorithm achieved 99.49% accuracy in SMFD. RMFD achieved 99.64% of accuracy, LFW achieved 100% of testing accuracy. Healthcare system is under crisis. List of precautionary measures is being taken care in order to reduce the spread of viruses in which wearing facemasks is one of them. [3] A system is created to find people not wearing facemasks in smart cities using Closed-Circuit-Television (CCTV) cameras.

The trained system achieved 98.7% accuracy in differentiating people with mask and without mask. [4] An efficient and high-accuracy detector called RetinaFaceMask detector is built to spot whether people are wearing facemasks. The framework is a one-stage detector with a new background attention module to concentrate on face mask identification and a pyramid network feature to combine high-level semantic data with several feature maps. An approach to the algorithm for the elimination of new background attention module artifacts to delete projections of high union intersections and poor confidences. Results of RetinaFaceMask achieves state-of-the-art results on facemask dataset with 2.3% and 1.5% higher than the standard result and mask detection precision, with 11.0% and 5.9% higher than the standard results. An architecture is built [5] to find whether people are wearing facemask in live streaming videos and even with human face images using Single Shot Detector (SSD) serves the purpose of object detection. Concepts of transfer learning in neural networks used in finding presence or absence of facemask in video streams and in images. Experimental findings indicate that the model performs well with 100% accuracy and 99% precision and recall, respectively. Simplified approach [6] towards detecting facial masks even in motion using basic machine learning packages such as Tensorlow, Keras, OpenCV, and Scikit-learn. The method attains accuracy up to 95.77% and 94.58% respectively on two different datasets. An automated process [7] for finding whether individuals wear facemask in public. The model is built by fine-tuning the pre-trained state-of-the-art deep learning models called InceptionV3. Simulated Face Mask Dataset (SMFD) dataset is used to train the dataset. Here, on the public face dataset mask is put and then it's simulated. This is used to better training and testing of the model. The image augmentation technique is used for improved training and testing of the model for restricted data usability The model reaches 99.9% precision during training and 100% accuracy during testing.

An efficient real-time system approach towards a computer vision based to detect both violation of wearing facemask and social distancing in public areas using the convergence of advanced deep learning algorithms with geometric techniques resulted in model creation which is robust in nature and covers the aspects of validation, detecting, tracking. Rasberry-pi4 is implanted in public surveillance cameras and the built robust model was allowed to run on raspberry-pi4. Paper uses lightweight neural network MobileNetV2 to analyse Real-Time Streaming Protocol (RSTP) video stream using OpenCV and transfer learning techniques with Single Shot Detector (SSD) used to achieve resource limitation and accuracy recognition in monitoring real-time video surveillance cameras in public areas to spot out violation of wearing facemask and maintaining social distancing.

Hybrid models are built in detection of facemask using classical machine learning and deep learning. The system has two components, one component is used for extraction of features using Resnet50 and the other component is used to classify facemask using Support Vector Machine (SVM), ensemble algorithm and decision tree. The Real World Masked Face Dataset (RMFD), the Labelled Faces in the Wild (LFW) and Simulated Masked Face Dataset (SMFD) are the three datasets used for experiment of considered three algorithms. SVM algorithm resulted more efficient than the other three algorithms with 99.64% accuracy in RMFD, 99.49% accuracy in SMFD and reached 100% accuracy in LFW.

A system is proposed to find absence of facial mask with the people in public areas in smart cities by monitoring with Closed-Circuit-Television (CCTV) cameras. A person without

a facial mask is detected and notified to higher authorities through the city network. The architecture is developed using Convolutional Neural Network (CNN) which helps for feature extraction from the dataset images as well the images captured by cameras in real-time. There is 98% of accuracy result by the trained architecture.

III. PROPOSED WORK

In this research work, we proposed an Optimistic Convolution Network that helps to ensure whether in public the people are wearing masks or not by monitoring automatically. Here in Fig 1 we have described an architecture that shows how our system functions automatically to prevent the spread of COVID19.

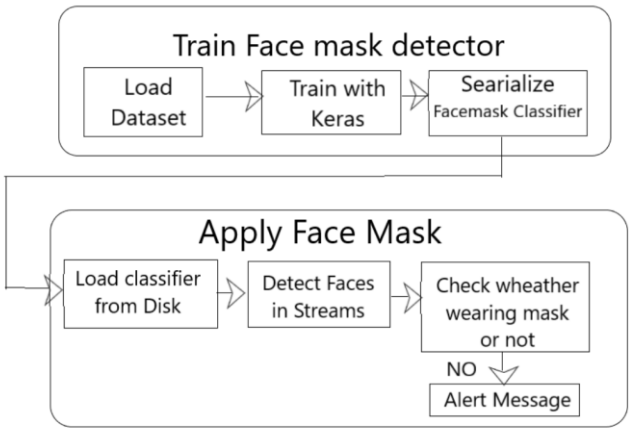

Fig. 1. Architecture of proposed system

Our systemuses the TensorFlow and Keras algorithm to detect whether an individual is wearing a face mask along with the Convolutional Neural network model. Here we first train the system with the Dataset from Kaggle and train it with Keras and TensorFlow, once the training is done then we will load face mask classifier from the disk, here faces are detected from a real time video stream. This process also involves use of MobileNet in order to train a huge collection of images and classification of high-quality images.

