Exploring Deep Learning Methods for Face Mask Detection

[1]MARITTA STEPHEN, [2]USHA K
[1,2]NSS College of Engineering, Palakkad, Kerala

Abstract:- The global epidemic COVID-19 has brought about a drastic change in the lives of mankind. Health and economic systems were severely impacted by the global epidemic. The world was under strict restrictions to fight and control the pandemic. Wearing a face mask is an essential protective measure to prevent the dissemination of virus in public. The reports suggests that a large population often ignore or tend to avoid wearing masks amid strict rules. The project's objective is to devise a face mask detector that can determine from visual inputs if an individual is wearing a mask. This work also focuses on face mask detection of a moving face and can perform real time face mask detection. Tensor Flow, Keras, OpenCV and Scikit-Learn were used to buld the detector along with Mobilenetv2 for face mask classification SSD (Single Shot Multibox detector) with Resnet as the base for face detection.

Keywords:- *Deep Learning, Object Detection, Face Recognition, TensorFlow, Keras, OpenCV, SSD, Mobilenet*

I. INTRODUCTION

The world is fighting against the largest pandemic never seen before in the history. Millions of people have been affected by the pandemic and the deeath rates at many parts of the word are alarmimg. Lockdowns and counter measures had been taken to control the spread of the virus which had great impacts on the economy and people. Unlocking phases are initiated at different regions and the lives are becoming back to normal but the precautions can never be ignored. The only way to outlive the situations is to continue to follow all safety measures to control the spread of viruses.

At this phase when the number of instances of covid is rising daily and holding our lives in huge risk, the face mask detector stands out as a step towards monitoring and maintaining a safe and healthy environment. Although the virus cannot be completely eradicated, we may take safe and healthy steps to live with it. Face masks are now required to be worn even in locations where lockdown restrictions have been lifted in order to prevent further mass propagation. In this regard, the initiative attempted to help the security procedures and systems for detecting whether the face mask is worn by every individual.

Wearing face mask is now a mandatory measure while stepping out anywhere. Strict measures were taken by the government to enforce the rules in this regard to avoid commuting the virus. A face mask detector arenow used for surveillance applications where people can be monitored on wearing a mask for the enforcement of the rules which can be implemented by embedding it to mobile applications and CCTV or other embedded systems.

Initial screen check for entry are now mandatory at places like examination halls , offices, malls with regard to the covid restrictions and safety measures. The entry od an individual is allowed only if the person wears a mask. The work is an useful safety precaution and efficient screening and monitoring of security checks at public places and gatherings.

II. RELATED WORKS

There are several studies on approaches for face mask which includes one stage detectors and multi stage detectors with various classifiers like fast- RCNN, Haar classifiers. Mohammed Loey in his work [1] developed a novel deep learning singlr stage detector model based on YOLO-v2 with ResNet-50 for the detection of medical face masks. For feature extraction, the ResNet-50 deep transfer learning model laid the foundation, while a component based on YOLO v2 was developed for the detection of medical face masks. Unlike models like RCNN and faster RCNN, which require two or more stages, the YOLO is an one stage detector, allowing detections to be completed in a single step. The proposed model had poor probability probability or confidence scores and the confidence of the predicted results was relatively low.

Zhang, Jun in his paper [2] proposed a novel detection framework based on the attention mechanism which explicitly balanced the multiple context features named Context-Attention R-CNN. The shortcoming of the system is that it used more hyper parameters and had higher training and inference time. The context information was also used by M. Jiang and X. Fan in their paper [3] which proposed RetinaMask detector based on the single-stage generic detector named RetinalNet. The RetinaMask also attempted to extract the robust feature while taking into account the context information. RetinaFaceMask combined the high-level semantic data using various feature maps and a feature pyramid network (FPN). The detector however had long training time and had complicated post processing.

B. Qin and D. Li [4] developed a new facemask-wearing condition identification method by combining image super-resolution and classification networks (SRCNet). However, there was a lengthy training process for the system, and it had a tendency to fail when the local face was clipped. The study [5] developed a two-stage method employing hybrid machine learning approaches to identify face masks. A broad learning system was utilised in the second stage to validate the actual facial masks, whereas the first stage was based on the transfer model of Faster-RCNN and InceptionV2. Small objects and the recognition of faces hidden by protective clothes and medical eyewear presented challenges for this technology. Real time facemask detection was not possible in this method.

