MASK DETECTION SYSTEM FOR COVID-19 SCENARIO USING COMPUTER VISION

Dr. Prakash Prasad
Head of Department(IT)
Priyadarshini College Of Engineering,
Nagpur, Maharashtra, India

Mukul Shende
Student
Priyadarshini College Of Engineering,
Nagpur, Maharashtra, India

Mayur Karemore
Student
Priyadarshini College Of Engineering,
Nagpur, Maharashtra, India

Lucky Khobragade
Student
Priyadarshini College Of Engineering,
Nagpur, Maharashtra, India

Amit Dravyakar
Student
Priyadarshini College Of Engineering,
Nagpur, Maharashtra, India

Davesh Bondre
Student
Priyadarshini College Of Engineering,
Nagpur, Maharashtra, India

Abstract—The new pandemic of (Coronavirus Disease-2019) COVID-19 continues to spread worldwide. Every potential sector is experiencing a decline in growth. (World Health Organization) WHO suggests that Wearing Face Mask can reduce the impact of COVID-19. So, This Paper Proposed a system that controls the growth of COVID-19 by finding individuals who don't wear masks in populated areas like malls, markets where all public places are under surveillance with closed-circuit television cameras (CCTV). When a person without a mask is found, the corresponding authority is informed by the CCTV network. And it can calculate the number of people that do not wear the mask and emit an audible signal to inform the authority. A deep learning module is trained on a dataset composed of images of people wearing different types of masks and people without masks collected from various sources. It also contains some confusing images that help the model to achieve greater precision than other models. This model will use the dataset to build a COVID-19 face mask detector with computer vision using Computer Vision. This approach allowed extracting even the details from the pixels.

Keywords-**COVID-19, Deep Learning, CCTV (Closed Circuit Television Camera), face mask, mask detection, Computer Vision**

I. INTRODUCTION

As the world is fighting against the COVID-19 pandemic there is so much vital equipment to combat this pandemic and one of them is the FACIAL MASK [3]. Before the pandemic people used a mask for air pollution. While other people often forgot to wear masks in public places. Scientists have demonstrated that wearing masks contributes to preventing the transmission of COVID-19. It is the latest epidemic virus that has plagued human health over the past century [3]. In 2020, the rapid expansion of COVID-19 pushed the WHO into announcing a global pandemic and Lockdowns in several countries.

According to the survey, over five million cases were infected with COVID-19 within six months in 188 countries. The virus travels through close contact into populated areas [3].

Government officials compel people to wear face masks in public places. These rules and laws were developed by the government authority to reduce the exponential growth of cases and fatalities in many areas. However, the process of tracking large groups of persons becomes more challenging. The monitoring process consists of detecting anyone without a mask. Here this paper is about, A mask recognition model built on computer vision and deep learning. The proposed model can be integrated with existing CCTV cameras to

hamper the transmission of COVID-19 by allowing the detection of individuals wearing a mask without wearing a mask. Advancements in the field of deep learning, particularly in convolutional neural networks (CNNs), have already shown remarkable improvement in the classification of images [2][4]. The key idea behind the CNNs is to create an artificial model, like a visualization area of the human brain. The biggest advantage of CNNs is that one can extract more important characteristics over the whole image, instead of just handcrafted attributes [2]. Researchers introduced different deep networks based on CNN, and these networks achieved the state of results in computer vision classification, segmentation, object detection, and localization. To achieve greater precision than other models, this module uses the dataset (which is used to form the model). That covers different images of masked or not masked faces. Including images of people wearing masks and people who don't wear masks. It also contains some confusing images That will lead to better outcomes than others. The model is an integration of deep learning and conventional machine learning techniques using OpenCV, TensorFlow, and Keras [1]. This approach is a deep learning methodology for feature extraction and combines it with the Convolutional Neural Network (CNN) algorithm [1]. In this research study, deep learning techniques are applied to construct a classifier to gather images of an individual wearing a mask and not wearing from the database and scan the video footage frame by frame to differentiate between those classes of facemask wearing and not facemask wearing and calculate the number of individuals who are wearing a mask and who isn't wearing then creates an alarm if it's not observed properly including the wearing of the facemask [2]. To let the authority know the person isn't wearing a mask. It'll contribute to the greater use of masks in crowded areas. The artificial neural network has proven to be an active process for extracting features from unprocessed data.

