An IoT Application Framework Using Deep Learning for Face Mask Detection

Pravat Kumar Routray
Department of Computer Science & Engineering, Institute of Technical Education and Research, Siksha 'O' Anusandhan, Deemed to be University, Bhubaneswar, Odisha, India, Email: pravat.routray@gmail.com
Binod Kumar Pattanayak*
Department of Computer Science & Engineering, Institute of Technical Education and Research, Siksha 'O' Anusandhan, Deemed to be University, Bhubaneswar, Odisha, India, Email: binodpattanayak@soa.ac.in
Mihir Narayan Mohanty
Department of Electronics and Communication Engineering, Institute of Technical Education and Research,Siksha 'O' Anusandhan, Deemed to be University, Bhubaneswar, Odisha, India, Email: mihir.n.mohanty@gmail.com

Abstract: A facemask covering nose and mouth is one of the most effective ways to protect against infection and to spread the coronavirus. This safeguard rule is applied by almost all governments. For automation, we have developed an approach based on deep learning to detect the mask. The approach has been extended to an Internet of Things (IoT) based framework that can be an element of a smart city to keep people safe. Since health safety is a major challenge, this paper proposes an automatic detection process. A convolutional neural network with transfer learning is used for the detection. This model is included in an IoT architecture for automation. The results of testing the approach are excellent in terms of accuracy. In addition, the module for IoT works well, as verified in the study, and also appears to be useful for the Internet of Medical Things.

Keyword: IoT, Deep Learning, Face Mask, COVID-19, Transfer Learning

1. Introduction

The novel corona virus (nCoV)causes sickness and symptoms that range from cold to deadly infections like Middle East Respiratory Syndrome (MERS) and severe acute respiratory syndrome (SARS) [1]. The earliest infected patient with the corona virus was found in Huang in China in December 2019. Since that period, COVID-19 has become a pandemic leading to challenges such as job loss, financial crises and sometimes loss of life. Since then, countless individuals have died, and this number is increasing day by day. The World Health Organization (WHO) has described the most important symptoms of the corona virus like fever, dry cough, tiredness, diarrhea, loss of taste, and smell [4] and people were advised to take precautions, such as cleaning their hands, maintaining social distance, wearing masks in public places, refraining from contact with eyes, nose, and mouth to fight against the virus. We can limit the spread of COVID-19 by strictly maintenance of social distance and the use of a facial mask. Unfortunately, not everyone follows these rules thoroughly, which is the main reason for the spread of the virus. Therefore, people who do not follow the rules should be identified by the corresponding authorities in an effort to reduce the spread of the corona virus. Numerous methods have been presented for object detection. Deep learning strategies are exceptionally utilized in medical applications. These techniques can be used in detecting the mask on a face. Additionally, a smart city implies a metropolitan region that comprises numerous Internet of Things (IoT) sensors to gather information.

In this paper, a system is designed and proposed for an automatic detection of an accurate (or missing/inaccurate) use of masks. The system consists of two parts: In one part, image processing for detection is performed using a deep learning approach. The other part is based on

IoT for the communication of information. If a person is found not using a mask properly, the system automatically informs the responsible authority in the smart city network. This is due to the utilization of real time video films from the CCTV cameras installed in public places in the city. Facial images from the real time videos are extracted and utilized to recognize the mask on the face. The features of the image are extracted by a deep learning algorithm using the Convolutional Neural Network (CNN) concept. Images are used for learning the parameters of the numerous hidden layers of the CNN. In this pandemic situation, an appropriate implementation of the law on individuals who are not following essential health rules can be guaranteed by utilization of the proposed framework.

The rest of the paper is organized as follows: Recent work on mask detection is depicted in Section 2. In Section 3, the proposed methodology for the entire framework is discussed. Section 4 discusses the outcomes of the developed framework and Section 5 presents the conclusion.

2. Related Works

In [3] the authors used a deep learning approach with TensorFlow, Keras, and OpenCV to detect face masks. For classification, the MobilenetV2 architecture has been used as a framework. To perform real-time mask detection, this model was used in embedded devices (like NVIDIA Jetson Nano, Raspberry pi). However, it was not suitable for the smart city use. MobilenetV2 was used in [4], where the model was saved after training for future use. The authors have used video as frame by frame for the face detection. If a face is detected, it proceeds to the next process. In [5], the authors proposed a model on real-time facemask recognition with an alarm system through deep learning techniques based on Convolutional Neural Networks. The image collected from an installed camera and passed through the trained model, which was loaded in Raspberry pi. They used VGG16 as a pre-trained model that achieved an accuracy of 96%. A comparison was made in [6] among three classifiers Support Vector Machines(SVM), Decision Trees, and a model developed by the authors. It was found that their model achieved 91.11% accuracy which is to be increased. A model based on YOLO v2 with a ResNet-50 was proposed in [7] for classification. They compared and fine-tuned different optimizers. Image details can be restored by Super-Resolution (SR) networks. Recently, SR networks have become more in-depth, and the ideas of auto-encoder and transfer learning we reincorporated for performance improvement. SR networks were also applied for image processing before image segmentation and classification[9], reconstructing images for higher resolution and restoring details. Moreover, SR networks can significantly improve the classification accuracy, especially when using a dataset with low-quality images, and they provide a feasible solution to improve the performance of identifying facemask-wearing. Therefore, the combination of an SR network with a classification network (SRCNet) could be utilized in facial image classification for accuracy improvement. This proposed model obtained an accuracy of 98.70% [8]. To prevent the spread of the corona virus,infected areascould be sanitized, e.g., by using deep learning and an automatically controlled quadcopter [10]. The disease could be detected, e.g., usinghuman chest X-ray images and applying Gated Recurrent Neural Networks [11,12]. By using IoT technologies, we can make people aware of Covid-19 and at the same time control its spread [13]. To detect the brain diseases fuzzy based clustering techniques are used for image segmentation[14]. With the implementation of R-CNN to OPOSFM dataset which contains 7826 images and achieved accuracy of 96 percentage [15].

