FOREX Publication
Open Access | Rapid and quality publishing

International Journal of
Electrical and Electronics Research (IJEER)
Research Article | Volume 10, Issue 3 | Pages 751-759 | e-ISSN: 2347-470X

IoT-Deep Learning Based Face Mask Detection System for Entrance and Exit Door

Simon Kasahun Bekele[1], Million Gonfa Gutema[2], Ebisa Damene Tujuba[3], Rituraj Jain[4*] and Yohannes Bekuma[5]

[1,2,3,4,5]*Department of Electrical and Computer Engineering, Wollega University, Nekemte – Ethiopia, leekasahun@gmail.com[1], milliongonfa29@gmail.com[2], eba12dem5@gmail.com[3], jainrituraj@yahoo.com[4], yohbek@gmail.com[5]*

***Correspondence:** Rituraj Jain; jainrituraj@yahoo.com

ABSTRACT- During the pandemic, it has been seen that the global population follows the guidelines issued by the health organization regarding wearing face masks, but some people do not take care of this and do not use masks. The objective of the proposed system, Wollega University Face Mask Detection System (WUFMDS), is to restrict people who are not wearing a mask on the door side by identifying the face mask from the face or open the door if the incoming person is wearing the mask. the This system is based on the Internet of Things (IoT) and a Deep Learning algorithm called Convolutional Neural Network (CNN). For this purpose, images with and without masks were collected as samples from the university. The CNN algorithm is used to detect the mask and classify it as with or without masks. The IoT module controls the door operation based on the classification response sent to the IoT module by the CNN algorithm. The system was tested lively with the dummy door system in order to ensure the functionality of the face mask detection system and developed software applications for the system model are working as defined objectives. Our model had 99.36% accuracy with the training dataset and 99.29% accuracy with the validation set. Hence, the proposed system could be used for the automatic identification and classification of masks on the face and to operate the door to allow the person who is wearing the mask to pass through while keeping it closed when no mask is found on the face.

Keywords: Convolutional neural network, deep learning, face mask, internet of things, tensor flow.

ARTICLE INFORMATION
Author(s): Simon Kasahun Bekele, Million Gonfa Gutema, Ebisa Damene Tujuba, Rituraj Jain and Yohannes Bekuma;
Received: 22/06/2022; **Accepted**: 20/09/2022; **Published**: 30/09/2022;
e-ISSN: 2347-470X;
Paper Id: IJEER180622;
Citation: 10.37391/IJEER.100356
Webpage-link:
https://ijeer.forexjournal.co.in/archive/volume-10/ijeer-100356.html

crossref member
CROSSREF.ORG
THE CITATION LINKING BACKBONE

Publisher's Note: FOREX Publication stays neutral with regard to Jurisdictional claims in Published maps and institutional affiliations.

1. INTRODUCTION

Before the pandemic (COVID-19), there was no such major supporting evidence that came into light to enforce the global community to wear the mask to restrict deadly viruses from spreading. Wearing the mask makes sense as this will help to some extent to control the spread of airborne respiratory infections that are primarily spread through the inhalation of respiratory aerosols resulting from sneezing, coughing, or talking [1]. The most effective ways suggested by all the health organizations are using face masks as well as sanitizers, which show good results in the transmission of such highly risky and deadly diseases. All over the world, governments of all countries have implemented so many initiatives to reduce such disease transmission [2]. In the current scenario, it was not a good choice to use available traditional systems like fingerprint or facial recognition base systems, as these systems can lead to the spread of deadly viruses. An automated process for detecting a person wearing or not wearing a mask is critical not only now but also in the future to protect against the spread of deadly viruses.

Machine learning methods provide the most exciting and convenient set of approaches, which are subsets of artificial intelligence (AI), where machines are required to train multiple times for the new dataset until the desired outcome is obtained. Deep learning (DL) is one of the merging techniques that uses artificial neural networks (ANN) as the backbone of its processing and comes as a subfield of machine learning (ML) technique. DL is based on different neural networks used to explore a wide domain in AI. The CNN is another DL approach used to train and classify the data set in a robust manner [3].

All governments try to ensure that face masks are worn in public places, but it's tough to manually determine those that don't seem to be using face masks in crowded places. Existing face recognition systems are very effective at identifying faces with or without masks. So, to face all the problems during pandemic time, it is necessary to develop a facial recognition system that is capable of identifying the availability of the mask on the face. So, we proposed and developed a smart system named Wollega University Face Mask Detection System (WUFMDS) that will check automatically whether a mask is worn by the person who wants to enter or not and send this response to an IoT module to take the decision to open the door if he/she has already worn a mask. The problem statement can also be concise as follows: the classification model must classify the input image, captured from the camera in a real scenario, into mask or without mask categories from the mask detection model and operate the entrance/exit door accordingly. IoT based system for managing the images captured by the device and deep learning processing for mask detection-driven face picture classification, the IoT module of the system will control the door by sending the required signal based on the classification done by the DL model. In the proposed research work, we have used MobileNetV2 with TensorFlow, TensorFlow Lite, Keras, and OpenCV libraries. The system has been balanced with a limited dataset and good recognition accuracy.

The following are the contributions of the proposed system developed through this research:

1. A new system has been developed with the fusion of IoT and CNN for an automatic identification of the facemask.
2. An in-house dataset was generated to support the research to overcome the scarcity problem of the dataset.
3. With this in-house collected dataset and good identification accuracy, the system is adequately balance for the defined problem statement.

2. CNN ARCHITECHTURE

A specific type of deep neural network known as convolutional neural networks (CNNs) can recognize and do the classification of particular features from image data sets. The CNN architecture is composed of 3 types of layers, which are the convolutional layer, the pooling layer, and the fully-connected layer.

