iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 9 **Issue:** V **Month of publication:** May 2021

DOI: https://doi.org/10.22214/ijraset.2021.34014

A Deep Learning-Based Assistive System to Classify COVID-19 Face Mask for Human Safety with MobileNetV2

Dudala Hemanth[1], SD Rashad Baba[2], Repala Santhosh[3]
[1, 2, 3]B-Tech Computer Science & Engineering, Lovely Professional University, Punjab

Abstract: Computer vision learning pays off attention due to the global epidemic COVID-19 to improve public safety. Since 2019 there has been a huge human loss cause of COVID-19 and the world is at risk. When it comes to human safety technology has always played a pivotal role to protect earth and its humanity from various calamities. In this era of COVID-19, Wearing masks is a very important thing for further transmission of this deadly virus from one person to another. But many of us are avoiding wearing masks in public places, to overcome this issue we have developed face mask detection software that detects if a person is wearing mask or not. Installing this software module in public places, in work environment and at various crowded areas will ease the surveillance of face detection than manual surveillance by man power.

Deep neural networks have always shown highly successful results when it comes to object detection.

MOBILENETV2 is one such perfect model to be used in facemask detection. Moreover, face mask detection done by MobileNetV2 is very much better when measuring real-time execution in relation to powerful GPU. And the ability to calculate with low memory MobileNetV2 command it is now possible for us to detect faces in real-time mode. The processing strength with low power memory MobileNetV2 command is acceptable for real-time processing. In our research project we have focused on detecting people with face mask and without facemask in real-time. Under the technical and practical conditions, the complete real-time video data detection is finalized with over-availability, localization, and recognition methods. Test results show normal loss is 0.0730 after training 20 epochs. After training 20 points for epochs MAP is 0.96. This is new unique Method of face masking system visual output with 96% categories and accurate detection.

Index Terms - MobileNetV2, DNN, face mask detection, mean standard accuracy, GPU, Computer Vision.

I. INTRODUCTION

Global epidemic situation of COVID-19 is emerged as the dangerous disease the world has ever seen in recent times. The situation all over the world is bad and is getting more worse in almost all the countries declared by WHO. In accordance with this pandemic more than 114 countries are being affected with the flu-like symptoms in under 6.4 days. Millions of sick people are becoming victim to this epidemic in one single day itself. In this time of crisis, everyone should raise awareness and something naturally should be done by your specific activities like self-isolation, maintaining safe social distance. In this case, the government, social authority, and workplace must strictly follow the rules to reduce the further spread of novel corona virus. With these viruses which are spreading to millions of people in a day and becoming viral everywhere by shaking hands, oral contraceptives, and exchanges accessories and others. To make this world a safer place technology has always played a vital role in defending our earth and its humanity from various calamities. For the same aim we have developed a MOBILENETV2 model-based face mask detection software that detects people with masks and without mask irrespective of angle of the face in outdoor and crowded areas. Computer vision surveillance is a real field of identification of an image, modifying descriptive image, output analysis and machine acquisition. Visual reading is closely related to the work done by MOBILENETV2 of any kind of image detection. In addition, MOBILENETV2 detection varies from place to place. Wearing a mask is being recorded by system software that use MobilnetV2 to check if it able to cover and identify only mouth part of the face. Computer vision is the next part of in-depth learning especially the area of the neural convolution network (CNN). CNN also supports very sophisticated configuration Processing Units (GPU), making real-time image or video rendering and recognition a difficult challenge. As we need people to wear a mask or not call the monitoring system as needed powerful validation as video analysis that is complemented by high-quality CNN. Now real-time video or video cloud system in the using R-CNN becomes complex within creating dissatisfaction event. As such, a new category has been added which contains 27 layers of CNN in all 24 convolutional layers fully integrated algorithm which is called as MOBILENETV2. As it costs less to target things, MOBILENETV2 algorithm at a time presented to well-covered experimental researcher's in various parts of the face to see the image detection model which completely captured by MOBILENETV2.