3. The Proposed Model

In the proposed method, an automated face mask detector is developed to screen people for proper use of face masks. Public places can be monitored by CCTV cameras. The system will process the images captured by the cameras. If a person fails to wear a mask correctly, the photo of that person will be mailed to the authority for immediate action. The developed framework is mentioned in the block diagram in Fig. 1.

The framework consists of the camera that is linked with the IoT module. The camera is set in the public place and connected to the IoT module which works on a Raspberry pi platform. The captured images are taken as the input to the classifier model. The classifier model is aCNN based on deep transfer learning for detection. In the final stage the transfer of information is made to the responsible authorities.

Fig. 2. showsas examples pictures of persons with mask and without mask, as they are considered in this work.The different steps for the suggested process are shown in Fig. 3.

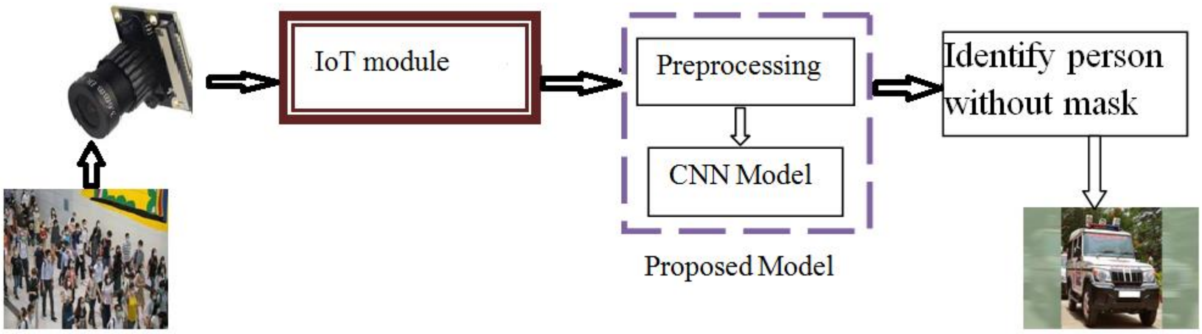

Fig.1: Proposed framework

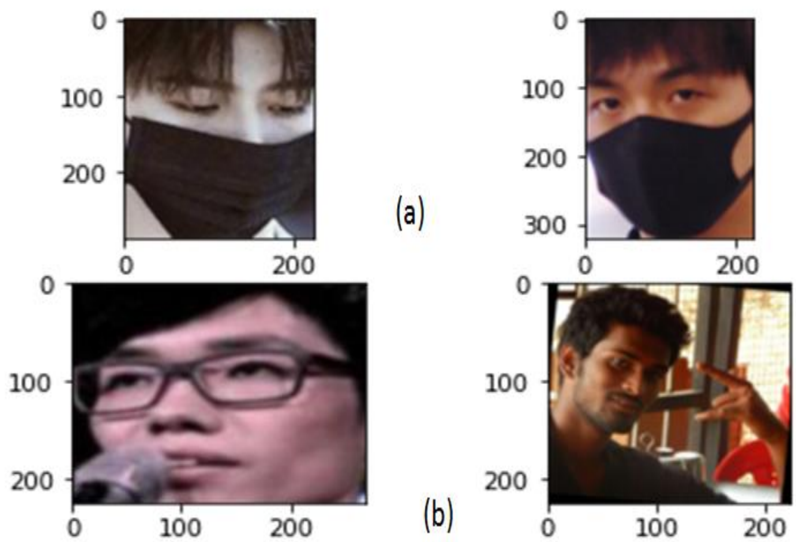

Fig. 2: a) two samples of a person with a mask, b) two samples of a person without a mask.

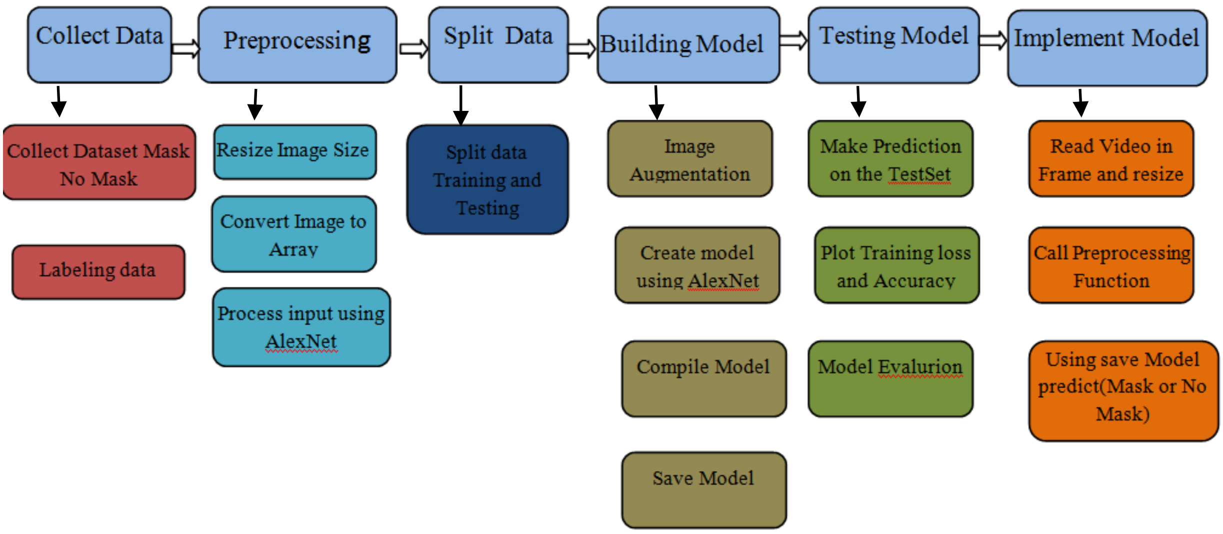

Fig. 3: Block diagram of the proposed model

The next subsections describe the different steps of the proposed model.