Here image dataset is loaded from Keras and then the images are converted into an array, later MobileNet is used to preprocess input image and to append image to the data list. In the proposed system the main contribution includes person face identification and face mask detection. These both are done in Real Time with the help of MobileNet and OpenCV. A square box is been displayed on every person's face with the color of red and green where red indicates the person is not wearing a mask and green indicates a person is wearing a mask.

We have used a face cropped dataset from Kaggle of about 3918 images of persons with masks and without masks. These images are used in order to train the model that classifies into two categories: that is, faces with masks and faces without masks. These datasets are then converted into arrays in order to create a Deep Learning Model. The result of the person from the video displays a person with a square bound box. This system monitors continuously, and whenever a person is identified without a mask then the person's face is been captured and then it is sent to the higher authorities, also to that person. Due to the outbreak of novel CoronaVirus this proposed model can be implemented in public at real-time for monitoring the people wearing face masks. Our model can be used for monitoring automatically in public places that would help for those who monitor people physically/manually, that is the reason we picked this architecture. Our system can be used in airports, schools, railway stations, shopping malls, offices and other public areas to make sure that in-public people are wearing masks.

A Convolutional Neural Network

CNN plays a significant part in computer vision related examples in recognizing patterns, on account of its less computation cost and also the ability of spatial extraction. CNN utilizes convolution portions to combine with the primary images in order to remove top-level features. The commencement network that is proposed in [9] permits the network to get familiar with the mix of kernels. Planning to build a good Convolutional Neural Network architecture actually remains as a primary inquiry. To prepare a lot further neural network, K. He et al. proposed Residual Network (ResNet) [10] that can take in personality planning from the past layer. As article locators are generally conveyed on portable or any embedded device, where the computing assets are extremely restricted, Mobile Network (MobileNet) [11] is proposed. This utilizes profundity shrewd convolution to remove highlights and channelised convolutions to change channel numbers, so that the computational expense of the MobileNet is a lot lower compared to networks utilizing standard convolutions. In Fig 2 we have shown a Schematic Diagram for Basic Convolution Neural Network.

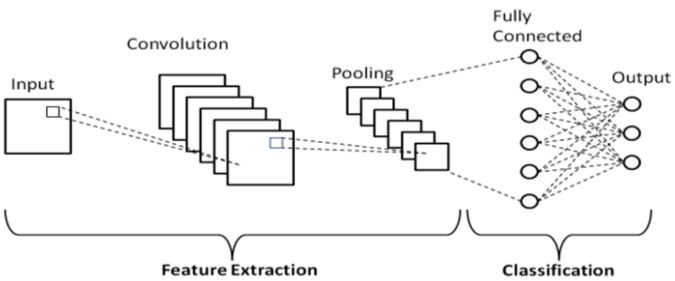

Fig. 2. Schematic Diagram for Basic Convolution Neural Network

B Dataset

The dataset consists of 3918 images, that is divided into two parts, faces with mask and faces without masks (Fig 3). These datasets are taken from Kaggle, face without masks includes faces with various skin colors, different angles, occlusion, etc. Faces with masks includes mask with hand, with masks and other objects that cover the face, that provides us an advantage to improve variants of the dataset.

IV. MATHEMATICAL NOTATION

FaceMask gives two outputs for the given input image that is localization offset prediction and classification prediction. $Ybloc \in Ra\times4$, $Ybc \in Ra\times b$, here a,b denotes number of anchors generated and number of classes. Also here we have $D \in Ra\times4$ default anchors, $Yloc \in Ro\times4$ the ground truth boxes, and $Yc \in Ro\times1$ classification label, here o refers to the number of objects.

At first default anchor D is being matched with the *Yloc* ground truth boxes and *Yb* classification label in order to obtain $Pml \in Ra\times4$ and $Pmc \in Ra\times1$, in which each row of *Pml* represents the offsets and *Pmc* represents top classification label, respectively.

We described positive localization prediction $Yc + loc \in Ra+\times4$ and positive matched default anchors $D + ml \in Ra+\times1$, here $a+$ indicates the number of default anchors. After that smooth loss is computed between $Yc + loc$ and $D + ml$, $Lloc(Yc + loc, D + ml)$.