II. RELATED WORK

To detect a mask on a person's face we have to extract the facial features and then determine if the mask is detected on the face or not. A lot of work has already been done on face detection techniques and is a very crucial step in our approach. The goal of the extraction of facial features is to determine the position of the face in a frame or image. Human faces can be difficult to detect and model as many variables that can interrupt during detection like lighting, face position, expressions, orientation, accessories on the face, etc [5].

There are two major ways to detect faces, which are feature-based and image-based. In the paper "Rapid Object Detection using a Boosted Cascade of Simple Features," Paul viola and Michael [6] suggest a machine learning approach for object detection technique which preprocess images very quickly and uses "Integral Image" allowing us to extract features which we required for detection. It also uses a method that combines more complex classifiers in "cascade" to remove background from the image more quickly and more computational power can be used in object detection. Viola and Jones used Haar-like features to detect faces in this algorithm [6], although this algorithm is fast and pretty accurate this method fails to tackle real-world problems like non-frontal image, face position, expressions, orientation, accessories wearable on the face which can influence results and some people can use this to fail the accuracy of our system. This issue made developers work on new methods which were based on deep learning.

There is another approach for face recognition based on a Modular Histogram [10] of oriented directional features. In this approach Almabrok Essa and Vijayan K. Asa. In their approach, they used edge detection algorithms like Kirsch mask, Prewitt mask, Gaussian derivation mask, and Nobel mask for edge detection and used binary coding strategy to retrieve a map of edge features [10]. They found that this method is very effective in uncontrolled conditions like mentioned illumination and different light effects on the face [10].

In the image-based approach, we generally use statistical approach and machine learning to find relevant characteristics in images containing faces and non-face objects and differentiate between them. We use different algorithms like Neural networks, AdaBoost [7], and MTCNN [11] to classify images and face detection. Many architectures are available like FaceNet[8], AlexNet, VGGNET, and Goolenet which uses parallel convolution kernels and concatenating the feature maps in 1*1,3*3 and 5*5 convolution and 3*3 max-pooling [9]. This approach is more accurate than feature extraction but requires more resources but mostly tackle real-world problems if we use more data and feed it to the trainer to get maximum accuracy. In the paper "Joint Face Detection and Alignment using Multi-task Cascaded Convolutional Networks" Kaipeng Zhang, Zhanpeng Zhang, Zhifeng Li [11] proposed to multitask cascade CNN-based framework for face detection and alignment and

this method outperformed many states of art algorithms and benchmarks for face detection [11]. So we decided on using the MTCNN framework for our approach.

Most people get confused between face recognition and face detection and use both terms interchangeably. Face detection is the part of face recognition, they are not the same. Many Face recognition algorithms are available to keep track of a single person like Facenet [8], YOLO, BlazeFace, and OpenCV [17] which are very helpful for constantly monitoring the same person for how often he removes masks and violet the laws. Although these algorithms are more accurate and can be used in different environments, they require more computational power and resources to implement. To tackle this issue cloud computing comes into the picture. Cloud computing gives more freedom to implement these algorithms and can be implemented and accessed from anywhere with enough hardware.

III. METHODOLOGY

Transfer learning is the method of taking a previously learned paradigm and adapting it to a new task, i.e. moving information from one task to the next. As opposed to models that do not use transfer learning, this is beneficial because the model does not have to learn from scratch and can reach higher precision with less time [14].

This approach is MTCNN [11] based. Because it has a faster detection speed. It requires lesser memory. MTCNN (Multi-task Convolutional Neural Networks) is an algorithm consisting of 3 stages, This identifies the bounding boxes of faces along with their 5 Point Face Landmarks in a picture. Each stage gradually enhances the identification of results bypassing its inputs through a CNN, Which restores candidate bounding boxes, followed by non-max suppression, with their ratings [14].