3.1. Data Augmentation

A set of data is taken from a Github repository [2] which consists of 3822 facial images with and without mask, of which 1918 images are with and the remaining 1914 are images without mask. The data was augmented to generate a maximum number of images for using the deep learning model. We considered geometric transformations, e.g., to support images taken from different angles and distances. Therefore, four types of geometric transformation are used: flipping, cropping rotation, and scaling. Further, the images are converted to gray scale and normalized. The size of the transformation images is 256 x 256. After center cropping of the images, the size was normalized to 224 x 224. Center cropping normalization is done by considering its mean and standard deviation. The augmented images are shown in Fig. 4.

Fig. 4: Augmented example data

3.2. Transfer Learning Approach in AlexNet

Transfer learning is to reuse the result of machine learning for some tasks with adaption for another connected task. It is generally used where sufficient training samples are not available to train a model, for example in medical image classification for emerging or uncommon diseases. This is especially true for models based on deep neural networks which have an enormous number of parameters to train. By using transfer learning, the model parameters start with effectively set initial values that need to be changed only slightly to be well determined for the new task. There are two primary ways by which the pre-trained model is used for a different task. In the first approach, the pre-trained model works as a feature extractor and the classification is carried out by a classifier which is trained on top of it. In the second approach, the entire network, or a subgroup thereof, is fine-tuned on the new task. Therefore, the pre-trained model parameters like the weights are treated as the initial values for the new task and are updated during the training stage [16].

Here we used AlexNet for transform learning which uses eight layers with learnable parameters. The model consists of five convolution layers and three fully connected layers. This model is using ReLU activation functions to accelerate the speed of the training process almost six times, and to minimizeOverfitting dropout layers (0.5) are used.In Table 1 the entire structure of the used AlexNet is explained. Fig5 shows the AlexNetused with details of each layer.

Table 1.Architecture of the used AlexNet

Layer	Filters/Neurons	Filter Size	Stride	Padding	Size of feature map	Activation
Input	-	-	-	-	227*227*3	-
Conv1	96	11*11	4		55*55*96	ReLU
MaxPool 1	-	3*3	2		27*27*96	-
Conv 2	256	5*5	1	2	27*27*256	ReLU
Max Pool 2	-	3*3*2	-	-	13*13*256	-
Conv 3	384	3*3	1	1	13*13*384	ReLU
Conv 4	384	3*3	1	1	13*13*384	ReLU
Conv 5	256	3*3	1	1	13*13*256	ReLU
Max Pool 3		3*3	2		6*6*256	-
Dropout 1	Rate=0.5	-	-	-	6*6*256	-
Fully connected1		-	-	-	9216	ReLU
Dropout 2	Rate=0.5	-	-	-	4096	-
Fully connected2	-	-	-	-	4096	ReLU
Fully connected3	-	-	-	-	1000	Softmax

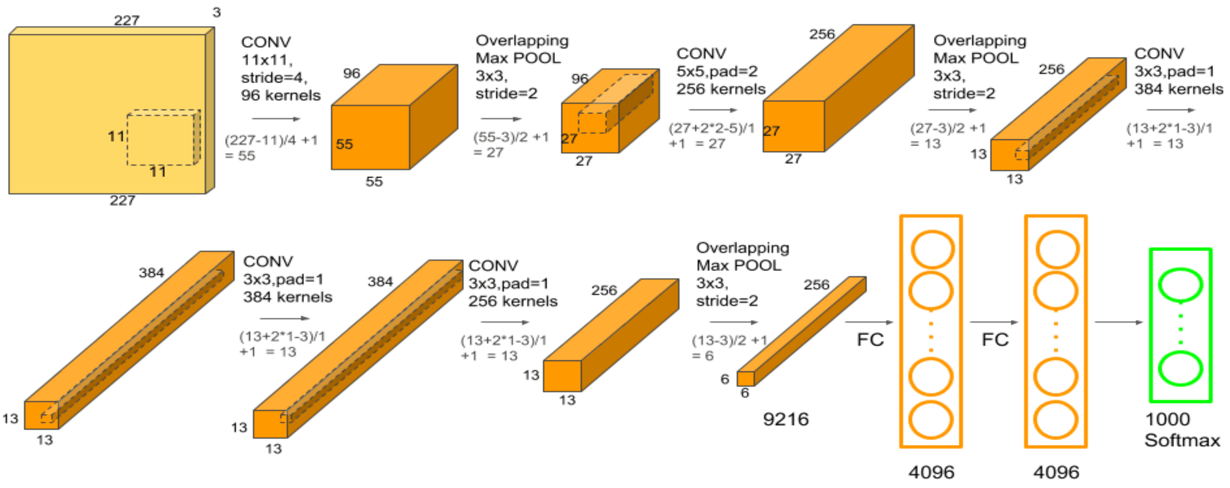

Fig. 5: AlexNet Architecture

3.3. Model Architecture

A CNN is used for image classification as it is the best suitable learning model. The hidden layers comprise various convolution layers that learn reasonable filters for significant feature extraction from the given samples. The features extracted by the CNN are used by various dense neural network layers for the purposes of classification. The over fitting of the neural network can be managed by a dropout layer by dropping out units.At last, a dense layer containing two neurons recognizes the classes.