Later hard negative mining, predictive anchors and the sampling negative is performed, $D- mb \in Ra-\times1$ and $Yc - b \in Ra-\times1$, here $a-$ indicates the number of negative sampling anchors. Then we calculated the confidence loss by $Lconf(Yc - b , D - mb) + Lconf(Yc + b , D + mb)$.

Hence , the total loss function is

$$L = 1 Z (Lconf(Yc - b , D - mb) + Lconf(Yc + b , D + mb) + aLloc(Yc + loc, D + ml)),$$

here *Z* is the matched default anchor.

Algorithm : Object removal
Require: choose face: D0 s, C 0 s; choose mask D0 n, C 0 n
for ps in face detection D0 s **do**
for pm in mask detection D0 n **do**
if IoU(ps, pn) > thresh then
remove objects that are lower confidence
end if
end for
end for

V. EXPERIMENTAL RESULTS AND ANALYSIS

Our system proposed is an Optimistic Convolutional Neural Network which is used to detect face masks in real time. In the proposed system we have followed these four steps:

- Collection of Data and pre-processing
- Building and training the Model
- Model Testing
- Implementing the Model

A Collection of Data and pre-processing:

The proposed system used face cropped data containing images with different angles and different poses of face with and without masks that are labelled and is used to train our model. The real time automated face mask detection has been done by MobileNet and OpenCV. The dataset consists of 3918 images that are used to train our proposed model. The data are divided into two different categories: Faces with mask and without mask. Faces with masks includes mask with hand, with masks and other objects that cover the face, that provides us an advantage to improve variants of the dataset.

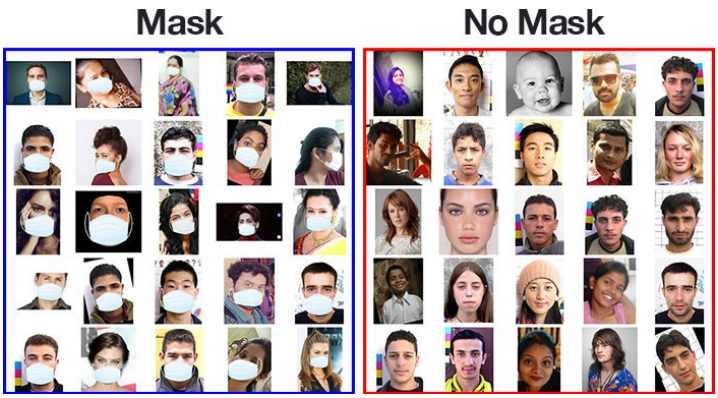

Fig. 3. Some Sample Images from the dataset

B Building and Training the Model:

In the proposed system the custom dataset is loaded and the algorithm is being trained based on labelled images. In this step the images are resized and its been converted into numpy array format. This model uses MobileNet performing as a backbone and train the model using TenserFlow. In Fig 4 we have shown Model training accuracy/loss curves. Parameters with a learning rate (initial) of INIT_LR = 1e-4, batch size BS = 32 and the number of epoch EPOCHS = 20. For the model webcam is used for face mask detection and once the person is found we mark the person with the square bounded box.

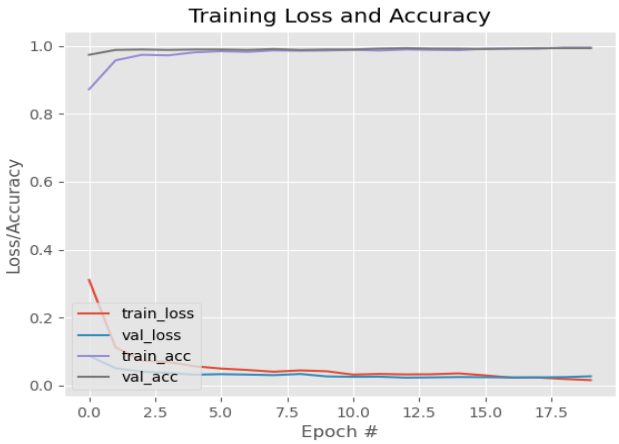

Fig. 4. Training Model accuracy and loss curves

C Model Testing:

Our system works in a automated manner that helps identify a person if he or she is wearing or not wearing a mask and notify the person's image to the person as well as to the higher authorities. Once our model is trained completely with the provided dataset we test by showing a bounded box with the confidence score on top of the bonded box. From the camera our proposed system identifies all the persons face with a green and red bounding box (Fig 5) that identifies whether a person is wearing a mask or not. If any of the people are not wearing the mask then the system will capture that person's image and send it to the victim as well as the higher authorities.