The convolution layer does the convolution operation to identify the major features from the input image data set. This layer contains convolution kernels, which have learnable filters (convolution kernels). Internal features of the images and deep valuable information from the data can be extracted by using these convolution kernels. The dot product between the kernel's weight and the local region will be used to verify the output of the neurons. This convolution process can be defined by *eq. 1*.

$$z = w_i^l * y^{(j)} \tag{1}$$

Where,

*is the dot product between the kernel's weight and local regions.

w_i^l = weights of the i^{th} filter kernel of layer l.

$y^{(j)}$ = j^{th} local region of convolutional layer l.

A Convolution operation must be followed by an activation function, i.e., sigmoid, Rectified Linear Units (ReLU), etc., to give a nonlinear factor to solve complex problems using neural networks. The outcome of this function is the activation map, which provides the responses of the filters. ReLU gives more advantages over common functions like sigmoid and tanh as it is used to train neural networks in a very small amount of time, which in turn reflects less calculation cost. The ReLU function can be defined as in *eq. 2* [4].

$$g(y) = \begin{cases} 0 \; for \; y < 0 \\ y \; for \; y \geq 0 \end{cases} \tag{2}$$

Similar to this, the SoftMax activation function is used to add the final layer of the neural network to improve its performance. *Eq. 3* [4] defines the SoftMax function.

$$\sigma(c)_j = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}} \tag{3}$$

With convolution layers, CNN also has pooling layers, which take the output of the convolution layers. The objective of the pooling layer is to cut some of the dimensions to compress the feature map and keep only useful information. It will lead to calculation acceleration and the prevention of overfitting. After many combinations of convolution and pooling layers, the high level of perceptive in CNN is completed with fully connected (FC) layers. These layers comprise weights and biases accompanied by the neurons usually used to interconnect neurons from two layers. The last layer of the fully connected layer has a number of output neurons equal to the number of the classes to be identified. The classification process uses this layer to take place.

3. LITERATURE REVIEW

Researchers presented an approach in [5], which used a 25000-image dataset of 224 x 224 pixels and got 96% accuracy. Their approach used artificial neural networks and Raspberry Pi to train systems regarding facemasks and alarm systems if a person comes without a mask, respectively. A method proposed in [6] uses a combination of two sets of datasets of medical facemasks as confined to one dataset to find the accuracy between them. By calculating the average precision of two datasets and combining YOLO-v2 with the ResNet-50 model, they were able to obtain a high-precision average. Their experimental results, which are based on the loss identified in each epoch as well as on the accuracy of validation, show the effectiveness of their approach. In [7], three phases, which are artificial face mask dataset creation, face mask detector training, and face mask detector testing-based system implementation, are described. They created their own dataset, which consists of 600 images of without mask cases, and obtained another dataset by applying an AI technique to put facemasks on these without mask datasets.

Another approach is defined by using Dlib and the OpenCV library in [8]. It uses the hog approach for searching and recognition to improve the efficiency of assessment of these libraries of computer vision. The coordinates of the face boundary were fetched by using the OpenFace library, and the results were divided into different facial groups by using 128 face-extracted features, which also helps them to get results with more accuracy. A method proposed in [9] uses three different datasets, namely the Real-World Masked Face Dataset (RMFD), the Simulated Masked Face Dataset (SMFD), and the Labeled Faces in the Wild (LFW), to compare the accuracy of these datasets by using them with the same algorithm set. This defined system is operated over Resnet50 and through known ML algorithms like decision trees, Support Vector Machine (SVM), and ensemble algorithms. While Resnet50 components are used for extracting the features, classification is done by using stated machine learning methods. According to them, by using this system they were able to achieve an accuracy of 99.64%, 99.49%, and 100% in RMFD, SMFD, and LFW, respectively.

Another system with availability of SMS based alert system was proposed in [10]. After screening a person without a mask via CCTV, SMS alerts are sent. This system uses CNN for the detection process of masks, and Amazon Web Services (AWS) is used to keep capered images properly. SMS alerts are generated by using the Twilio messaging API. An ARM-Cortex M7 microcontroller and OpenMV cam H7 controller camera-based architecture with a CNN based solution has been defined in [11] for face mask detection on edge devices with high resource constraints. This proposed model makes use of a 12232-image dataset from Kaggle, as well as a 1979 dataset set generated by the authors using an OpenMV cam H7 controller camera. To increase the dataset volume, this dataset was augmented and resized to 1,31,055 and 3232 pixels, respectively and made them the optimal size suitable for the microcontroller. According to the authors, their approach resulted in 99.79% accuracy and is supposed to be the best fit model for RAM-constrained microcontrollers.

An improved Mask R-CNN named G-Mask framework is proposed in [12], which uses ResNet-101 for extracting the features, RPN to generate RoIs, and RoIAlign to preserve spatial locations to generate a binary mask through a fully convolution network (FCN). This framework was trained by using the Keras framework with the help of a 5115-image sample dataset, which was resized to 1024x800, 3000 steps per epoch with a 0.001 model learning rate. A mobile microscope-based detection system has been developed in [13]. It takes micro-photos of the facemask used by a person. Feature extraction is done by using the grey level co-occurrence matrix (GLCM) of the facemask's micro photos, which will be used by the KNN algorithm for result detection. A method involving using ML packages like OpenCV, TensorFlow, Scikit-Learn, and Keras is defined in [14]. It uses two datasets to compare the loss and accuracy between the two datasets. In their approach, they used one dataset of 1376 (690 with mask and 686 without mask) images and another dataset comprising of 853 images (classified as with mask and without mask) from Kaggle. The system works with training of 20 epochs and 90–10% data splitting of the datasets.