3.3.1. Convolutional Layers

The main block of the CNNare the convolution layers. The convolution function uses a sliding window strategy that helps extract features from original images[3]. This convolution function helps generate feature maps. The convolutional output $(C)$from a convolutional layer is represented as in Eq. (1).

$$C[m,n] = (A * B)[m,n] = \sum_{i} \sum_{k} B[j,k]A[m-j,n-k] \tag{1}$$

In (1), $A[m,n]$ is the input image of size$m\times n$ whereas $B[j,k]$ is the kernel of size $j\times k$.

3.3.2. Pooling Layers

The pooling activities make computations quicker by allowing a decrease of the size of the input matrix without losing numerous features. Some of the pooling operations are explained below:

i.) Max Pooling: It uses the greatest value present in the selected region where the kernel is right now present as the value for the output matrix for that cell.

ii.) Average Pooling: It takes the average value present in the selected region where the kernel is currently present as the value for the output matrix for that cell.

iii) Min Pooling: It takes the minimum value present in the selected region where the kernel is right now presentas the value for the output matrix for that cell.

3.3.3. Dropout Layers

Dropout layers are used to reduce problems related to neural network over fitting that may occur during training by dropping arbitrarily biased neurons from the model. These neurons can be a part of visible layers as well as hidden layers [3]. The probability for a neuron to be dropped can be changed by changing the dropout proportion.

3.3.4. Non-Linear Layers / Activation Functions

An activation functions uses the sum of the product of the different weights and inputs with the bias to determine the last output value for the current hidden layer, which then becomes the input for the next layer [17]. The proposed model used two activation functions which are as follows:

i.) Softmax

The Softmax activation standardizes the input values in such a way that all the output values lie between 0 and 1, with their sum equal to 1. The probability of each class being determined eithertrue or falseis specified by this activation function.

If the input is $x = \begin{bmatrix} 1.2 \\ 0.7 \\ 0.6 \end{bmatrix}$, then $F(x) = \begin{bmatrix} 0.48 \\ 0.28 \\ 0.24 \end{bmatrix}$, that is, if x is the vector of inputs, the Softmax function F(x) isspecified by Equation (2).

$$F(X_j) = \frac{x_j}{\sum_{i=1}^{n} x_i} \qquad (2)$$

In this equation, x_j is the *j*th value of the vector and *n* is the number of classes.

ii.) Rectified Linear Unit (ReLU)

The rectified linear activation function or ReLU is a piecewise linear function that will yield the input directly in case it is positive, else, it will yield zero:

$$F(x) = \begin{cases} x, & \text{if x is positive} \\ 0, & \text{otherwise} \end{cases}$$

Traditionally, there are 256nodes in the convolution 2D layers in AlexNet. In this study, the output consists of two layers whose output corresponds to two classes, i.e. wearing a mask or not wearing a mask, and are fine-tuned for connections having 9216nodes, which results from two fully connected layers in AlexNet with 256 nodes each and 2output nodes as we require binary classification. Overfitting happens when a model learns the undesirable patterns of the training data. Hence, training accuracy increases whereas test accuracy decreases in case of over fitting. During the training of the model, to avoid over fitting, a dropout of 0.5 is added as this learning rate is appropriate for our dataset, and "CrossEntropyLoss" is used as the loss parameters and "Adam" as optimizer [5]. The training is performed over 20 epochs.

3.4.Training, Test, and Validation Data

A dataset from GitHub [2] is taken for training and testing purposes. From the data, 80%of the images of each class are used for training purposes, 15%for the testing purpose and5%for the validation purpose. Fig. 6 shows the splitting of the dataset, which is plotted using Python.

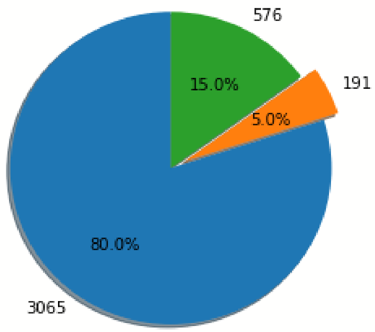

Fig.6: Sample dataset; blue = data the training; green = data for testing; orange = data for validation

We have created a new dataset taking the similar to the face images . Total 3975number of images are present in the dataset. Here also 80% of the images of each class are used for training purposes, 15%for the testing purpose and 5% for the validation purpose. Fig. 7 shows the splitting of the dataset.

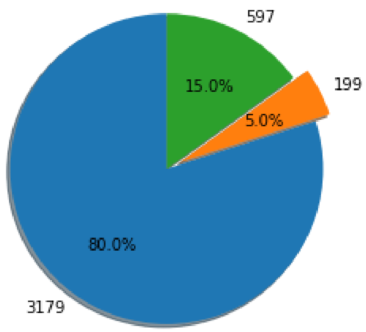

Fig.7: Sample dataset; blue = data the training; green = data for testing; orange = data for validation

4. Results and Discussion

Table 2 and Table 3 show the confusion matrix which allows to calculate accuracy, where as precision and F-score as presented in Table 4 and Table 5. For our implementation, we used MobileNet, VGG16, and AlexNet. Among these three models, we achieve the highest accuracy and the lowest loss with AlexNet data batch size set of 32 and 20 iterations for epochs. The detailed results from the performance tests for accuracy and loss are illustrated in Fig. 7 and Fig.8.The graph of the accuracy and loss for AlexNet for both dataset are shown in Fig9 and Fig 10. Figs.12,13, 14,and 15 display different test results on the performance of the model in detecting persons wearing or not wearing properly a facemask correctly. The images in Figs.12 and 14 show results of the detection of a person in the image not wearing a facemask with a 100% accuracy. Fig.15 shows an example of a person wearing a face mask, but the mask is not covering either mouth or nose, which is identified as a person not wearing a facemask with an accuracy rate of 78.40%. Fig.13 illustrates a person wearing a facemask with an accuracy of 100%.

The used metrics precision, recall, F1-score, and accuracy are shown in Eq. (3)-(6).

$$\text{Precision}= \frac{N(TP)}{N(TP)+N(FP)} \tag{3}$$

$$\text{Recall}= \frac{N(TP)}{N(TP)+N(FN)} \tag{4}$$

$$F_1 Score = 2\,\frac{precission*recall}{precission+recall} \tag{5}$$

Table 2.Confusion matrix of the proposed model

True Class	Predicted	
	Abnormal	**Normal**
Abnormal(288)	287	1
Normal(288)	3	285

Table 3.Confusion matrix of the proposed model on New Dataset

True Class	Predicted	
	Abnormal	**Normal**
Abnormal(288)	304	5
Normal(288)	3	285

Table 4.Performance analysis of the proposed model for detection and classification of face mask images

Class	Precision	Recall	$F1_{score}$	Support
Abnormal	1	0.99	0.99	288
Normal	0.99	1	0.99	288
Average	0.99	0.99	0.99	576

Table 5.Performance analysis of the proposed model for detection and classification of face mask images on New Dataset

Class	Precision	Recall	$F1_{score}$	Support

Abnormal	0.99	0.98	0.99	309
Normal	0.98	0.99	0.99	288
Average	0.99	0.99	0.99	597

$$Accuracy = \frac{N(TP) + N(TN)}{N(TP) + N(TN) + N(FP) + N(FN)} \tag{6}$$

In these equations, N(TP) indicates the total true positives, N(FP) are the total false positives, N(TN) are the total true negatives and N(FN) indicates the false negatives. All these measures are computed for each class, and an overall measure of the algorithm is computed by taking the average of all these measures across the two classes. From the analysis, it is observed that the developed mask classification framework achieved an average classification accuracy of 99.306%. Out of the 288 images representing "with_mask", 285are classified correctly while the other 3 images are classified as "without mask". Similarly, from the 288 "without mask" images, 287 are classified correctly while the remaining1 image is classified as "with mask" which is a wrong classification. Table-2 and Table-3 show the confusion matrix obtained when the developed framework is validated using the developed dataset. Table-4 and Table-5 provides the F-score, precision, and recall. The highest precision of 100% is achieved for the category "with mask" followed by 100% for "without mask". On an average, the proposed algorithm offers a precision of 99.3%. Regarding recall, the highest is achieved for the abnormal class with 100%; the lowest recall of 99% is obtained for the normal class and the average obtained recall is 99%. Finally, the average F-score is observed to be 99%, where the maximum is 99%for the normal class and 99% the minimum for the abnormal class. This proposed model requires some amount of energy for cameras to transmit photos and to e-mail to the responsible authorities [19]. The 99.306% testing accuracy was achieved during the training of the CNN model. The description of the considered performance accuracy is given below.

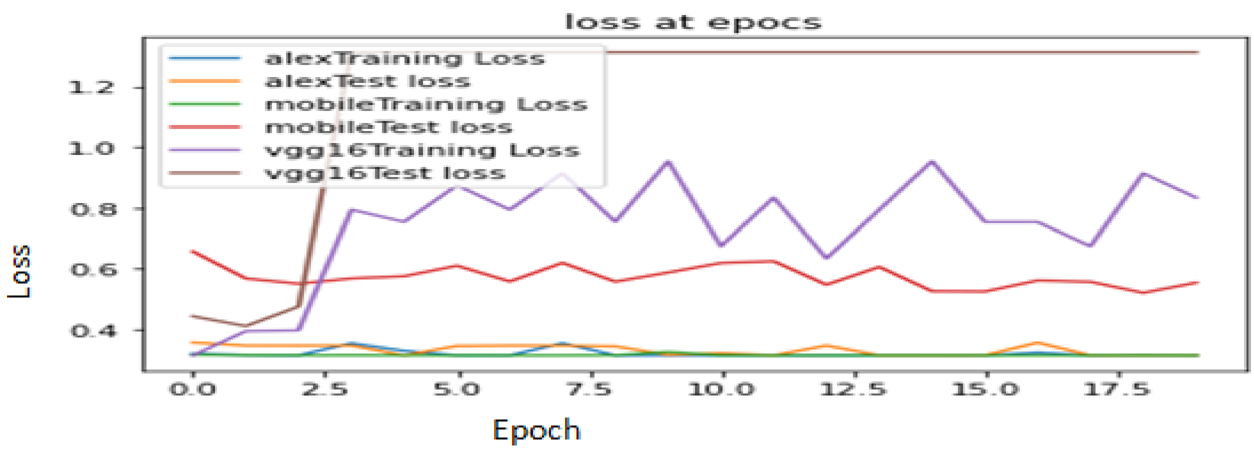

Fig. 7:The loss graph for all models

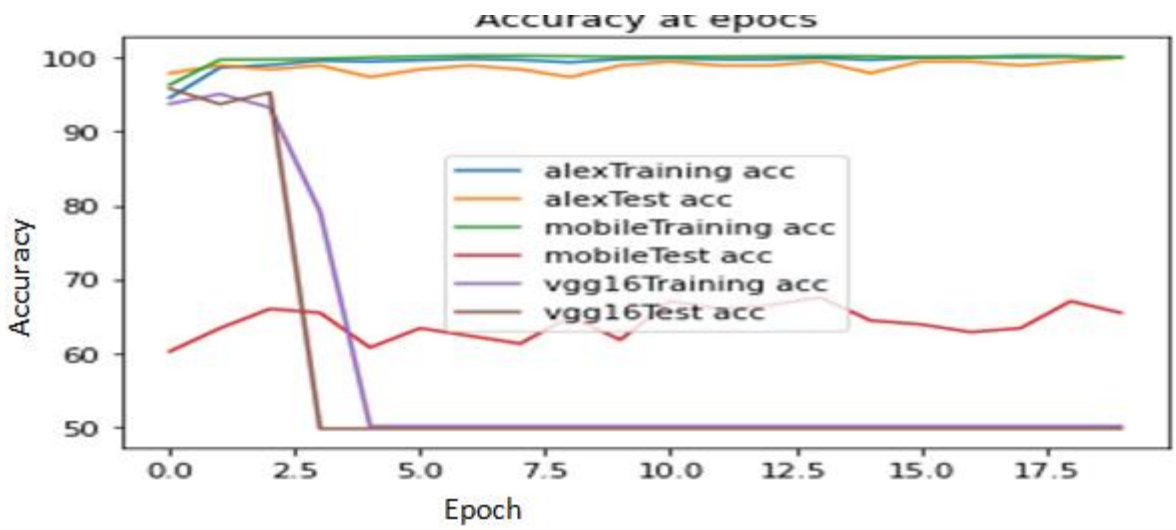

Fig. 8: The accuracy curve of training and validation dataset used for all models

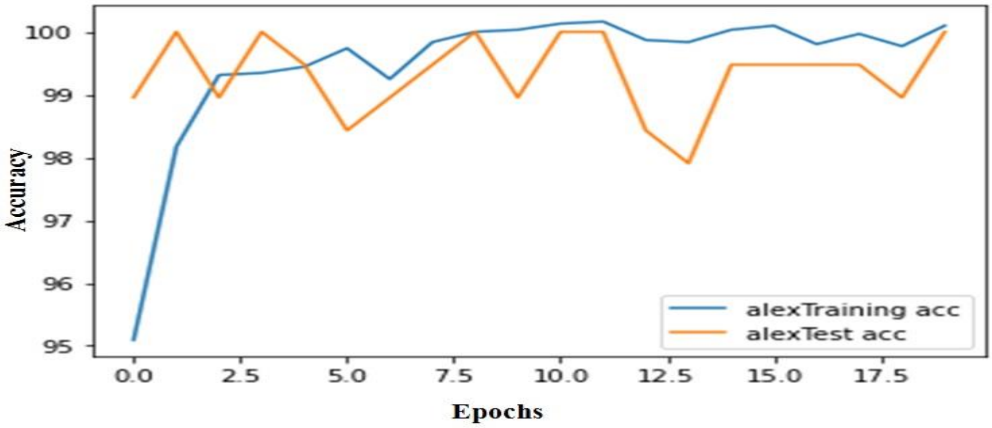

Fig. 9: The accuracy curve of the training and validation dataset used in the proposed model (AlexNet)

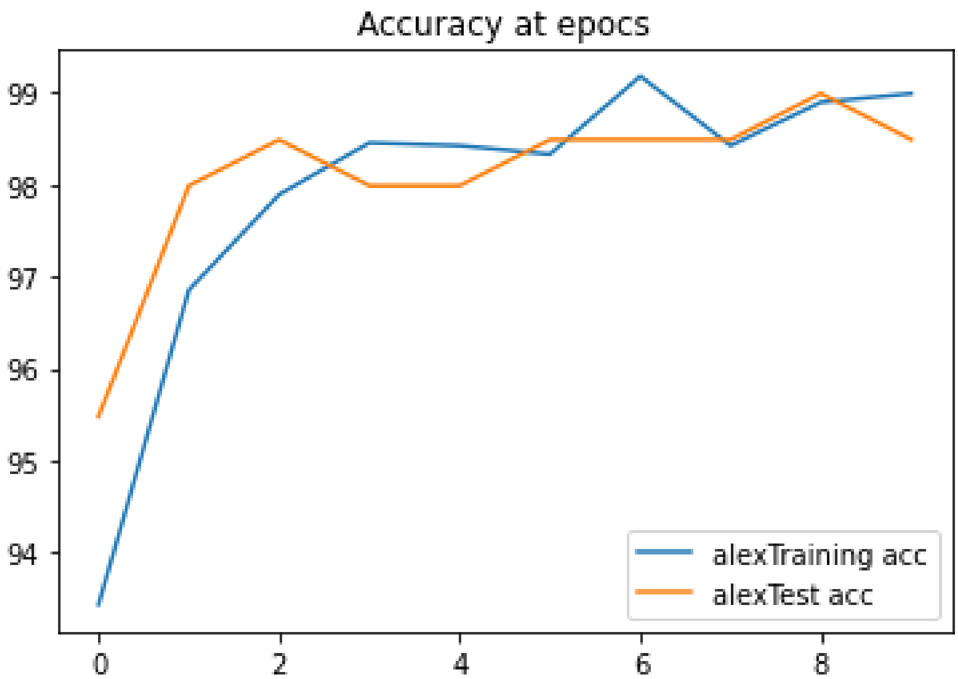

Fig. 10: The accuracy curve of the training and validation dataset used in the proposed model on new dataset

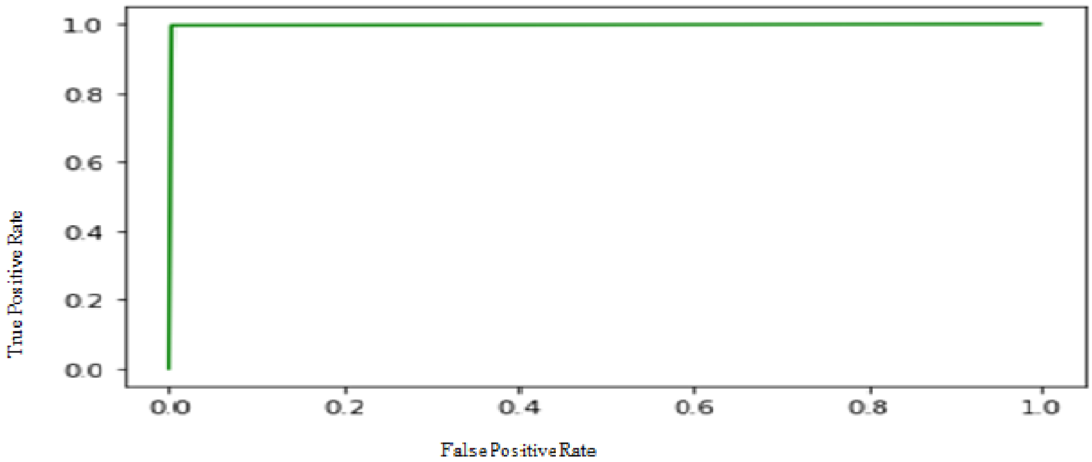

Fig.11: The Receiver Operating Characteristic (ROC) curve

Fig. 10represents the receiver operating characteristic (ROC) curve of the suggested model. The ROC curve shows the diagnosis ability of the binary classifier for varying discrimination. The ROC curve is plotted using two parameters, the true positive rate (TPR) and the false positive rate (FPR) which is calculated according to Eqn. (7) and Eqn. (8). TPR and FPR values of the ROC curveare measured for different threshold values. The area under the ROC curve (AUC) measures the performance of the binary classifier for all possible thresholds[20]. The value of AUC ranges from 0 to 1. AUC is 1 when a model predicts 100% correct and AUC is 0 when it predicts 100% wrong. The AUC achieved from our classifier model is 0.993, indicating a decent classifier.

$$\text{True Positive Rate}=\frac{TruePositive}{TruePositive+FalseNegative} \qquad (7)$$

$$\text{False Positive Rate}=\frac{FalsePositive}{TrueNegative+FalsePositive} \qquad (8)$$

The suggested framework analyses the video from frame to frame before the face detection algorithms is activated. Here the Caffe model [21]is used for face detection. After face detection, the image is processed, including resizing of the image, and then converting it to an array for execution [22]. This processed image is then provided as input for the proposed model and the video frame is also labeled. The output of the model predicts whether a person wears a mask or not along with the predicted percentage. This model will predict "no mask" if a person will not cover both mouth and nose with a facemask.

Fig.12: Person without mask 100%

Fig.13: Person with a mask 100%


Fig.14: Person without a mask 100% Fig.15: Person without mask 78.40%

The algorithm implemented will generate an automatic e-mail alert along with a photo of the person to the responsible authority to initiate action if the level of coverage of the face with the mask is less than 80%.Our proposed algorithm uses Twilio to send an e-mail, which is displayed in Fig. 15.

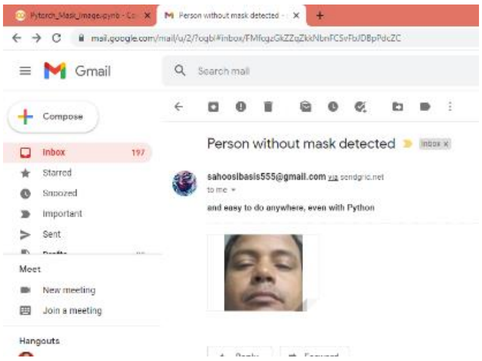

Fig.16:E-mail to the authority together with a photo

Table 7. External validation

Author	Model	Accuracy(%)
[8]Bosheng Qin and Dongxiao Li	SRCNet	98.70
[4] S. A. Sanjaya and S. A. Rakhmawan	MobileNetV2	92
[15] Gabriel T S Draughon *et al.*	R-CNN	96
[7] Mohamed Loey *et al.*	ResNet 50	81
[5] S. V. Militante and N. V. Dionisio	VGG-16	96
[6] G K Jakir Hussain *et al.*	Proposed Method CNN	91.11
[3] PreetiNagratha *etal.*,	MobileNetV2	92.64
ProposedModel	**Proposed CNN Model**	**99**

5. Conclusions

Our work suggests an integrated IoT framework where each part performs better in our study compared to previous work. The detection procedure based on a deep learning method is approximately 100 percent accurate. In addition, the communication through the internet is verified.

Future research should address the following aspects: In terms of accuracy, various simple architectures may be used with respect to processing time. In addition, the IoT model can be improved from the communication point of view to enable faster communication with greater band width. In addition, it might be useful to consider more specified mask types, e.g. masks that are more similar to faces or cover a larger part of the face.

References

[1] WHO EMRO | About COVID-19 | COVID-19 | Health topics. [Online]. Available: http://www.emro.who.int/health-topics/coronavirus/about-covid-19.html, accessed on: Jul. 26, 2020.
[2] https://github.com/balajisrinivas/Face-Mask-Detection
[3] P. Nagrath, R. Jain, A. Madan, R. Arora, P. Kataria, and J. Hemanth, "SSDMNV2: A real time DNN-based face mask detection system using single shot multibox detector and MobileNetV2," Sustain. Cities Soc., vol. 66, no. December 2020, p. 102692, 2021, doi: 10.1016/j.scs.2020.102692.
[4] S. A. Sanjaya and S. A. Rakhmawan, "Face Mask Detection Using MobileNetV2 in the Era of COVID-19 Pandemic," 2020 Int. Conf. Data Anal. Bus. Ind. W. Towar. a Sustain. Econ. ICDABI 2020, 2020, doi: 10.1109/ICDABI51230.2020.9325631.
[5] 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, doi: 10.1109/ICSGRC49013.2020.9232610.
[6] G. K. J. Hussain, R. Priya, S. Rajarajeswari, P. Prasanth, and N. Niyazuddeen, "The Face Mask Detection Technology for Image Analysis in the Covid-19 Surveillance System," *J. Phys. Conf. Ser.*, vol. 1916, no. 1, 2021, doi: 10.1088/1742-6596/1916/1/012084.
[7] M. Loey, G. Manogaran, M. H. N. Taha, and N. E. M. Khalifa, "Fighting against COVID-19: A novel deep learning model based on YOLO-v2 with ResNet-50 for medical face mask detection," *Sustain. Cities Soc.*, vol. 65, p. 102600, 2021, doi: 10.1016/j.scs.2020.102600.
[8] Q. Bosheng and D. Li, "Identifying Facemask-Wearing Condition Using Image Super-Resolution with Classification Network," *Mdpi Sens.*, pp. 1–23, 2020.
[9] Das, A., & Mohanty, M. N. Covid-19 detection from x-ray images using convolutional neural network. *International Journal of Advanced Science and Technology*, *29*(8 Special Issue), 2020.
[10] Mohanty, M. D., Mallick, P. K., & Mohanty, M. N. (2021). Design of an Intelligent Controller of a Self-Tuned Quadcopter for IoT-Based Application on COVID Infected Area Sanitization. *IEEE Internet of Things Magazine*, *4*(3), 24-29.
[11] Das, A., Mohapatra, S. K., Subudhi, A., & Mohanty, M. N. (2021). Classification of COVID-19 InfectedX-Ray Image using deep Learning Techniques. *Annals of the Romanian Society for Cell Biology*, 2736-2743.
[12] Patra, G. R., & Mohanty, M. N. (2021). Prediction of COVID Patients using Gated Recurrent Unit Neural Networks. *Turkish Journal of Computer and Mathematics Education*, 2333-2340.
[13] Bhoi, A. K., Mallick, P. K., Mohanty, M. N., & de Albuquerque, V. H. C. (2021). Hybrid Artificial Intelligence andIoT in Healthcare.
[14] Mallick, P. K., Satapathy, B. S., Mohanty, M. N., & Kumar, S. S. Intelligent technique for CT brain image segmentation. In 2015 2ndInternational Conference on Electronics and Communication Systems (ICECS) (pp. 1269-1277). IEEE, 2015. **DOI:** 10.1109/ECS.2015.7124789
[15] Draughon, G. T., Sun, P., & Lynch, J. P. (2020, September). Implementation of a computer vision framework for tracking and visualizing face mask usage in urban environments. In 2020 IEEE International Smart Cities Conference (ISC2) (pp. 1-8). IEEE.

[16] Jagadev A. K., Pattanayak B. K., Mishra M. K. and Nayak M., Power and Delay Aware On-Demand Routing for Ad Hoc Networks, International Journal of Computer Science and Engineering, Vol.2, No.4, pp.917-923, 2010.

[17]Jyoti, A., Mohanty, M. N., & Kumar, M. P. Morphological based segmentation of brain image for tumor detection. In 2014 International Conference on Electronics and Communication Systems (ICECS) (pp. 1-5). IEEE, 2014. **DOI:** 10.1109/ECS.2014.6892750

[18]Biswal, A. K., Singh, D., Pattanayak, B. K., Samanta, D., & Yang, M. H. IoT-Based Smart Alert System for Drowsy Driver Detection. Wireless Communications and Mobile Computing, 2021. Volume 2021 |Article ID 6627217 | https://doi.org/10.1155/2021/6627217

[19]Panda, N., Sahu, P. K., Parhi, M., &Pattanayak, B. K. A Survey on Energy Awareness Mechanisms in ACO-Based Routing Protocols for MANETs. Intelligent and Cloud Computing, 791-800, 2021. https://doi.org/10.1007/978-981-15-5971-6_85

[20]Rath, M., &Pattanayak, B. “Technological improvement in modern health care applications using Internet of Things (IoT) and proposal of novel health care approach”, International Journal of Human Rights in Healthcare, pp.148-162, vol-12, issue-2, 2019, doi:10.1108/ijhrh-01-2018-0007

[21]Reddy, D.S., and Rajalakshmi, P., “A Novel web application framework for ubiquitous classification of fatty liver using ultrasound images,” IEEE 5th World Forum Internet Things, WF-IoT 2019 - Conf. Proc., pp. 502–506, 2019, doi: 10.1109/WF-IoT.2019.8767283.

[22] SCICS: A Soft Computing Based Intelligent Communication System in VANET, Rath M. and Pattanayak B. K., Proceedings of the International Conference on Intelligent Information technologies, pp.255-261, 2017.