Fig. 5. Test result of system/model

D Implementing the Model:

Our system uses a custom dataset with the input video taken from any camera device. The system feeds with a real time video in public places which automatically monitors and detects whether or not people are wearing face masks. Whenever a person is found without wearing a mask then his/her photo is captured, then it is been sent to the higher officials/authorities as well the victim so that they can take any further actions.

VI. CONCLUSION

In our research we have proposed a system that automatically identifies whether or not a person is wearing a face mask and notify the higher authorities if not wearing a mask. This proposed system uses Computer Vision and MobileNet to help the public ensure that they are wearing face masks and to keep away from the spread of COVID-19 virus. Our research also helps police or higher authorities that makes it easier to identify whether a person is wearing a mask, if not then they will be also having the victim's photo by which they can take further actions. The proposed system can be implemented in places like railway stations, shopping malls, offices, schools, airports, etc.

VII. FUTURE WORK

There are many more different cases in which this model can be integrated for the safety of the public:

- Identify a person if he is doing any crime by wearing face mask.
- Identify what type of mask is the person wearing.
- Coughing and Sneezing Detection.
- Temperature Screening

REFERENCES

[1] Yadav S. "Deep learning based safe social distancing and facemask detection in public areas for COVID-19 safety guidelines adherence." *Int J Res Appl Sci Eng Technol*. 2020;8(7):1368-1375.

[2] Loey M, Manogaran G, Taha MHN, Khalifa NEM. "A hybrid deep transfer learning model with machine learning methods for face mask detection in the era of the COVID-19 pandemic." *Measurement (Lond)*. 2021;167:108288.

[3] Suresh. K and Pattabiraman. V, "An improved utility itemsets mining with respect to positive and negative values using mathematical model", International Journal of Pure and Applied Mathematics, Volume-101 No-5, Page No-763772, 2015.

[4] Suresh. K and Pattabiraman. V, "Developing a customer model for targeted marketing using association graph mining", International Journal of Recent Technology and Engineering, Volume 8, Issue 2S4, July 2019, Pages 292-296. DOI:10.35940/ijrte.B1055.0782S419

[5] K. Suresh and O. Praveen, "Extracting of Patterns Using Mining Methods Over Damped Window," 2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA), Coimbatore, India, 2020, pp. 235-241, doi: 10.1109/ICIRCA48905.2020.9182893.

[7] S, Akshay & Bhat, Mandara & Rao, Aishwarya. (2019). "Facial Expression Recognition using Compressed Images." International Journal of Recent Technology and Engineering. DOI:10.35940/ijrte.B1041.078219.

[8] S. Akshay and P. Apoorva, "Segmentation and classification of FMM compressed retinal images using watershed and canny segmentation and support vector machine," 2017 International Conference on Communication and Signal Processing (ICCSP), Chennai, 2017, pp. 1035-1039, doi:10.1109/ICCSP.2017.8286531.

[9] Khan MKJ, Ud Din N, B.S.Y.J.: "Interactive removal of microphone object in facial images." Electronics 8(10) (2019)

[10] B J, Bipin & Nihar, K & Adarsh, C. (2020). "A Comparative Binarization Approach for Degraded Agreement Document Image from Various Pharmacies." Vol 12. 806-814. 10.31838/ijpr/2020.12.02.0124.

[11] B J, Bipin & Khamarudheen, K.S. & Ranjitha, H.s. (2016). "An approach for identifying the presence of factor IX gene in DNA sequences using position vector ANN." 87. 396-403.

[12] Szegedy C, Liu W, Jia Y, et al. "Going deeper with convolutions. In: 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)." IEEE; 2015.

[13] Suresh.K and Pattabiraman. V, "Reduction of large database and identifying frequent pat terns using enhanced high utility mining", International Journal of Pure and Applied Mathematics, Volume-109, No-5, Page No-161-169, 2016.

[16] Suresh. K and Devika Mohan, "Development of High Utility Itemsets In Streaming Database", Test Engineering and Management, Volume No-82, Page No-13052 - 13056, ISSN:0193 – 4120, 2020.

[17] Suresh. K and Kashyap. C, "Effectively Mining on Utility Itemset by Using Conventional Method", Test Engineering and Management, Volume No-82, Page No-13062 - 13068, ISSN: 0193 – 4120, 2020.

[18] Chollet, F.: Xception: "Deep learning with depthwise separable convolutions. CoRR" abs/1610.02357 (2016), http://arxiv.org/abs/1610.02357

[19] Sandler, M., Howard, A.G., Zhu, M., Zhmoginov, A., Chen, L.: "Inverted residuals and linear bottlenecks: Mobile networks for classification, detection and segmentation." CoRRabs/1801.04381 (2018), http://arxiv.org/abs/1801.04381

[20] "Howard AG, Zhu M, Chen B, et al. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications." arXiv [csCV]. Published online 2017. http://arxiv.org/abs/1704.04861