In stage one, the image which is given as input is down-scaled several times to create an image pyramid, and each scaled image version is
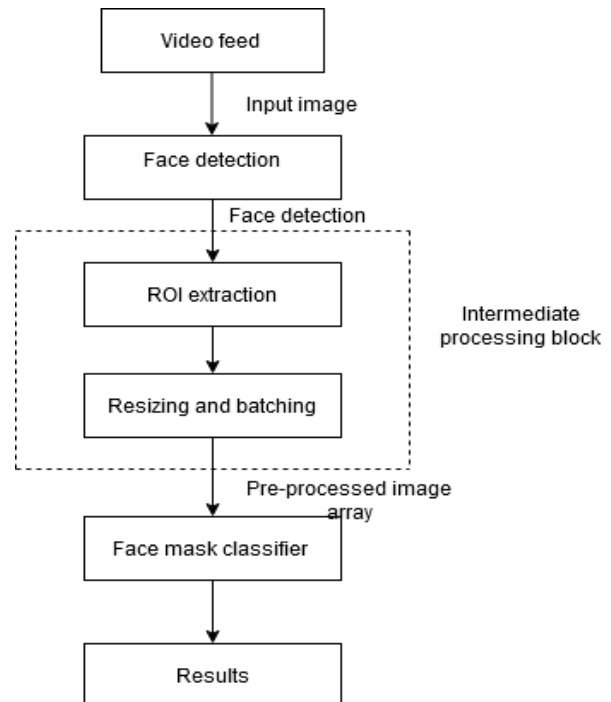

Figure 1 : Flowchart showing different phase of Face Mask detection Model

transferred into the CNN version. Extract the image patches for each bounding and enclose the stages a pair and three and scale them (24x24 in the second stage and 48x48 in the third stage) and send them to the CNN stage.

In addition to bounding boxes and ratings, stage 3 additionally calculates 5 points of face landmarks for each bounding box. After fiddling around with some MTCNN implementations, it turns out that at far lower inference times compared to SSD Mobilenet [12], even by running inference on the CPU, you can actually get very good detection results. As an additional perk, we get facial alignment from the 5 Point Face Landmarks. This way, before computing a face descriptor, we do not perform 68 Point Face Landmark detection as an intermediate stage. The following figure represents how our approach will work.

This approach has 3 main phases:
The first phase is face detection in which this needs to optimize different types of faces in images of varying sizes and detect faces even in overlapping scenarios [13]. The extracted detected faces (regions of interest) are then merged and transported to the next phase of this method., which is the intermediate block of processing.

The second phase is the intermediate processing block that performs the processing of the faces

recognized and batches them for classification together. In this phase, this module needs the individual's entire head to correctly identify the faces as masked or unmasked. This phase involves cutting the extended bounding boxes out of the picture for each detected face to remove the ROI [13].

The third phase or final phase of our approach is the Face mask Classifier. The processed ROI is taken from the Intermediate Processing Block at this phase and categorized as either Mask or No Mask. Based on different types of images, a CNN-based classifier was trained for this process. A picture of distinct faces that can be identified as masked or unmasked [13].

IV. PROPOSED METHOD

In 2020, Face Mask Detection using Transfer Learning of InceptionV3 [15] proposed by Narinder Punn stated that some sort of statistics-based classifier was then helpful in differentiating between facial and non-facial regions. Human faces, on the other hand, are different. Textures that can be used to tell the difference between a mask and other objects Furthermore, The edge of the role will help separate the objects from the face. We'll add a new feature. The proposed approach consists of a pre-trained CNN consisting of two layers of 2D convolution Linked to dense neuron layers [16]. The approach for face mask identification is as follows:

4.1 Dataset Collection

From the internet, the set of photographs used to train and validate the model was compiled. The dataset used to train the model in each method was a blend of open-source datasets and photographs, including data from Mikolaj Witkowski's Medical Mask Datasets on Kaggle and the Prajna Bhandary dataset on PyImageSearch [17].

Using 2 classes, face mask wearing, and no-face mask-wearing, the dataset was capable of producing 2,000 images. The resolution of the dataset's image size varies from 800 to 1200 pixels. The images were all formatted in JPEG format. The sample of the image data collection obtained is seen showing persons wearing a facemask and not wearing a facemask.