A neural network was all that was put into the one-stage detector to look find objects. This can be accomplished by predefining some anchor boxes that specify the proportion of object widths to object heights. One-stage detectors improve performance in comparison to two-stage detectors to greatly increase detection speed. In order to achieve the goal, YOLO divided the image up into different cells. Thereafter, they sought to match the anchor boxes to the items in each cell, however this strategy did not perform well for small objects.

In order to detect faces of various sizes, the multi-scale detection strategy introduced in SSD was adopted. This method conducts detection on a number of feature maps and allows for the detection of faces of various sizes. The proposed system in the project used SSD as the face detection model.

III. PROPOSED WORK

The proposed system Face Mask Detector detects the face from the image before determining whether or not the subject has been wearing a mask. A face and a mask can both be detected in real time. If a mask is being worn, it is indicated by a bounding box drawn over the person's face. Packages like Tensor Flow, Keras, OpenCV, and Scikit-Learn were used for the facemask detector. With Resnet as the basis architecture model, the face detection model utilized was SSD (Single Shot Multibox detector), with the face mask classifier Mobilenetv2. The mask classification and detection was performed on various input formats which includes image input, video input, real time video and images taken from a real time video. The system architecture is shown in Figure 1

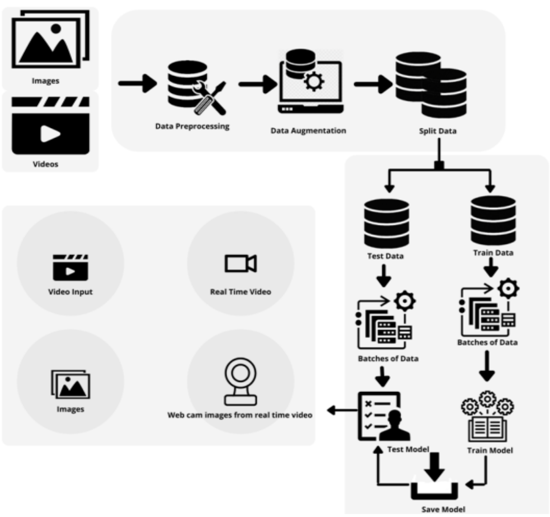

Fig 1: System architecture

The Imagenet dataset, which comprises of 3833 images and 1915 images with masks and 1918 images without mask, was used to train the face mask classification model. The modules involved are discussed below.

- *Data Preprocessing*

The data preprocessing step prepares the raw data and make it suitable for training the model. The data cleaning was done manually.The preprocessing module of keras was used for preprocessing the images before training the model.Initially the dataset was saved in the directory as two folders named with mask and without mask.The aim at the end was to convert all images into arrays and to obtain the train mask detector file.Two list was created as data and label which was used to to append image arrays and append with mask or without mask labels.The categories with mask and without mask was looped into and the images was listed down which then followed by joining paths to list corresponding images.The images was saved as array.The array was appended to the list.The preprocessed data and labels was obtained.

- *Data Augmentation*

Data augmentation on the dataset was performed to meet the inadequate dataset.Various operations on the existing dataset was done to expand the dataset.Image datagenerator module of keras was used for data augmentation.The operations performed on the dataset were rotation, shift,shear and flip.

- *System Design*

The face mask detector was developed in two phases.The detailed design of the detector is shown in figure 2. Training and face mask detector deployment are the two phases of the face mask detector.

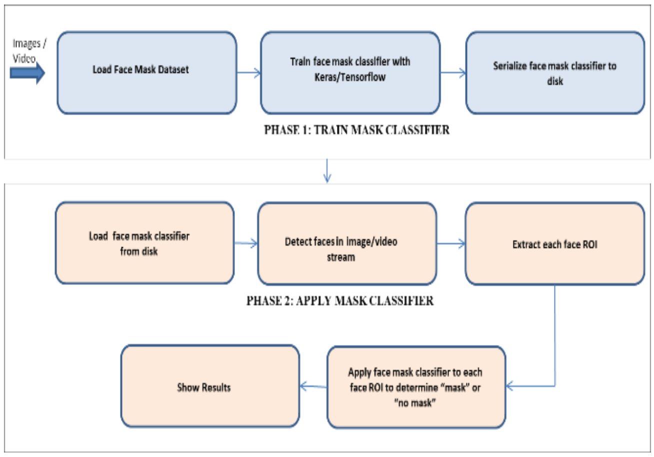

Fig 2: System design

The dimensions of the frame from input images were acquired, a blob was created from it, and the blob was then put through the network to obtain the face detections. This allowed the face mask classifier model to be applied to video streams and photos. The face detection network used here was SSD with Resnet as its base architecture. Blob was an image or collection of images with some spatial dimensions and it took an image as an input and performed preprocessing on it by scaling and performing mean subtraction to it. The images were normalized by performing mean subtraction which later on was converted to RGB format.

ISSN No:-2456-2165

The faces, or a list of faces with their corresponding locations, and the list of predictions from the face mask network were obtained for the purpose of face detection. The confidence which was the probability associated with the detections was extracted by looping over all the detections for video and just on the detection for image. By making sure the confidence was higher than the minimum confidence, the weak detections were eliminated. Every detection that fell below the confidence level was ignored.

The face mask detector detected faces and classified them with mask and without mask and drew bounding boxes over them. Retrieved the object's bounding box's x and y coordinates and made sure the bounding boxes were contained within the frame's boundaries. Bounding boxes were drawn over the faces. The ROI was extracted and converted it from BGR to RGB channel as it was initially preprocessed by performing mean subtraction. The preprocessing was done again by ordering, resizing it to 224x224.The faces and bounding boxes were added to the list .The predictions were made only when atleast a face was detected otherwise no predictions were made. The end result was a 2-tuple of the locations of the faces and their associated locations.

The predictions and locations obtained were then used for face mask detection. Serialized face detector model saved in the disk was loaded. The face mask classifier model was also loaded from disk along with it. The video stream was initialised and looped through the frames in order to apply the model. The threaded video stream's frames were extracted and resized. Faces were detected in the frame and determined if they had worn mask or not. The detected face locations were looped and their associated locations were found which was followed by unpacking the bounding box and predictions. The class label and probability were included in the output and the output frames and images were displayed.

IV. RESULTS

The evaluation metrics accuracy , precision and F1 score were used to evaluate the performance of face mask detector. The classes were balanced, thus the accuracy metric is an excellent measure to start. Precision provided a measure of the values that the mask predicted. The classifier could discover all positive samples using recall, and test accuracy was determined by the f1 score. The best results in a balanced dataset were obtained using these evaluation metrics, so that they were chosen. The classification report for the model is as given below in Table 1

Table 1: Classification Report

	Precision	Recall	F1 Score	Support
With mask	0.99	0.99	0.99	383
Without mask	0.99	0.99	0.99	384
Accuracy			0.99	767
Macro avg	0.99	0.99	0.99	767
Weighted avg	0.99	0.99	0.99	767

The model's average accuracy in predicting whether a person wearing a face masks or not on a validation dataset is 99%. The plot in Figure 3 depicts the training loss and accuracy. The predictions on sample images using the face mask detector is shown in Figure 4.

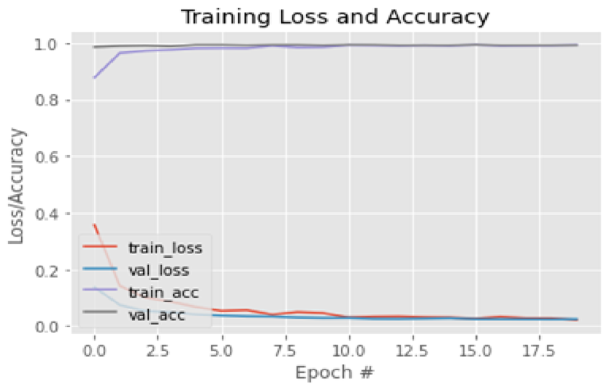

Fig 3: Training loss and accuracy

The rectangle red box depicts persons without a mask, while the rectangular green box displays detections and predictions of individuals wearing masks and includes a probability score on the upper left. The prediction on images obtained from real time video taken from web camera is shown in the figure 5.

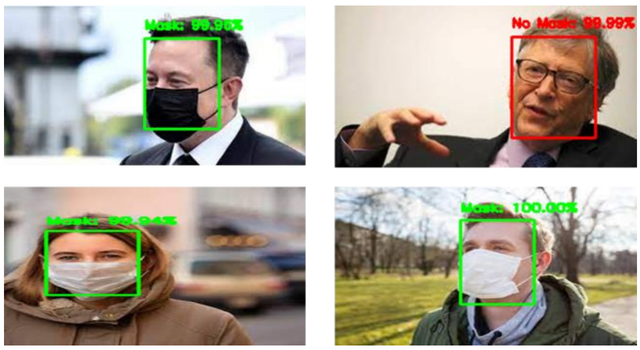

Fig 4: Predictions on test images

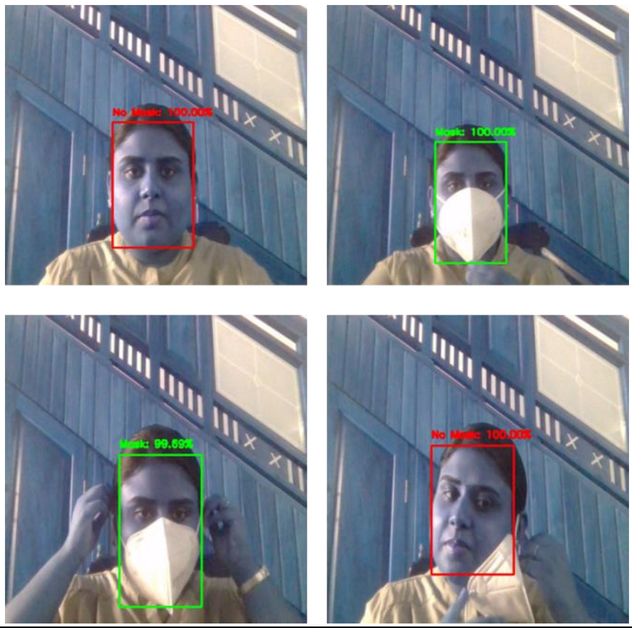

Fig 5: Predictions on real time video.

The face mask detector worked well with images and video stream inputs. It showed a good accuracy too. The images and outputs obtained had the face detected and bounding box drawn over it. This could be further improved upon drawing the bounding boxes over the mouth rather than the entire face.

V. CONCLUSION

Health experts suggests that the Covid virus cannot be eliminated entirely from the world. The propaganda preached is to take safe and healthy precautions to live with it. Wearing face mask has become mandatory even at places where safety restrictions are removed so as to avoid further mass commutation and is taken as precautionary step.The project aimed to assist the security systems and practices in this regard.

The proposed system is able to detect face with mask and without masks successfully using SSD with Resnet as backbone architecture and Mobilenetv2 as mask classifier. The transfer learning by using a pretrained model was done here as to take advantages and weights of the model pretrained on large GPU and dataset. The Mobilenetv2 was selected along with face detection model SSD Multibox because of their ease of deployment in embedded systems and mobile applications.

The keras application model Mobilenetv2 which had lesser top1 and top5 accuracy than models like inception,resnet,VGG16 was selected as their size in hard disk is very much about four times lesser than other keras application models so that they requires less runtime memory and was a best choice for mobile applications SSD used for face detection makes more predictions and had better coverage on location, scale, and aspect ratios.

REFERENCES

[1]. Nagrath, Preeti, Rachna Jain, Agam Madan, Rohan Arora, Piyush Kataria, and Jude Hemanth. "SSDMNV2: A real time DNN-based face mask detection system using single shot multibox detector and MobileNetV2." Sustainable cities and society 66 (2021): 102692.

[2]. M. Loey, G. Manogaran, M. H. N. Taha, and N. E. M. Khalifa, ‘‘Fighting against COVID-19: A novel deep learning model based on YOLO-v2 with ResNet-50 for medical face mask detection,’’ Sustain. Cities Soc., vol. 65, Feb. 2021, Art. no. 102600.

[3]. Zhang, Jun, Feiteng Han, Yutong Chun, and Wang Chen. "A Novel Detection Framework About Conditions of Wearing Face Mask for Helping Control the Spread of COVID-19." IEEE Access 9 (2021): 42975-42984.

[4]. M. Loey, G. Manogaran, M. H. N. Taha, and N. E. M. Khalifa, ‘‘A hybrid deep transfer learning model with machine learning methods for face mask detection in the era of the COVID-19 pandemic,’’ Measurement, vol. 167, Jan. 2021, Art. no. 108288

[5]. Comparative Review of YOLO & MobileNet Versions. (2021). International Research Journal of Engineering and Technology (IRJET), 7(4), 1–5. https://www.irjet.net/archives/V7/i4/IRJET-V7I4970.pdf