An approach defined in [15] is used to detect the person's face from a video stream using adaboost, a haar-like algorithm, and OpenCV. Objects are detected by using OpenCV, sample images' training is done by using Adaboost, and bounded coordinates of the face are extracted by using a Haar-like algorithm. An android-based system has been developed [16] for gender, age, and face recognition. This system used the OpenCV library for detection and recognition. For face detection and face recognition, this system uses the LBP face feature classifier and the LBPH model, respectively. Gender and age are identified by using deep neural networks without taking images into account and training the respective model. Featured are inputs to the 3-layer CNN architecture, where each layer has 96 filters, 256 filters, and 384 filters, respectively. The efficiency of this system is influenced by factors such as human gestures, facial expressions, illumination, face coverage, and so on.

Researcher in [17] proposed a Deep Learning-based Safe Social Distancing and Face Mask Detection in Public Areas for COVID-19. This computer vision-based system is a transfer learning implementation with a Single Shot Detector (SSD) and a lightweight neural network, MobilenetV2. Hence, it achieves a balance between resource limitations and recognition accuracy. Authors proposed a run-time facial mask detection system based on convolutional neural networks (CNN) and computer vision in [18]. Face masks are crucial in times of a COVID-19 pandemic. The proposed model does image preprocessing to convert the RGB image into a grayscale image. A deep learning architecture called CNN is used to identify whether people are wearing masks or not. This approach helped in dealing with the preprocessing aspect while training a model. In the face detection technique used in [19–20], several attributes from the given input image can be used, including face recognition, pose estimation, face expression, and pose estimation. Due to variable changes in many of the attributes, like colour or structure, it was a tough task. The challenging task of this process was to correctly identify the person's face with the mask in the image. Even if only used for scrutiny purposes, the system must be able to identify the human face and the mask. With the fusion of deep learning and computer vision, another system proposed for the detection of facemasks was presented in [21]. The proposed system identifies faces of individuals wearing masks from video/image frames using image processing, OpenCV, and deep learning frameworks like TensorFlow, Keras, and PyTorch. A model trained with MobileNetV2 was applied to identify the mask from the images or video streams.

4. METHODOLOGY

Online and offline processing are the two defined approaches in the proposed approach. While offline processing's responsibility is to defined a DL model to classify the captured image with or without facemasks, the online phase is responsible for deploying the DL model to the application created for the proposed system and IoT system to capture the images from the installed camera device and provide them to the offline DL model for further classification. *Figure 1* depicts the two phases of the proposed approach. The two sets are prepared for training and testing purposes. They are customized sets of on-face masks and off-face masks. The goal of this experimentation is to train the first set and test the second set to detect the face masks for safety. Finally, with the help of some libraries such as Keras and Tensorflow, we will detect some real-time face masks, whether they are on/off.

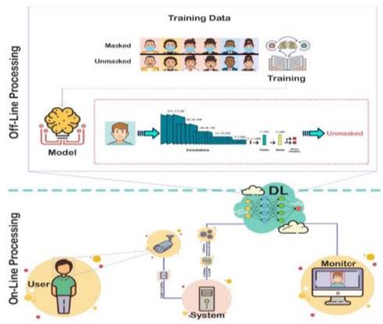


Figure 1: Two phases of the proposed approach [21]

4.1 Facemask Detection Model

In this proposed system, a deep learning algorithm called convolution neural networks (CNN) classification is used for the detection of face masks. A CNN may be a variety of artificial neural network that's specifically engineered to interpret element input and is principally used for image recognition and analysis, during which every layer applies to a special set of filters. Results for each layer are obtained by combining 100 to 1000 filters. These obtained results are finally sent to the neural network's next layer. As shown in *figure* 2, proposed approach has been evolved for the facemask detection model with the help of Keras and Tensorflow software libraries for training. Dataset collection, preprocessing, splitting, training, and testing/evaluation are the different steps to be used by the defined model.

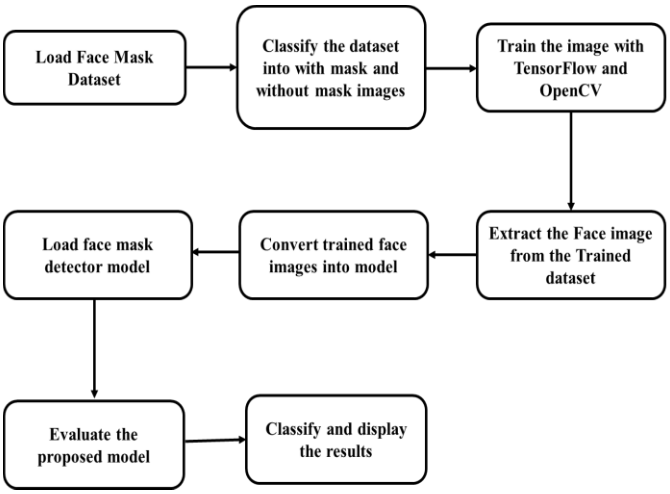


Figure 2: Flow diagram of the proposed system

The proposed model takes sample images of people wearing masks and without face masks as input. The first convolutional layer takes input in the form of images with some activation function. The pooling layer is stacked up after the activation function layer. For the reduction in the spatial size of representation, a max pooling layer is added. The proposed model has a second convolutional layer with the same sequence of layers as the first, viz., an activation layer, a pooling layer, and a max-pooling layer.

As stated in *section 2* and shown in *figure 3*, the CNN architecture comprises an input, an output, and hidden layers consisting of convolutional with ReLU, pooling, and fully connected layers. This network's input data sets include images with and without masks, as well as three RGB channels of the input data set is 224*224. This input data set is given to the first layer, i.e., the convolution layer added with the ReLU activation function. The first layer has a kernel size of 224*224 and 32 output channels. The output of the ReLU function is 0 on getting a negative input, while it returns y for a positive x. The next layer is the max pooling where the defined window size is 112 *112 and there are 32 output channels to reduce the feature map. For each window, max values are going to be recorded while the window is filtered along the stride. This action will be repeated for all the filtered images. The third layer of the network consists of a convolution layer which has a kernel size of 112*112, as well as a pooling layer with a kernel size of 7*7 preceding the fully connected layer. Additionally, a flattened and fully connected layer in the network. A fully connected layer is added with the SoftMax function. Prediction of the class of an image (with or without mask) will be done at the last layer.

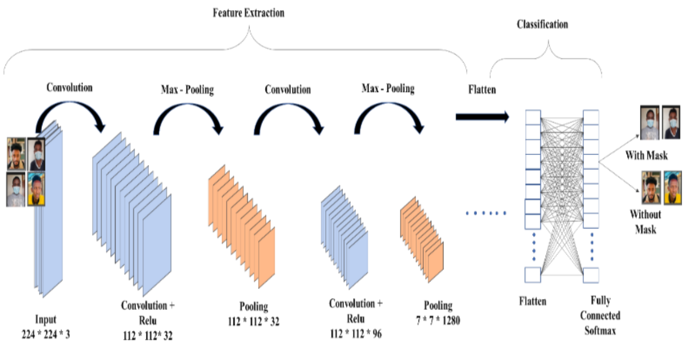


Figure 3: Architecture of used convolutional neural network model

4.2 Operational Technology

Figure 4 and *figure 5* show the hardware and software architecture models, respectively, of the proposed system. *Figures 6* and *7* show the working flow graph of the system at the entry door and exit door, respectively. While designing this system, the selection of appropriate elements, whether hardware or software, is required to achieve the required set points. This system is mainly composed of the Ardiuno Mega, Passive Infrared Receiver (PIR) sensor, ultrasonic sensors, camera, servo motor, Bluetooth module, dummy doors etc.

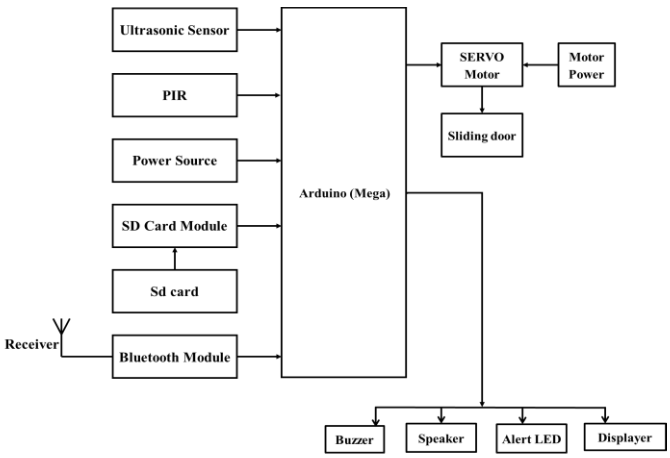


Figure 4: Hardware architecture model for the face mask detection system

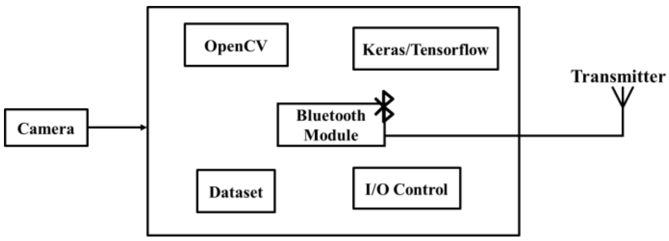


Figure 5: Software configuration architecture model for face mask detection system

One ultrasonic sensor and one camera will be used in the entrance door, where the ultrasonic sensor will be used to measure distance and the camera will detect individual people. One more PIR (passive infrared receiver) sensor will be used in the exit door along with one more ultrasonic sensor. In the exit

FOREX Publication
Open Access | Rapid and quality publishing

International Journal of
Electrical and Electronics Research (IJEER)
Research Article | Volume 10, Issue 3 | Pages 751-759 | e-ISSN: 2347-470X

door, PIR is used for detecting human presence, and ultrasonic is for measuring a distance from inside. A live video stream camera is interfaced with an Arduino microcontroller for detection of live image feed and, based on the deep learning process, will do analysis and verify that a person coming in is actually wearing a mask or not. Based on the classification result, the door will be open as well as display in text on the displayer if the person is deaf or to tell him/her with a speaker that the person is blind. When someone arrives without a mask, a buzzer and a light are used as an alert system.

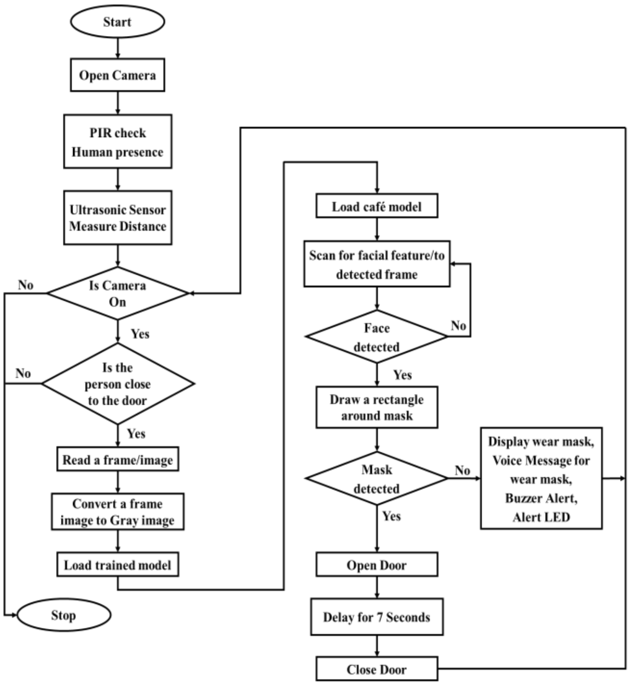


Figure 6: Flow chart of the system working at the entrance door

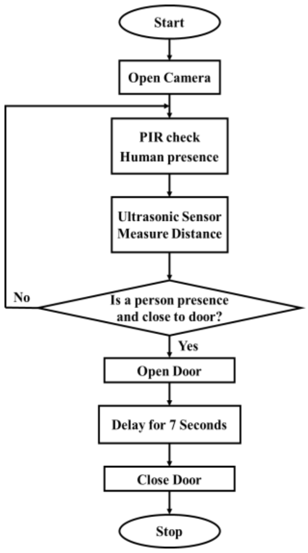


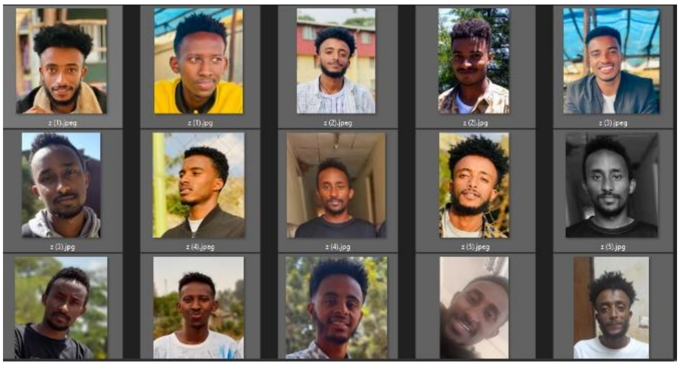
Figure 7: Flow chart of the system working at the exit door

4.3 Dataset Collection

The dataset used for this research was collected from our own university. This dataset has images of people wearing masks and those not wearing masks. Zooming and rotating were augmentation techniques applied to the dataset to increase its size, as more datasets are required for better accuracy of the system performance. The final set of images used in system evaluation are 5670, which are categorized into two classes, namely "with mask" and "without mask". *Table 1* shows the details of the dataset used in this research. Images from our university are shown in *figures 8(a)* and *8(b)*, respectively.

Table 1: Dataset and respective sources details used in the system

Source of Dataset	With Mask	Without Mask	Total Images
From our university	2968	2702	5670



(a)

(b)

Figure 8: Dataset images collected from our university (a) without mask images (b) with mask images

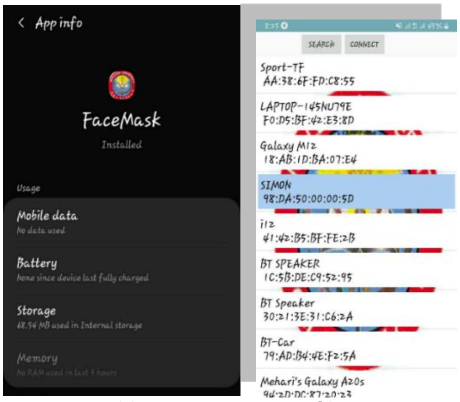
From the collected dataset, 80% of the images of each class are used for training the defined system, and the remaining 20% of the dataset is used for testing and validating the system. The total number of images (with mask or without mask) from the dataset used for training and testing purposes is listed in *table 2*. After transforming all the images to array form, we used the *tensorflow.keras.mobilenetv2* module library for preprocessing.

Table 2: Category-wise dataset used for training and testing process in the system

Type of Images	Training Images	Testing Images	Total Images
With Mask	2374	594	2968
Without Mask	2162	540	2702
Total	**4536**	**1134**	**5670**

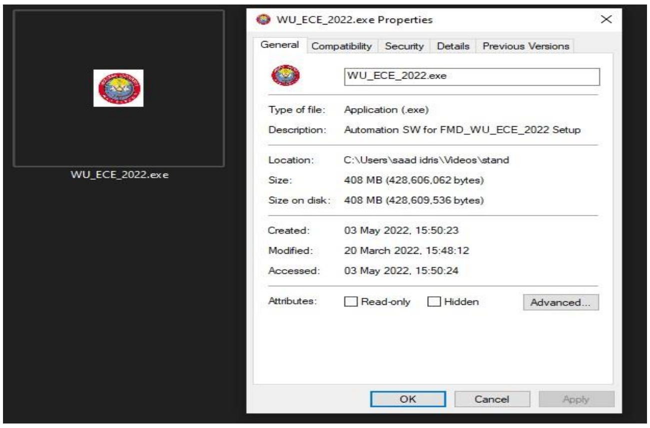
FOREX Publication
Open Access | Rapid and quality publishing
International Journal of
Electrical and Electronics Research (IJEER)
Research Article | Volume 10, Issue 3 | Pages 751-759 | e-ISSN: 2347-470X

4.4 System Realization



(a) (b)

Figure 9: Android app developed for the system (a) installed app interface (b) searching for online enabled system device

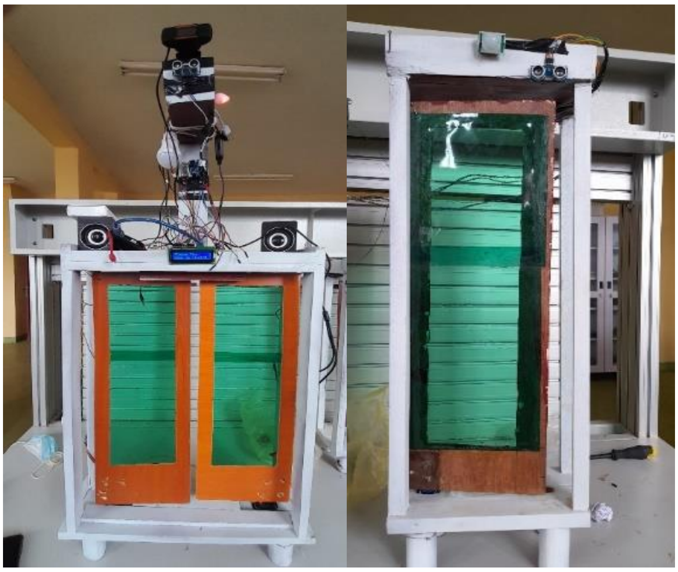


(a)

(b)

Figure 10: Developed model application for the system (a) installed app interface (b) execution mode for searching device to connect

The system configuration for this research experiment is an Intel(R) Core(TM) i3-5005U CPU @ 2.00GHz, 2.00GHz, with 8 GB of RAM installed. Each hardware and component's functionality were tested and displayed through an application package created on a computer as well as an Android application created through the Arduino integrated development environment (IDE). We have used Tensorflow.js library for deploying the developed DL model on the web and Tesnorflow-Lite API to deploy the same model onto android-based application. The system was tested with the dummy door system in order to make sure the functionality of the face mask detection and the developed model software application worked as defined. The Android applications created for this research are shown in *figures 9 (a)* and *(b)*, while the application package created for computers is shown in *figures 10 (a)* and *(b)*. The hardware setup of the entrance door and exit door installed for the research is shown in *figures 11 (a)* and *(b)* respectively.

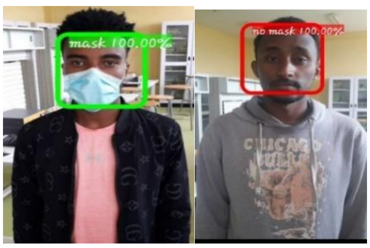


(a) (b)

Figure 11: Hardware setup for testing of system working (a) entrance door setup (b) exit door setup

5. RESULT AND DISCUSSION

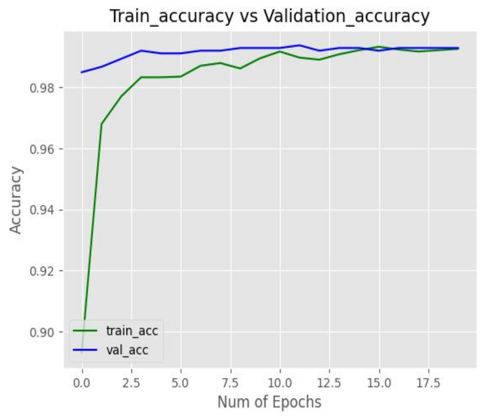
As the person approaches the entrance gate, the system measures the distance of a person is close enough to 2 meters and checks whether the person is wearing a mask. As a camera is installed to capture the face of the person, the system will get the captured image and show the detected face with a green frame with a match percentage if he/she is already wearing a mask, and a red frame with a match percentage for a face without a mask, as shown in *figures 12 (a)* and *(b)* respectively. If the person wears a face mask, the face mask detector code sends a digital signal to an Arduino setup, which in turn processes the signal to open the gate as well as turns the green indicator on and displays a message to the user that says "opening..." while also displaying the person's distance from the door in 'meters'. As the person approaches the exit gate, the PIR checks the presence of the person at least 0.5 m away and then proceeds to open the exit door, otherwise it will remain closed.



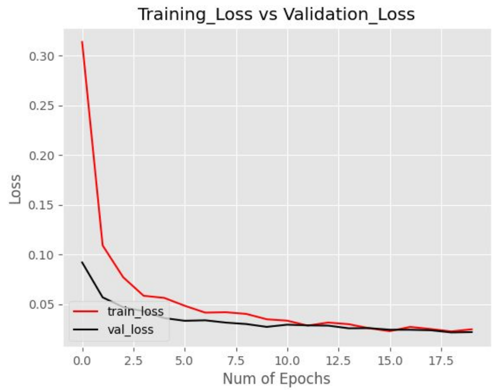
(a) (b)

Figure 12: Face captured and identification of mask from the system (a) face with mask (b) face without mask

The model in our system was trained over 20 epochs for classify the images with masks and without mask categories. *Figures 13(a)* and *(b)* show the accuracy and loss curves of the MobileNetV2 model during training and validation over 20 epochs. Details of accuracy and loss for training and validation over 20 epochs are shown in *table 3*. As it can be seen from *figure 13 (b)*, there is very little fap between training and validation loss curves, which means that convergence of the defined model is good and there are no overfitting issues that will be identified during training and validation.



(a)



(b)

Figure 13: Analysis graphs over 20 epochs (a) training / validation accuracy (b) training / validation loss

Table 3. Accuracy % and loss in training and validation over 20 epochs of MobileNetV2 model

	Accuracy	Loss
Training	99.36 %	0.0229
Validation	99.29 %	0.0198

The confusion matrix generated for the system is shown in *figure 14*. This matrix is used to visualize performance measures for the system. Based on this information, we can compute more evaluation metrics, which are useful to highlight the performance of the defined system. The concluded values we can get from the confusion matrix are True Positive, False Positive, True Negative, and False Negative. *Table 4* describes each of these values.

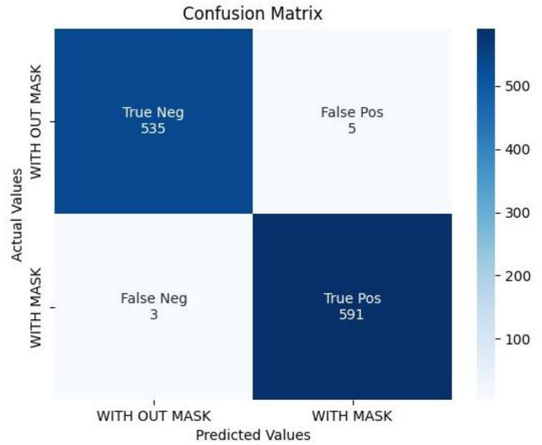


Figure 14: Confusion Metrix of evaluated accuracy for the proposed system

Table 4: Description of concluded values from confusion matrix

Concluded Values	Description
True Positive (Tpos)	Person with the mask and the system detected mask on the face
False Positive (Fpos)	Person without the mask and the system detected mask on the face
True Negative (Tneg)	Person without the mask and the system detected no mask on the face
False Negative (Fneg)	Person with the mask and the system detected no mask on the face

For the performance evaluation of the classification model defined here, we evaluated some other derived parameters shown in equations 4–7 [17], as well as some performance metrics such as precision, accuracy, and F1 score shown in equations 8–10 [8]. Recall and Specificity are the other two metrics, which are the same as True Positive Rate (*eq. 4*) and True Negative Rate (*eq. 6*) respectively.

$$True\ Positive\ Rate\,(TPR) = \frac{Tpos}{Tpos + Fneg} \tag{4}$$

$$False\ Negative\ Rate\,(FNR) = \frac{Fneg}{Tpos + Fneg} \tag{5}$$

$$True\ Negative\ Rate\,(TNR) = \frac{Tneg}{Tneg + Fpos} \tag{6}$$

$$False\ Positive\ Rate\,(FPR) = \frac{Fpos}{Tneg + Fpos} \tag{7}$$

$$\Pr ecision = \frac{Tpos}{Tpos + Fpos} \tag{8}$$

$$Accuracy = \frac{Tpos + Tneg}{Tpos + Tneg + Fpos + Fneg} \tag{9}$$

$$F1\ Score = \frac{2 * \Pr ecision * \mathrm{Re}\,call}{\Pr ecision + \mathrm{Re}\,call} \tag{10}$$

The accuracy metric is used to identify the effectiveness and usefulness of the proposed system. The precision metric is used to identify the actual rate of positive prediction. Recall, also known as the sensitivity metric, is used to identify the rate of actual true prediction results from the proposed model. The F1-score is the harmonic mean of the evaluated recall and precision from the model. Based on the above stated equations, we can get more insight into our proposed system. *Table 5* shows the classification report of the proposed system. From the result, we can see that this model achieved a higher score for all evaluated metrics, which means that this model performed well on the dataset selected for the system.

Table 5: Performance measures of different metrics for the proposed system

Metrices	Evaluated Values
TPR	0.9949
FNR	0.0051
TNR	0.9907
FPR	0.0093
Precision	0.9916
Recall same as TPR	0.9949
Specificity same as TNR	0.9907
Accuracy	0.9929
F1-Score	0.9933

6. CONCLUSION

In the last 3 years, COVID-19 has been identified as one of the major threats to the whole world. It has a high spread rate, which makes the scenario more difficult to control. Using a face mask is one of the most important precautions one can take to avoid being infected with this virus. Wearing of the mask will give safety to himself/herself as well as others too. So, this research project was taken into account to develop a model with fusion IoT and deep learning CNN architecture using MobileNetV2 for automatic detection of masks on one's face and controlling the entrance and exit door according to the availability of mask on the face. The model uses a real-time deep learning system which, uses a camera to detect face masks and transmit signals to the Arduino Mega for controlling doors and other alarm systems. To train, validate, and test the model, we have used a real-time dataset containing 5700 images, which includes datasets generated at our own university. This dataset comprises 2968 images with face masks, while the remaining 2702 images are without face masks. The performance metrics such as precision, accuracy, recall, specificity, and F1 score were performed to analyze the performance of the proposed system. The results showed that the trained model was able to achieve a result of 99.36% percent train accuracy and 99.29% validation accuracy.

Although the developed system is tested in a lively environment, it has some limitations, like it can't detect or identify the social distance, can't measure the temperature of the coming person who is wearing the mask, and can't count the people entering together. So, for further research, we can use relevant components to enable temperature sensing, identify the social distance and count people. We can also test the in-house collected dataset on the other CNN models with a greater number of epochs to get more comparable results.

7. ACKNOWLEDGEMENTS

Our thanks to the all the students and faculties who have been ready to give their images to use as our dataset.

REFERENCES

[1] Asghar MZ, Albogamy FR, Al-Rakhami MS, Asghar J, Rahmat MK, Alam MM, Lajis A and Nasir HM (2022) Facial Mask Detection Using Depthwise Separable Convolutional Neural Network Model During COVID-19 Pandemic. Front. Public Health 10:855254. https://doi.org/10.3389/fpubh.2022.855254

[2] B Varshini, HR Yogesh, Syed Danish Pasha, Maaz Suhail, V Madhumitha, Archana Sasi, IoT-Enabled smart doors for monitoring body temperature and face mask detection, Global Transitions Proceedings, Volume 2, Issue 2, 2021, Pages 246-254, ISSN 2666-285X, https://doi.org/10.1016/j.gltp.2021.08.071

[3] Munjal, P., Rattan, V. ., Dua, R., & Malik, V. . (2021). Real-Time Face Mask Detection using Deep Learning. Journal of Technology Management for Growing Economies, 12(1), 25–31. https://doi.org/10.15415/jtmge.2021.121003

[4] S. V. Militante and N. V. Dionisio, "Real-Time Facemask Recognition with Alarm System using Deep Learning," 2020 11th IEEE Control and System Graduate Research Colloquium (ICSGRC), 2020, pp. 106-110, https://doi.org/10.1109/ICSGRC49013.2020.9232610

[5] Loey, Mohamed & Manogaran, Gunasekaran & Taha, Mohamed & Khalifa, Nour Eldeen. (2020). Fighting against COVID-19: A novel deep learning model based on YOLO-v2 with ResNet-50 for medical face mask detection. Sustainable Cities and Society. https://doi.org/65. 10.1016/j.scs.2020.102600

[6] Guillermo, M., Pascua, A. R. A., Billones, R. K., Sybingco, E., Fillone, A., & Dadios, E. (2020). COVID- 19 Risk Assessment through Multiple Face Mask Detection using MobileNetV2 DNN. The 9th International Symposium on Computational Intelligence and Industrial Applications (ISCIIA2020), Beijing, China. https://isciia2020.bit.edu.cn/docs/20201114082420135149.pdf

[7] N. Boyko, O. Basystiuk and N. Shakhovska, "Performance Evaluation and Comparison of Software for Face Recognition, Based on Dlib and Opencv Library," 2018 IEEE Second International Conference on Data Stream Mining & Processing (DSMP), 2018, pp. 478-482, https://doi.org/10.1109/DSMP.2018.8478556

[8] Loey, M., Manogaran, G., Taha, M. H. N., & Khalifa, N. E. M. (2021). A Hybrid Deep Transfer Learning Model with Machine Learning Methods for Face Mask Detection in the Era of the COVID-19 Pandemic. Measurement, 167, 108288. https://doi.org/10.1016/j.measurement.2020.108288

[9] Pandiyan, P. (2020, December 17). Social Distance Monitoring and Face Mask Detection Using Deep Neural Network. Retrieved from: https://www.researchgate.net/publication/347439579_Social_Distance_Monitoring_and_Face_Mask_Detection_Using_Deep_Neural_Network

[10] Mohan, P., Paul, A.J., Chirania, A. (2021). A Tiny CNN Architecture for Medical Face Mask Detection for Resource-Constrained Endpoints. In: Mekhilef, S., Favorskaya, M., Pandey, R.K., Shaw, R.N. (eds) Innovations in Electrical and Electronic Engineering. Lecture Notes in Electrical Engineering, vol 756. Springer, Singapore. https://doi.org/10.1007/978-981-16-0749-3_52

[11] Lin, K., Zhao, H., Lv, J., Li, C., Liu, X., Chen, R., & Zhao, R. (2020). Face Detection and Segmentation Based on Improved Mask R-CNN. Discrete Dynamics in Nature and Society, 2020, 9242917. https://doi.org/10.1155/2020/9242917

[12] Chen, Y., Hu, M., Hua, C., Zhai, G., Zhang, J., Li, Q., & Yang, S. X. (2020). Face Mask Assistant: Detection of Face Mask Service Stage

Based on Mobile Phone. Preprint arXiv:2010.06421. https://arxiv.org/abs/2010.06421

[13] A. Das, M. Wasif Ansari and R. Basak, "Covid-19 Face Mask Detection Using TensorFlow, Keras and OpenCV," 2020 IEEE 17th India Council International Conference (INDICON), 2020, pp. 1-5, https://doi.org/10.1109/INDICON49873.2020.9342585

[14] Mamata S.Kalas , (2014) "Real Time Face Detection And Tracking Using Opencv " , International Journal of Soft Computing And Artificial Intelligence (IJSCAI) , pp. 41-44, Volume-2,Issue-1

[15] A. Salihbašić and T. Orehovački, "Development of Android Application for Gender, Age and Face Recognition Using OpenCV," 2019 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), 2019, pp. 1635-1640, https://doi.org/10.23919/MIPRO.2019.8756700

[16] Yadav, Shashi. (2020). Deep Learning based Safe Social Distancing and Face Mask Detection in Public Areas for COVID-19 Safety Guidelines Adherence. International Journal for Research in Applied Science and Engineering Technology. 8. 1368-1375. https://doi.org/10.22214/ijraset.2020.30560

[17] M. M. Rahman, M. M. H. Manik, M. M. Islam, S. Mahmud and J. -H. Kim, "An Automated System to Limit COVID-19 Using Facial Mask Detection in Smart City Network," 2020 IEEE International IOT, Electronics and Mechatronics Conference (IEMTRONICS), 2020, pp. 1-5, https://10.1109/IEMTRONICS51293.2020.9216386

[18] Ashu Kumar, Amandeep Kaur, and Munish Kumar. 2019. Face detection techniques: a review. Artif. Intell. Rev. 52, 2 (August 2019), 927–948. https://doi.org/10.1007/s10462-018-9650-2

[19] A. Khodabakhsh, R. Ramachandra, K. Raja, P. Wasnik and C. Busch, "Fake Face Detection Methods: Can They Be Generalized?," 2018 International Conference of the Biometrics Special Interest Group (BIOSIG), 2018, pp. 1-6, https://10.23919/BIOSIG.2018.8553251

[20] Dr. Anil Kumar Yaramala, Dr. Sohail Imran Khan, N Vasanthakumar, Kolli Koteswararao, D. Sridhar and Dr. Mohammed Saleh Al Ansari (2022), Application of Internet of Things (IoT) and Artificial Intelligence in Unmanned Aerial Vehicles. IJEER 10(2), 276-281. DOI: 10.37391/IJEER.100237.

[21] W. Boulila, A. Alzahem, A. Almoudi, M. Afifi, I. Alturki and M. Driss, "A Deep Learning-based Approach for Real-time Facemask Detection," 2021 20th IEEE International Conference on Machine Learning and Applications (ICMLA), 2021, pp. 1478-1481, https://10.1109/ICMLA52953.2021.00238

[22] Hyogeun An, Sudong Kang, Guard Kanda and Prof. Kwangki Ryoo (2022), 128-Bit LEA Block Encryption Architecture to Improve the Security of IoT Systems with Limited Resources and Area. IJEER 10(2), 245-249. DOI: 10.37391/IJEER.100232.

[23] Meeta Dubey, Prof. Prashant Jain (2013), Face Recognition using PCA and LDA Technique for Noisy Faces. IJEER 1(2), 22-28. DOI: 10.37391/IJEER.010201. https://ijeer.forexjournal.co.in/papers-pdf/ijeer-010201.pdf