4.2 Splitting the data

In 2021, A real-time DNN-based face mask detection system using a single shot multibox detector and MobileNetV2 [18] proposed by, Kataria P, Jain R, Madan A, Hemanth, Arora R,

suggested dividing the dataset into the training set that will include the images that will be trained on the CNN model and the test set of the images that will be evaluated on the model. They took split-size=0.8 in this, which means that 80% of the overall pictures will go to the training set and the test collection would receive the remaining 20% of the images from the training set. After breaking, they observed that, as described above, the optimal percentage of images has been allocated to both the training set and the test set.

4.3 Building the Model

After the classification, the image data will be given as the input to keep the model creation procedure on stand-by. To output a vector that gives the likelihood of either of the two classes [19]. For greater precision, we can also use MobileNetV2[18].

4.4 Training the CNN model

This step is the key step in the training set where the model matches dataset images and the test set to our Sequential model that was developed using the Keras library. As per Deep Learning with Python proposed by François Chollet [20] This may, however, prepare for further epochs to gain greater precision if there is no over-fitting. If the accuracy rate is high then it indicates that without any overfitting, the model is well-trained and accurate.

4.5 Labeling the Information

After building the model, there is a tendency to label 2 possibilities for our results. ['0' as 'without_mask' and '1' as 'with_mask']. We are additionally setting the boundary parallelogram color mistreatment of the RGB values.['RED' for 'without_mask' and 'GREEN' for 'with_mask']

4.6 Importing the Face detection Model

As it was proposed by Anjith George in Biometric Face Presentation Attack Detection with Multi-Channel Convolutional Neural Network,2019 [21] the model intends to use it after that to detect whether we are wearing a face mask using the camera of our PC. Second, we need to introduce face recognition for this. In this, MTCNN Classifiers are used to detect the face characteristics and pass the extracted image to the classifier.

4.7 Detecting the Faces with and without Masks

To run an endless loop in the final stage,we pass the extracted image to the model. The model

would estimate the probability ([without mask, with mask]) of each of the two classes. The label will be picked and shown around our ears, depending on the high probability.

V. CONCLUSION

We researched and compared the methods that accurately detects faces and decides whether an individual is wearing a mask. The results of this approach are very accurate and give the best performance as compared with old systems. Extensive tests show that the proposed approach outperforms certain state-of-the-art methods while being more consistent on multiple benchmarks. The proposed system can detect faces while tackling different variables which can give problems to convention feature extraction systems.

VI.   REFERENCES

[1] Militante S. V. and Dionisio N. V.(2020), "Deep Learning Implementation of Facemask and Physical Distancing Detection with Alarm Systems," 2020 Third International Conference on Vocational Education and Electrical Engineering (ICVEE), Surabaya, Indonesia,, (pp. 1-5), doi: 10.1109/ICVEE50212.2020.9243183.

[2] Vinitha. V1, Velantina. V2,2020,COVID-19 FACEMASK DETECTION WITH DEEP LEARNING AND COMPUTER VISION,(pp. 3127-3132) : .www.irjet.net

[3] https://www.who.int/news/item/31-08-2020-in-who-global-pulse-survey-90-of-countries-report-disruptions-to-essential-health-services-since-covid-19-pandemic

[4] Ejaz S. and Islam M. R. (2019), "Masked Face Recognition Using Convolutional Neural Network," 2019 International Conference on Sustainable Technologies for Industry 4.0 (STI), Dhaka, Bangladesh,, (pp. 1-6), doi: 10.1109/STI47673.2019.9068044.

[5] Kumar, Ashu & Kaur, Amandeep & Kumar, Munish. (2019). "Face Detection Techniques: A Review". Artificial Intelligence Review. 52. 10.1007/s10462-018- 9650-2.

[6] Viola, Paul & Jones, Michael. (2001). Rapid Object Detection using a Boosted Cascade of Simple Features. IEEE Conf Comput Vis Pattern Recognit. 1. I-511. 10.1109/CVPR.2001.990517.

[7] Zakaria Z. and Suandi S. A. (2011), "Face detection using combination of Neural Network and Adaboost," TENCON 2011 - 2011 IEEE Region 10 Conference, Bali, Indonesia, (pp. 335-338), doi: 10.1109/TENCON.2011.6129120.

[8] William, D. R. Ignatius Moses Setiadi, E. H. Rachmawanto, H. A. Santoso and C. A. Sari,(2019) "Face Recognition using FaceNet (Survey, Performance Test, and Comparison)," 2019 Fourth International Conference on Informatics and Computing (ICIC), Semarang, Indonesia,   (pp.   1-6),   doi: 10.1109/ICIC47613.2019.8985786.

[9] Deore G., Bodhula R., Udpikar V. and More V. (2016), "Study of masked face detection approach in video analytics," Conference on Advances in Signal Processing (CASP), Pune, India,   2016,   pp.   196-200,   doi: 10.1109/CASP.2016.7746164.

[10] Essa A. and Asari V. K. (2016), "Face recognition based on modular histogram of oriented directional features," IEEE National Aerospace   and   Electronics   Conference (NAECON) and Ohio Innovation Summit (OIS), Dayton,   OH,   2016,   pp.   49-53,   doi: 10.1109/NAECON.2016.7856773.

[11] Zhang K. , Zhang Z., Li Z. and Qiao Y. (Oct.   2016),   "Joint   Face   Detection   and Alignment   Using   Multitask   Cascaded Convolutional   Networks,"   in   IEEE   Signal Processing Letters, (vol. 23, no. 10), pp. 1499-1503, doi: 10.1109/LSP.2016.2603342.

[12] Howard Andrew ,Menglong Zhu Bo Chen, Kalenichenko Dmitry,Wang Weijun, Tobias Weyand,   Andreetto   Marco,(2017) Ad"MobileNets: Efficient Convolutional Neural Networks for Mobile Vision ",

arXiv:1704.04861v1

[13] Chavda, Amit & Dsouza, Jason & Badgujar, Sumeet & Damani, Ankit. (2020).   Multi-Stage CNN Architecture for Face Mask Detection.

[14] Jiang Mingjie, Fan Xinqui, Yan Hong (2020), RetinaMask: A Face Mask detector: arXiv:2005.03950 ,(pp. 1-4)

[15] Chowdary Jignesh, Singh Narinder, Agarwal Sonali(2020), Face Mask Detection using Transfer :Learning of InceptionV3: arXiv:2009.08369

[16] Yang, Ming-Hsuan & Kriegman, David & Ahuja, Narendra. (2002). Detecting Faces in Images: A Survey. Pattern Analysis and Machine Intelligence, IEEE Transactions on. 24.(34 - 58.) 10.1109/34.982883.

[17] Das, Arjya & Ansari, Mohammad & Basak, Rohini. (2020). Covid-19 Face Mask Detection Using TensorFlow, Keras and OpenCV. 10.13140/RG.2.2.21473.94565.

[18] Nagrath P, Jain R, Madan A, Arora R, Kataria P, Hemanth J.(2021) SSDMNV2: A real time DNN-based face mask detection system using single shot multibox detector and MobileNetV2. *Sustain Cities Soc*.2021;66:102692. doi:10.1016/j.scs.2020.102692

[19] Ning Jiang, Yijun Lu, Shaopeng Tang and S. Goto (2010), "Rapid face detection using a multi-mode cascade and Separate Haar Feature," 2010 International Symposium on Intelligent Signal Processing and Communication Systems, Chengdu, China,, (pp. 1-4.) doi: 10.1109/ISPACS.2010.5704623

[20] Chollet, F., 2018. *Deep learning with Python* (Vol. 361). New York: Manning

[21] George A., Mostaani Z., Geissenbuhler D., Nikisins O., Anjos A. and Marcel S. (2020), "Biometric Face Presentation Attack Detection With Multi-Channel Convolutional Neural Network," in IEEE Transactions on Information Forensics and Security, vol. 15, (pp. 42-55), doi: 10.1109/TIFS.2019.2916652.