II. RELATED WORK

Initially, small researches from Wuhan and Hubei province also comes from other regions of China show variable rates of diabetes among hospitalized patients with COVID-19, the difference may indicate the age of the sample, the state of the hospital, and the location of the facilities. In Two multicentre, nationwide reports, diabetes was available at 7.4% of 1099 (middle age 47) [1] and 8.2% of 1,590 (age, 48.9 years) [2] people in the hospital. In addition, of 7337 people admitted to nineteen hospitals at Hubei Province (middle age 54), 952 (13.0%) he had type 2 diabetes, and a Chinese report Disease Control and Prevention Centre (China The CDC), which includes those who are not in the hospital, showed low diabetes mellitus (5.3%) among 44,672 confirmed the charges of COVID-19 on February 11, 2020. Many meta- analyses include Chinese COVID-19 patients have confirmed an increase in diabetes is about 8-10%, that is not more, if any rather than the usual,12.9% in 2013 among people aged 40–59 years. In addition, the increase was 11.2% (9.8% after adjustment for heterogeneity) in meta-3 the analysis includes two studies from the US and one study from France and accounted for 23.8% in another meta-analysis where more than half of patients were from a US research centre.

Another meta-analysis in which more than half of the patients were from a study conducted in America. Regarding surveys outside China, a single-centre study in Padua, Italy showed that among 146 hospitalized patients (mean age 65.3 years), the universality of diabetes was 8.9%, again from the same region. The population was not higher than that of the general population. (11.0%). Another single-centre Italian survey from Milan reported a universality of diabetes of 14.9% among 410 hospitalized individuals with COVID-19 (mean age 65 years). In contrast, diabetes prevalence was higher among 1339 COVID-19s from seven hospitals in Madrid, Spain, (mean 69.1 years), compared with 13,390 matched controls (27.2% vs. 20.3%); Raw odds ratio, or, 1.50 [95%] confidence interval, CI, 1.30–1.val3]). Prevalence rates in US patients hospitalized with COVID-19 ranged from 22.6 to 37.2%. These figures are much higher than those reported in normal America, that is, 13.0%, and even higher than those seen among US individuals aged 45–64, i.e. 17.5%, although the median age in most studies was close to the upper limit of this range. However, 7162 COVID-19 cases were reported to the US Centres for Disease Control and Prevention (CDC) with complete information and also included non-hospitalized patients, with overall diabetes prevalence being 10.9%.

Numerous studies and meta-analyses have investigated the effect of diabetes on COVID-19 severity, requiring no hospitalization vs. no need, significant vs. non-serious illness, progression vs. non-progression, ARDS vs. non-ARDS, Defined as need vs. not. ICU admission or the need for mechanical ventilation, fatal versus non-fatal disease, or non-occurrence of combination versus combination occurrence.

The novel Coronavirus 2019 (COVID-2019), which first appeared in the Chinese city of Wuhan in December 2019, quickly spread across the world and became an outbreak. This has had a devastating effect on everyday life, as well as public health and the world economy. Early detection of positive cases is important to prevent the transmission of this disease and for early treatment of affected patients. Since there are no accurate automatic toolkits available, the need for assistive diagnostic equipment has grown. Recent observations from radiology imaging techniques indicate that such photographs provide important details about the COVID-19 virus. The use of sophisticated artificial intelligence (AI) techniques in conjunction with radiological imaging will aid in the correct diagnosis of disease and in overcoming the issue of a lack of specialised physicians in remote villages. A new paradigm for automated COVID-19 identification using raw chest X-ray images is introduced in this report.

The proposed model was created to provide a precise diagnostic for binary classification (COVID vs. No-Findings) and multi-class classification (COVID vs. No-Findings vs. Pneumonia). For binary classes, our model had a classification accuracy of 98.08 percent, and for multi-class scenarios, it had a classification accuracy of 87.02 percent. The Darknet model was used as a classifier in our study so that you can see the (YOLO) real- time object recognition system only once. We applied 17 narrow layers and introduced separate filtering on each layer. Our models https://github.com/muhammedtalo/COVID-19) can be used to support radiologists in invalidating their original investigations, as well as to scan patients directly using the cloud. may leave.

III. PROPOSED METHODOLOGY

COVID-19 is a global pandemic whose conditions have emerged as epidemic in every country across the world. The situation is now under attack in all the countries declared by the WHO and is growing badly. According to this outbreak, the virus as affected more than 114 countries in under 6.4 days. Millions of people are getting sick in a day. During the period of this disaster all should increase consciousness and, if possible, engage in self- isolation. As a result of this problem, the country's governing body and social organizations and the workplace must strictly follow the necessary rules to protect public health. The virus has the ability reach any part of the world and humans being the major part of transmission is what scares the world. Wearing masks, sanitizing hands, workplace is must these days to reduce further more transmission of this deadly virus.

Because of the sensitivity of the subject, we demonstrate our work by identifying who is wearing masks and who isn't, both outside and in crowded environments. Computer vision learning is the real field of identifying images, converting descriptive images, analysing output, and acquiring machines. For the capacity to recognise people, image recognition identifies it by a numeric number. For any form of image recognition, vision learning is inextricably linked to MobileNetV2. Furthermore, MobileNetV2 detection is adaptable in every environment such as family gatherings, workplace, public places and friends in a playground.

Computer vision is a branch of Deep Learning that focuses on convolutional neural networks (CNN). In addition, CNN supports very large configuration Graphic Processing Units (GPU), making real-time image or video retrieval and simulation a difficult challenge. As we need people to wear masks or not, this is referred to as a monitoring scheme there is a need for efficient validation, such as video stream analysis, which advanced CNN can provide.

Cloud machine real-time video or image illustration among R-CNN is becoming complicated, resulting in an unsatisfied event. As a result, the new step included 27 CNN layers across 24 convolutional layers that are entirely connected algorithm is MobileNetV2 as it is affordable to identify tiny artefacts. When the MobileNetV2 algorithm was first presented to researchers, it quickly captured the attention of many who covered a wide range of image detection applications, especially face detection.

IV. METHODOLOGY

In this section we have divided our work into three parts that is data acquisition, data annotation and then MOBILENETV2 setup and training of data. Firstly, data acquisition, this part covers the collection of raw data i.e. images of people with face masks and without masks in different angles. The data collected is from various internet resources like Pinterest. After data is collected pre-processing of data is required followed by data annotation. Secondly, MOBILENETV2 is setup using jupyter notebook and the data is loaded for training and testing using MOBILENETV2 deep learning algorithms

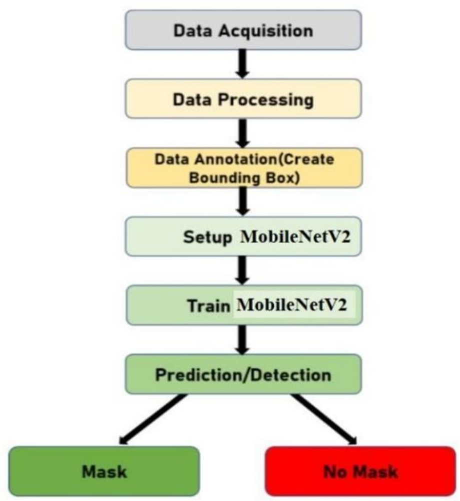


Fig. 1. Workflow Structure.

Fig. 1. Workflow Structure

A. Data Aquisition

Data is an important entity when it comes to data- based technologies like deep learning and machine learning. Image detection is completely based on raw data we collect i.e. images and a lot of it. The more the data is, the well trained our machine or system will be. For instance, the machine trained by using 1000 unique images is well trained than data trained by using 500 unique images.

So, we had to collected more than 3800 approx. images for our system to understand how to detect a face with mask and without mask. We have used web-wrapping tool from our page to collect 3800+ different with face and without face mask images that are captured from different angles.

Next challenge is, these images are not suitable for us to directly load them into algorithm and train our model. So, we need to do some pre-processing cause there are some irrelevant images that are not required for our model and also keeping such images in the model for training will train the model incorrectly so we remove all such images in pre- processing section.

After pre-processing we are left with 3000 images. Half of which are with face and rest are without face

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue V May 2021- Available at www.ijraset.com

B. Data Annotation

Data annotation is also called as image annotation. In deep learning and machine learning techniques data annotation is really important for the model to easily detect the object present in the image loaded.

We used label IMG. To identify images among several present in the dataset which are broadly divided into three categories

1) If the image contains single object, draw a rectangular box around it and identify it as mask or without mask depending upon the image.
2) If the image contains multiple objects, then draw multiple rectangular boxes around it. And identify them as with mask and without mask faces depending upon the image.
3) If the image is shaken or blurry or pic dissertation is observed then identify if object exist and annotate it.

C. Mobilenetv2 Setup And Model Train

MobilenetV2 is setup on jupyter notebook after importing all the required library functions, dependencies and frameworks from TensorFlow and keras, (TensorFlow and Keras are open cv frameworks used in Deep learning and Machine learning techniques).

1) Learning rate, number of epochs to be trained are all batched. We have created 20 epochs each of 96 steps for training the data set. [Fig 3.1]
2) Image is then converted to NumPy arrays for model understanding.
3) Data is then split into train and test data in 4:1 ratio for better model learning so as to result in better image detection with higher accuracy.
4) Loading MOBILENETV2 for model training. A picture is fed into the MOBILENETV2 model as an input. This object detector searches through a picture to locate the coordinates that are present. It essentially separates the data into a grid and analyses the target object's features based on the grid. The features observed with a high confidence rate in neighbouring cells are combined in one position to achieve model performance.
5) Data is then trained which are batched in 20 epochs (20epochs gave us the maximum accuracy of image detection).
6) Predictions are made on tested data and then training loss accuracy is plotted on graph for analysis [Fig 3.4]
7) Trained data is then stored in the form of data test. This trained data is then used by python program for live image detection which is observed in figures [Fig.4.1, Fig 4.2. Fig 4.3. Fig 4.4.]

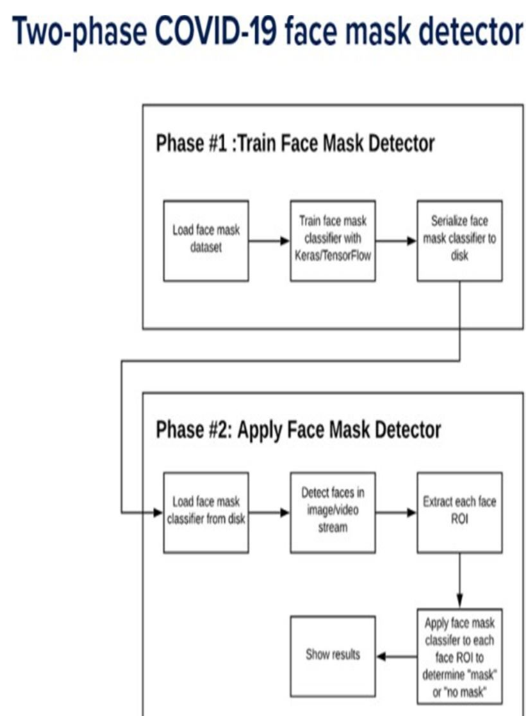


Fig 2. Software Flow

V. EXPERIMENT RESULTS

The model tends to run with Tensor Flow, Keras MobileNetV2 and produced the results with the best accuracy and good learning rate.

Training the head of the network

```
In [7]: # training the head of the network

H = model.fit(
    aug.flow(trainX, trainY, batch_size=BS),
    steps_per_epoch=len(trainX) // BS,
    validation_data=(testX, testY),
    validation_steps=len(testX) // BS,
    epochs=EPOCHS)
94/95 [============================>.] - ETA: 8s - loss: 0.1718 - acc: 0.9374 Epoch 1/20
767/95 [==========================================================================================
==========================================================================================
==] - 59s 77ms/sample - loss: 0.2242 - acc: 0.9061
95/95 [==============================] - 867s 9s/step - loss: 0.1716 - acc: 0.9377 - val_loss: 0.2770 - val_acc: 0.9061
Epoch 5/20
94/95 [============================>.] - ETA: 8s - loss: 0.1623 - acc: 0.9377 Epoch 1/20
767/95 [==========================================================================================
==========================================================================================
==] - 59s 77ms/sample - loss: 0.2048 - acc: 0.9074
95/95 [==============================] - 860s 9s/step - loss: 0.1620 - acc: 0.9380 - val_loss: 0.2587 - val_acc: 0.9074
Epoch 6/20
94/95 [============================>.] - ETA: 8s - loss: 0.1561 - acc: 0.9440 Epoch 1/20
767/95 [==========================================================================================
==========================================================================================
==] - 59s 77ms/sample - loss: 0.2160 - acc: 0.9061
95/95 [==============================] - 858s 9s/step - loss: 0.1549 - acc: 0.9446 - val_loss: 0.2765 - val_acc: 0.9061
Epoch 7/20
94/95 [============================>.] - ETA: 9s - loss: 0.1465 - acc: 0.9457 Epoch 1/20
767/95 [==========================================================================================
```

Fig.3.1

Fig 3.1. The learning rate used with epoch value of 20 and model trained is perfect.

```
94/95 [============================>.] - ETA: 14s - loss: 0.0946 - acc: 0.9647Epoch 1/20
767/95 [==========================================================================================
==========================================================================================
==] - 81s 105ms/sample - loss: 0.2075 - acc: 0.9166
95/95 [==============================] - 1430s 15s/step - loss: 0.0939 - acc: 0.9651 - val_loss: 0.2739 - val_acc: 0.9166
Epoch 19/20
94/95 [============================>.] - ETA: 12s - loss: 0.0852 - acc: 0.9720Epoch 1/20
767/95 [==========================================================================================
==========================================================================================
==] - 98s 128ms/sample - loss: 0.2597 - acc: 0.9035
95/95 [==============================] - 1305s 14s/step - loss: 0.0849 - acc: 0.9720 - val_loss: 0.3249 - val_acc: 0.9035
Epoch 20/20
94/95 [============================>.] - ETA: 13s - loss: 0.0915 - acc: 0.9663Epoch 1/20
767/95 [==========================================================================================
==========================================================================================
==] - 112s 146ms/sample - loss: 0.2006 - acc: 0.9140
95/95 [==============================] - 1372s 14s/step - loss: 0.0913 - acc: 0.9663 - val_loss: 0.2638 - val_acc: 0.9140
```

Fig.3.2

Fig.3.2 The model tends to run with TensorFlow Keras MobileNetV2 and produced the results with the best

	precision	recall	f1-score	support
with_mask	0.99	0.84	0.91	383
without_mask	0.86	0.99	0.92	384
micro avg	0.91	0.91	0.91	767
macro avg	0.92	0.91	0.91	767
weighted avg	0.92	0.91	0.91	767

Fig 3.3

Fig 3.3 The model with precision-recall f1score support values

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue V May 2021- Available at www.ijraset.com

Fig.3.4
Fig.3.4 Training Loss and Accuracy of the model with epoch values of 20

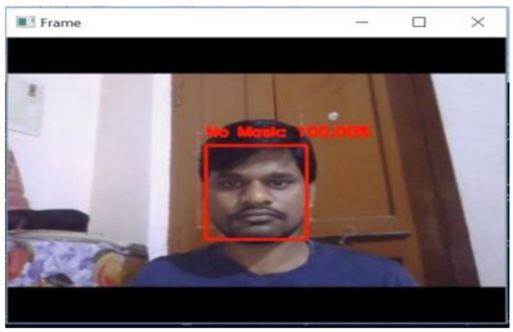


Fig.4.1
Fig.4.1 Identifying the person as no mask

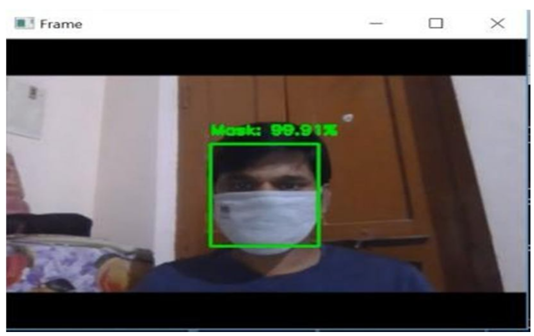


Fig.4.2
Fig.4.2 Mask with accuracy

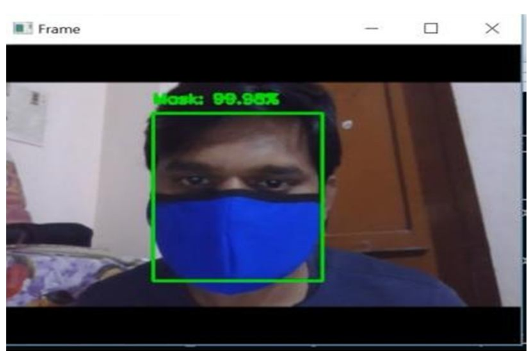


Fig. 4.3
Fig.4.3 Identifying the person with mask and accuracy results with the best

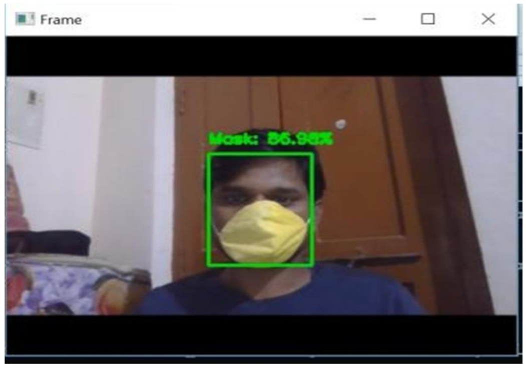


Fig.4.4
Fig.4.4 Identifying the person with mask and accuracy results with the best results with half covered mask

VI. CONCLUSION

In our research, we used state-of-the-art MobileNetV2 architecture to detect whether a human is wearing a mask or not. It works well in photographs, and our detection results were also very good. Later, we applied this concept to a real- time video to see if our model's fps rate within the video and detection accuracy with two groups mask/no mask is appropriate. Inside the video, our model produces amazing results, with an average frame rate of 17. This study focuses on creating a custom object detection model using MobileNetV2 rather than creating the whole architecture. Despite the fact that the dataset we gathered is not very diverse, it provides us with promising consistency in testing with real-world results.

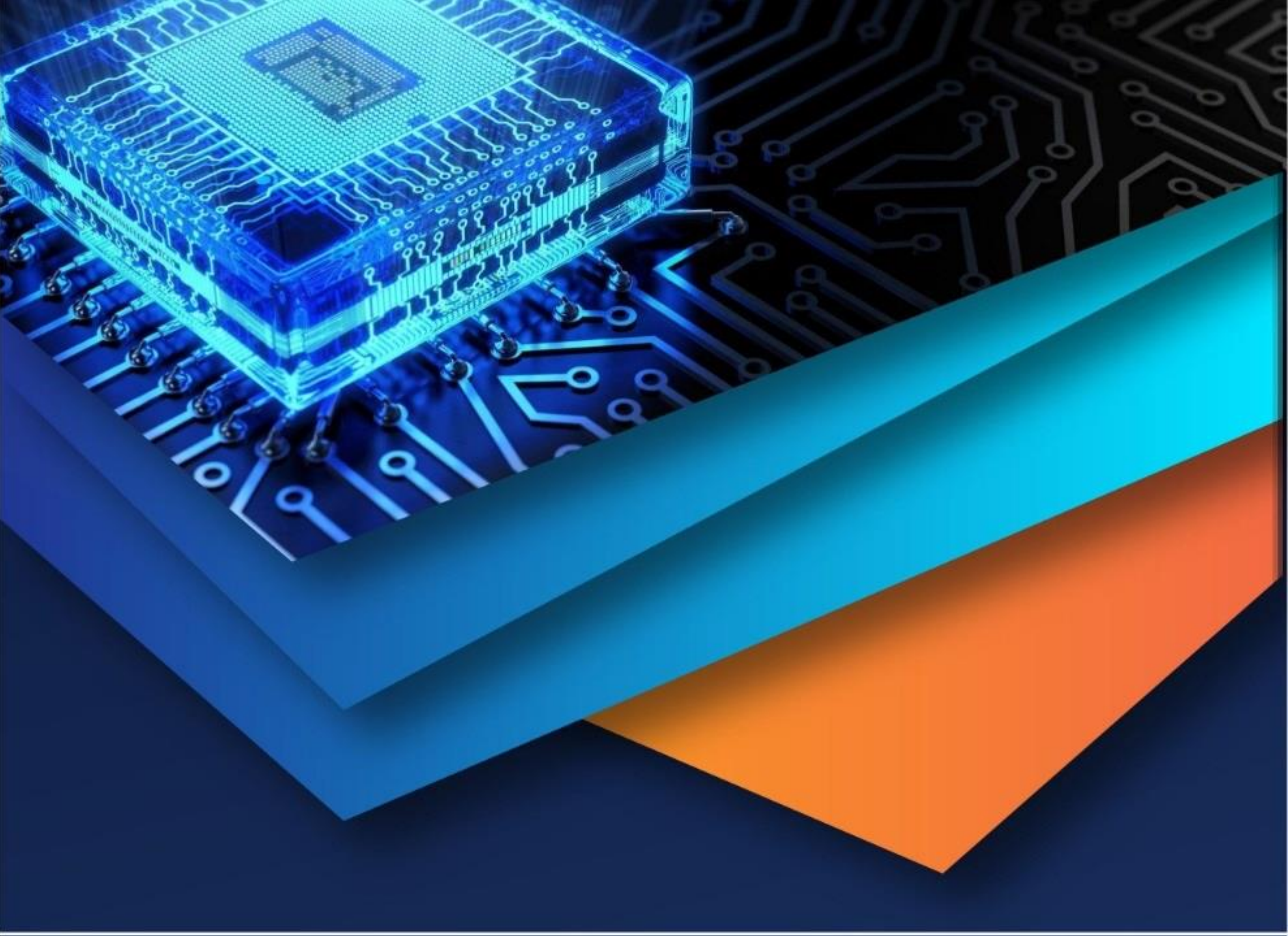
VII. FUTURE WORK

In the future, we will add more data to provide more reliable detection results. Since our resources are restricted, we cannot achieve a higher frame rate in video. In the future, we will train and test our model to make it a better system software to detect faces with more precision and accuracy. Better recognition architectures, such as Mask RCNN and Faster RCNN, have recently been implemented. A new version of MOBILENETV3 has also been released in the last few days. We will use these models to compare their results.

REFERENCES

[1] Hongzhou Lu, Charles W. Stratton, and Yi-Wei Tang. An outbreak of pneumonia of unknown etiology in Wuhan, china: The mystery and the miracle. Journal of Medical Virology, 2020
[2] Chih-Cheng Lai, Tzu-Ping Shih, Wen-Chien Ko, HungJen Tang, and PoRen Hsueh. Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and coronavirus disease-2019 (COVID-19): The epidemic and the challenges. International Journal of Antimicrobial Agents, 55(3):105924, 2020
[3] Hussian A. Rothan and Siddappa N. Byrareddy. The epidemiology and pathogenesis of coronavirus disease (COVID-19) outbreak. Journal of Autoimmunity, page 102433, 2020.
[4] R. Girshick, J. Donahue, T. Darrell, and J. Malik, “Rich feature hierarchies for accurate object detection and semantic segmentation,” In Conference on Computer Vision and Pattern Recognition, 2014
[5] Reinhard Klette, “Concise Computer Vision”. Springer, 2014.
[6] Tulin Ozturk, Muhammed Talo, Eylul Azra Yildirim, Ulas Baran Baloglud, Ozal Yildirim, U. Rajendra Acharya”, “Automated detection of COVID-19 cases using deep neural networks with X-ray images”, 2020.
[7] Joao Carlos Virgolino Soares, Marcelo Gattass, Marco Antonio Meggiolaro, “Visual SLAM in Human Populated Environments: Exploring the Trade-off between Accuracy and Speed of YOLO and Mask R-CNN”, 19th International Conference on Advanced Robotics (ICAR), 2019.
[8] L. Hensley, “Social distancing is out, physical distancing is inheres how to do it,” Global News– Canada (27 March 2020), 2020.
[9] S. Wan, Y. Liang, Y. Zhang, “Deep convolutional neural networks for diabetic retinopathy detection by image classification”, Comput. Electr. Eng. 72 (2018)
[10] A. Koubaa, B. Qureshi, M. Sriti, Y. Javed, and E. Tovar, “A service-oriented cloud-based management system for the internet-of-drones”, 2017 IEEE International Conference on Autonomous Robot Systems and Competitions, ICARSC 2017, April 26-28, 2017, pp. 329–335, 2017.
[11] S. Ren, K. He, R.B. Girshick, J. Sun, “Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks”, CoRR abs/1506.01497, 2015.
[12] E. Dong, Y. Zhu, Y. Ji, S. Du, An improved convolution neural network forobjectdetectionusingYOLOv2,2018IEEEInterna tionalConference on Mechatronics and Automation (ICMA), pp. 1184–1188, 2018.
[13] M.H. Putra, Z.M. Yussof, K.C. Lim, S.I. Salim, “Convolutional neural network for a person and car detection using YOLO framework", 67– 713, 2018.
[14] Bin Wang, S. T.-B.-F.-D., “Detection and tracking based tubelet generation for video object detection. Journal of Visual Communication and Image Representation”, 102-111, 2019.
[15] Dongqing Shen, X. C., “Flame detection Using deep learning”, 2018 4th International Conference Control, Automation and Robotics, IEEE, 2018.
[16] First A. Jiangyun Li, S. B., “Real-time Detection of Steel Strip Surface Defects Based on Improved YOLO Detection Network”, IFACPapersOnLine, (pp. 76-81), 2018.

[17] Jiajun Lu, H. S., "No Need to Worry about Adversarial Examples in Object Detection in Autonomous Vehicles", Computer Vision and Pattern Recognition (cs. CV), 2017.
[18] Joseph Redmon, A. F., "YOLO9000 :Better, Faster, Stronger.", arXiv:1612.0824v1 [cs.CV], Washington: IEEE, 2016.
[19] Joseph Redmon, S. D., "You Only Look Once: Unified, Real-Time Object Detection" IEEE, 2016.
[20] Mennatullah Siam, H. M., "MODNet: Motion and appearance-based Moving Object Detection Network for Autonomous Driving", Computer Vision and Pattern Recognition (cs.CV), IEEE, 2017.
[21] Rayson Laroca, E. S., "A Robust Real-Time Automatic License Plate Recognition Based on the YOLO Detector", 2018 International Joint Conference on Neural Networks (IJCNN), 2018.



10.22214/IJRASET

INDEX COPERNICUS 45.98

IMPACT FACTOR: 7.129

TOGETHER WE REACH THE GOAL

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY