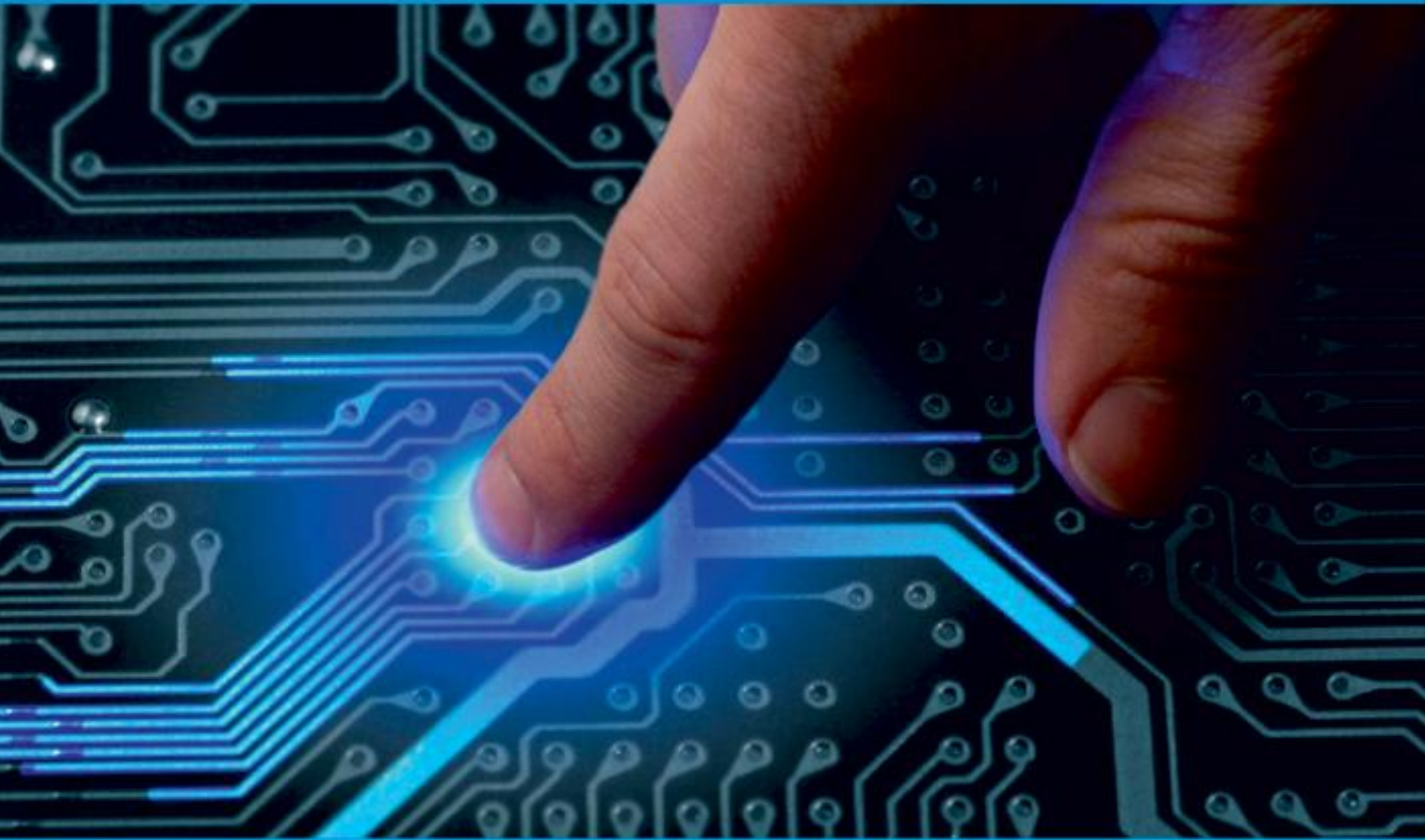
IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 9, Issue 7, July 2021

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.542

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542 |

|| Volume 9, Issue 7, July 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0907052|

Face Mask Detection Using Machine Learning

A.Sudhasan Reddy[1], P. Sai Chandrika[2], P.N.L.Tirupatamma[3], S.Naga Deekshitha[4], P.Taruni Anjani[5]

Associate Professor, Department of Information Technology, Vasireddy Venkatadri Institute of Technology, Nambur, Guntur, Andhra Pradesh, India[1]

UG Students, Department of Information Technology, Vasireddy Venkatadri Institute of Technology, Nambur, Guntur, Andhra Pradesh, India[2,3,4,5]

ABSTRACT: In the present scenario due to Covid-19, there are no efficient face mask detection applications which are now in high demand for transportation means, densely populated areas like shopping malls, large-scale manufacturers and other enterprises to ensure safety. This system can therefore be used in real-time applications which require face-mask detection for safety purposes. This project can be integrated with embedded systems for application in airports, railway stations, offices, schools, and public places to ensure that public safety guidelines are followed. Today everyone is aware of taking precaution and safety measures regarding covid-19, so face mask detection will play a huge role to avoid such pandemic situations. We use machine learning algorithms where it predicts the accuracy or percentage of the person with mask and without mask. If the accuracy of the person without a mask is high then we will alert with the sound. We will use Python scripts to Detect face masks in real-time video streams.

KEYWORDS*:* Deep Learning, CNN, face mask, CCTV cameras, detecting.

I.INTRODUCTION

Droplets may also land on surfaces where the virus could remain viable; thus, the immediate environment of an infected individual can serve as a source of transmission (contact transmission). The study analyses a set of video streams/images to identify people who are compliant with the government rule of wearing medical masks. This could help the government to take appropriate action against people who are non-compliant. The raw images used for the current study were downloaded from Kaggle and the majority of the images were augmented by OpenCV. The set of images were already labeled "mask" and "no mask". The images that were present were of different sizes and resolutions, probably extracted from different sources or from machines (cameras) of different resolutions. Here, we propose a two-stage CNN architecture, where the first stage detects human faces, while the second stage uses a lightweight image classifier to classify the faces detected in the first stage as either 'With Mask' or 'Without Mask' faces.

This project can be integrated with embedded systems for application in airports, railway stations, offices, schools, and public places to ensure that public safety guidelines are followed.

II.RELATED WORK

Object detection is one of the trending topics in the field of image processing and computer vision. Some examples include image retrieval, security and intelligence. A breakthrough face detection technology then was developed named as Viola Jones detector that was an optimized technique of using Haar, digital image features used in object recognition. However, it failed because it did not perform well on faces in dark areas and non-frontal faces. There are various methods of object detection based on deep learning which are divided into two categories: one stage and two stage object detectors.

Two stage detectors use two neural networks to detect objects, for instance region-based convolutional neural networks (R-CNN) and faster R-CNN. This strategy results in high detection performance compromising on speed. The seminal work R-CNN is proposed by R. Girshick. R-CNN uses selective search to propose some candidate regions which may contain objects. After that, the proposals are fed into a CNN model to extract features, and a support vector machine (SVM) is used to recognize classes of objects. However, the second stage of R-CNN is computationally expensive since the network has to detect proposals on a one-by-one manner and uses a separate SVM for final classification. Fast R-CNN solves this problem by introducing a region of interest (ROI) pooling layer to input all proposal regions at once. Faster RCNN is the evolution of R-CNN and Fast R-CNN, and as the name implies its training and testing speed is greater than those of its predecessors. While R-CNN and Fast R-CNN use selective search algorithms limiting the

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542 |

|| Volume 9, Issue 7, July 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0907052|

detection speed, Faster R-CNN learns the proposed object regions itself using a region proposal network (RPN). On the other hand, a one stage detector utilizes only a single neural network for region proposals and for detection; some primary ones being SSD (Single Shot Detection) and YOLO (You Only Look Once). To achieve this, the bounding boxes should be predefined. YOLO divides the image into several cells and then matches the bounding boxes to objects for each cell. This, however, is not good for small sized objects. One-stage detectors have higher speed but trades off the detection performance but then only are preferred over two-stage detectors. However, there is occasional research focusing on face mask detection. Some already existing face mask detectors have been modeled using OpenCV, Pytorch Lightning, Mobile Net, Retina Net and Support Vector Machines.

III.PROPOSED SYSTEM

The proposed system focuses on how to identify the person on image/video stream wearing face mask with the help of computer vision and deep learning algorithm by using the OpenCV, Tensor flow, Keras and MobileNetV2.A face detector acts as the first stage of our system. A raw RGB image is passed as the input to this stage. The face detector extracts and outputs all the faces detected in the image with their bounding box coordinates.This block carries out the processing of the detected faces and batches them together for classification.The first step involves expanding the bounding boxes in height and width by 20%, which covers the required Region of Interest (ROI) with minimal overlap with other faces in most situations.The second step involves cropping out the expanded bounding boxes from the image to extract the ROI for each detected face.The second stage of our system is a face mask classifier. This stage takes the processed ROI from the Intermediate Processing Block and classifies it as either Mask or No Mask.

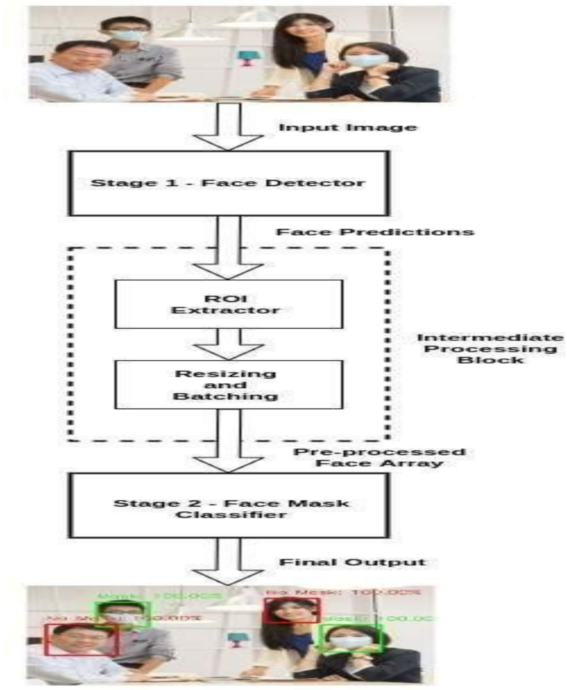


Fig 1: Architectural Design

4.1. Importing Libraries

Tensor flow→ 1.15.2

Tensor Flow is a free and open-source software library for dataflow and differentiable programming across a range of tasks. It is a symbolic math library, and is also used for machine learning applications such as neural networks.

Keras→ 2.3.1

Keras is a powerful and easy-to-use free open source Python library for developing and evaluating deep learning models. It wraps the efficient numerical computation.

Imutils→0.5.3

A series of convenience functions to make basic image processing functions such as translation, rotation, resizing, skeletonization, displaying Matplotlib images, sorting contours, detecting edges.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542 |

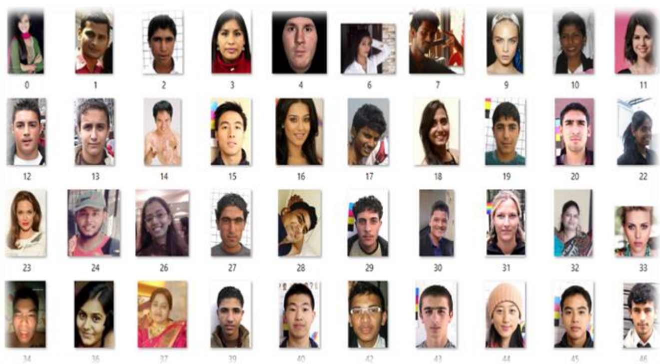
|| Volume 9, Issue 7, July 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0907052|

Numpy→ 1.18.2
NumPy or Numerical Python is an open-source Python library that makes it easy to complex numerical operations. Working with machine learning and deep learning applications involve complex numerical operations with large datasets.
Opencv-python→ 4.2.0.*
OpenCV (Open Source Computer Vision Library) is an open source computer vision and machine learning software library. The library has more than 2500 optimized algorithms. These algorithms can be used to detect and recognize faces, identify objects, classify human actions in videos, track camera movements, track moving objects.
Matplotlib→ 3.2.1
Matplotlib is one of the most popular library in Python which is used in Machine Learning. It helps to understand the huge amount of data through different visualizations.

4.2. Reading the Dataset

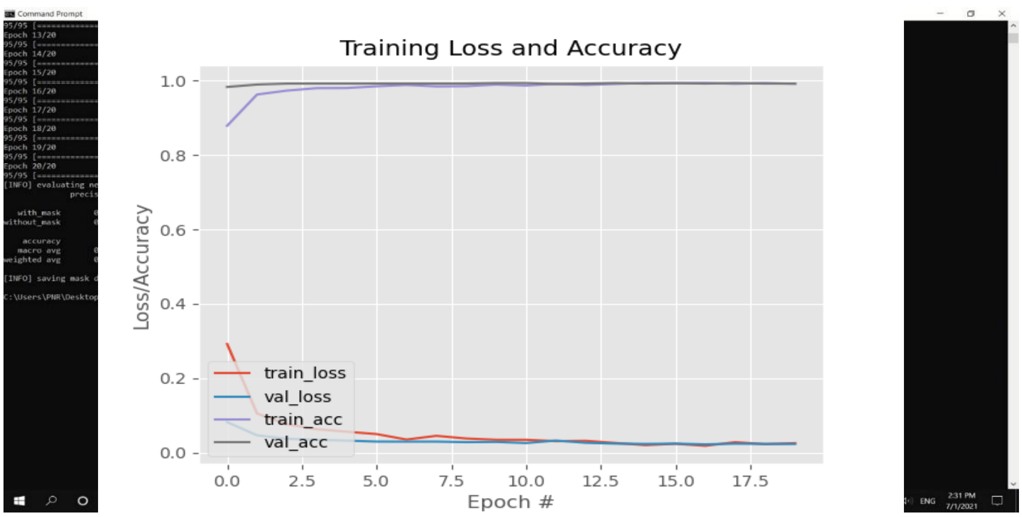
The dataset contains a total of 2136 images of people with masks and 2136 images of people without a mask. For training purposes, 80% images of each class are used for training and the rest of the images are utilized for testing purposes. Dataset is collected from Kaggle.



4.3. Training

At the training time, for each pixel, we compare the default bounding boxes having different sizes and aspect ratios with ground truth boxes and finally use Intersection over Union (IoU) method to select the best matching box. IoU evaluates how much part of our predicted box match with the ground reality. The values range from 0 to 1 and increasing values of IoU determine the accuracies in theprediction;the best value being the highest value of IoU. The equation and pictorial description of IoU is given as follow:

$$IoU(B1, B2) = B1 \cap B2 / B1 \cup B2$$



| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542 |

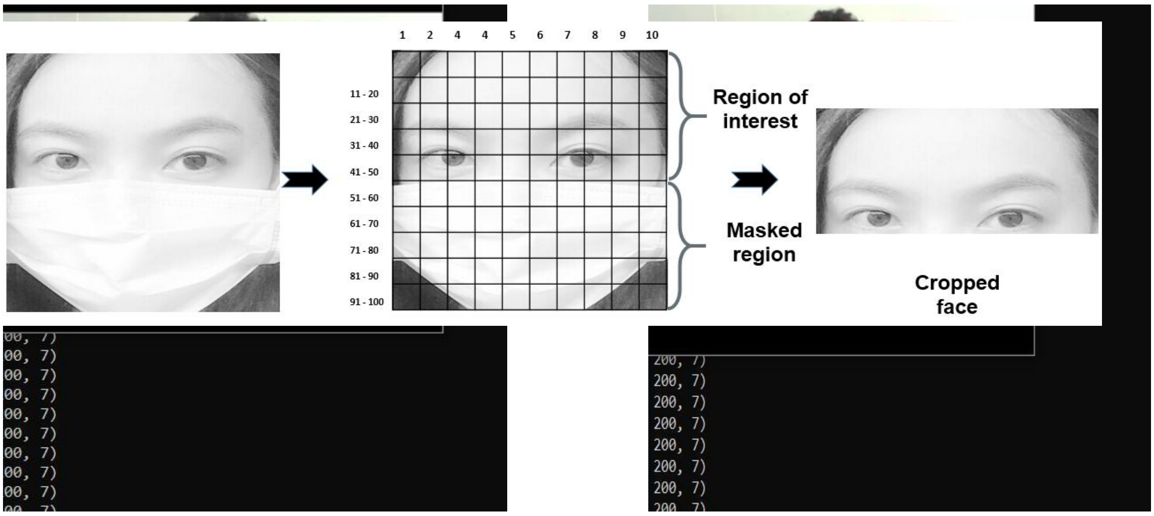
|| Volume 9, Issue 7, July 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0907052|

4.4.Image Pre-Processing

In the pre-processing step, the image is transformed into a grayscale image.It carries out the processing of detected faces and batches them together for classification.Next it involves cropping out the expanded bounding boxes from the image to extract the ROI for each detected faces.Then, we reshaped the images into (64×64) shape to maintain uniformity of the input images to the architecture.Then, the images are resized and normalized as required.

4.5. Output screens



V.CONCLUSION

The article proposed an efficient real-time deep learning-based technique to automate the process of detecting masked faces, where each masked face is identified in real-time with the help of bounding boxes. The extensive trials were conducted with popular models, namely, Faster RCNN and YOLO v3. F-RCNN has better precision, but for applying this in real-world surveillance cameras, it would be preferred to use the model with YOLO algorithm as it performs single-shot detection and has a much higher frame rate than Faster-RCNN or any other state-of-the-art object detection algorithms. One can tell YOLOv3 is better than faster R-CNN. Which model to use, also depends on the resources available. If high-end GPUs are available on the deployed devices, faster R-CNN must be used. YOLOv3 can be deployed on mobile phones also. Since this approach is highly sensitive to the spatial location of the camera, the same approach can be fine-tuned to better adjust with the corresponding field of view. These models can be used along with surveillance cameras in offices, metros, railway stations and crowded public areas to check if people are following rules and wearing masks.

VI.FUTURE ENHANCEMENTS

More than fifty countries around the world have recently initiated wearing face masks compulsory. People have to cover their faces in public, supermarkets, public transports, offices, and stores. Retail companies often use software to count the number of people entering their stores. They may also like to measure impressions on digital displays and promotional screens. We are planning to improve our Face Mask Detection tool and release it as an open-source project. Our software can be equated to any existing USB, IP cameras, and CCTV cameras to detect people without a mask. This detection live video feed can be implemented in web and desktop applications so that the operator can see notice messages. Software operators can also get an image in case someone is not wearing a mask. Furthermore, an alarm system can also be implemented to sound a beep when someone without a mask enters the area.

REFERENCES

1. Deng J, Dong W, Socher R, Li L-J, Li K, Fei-Fei L (2009) ImageNet: A large-scale hierarchical image database. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, pp 248–255

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542 |

|| Volume 9, Issue 7, July 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0907052|

2. He K, Zhang X, Ren S, Sun J (2016) deep residual learning for image recognition. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, pp 770–778

3. Howard AG, Zhu M, Chen B, Kalenichenko D, Wang W, Weyand T, Andreetto M, Adam H (2017) Mobile nets: efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv:1704.04861, 2017

4. Jason B A Gentle Introduction to Transfer Learning for Deep Learning.https://machinelearningmastery.com/transfer-learning-for-deep-learning/

5. Matuschek C, Moll F, Fangerau H, Fischer JC, Zänker K, van Griensven M, Schneider M, Kindgen-Milles D, Knoefel WT, Lichtenberg A, Tamaskovics B, Djiepmo-Njanang FJ, Budach W, Corradini S, Häussinger D, Feldt T, Jensen B, Pelka R, Orth K, Peiper M, Grebe O, Maas K, Gerber PA, Pedoto A, Bölke E, Haussmann J (2020) Face masks: benefits and risks during the COVID-19 crisis. Eur J Med Res 25:32 (2020).https://doi.org/10.1186/s40001-020-00430-5

6. Redmon J, Divvala S, Girshick R, Farhadi A (2016) You only look once: unified, real-time object detection, vol 2016. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, pp 779–788.https://doi.org/10.1109/CVPR.2016.91

INNO SPACE
SJIF Scientific Journal Impact Factor
Impact Factor: 7.542

doi crossref

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर
NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING