Real time face mask detection using alarm system

Bhoomika H J[1], Soundarya P[2], Surabhi S[3], Vidya T S[4], Dr. Krishnan Bandyopadhyay[5], Dr. Murigendrayya Hiremath[6]

[1,2,3,4] *BE(Medical Electronics), Dayananda Sagar College of Engineering, Bangalore, KA, India*

[5, 6]*Assistant Professor, Dayananda Sagar College of Engineering, Bangalore, KA, India*

***Abstract*—In the context of the COVID-19 epidemic, establishments like the academy square measure in danger of being effectively closed worldwide if the present situation doesn't improve.Coronavirus-2 is a contagious disease that spreads through metabolic process droplets from the diseased person who talk without mask , coughs, or mainly he the person sneezes. This virus catches quickly through close contact with the diseased person or by contact with the infected goods or things. The only way to protect ourselves from COVID-19 appears to be to avoid getting infected with the virus. Preventing infection by using a protective mask that mainly covers the nose and mouth in public areas. Deep Learning has being accepted that its usefulness in image detection and characterization as technology progresses. The study provides additional evidence to use deep learning algorithms for facial recognition and identifies whether or not the subject is wearing a facemask. The dataset gathered consists of 3833 images with a resolution of 224x224 pixels, with a 96% accuracy rate on the trained model's performance. If the person spotted isn't using a facemask, the system develops an amount of time for facemask detection and takes the facial image. This research is useful in preventing virus spread and avoiding eye contact with the virus.**

Keywords- Alarm System, COVID-19, CNN OpenCV, Raspberry Pi, Tensorflow.

I.INTRODUCTION

On a worldwide level, everyone has been aware of a COVID-19 coronavirus epidemic. It is the global economic process of the entire nation. The 2019 Coronavirus Sickness (COVID-19) is linked to an increase in respiratory illness. As of June 10th, 2020, the virus had infected nearly eight million people and killed half a million. To prevent the virus from spreading, they've established WHO recommended practices such as wearing a facemask at all times, keeping a strict social distance from public venues, and cleaning hands. There has been work that show that wearing a facemask is critical in preventing the virus from spreading. According to research, the efficiency of N95 and surgical masks in limiting virus spreading is ninety one percent and 68 percent, respectively. These masks may successfully interrupt mobile viruses, preventing infections from reaching a person's system, and they are inexpensive due to the reduction in fatalities and infection problems. Nonetheless, due to insufficient facemask use, the overall efficacy of facemasks in reducing the transmission of pathogens in public has been reduced. Deep neural networks (DNNs) has been used because the major element of deep learning strategies offers everything together with detection of the object, classifying the images, and segmentation. Convolutional neural networks (CNNs) is the most common deep network, but they're commonly used in computer vision tasks. Mobile-Net a class of CNNs will build and classify face images even with little differences after being trained, given their enormous extracting features ability, and maintain image pattern details once the model has been trained. Deep learning methods have been used in this research analysis to build a classifier that will collect photographs of people wearing masks and not wearing masks from the data and distinguish between the two categories of facemask wearers and non-face mask wearers. The neural network is the excellent way to extract features from unprocessed data. This study develops using a convolutional neural network to style the facial mask classification and incorporating the effect of the neural network layers number on the accuracy. This application is done on a Raspberry Pi using Open Computer vision, Tensor Flow and Python language. It's a Python programming language that comes pre-installed.

II.RELATED WORK

Pattern Haar cascade classifier algorithms were used to detect human faces in the main eyes and mouth. Three more weak classifiers were later added to the fundamental Haar-like choices-based cascaded classifiers. A decision node that allowed human skin color bar graph matching could be one example. The eyes and mouth detections were supported by the second and third weak classifiers units of measurement, respectively.

A simplified approach to mask detecting patterns using TensorFlow, Keras, Open CV, and Scikit-Learn, as well as other machine learning packages. The planned methodology correctly detects the human face from the image and determines even if there isn't a mask on it. It'll even notice a face with a mask moving as an investigative task entertainer. It will be utilized for a variety of purposes. Because of the Covid-19 problem, wearing a mask is also required as part of future precautions. The ideas and innovations will make a significant contribution to a final public health care system.

Throughout this project, an AI-based device (Raspberry Pi with AI model and camera) will be used to determine whether or not a private is wearing a mask and send a corresponding alarm message to the United States of America (via mobile app). A smartphone app is included with this gadget. When people do not seem to be really present in their home, mobile apps detect if someone enters. If people wear masks, this clever technology will instantly unlock the door. This device works 24 hours a day, 7 days a week. It will be used in a variety of places, including malls, businesses, hospitals, and temples.

The YOLO V4 rule was used to create a mask detection. The YOLO V4 rule is based on a deep learning technology that can correctly detect the factor. This device has already been installed at Politeknik Negeri Batam as a real-time application to prevent COVID-19 from spreading throughout the sector. From their experiment results, the rule was able to notice and distinguish between a non-wearing and a wearing mask specifically with any condition of the skirting surroundings.

III.METHODOLOGY

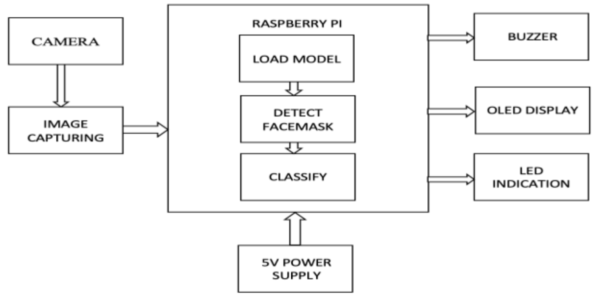


Figure.1.Block diagram of the current work

Real-time images are captured by camera. The captured images are then fed into raspberry pi for processing. The model detects and classifies the images with mask and without mask. The output will be a message 'Please wear the mask' will be displayed on the OLED screen coupled with a glowing red LED and If the user is not wearing a mask, a beep will sound as an alert. If the person is wearing the facial mask, the words 'Mask detected' will appear on the screen, along with a green led indicating that the person is permitted to enter.

A. Materials and methods

1.Raspberry Pi and Pi camera

The Raspberry Pi may be a tiny PC that may be used for a range of functions, together with AI, homes, networked AI cores, industrial plant controllers, and more. The last Raspberry Pi four model has the utmost effective hardware. The camera pi is intended to figure along the preceding board. The Pi Camera connects via USB to the Raspberry Pi's CSI association. It's a 5MP resolution and may provide crystal clean pictures. It's easy to line up and utilize.

2.OLED

Organic Light-Emitting Diodes (OLEDs) are a flat light-emitting device comprised of a sequence of organic thin sheets placed between conductors. While electrical power is supplied, a brightly illuminated light is emitted. OLED displays are emissive displays that do not require a backlight, allowing them to be smaller and more energy-efficient than LCDs. The LCDs is utilized in the output stage to show the message based on whether or not it is wearing the masks.

3.Buzzer

A buzzer or electrical audio device that could be electromechanical, mechanical, or piezoelectric in nature. Alarm clocks and timers are examples of common buzzer applications.

4.Dataset

The data set comes from Kaggle and consists of a large number of images of people's faces with or without masks. Face collections include head turns, tilts, and slants with multiple faces in the frame, as well as numerous masks in a variety of colors.

Incorporated packages

1.Tensor flow

TensorFlow is a platform for expressing machine learning algorithms that is used to fabricate ML systems in a variety of computer science fields, including sentiment analysis, image recognition, machine translation, recognizing voice, information extraction regarding geographic, computer vision, summarization, gathering information, drug discovery using computational power, and flaw detection. Tensor Flow is used as the backend for the entire architecture in the designed model (which consists of numerous layers). It's also used in data processing to reformat information.

2.Keras

Keras delivers key insights and building blocks for the rapid generation and transmission of ML setups. Tensor Flow's scalability and cross-platform features are fully utilised .Keras primary data structures are layers and models. Keras is used to implement all of the layers in the CNN model. It aids in the overall model compilation as well as the data processing converting the vectors into the matrix in the binary form.

3.OpenCV

OpenCV (Open-Source Computer Vision Library) is a computer vision and machine learning software library that can be used to recognise and distinguish faces, in recordings, tracing the modules, analyze the eye movements, camera actions can be tracked, perceive landscape and overlay it with enhanced reality, and so on.

Our work makes use of Open Computer Vision's features to resize and color transform data images.

B. Software flowchart/algorithm

The process is to mainly classify from learned data and consists of 5 major steps.

1.Collecting the Data

Capturing video and recognizing the face is the initial step in our project. Pi camera is used to capture high-quality live stream video of the people with and without facemasks.

2.Image Pre-Processing

In the pre-processing step there are four main steps such as resizing the image, convertion from image to array, processing using mobilenetv2, and then finally encoding the labels using one hot encoding.

The face detection model works mainly with the gray-scaled images instead of dealing with RGB images which gives more impact in accuracy and the precision of the project. Gray-scaled images are used for feature extraction and will be followed by a pre-processing step in order to increase the specificity of the features in the Image. Image segmentation process is involved in which the pre-processed images undergo dividing these images into a number of segments and emphasizing the facemask-covered sections.

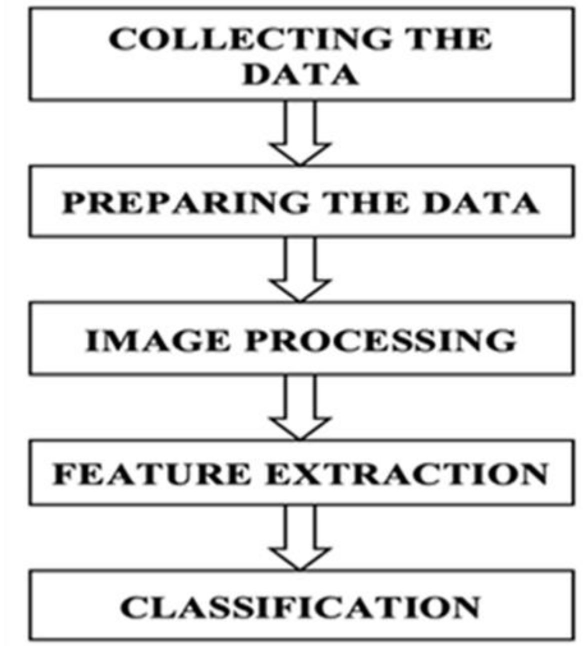


Figure.2.Flow chart of the proposed work

Image reshaping is another important phase along with convertion to gray-scaled image in which the classification model requires a specific same size for all the images in the dataset. The images are resized to 244 x 244 using the function cv2.resize().

The input to the model is the three-dimensional tensor with all the channels containing a unique pixel values. Therefore, they are adjusted in such a way that pixel value must range between 0 and 1 which is technically called normalization. The data is converted as categorical representation. Hot encoding is performed on labels because no machine learning algorithm can directly operate on labels. It is required that all the input and output variables must be only in the numeric form.

3.Feature Extraction

Feature extraction involves obtaining the important features from the processed images and is combined with each of the layers along with ReLU. To reduce the computational power both deep wise convolution and point convolution. The layers of the mobileNet model converts pixels of the image into the features that represent the image and pass to the next layer.

4.Classification

In order to classify the images the model has to be trained perfectly with the properly annotated images on how to perfectly classify with and without wearing the mask. Therefore, classification is the very crucial step in any machine learning techniques.

Using Mobile-Net architecture to classify the images in order for the model to classify properly it has to be trained precisely. The first mobile model for computer vision is the Mobile-Net model a class of CNN which is designed mainly for mobile applications. This is mainly preferred as it reduces the number of parameters and thereby results in lightweight networks, increase the accuracy. Mobile-Nets are usually small, have less latency they are build for detection and classification.

Mobile-Net contains two main operations namely Depth-Wise convolution and the pointwise convolution.

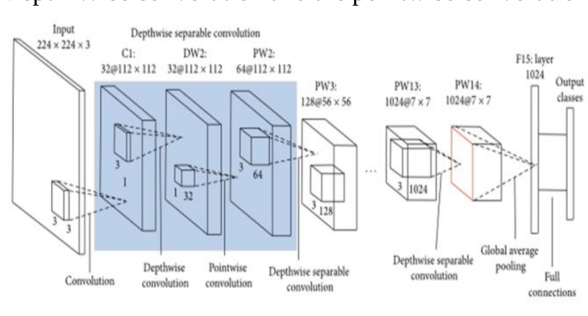


Figure.3.Convolutional layers

It is basically the channel wise special convolution that is DK*DK. It is process of mapping the single convolution on each and every input channel individually. Hence the total count of the output channel is always same as the total number of input channels. Point wise convolution are the kernels with the size 1*1 that just connects the features obtained from the depth wise layer and builds new features by performing the input channels in a linear combination. Mobile-Net splits the convolution layers into 3*3 depth wise convolution layers and then followed by batch normalization, ReLU and then the 1*1 pointwise convolution layer along with batch normalization and ReLU instead of a single 3*3 convolution layer and then individually followed by the batch normalization and rectified linear unit. And finally, the global average pooling is used in order to obtain one feature map for each of the corresponding label category of the classification.

The main idea is instead of summing the fully connected layers on to the feature maps we simply take average of each of the map and pass it on to the output layer.

Generally, Depth wise separable convolutions are the replacement of the traditional layer of CNN as the effectively reduce the computation power compared to other layers so it can be said that the computational cost will be 9 times smaller than that of regular layer of convolutions.

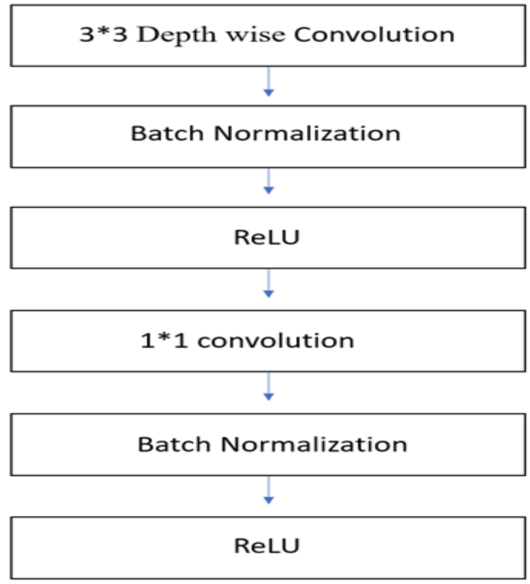


Figure.4.Depth-wise convolution and point-wise convolution

Model training

Transfer learning is a technique used to basically speed up the network to train the model. It is the efficient training process using the pre-trained model which is been already learned on the similar dataset. Here image-net dataset which is widely used vision technique is applied to the model as the weights. Therefore, using the Mobile-Net as the backbone and transfer learning the performance and accuracy can be increased. One of the advantages of using this type of weights is the feature extraction capacity.

We used tensorflow libraries to train the mobilenet model. Adam optimizer with learning rate 1e-4, 20 epochs and 32 batch size is used. Batch normalization is used after every layers as shown in the figure.

IV.RESULTS

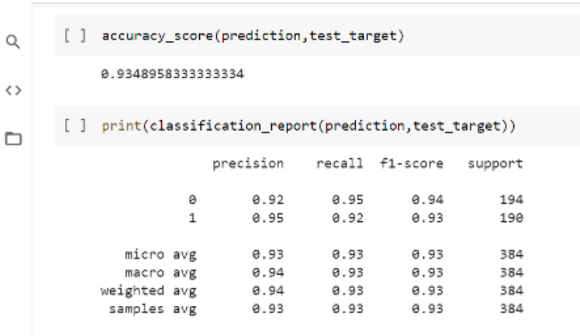
During the MobileNet model's training, it attained a validation accuracy of 93%. As indicated in the figures below for performance testing from visualization to accuracy and loss, After multiple testing with a batch size of 100 iterations for each epoch, this is the greatest percentage seen.

When these factors are evaluated, we can conclude that hardware processing power is critical and has a direct impact on the deep learning model's performance. Particularly in real-time scenarios.

The suggested system determines whether or not the user is wearing a face mask.

The message 'Please wear the mask' will be displayed on the OLED screen coupled with a glowing red LED and a beep sound as an alert if the person is not covering the face with the mask.

If the person is wearing a mask, the statement 'Mask identified' will appear on the screen, along with a green led that signals the person is allowed to enter.



```
[ ] accuracy_score(prediction,test_target)

    0.9348958333333334

[ ] print(classification_report(prediction,test_target))

                  precision    recall  f1-score   support

               0       0.92      0.95      0.94       194
               1       0.95      0.92      0.93       190

       micro avg       0.93      0.93      0.93       384
       macro avg       0.94      0.93      0.93       384
    weighted avg       0.94      0.93      0.93       384
     samples avg       0.93      0.93      0.93       384
```

Figure.5.Depth-wise convolution and point-wise convolution

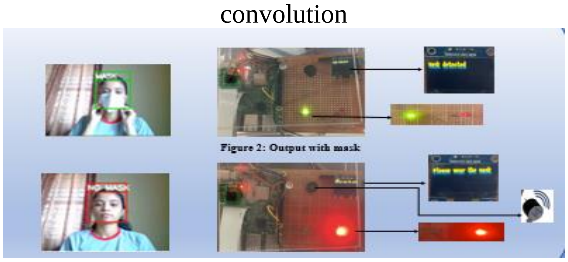


Figure.6. Output with and without mask

V.CONCLUSION

To summarize, in our study, we developed a facemask detection prototype employing with the related devices and sophisticated Artificial intelligent models. We fine-tuned, trained, and evaluated the facemask detection system using MobileNet pre-trained face identification models. Using a Raspberry Pi board and a pi camera, we were able to detect people's faces and determine whether or not they were wearing masks. When a detected face is not wearing a mask, a real-time beep sound is produced and the message is shown.

Future work could include the combination of physical distance measurement, in which the camera detects whether or not the person wears a facemask while continuously taking measurements between each person and sounding an alarm if the physical distance is not maintained. The system can be Integrated also with a contactless temperature sensor to check the body temperature. An automated gateway system can be implemented to restrict the entry based on conditions provided. Additionally, vocal alerts can be customized depending on the language of the place.

When it comes to identifying the facemask, our system delivers precise results. Using the MobileNet model, the trained model was able to complete its task with 93% accuracy. Furthermore, the study proposes a necessary tool in combating the propagation of the Corona virus by recognizing whether or not a person wears a facemask and displaying the appropriate message.

REFERENCES

[1] Sammy V. Militante, Nanette V. DionisioReal-Time Facemask Recognition with Alarm System using Deep Learning 2020 IEEE

[2] Ikram Ben Abdel Ouahab, LotfiElaachakReal-time Facemask Detector using Deep Learning and Raspberry Pi 2021 IEEE

[3] Baluprithviraj.K.N, Bharathi.K.R Artificial Intelligence based Smart Door with Face Mask Detection 2021 IEEE

[4] World Health Organization. Coronavirus disease 2019 (COVID-19) Situation Report– 142, 2020.