An Intelligent Health Control Security Robotic System

Syed Danish Ahmad Sharifi [a] ✉ , *Muhammand Usman* [a] , *and Eman Gul* [b]

[a] Faculty for adnvanced research and integrated system (FARIS), Pakistan

[b] Department of Electrical Department, University of Engineering and Technology Taxila, Pakistan

✉, [a] syed.dahmed13@gmail.com
[a] Facility for advanced research and integrated system (FARIS), Pakistan

ABSTRACT

Nowadays, scientific knowledge is constantly bringing comfort and change to everyday life but the entire world is facing a major health crisis as a result of coronavirus disease transmission. According to the World Health Organization (WHO), wearing a mask on the face in public areas is one effective method of protection against COVID. Autonomous robots have become a prominent technology in recent years, with applications in a variety of fields. Robots are utilized to complete tasks more quickly than humans. Generally, robots are smarter with endless energy levels, and are precise in task management. Therefore, propose the methodology of this work is focused on developing a robot-based COVID-19 protection system. The proposed robot-based COVID-19 protection system consists of five core steps: 1) Person Identification, 2) Vaccination Checking 3) Face Recognition 4) Face Mask Detection, and 5) Temperature Checking. Person identification is performed using HOG to detect faces and machine learning classifier SVM for identification of the person. Then vaccination status is checked. To check vaccination status, it gets the name from face recognition and matches with the names in the database and gets a value from the vaccination column to show if the person is vaccinated or not and facemask detection is performed by facemask detector using Keras Mobile-Net architecture. The temperature is checked using the MLX90614 temperature sensor. The robot performs all these functions by speaking and by displaying them on the LCD screen. Artificial Intelligence on basis of deep learning and neural networks could help in fighting Corona Virus in many ways. The purpose of this system is to use a security robot instead of a security guard in organizations such as universities, colleges, schools, offices, software houses, and other organizations. The proposed system performed better compared to the existing systems as it achieves 99% precision and a 0.01% error rate.

Keywords: *Face Detection, Facemask Detection, Audio Features, Robot*

1 INTRODUCTION

Intelligent Health Control Security Robot is built to provide security i.e. person identification and health facility. Technology has come a long way from the development of the telephone to artificial intelligence. And the fact that we all have to accept is that we have extended our arms and adapted the technology to make the task easier [1]. Facial recognition is a biometric method that recognizes human faces without the need for physical contact. This method involves algorithms matching the facial nodes on a person's face to images stored in a database. Facial recognition can enhance the security of organizations. The versatility of face recognition makes it a preferred option for enhanced security [2]. Regarding the transmission of coronavirus disease (COVID-19)[3] , WHO has advised different countries to ensure their citizens wear masks in public places. Before COVID-19, only some people wore masks for their health protection from pollution of air, and medical professionals when practicing in hospitals. By the rapid spread of COVID-19, WHO has acknowledged it as a global pandemic [4]. In accordance to WHO, worldwide infections are around 424.5 million. Mostly in crowded and overcrowded areas, positive cases are found. Therefore, scientists have prescribed that wearing masks in public areas could prevent disease transmission. Technologies of Artificial intelligence (AI) like deep learning and machine learning can be used in a variety of ways to prevent COVID-19 infection [5]. Vaccination is necessary to protect from Corona Virus and temperature is a major coronavirus symptom. Many organizations have made vaccination mandatory for all people, especially people aged eighteen or above eighteen. This manual vaccination and temperature check up will be replaced by our system. Intelligent Health Control Security Robot checks whether the person belongs to the organization or not and then checks if he is wearing a mask or not. Then it checks person is vaccinated or not and then checks the temperature at the end [6].

The main aim of this work is the development of an Intelligent Health Control Security Robot by using Artificial Intelligence to detect the face of a person and tell the name of a person, check the facemask and vaccination status, and also check the temperature by using a contactless temperature sensor and when all verification is completed then automatically opens the gate.

The article's organization is as follows: Section II discusses the related work, the proposed method steps are explained in Section III, the Result and discussion are elaborated in section IV, and finally, the conclusion is written in Section V.

2 RELATED WORK

Deep learning has made significant advances in object detection and recognition in a variety of application domains over the years [7]. In general, these methods focus on identification and facial structure when wearing a mask. To stop the spread of COVID-19, all study aims to identify those who do not use masks. Scientists and researchers have proven that wearing masks can help minimize the spread of the coronavirus. The main disadvantage is that most of the classic machine learning methods do not achieve the lowest, highest, and most time-consuming correctness. Utilizing deep transfer learning techniques for feature extraction [8] and the neuromorphic domain has demonstrated promising potential in classification and detection issues [9]. Face mask detection has become a widely used application as a result of the Covid-19 virus, which mandates wearing face masks, maintaining social distancing, and using hand sanitizer to wash one's hands. While other issues such as social distance and hand sanitization have been addressed, facemask detection has yet to be addressed. A key preventive measure during this pandemic is wearing a mask, especially during periods when maintaining social distance is challenging. It is imperative to wear a mask, especially for people who are more susceptible to developing a serious disease from COVID-19 infections [10]. Face recognition has grown in popularity as a problem in computer vision and image processing. Convolutional architectures are being used in the development of several new algorithms to maximize their accuracy. Even pixel information may be extracted using convolutional structures. Meenpal et al. used RGB channel images with localized objects to construct accurate face masks for human objects. That solved the issue of error predictions, produced results on the Multi Human Parsing Dataset with mean pixel-level accuracy, and correctly limit the segmented area. The suggested network is capable of identifying numerous faces as well as non-frontal faces from a single picture. The technique may be used for complex tasks like face part detection [11]. Coronavirus disease (COVID-19) emerged at the end of 2020 and will continue to kill millions of people and businesses in 2021. However, manually tracking the implementation of this strategy is not possible. The key here is technology. Chavda et al. developed a deep learning-based system that can identify instances of improper mask usage This system is integrated with the built-in camera and is composed of a two-stage CNN architecture that distinguish between uncovered and masked faces. This is encourage the use of masks, helps track security problems, and maintains a secure atmosphere. The world has been greatly affected by the coronavirus illness in 2019. To safeguard individuals, masks should be worn in public [12]. Yadav et al. propose an efficient method based on computer vision. The method focuses on automatic real-time monitoring of a person. By implementing the model on a Raspberry Pi-4 to monitor the activities and violations using the camera, it is possible to become aware of social distancing and masks in public places. When the Raspberry Pi-4 finds infractions, it notifies the control center at the state police headquarters and the general public. In this suggested system, modern deep learning algorithms are integrated with geometrical methods to provide a robust model that addresses three elements of detection, tracking, and validation. By reducing the transmission of the Coronavirus and conserving time, the suggested solution helps society. When the lockdown is reduced in the current scenario, it may be utilized successfully to check persons in public places like malls and shopping centers. The use of automated testing may be done anywhere and lowers the need for human analysts [13].

3 METHODOLOGY

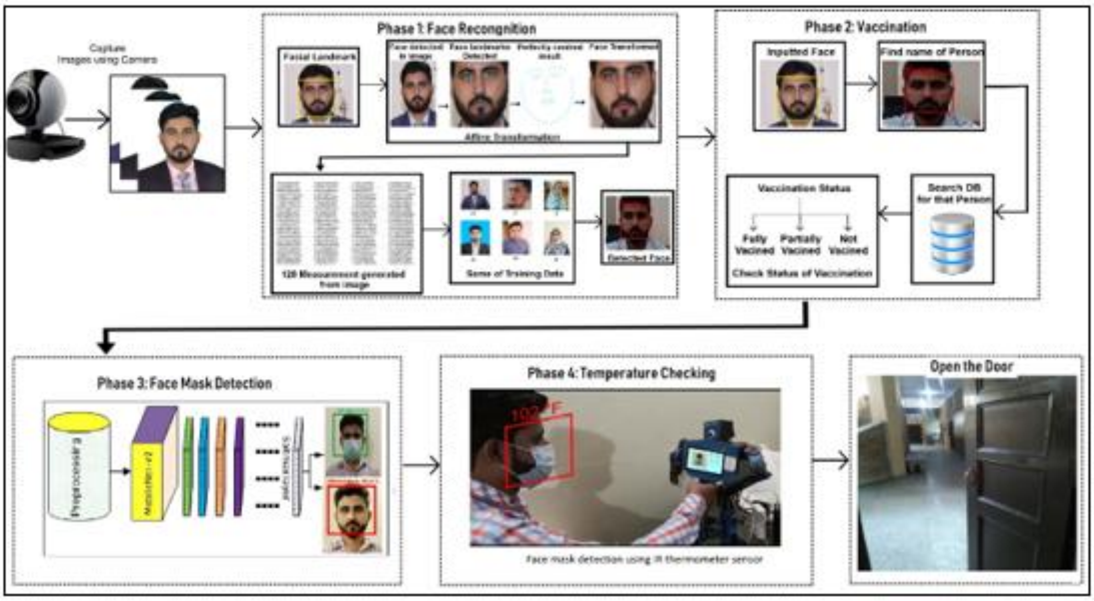


Fig. 1. Intelligent Health Control Security Robotic System

There are high risks of infection for security guards during manually face-to-face COVID-19 checking. To address this issue, we propose an Intelligent Health Control Security Robot that helps in healthcare society. Intelligent Health Control Security Robot provides the security and health facility to protect against the COVID-19 virus. The robot performs all these functions by speaking and displaying on an LCD screen as shown in Fig.1. This system does person identification by face recognition, checks the facemask, checks if the person is vaccinated or not, checks the temperature and if all conditions are fulfilled then it opens the gate.

3.1 FACE RECOGNITION

Face recognition is a combination of several interconnected issues [14].

- First, it looks at the person and takes the photo from live video, then searches the face in its database.
- Then focuses on each face and recognizes that they are the same person, even if the faces rotate in strange directions or have poor lighting.
- Third, enable the recognition of unique facial features that can be used to distinguish them from others' eye size, face length, etc.
- Finally, compares the face's unique features to find the perfect match and determines the name of the person and his detail.

Our brain is programmed to do all of this automatically and rapidly as a human beings. Humans are far too good at face recognition and see the faces of objects every day. This type of high-level generalization is not supported by computers. Therefore, this document describes how to perform each step of this process individually.

We need to resolve each face recognition step individually and build a pipeline in which the current step's results are passed to the next step.

- Find a face in a photo that is taken by the camera during a live video
- Analyze facial features
- Compare with a familiar face
- Predict the face and determine the details of the person

3.2 STEP 1: FIND A FACE IN A PHOTO

The initial step in the pipeline is the detection of the face. Of course, before distinguishing faces, you need to find them in your photos and videos. First, the robot takes picture of the person from a live video and then detects the face of the person [15]. You can see face detection as shown in Fig.2.

Fig. 2. Face Detection

Fig. 2 presents, a face finding as an excellent camera feature. In the proposed model, the camera automatically selects faces, before taking a photograph, ensuring that all of the faces are in focus to find the image area and proceed to the next stage of the pipeline. It uses a method named Histogram of Oriented Gradients (HOG), which was invented in the year 2005 [16]. To find a face in an image, first make the image black and white, as there is no need for color data for finding the face as depicted in the Fig.3.

Fig. 3. Convert Color Image into Black and White

Then examine each image pixel individually. For each pixel, examine the pixels that directly surround it as shown in Fig. 4 and Fig. 5.

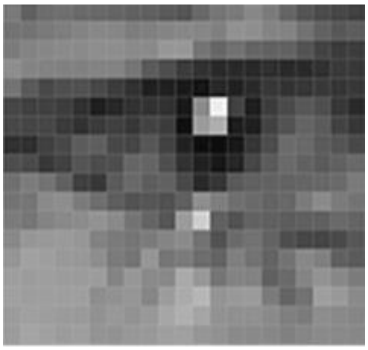


Fig. 4. Pixel-based Zoom in and Look at Every Pixel

Fig. 4 presents the current pixel, which is how much dark in comparison to those pixels that directly surround it. Next, draw an arrow for indicating the direction of the image in which it will darken. Repeat this procedure for each pixel of the image, every pixel will be exchanged by an arrow. Arrows, known as gradients, indicate the progression all over the image from light to dark. This may seem random, but there are really good reasons to replace pixels with gradients. When you analyze the pixels directly, the values of the pixel of a very dark image and a very bright image of the same person are very different. However, considering only the direction where the brightness varies, both a dark image and a bright image get the same accurate representation. This makes problem-solving much easier. If you see at a high level the basic flow of light and dark, you should be able to see the basic patterns of the image. To perform this, image is divided into small 16x16 pixels. For every count, square the gradients that pointing in every principal path (up, up, right, etc.). Then in the image replace this square in the strongest arrow direction. Finally, it will transform the image into representation, making it easy to capture the face basic structure. In this HOG image to discover the face, you need to discover the image part, that most closely resembles the HOG pattern and is extracted after other training face sets. The face part is localized using HOG as depicted in Fig. 5.

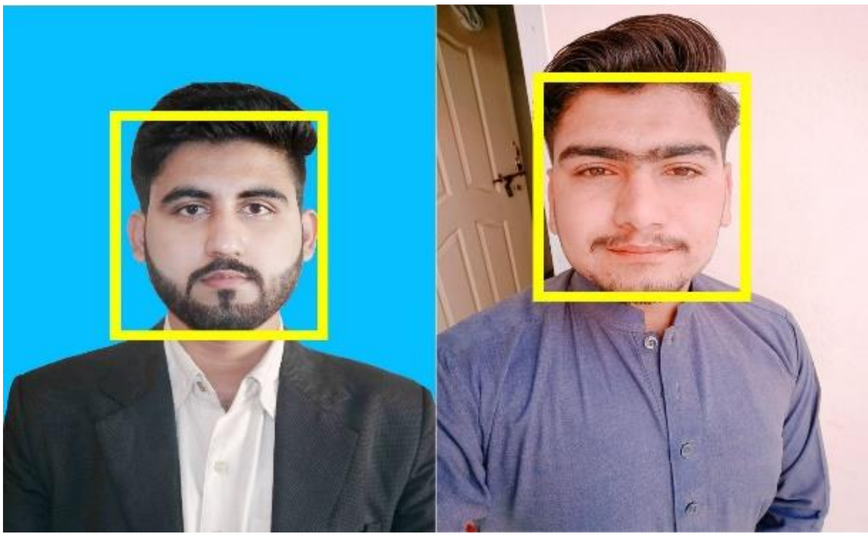


Fig. 5. Find Faces in the Object

3.3 STEP 2: ANALYZE FACIAL FEATURES

The face part being separated from the whole object is a challenging task because [17] surfaces rotated in different locations look completely distinct from other regions as shown in Fig. 6.

Fig. 6. Completely Distinct Persons

The facial landmark is applied for image wrapping from eyes to lips, in which identify the 68 specific points (called landmarks) on each face, such as the top of the chin, the inner edge of each eyebrow, the outside edge of each eye, and so on. Then machine learning model is trained to locate these 68 specific points that are presented in Fig. 7.

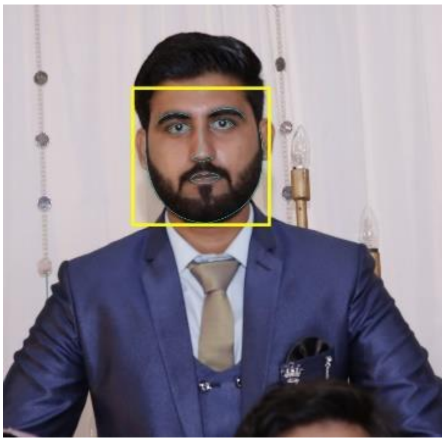


Fig. 7. Locate 68 Landmarks on Each Face.

This is the outcome of locating 68 facial landmarks in the test image as shown in Fig. 8. [18].

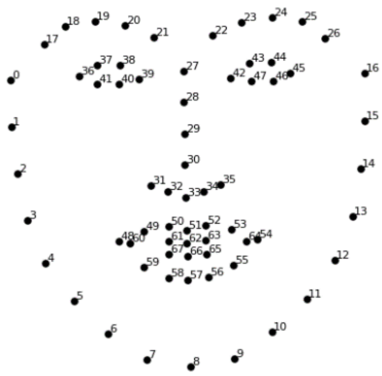


Fig. 8. Face Landmarks Identification

Now that you know where your mouth and eyes are, scale, shear, and rotate the image to get them as centered as possible. No flashy 3D distortion will occur, as the image will be distorted. Using only basic image transformations like rotation and scaling keeps parallel lines parallel (called affine transformation) as shown in Fig. 9 [19].

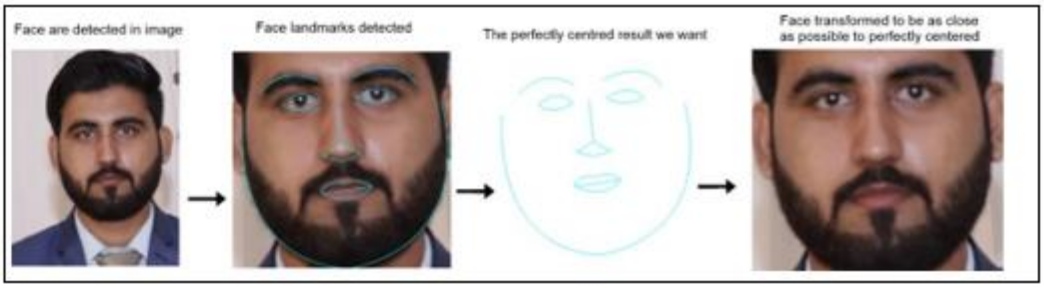


Fig. 9. Affine Transformation

Now, no matter how you rotate your face, within the frame, your mouth and eyes will be in roughly the same position. This greatly improves the accuracy of the following step [19].

3.4 STEP 3: ENCODING FACES

We've reached the problem's core: identifying faces. Comparing the unidentified face found in step 2 with all of the images of people who have previously been labeled as the most basic face recognition technique. This approach has a serious issue since it assumes that if the previously tagged face and the unknown face are very similar, they belong to the same individual. It takes far too long. In milliseconds, not hours, you must be able to identify your face. A quick approach has been proposed for taking some fundamental measures from either side. Then, using the same technique, measure the unidentified face and use the precise measurement to identify the familiar facial expression.

Deep learning can determine which part of the face is better measured than humans. As a solution, the Deep Convolutional Neural Network is trained. The network is now trained to provide 128 measures per face rather than training it to detect visual objects as it did before.

The training procedure involves simultaneously viewing three facial images.

- Load training face images of known people.
- Load another photo of the same person.
- Import a photo of a completely different person

To make the measurements for numbers 1 and 2 slightly closer and numbers 2 and 3 slightly further apart, the approach checks the current readings for each of these three images before making a little adjustment to the neural network as shown in Fig.10.

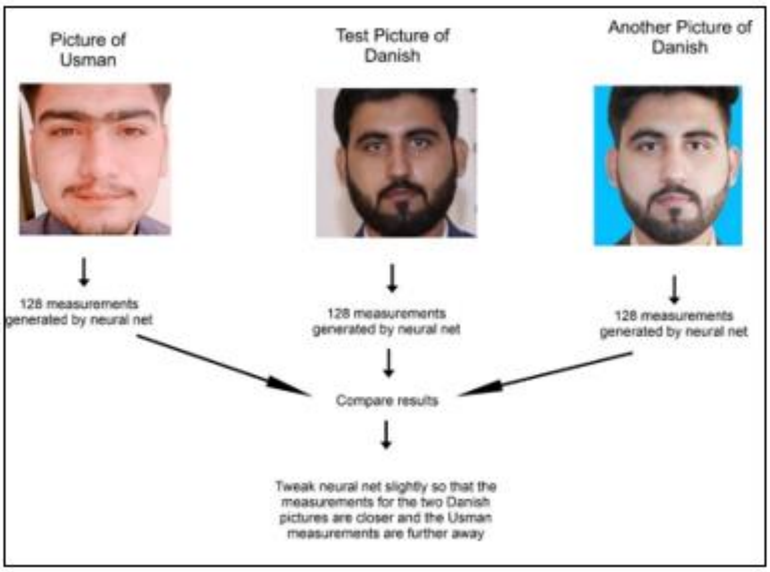


Fig. 10. Compare the Result

In Fig. 10, neural network learns to create 128 measurements for each individual after millions of repetitions for millions of images from innumerable distinct persons. The dimensions of the same individual are similar throughout all ten images. Each face 128 measurements in machine learning are known as embedding [20]. There are many similar methods.

3.5 ENCODING FACE IMAGE

The CNN model is trained to produce face embedding, to obtain 128 measurements across each face. The dimensions of the test image are shown in Fig.11.

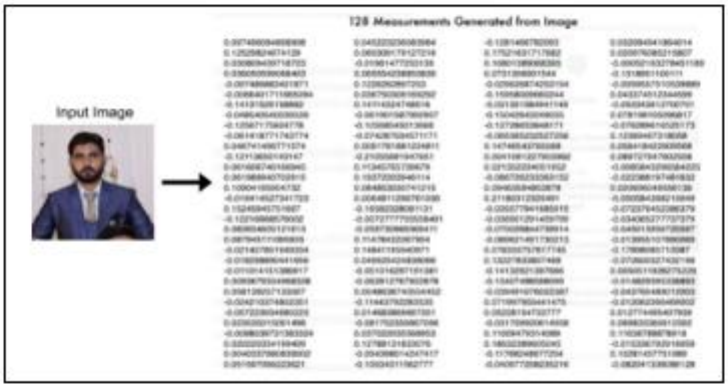


Fig.11. Measurement of the Test Image

Which part of the face do these 128 numbers accurately measure? We don't care. What we care about is that when two different images of the same person are compared, the network generates roughly the same number [21].

3.6 STEP 4: FINDING THE PERSON'S NAME FROM THE ENCODING

The entire process is simplified in this final phase. To choose the individual who most closely resembles the test image, we simply search a database of well-known individuals. Basic machine-learning classification methods are used for this. Although it uses a simple SVM linear classifier, any classification technique will do. In Simplest terms, we train a classifier to identify the known individual who is most similar to the new test image based on readings from the image. The processing time of this classifier is milliseconds. The output of the classifier is the name of the genre. To begin, we trained the classifier by embedding approximately 20 images each from Danish, Usman, and Dr. Javeria as shown in Fig. 12.

Fig. 12. Some of the Training Data

In Fig.12, for images, encoding utilizes the HOG algorithm to create a simple version of the person's image. Locate the image part that most closely resembles the general HOG face encoding using this simplified image. Search the main face landmarks and discover the Face pose. After we have located those landmarks, we'll use them to distort the image so that the mouth and eyes are in the center. Pass the facial picture that has been central to the neural network that measures facial features and saves 128 measurements. We could see which person is close to our measurements of the face by having to look at all the faces that we measured previously [22].

3.7 Face Mask Detection

The dataset consists of two types of images: a face without a mask and a face with a mask [23]. The dataset was designed specifically for facemask detection and classification as presented in Table 1.

Table 1. Dataset with Face Masks and without Face Masks

Face with masks	5484
Face without masks	5510
Total	10994

This dataset has already divided the two types of images into two folders. After loading the images from both folders, the preprocessing transforms the images into a NumPy array and pairs the images with their category labels. Then all the images and their corresponding labels are saved in the Python list. Finally, the dataset is divided: the training set is divided into 80% and the test set is divided into 20%. The libraries used for dataset preprocessing include NumPy, Scikit-learn, and Keras. After the first attempt of implementing a facemask detector based on YOLO v3, the results were slower and less accurate because the device did not have a GPU implementation to better support the model [24-43]. Therefore, in this project, we decided to build a facemask detector using MobileNet, a "convolutional neural network architecture aimed at running very efficiently on mobile devices." MobileNet has 3,538,984 parameters, which is much less than other pre-trained models like VGG19 and Xception, with 143,667,240 and 23,851,784 parameters, respectively. Therefore, MobileNet performs fast and uses less storage space [44]. As shown in Fig.13, the mask detection model has three basic layers: MobileNet, the fully connected layer, and the pooling layer. The architecture of the mask model was created using CNN's structure as a basis. However, the mask model in this replaces the convolution layer with MobileNet [45].

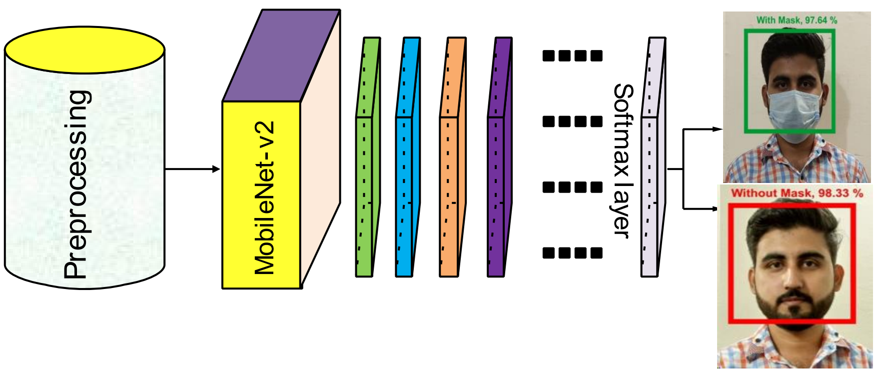


Fig. 13. Mask Detection Model

In Fig.13, the masked model is implemented using Keras API. First, the MobileNet [46] layer is created by Keras MobileNetv2 function. The weights and the input size are set to default, “ImageNet” and (224, 224, 3), respectively. We will construct the fully connected layer by ourselves. Thus, the fully-connected layer of MobileNet is disabled. The Keras AveragePooling2D function generates the pooling layer with the input window setting (7, 7). The average pooling method is selected in this layer because it would smooth out the images. Then, the Keras Flatten function is applied to make the data for the fully connected layer. The FCL is constructed by Keras Dense function with 128 units, and the activation function is "ReLU", a suitable activation function for non-linear problems.

$$f(x) = \text{maximum}(0, x) \quad (1)$$

$$f(x) = \begin{cases} 1, & \text{if } x > 0 \\ 0, & \text{Otherwise} \end{cases} \quad (2)$$

A Keras dropout function is involved after the fully connected layer to avoid overfitting. Then, the output is represented by Keras Dense function with 2 units: one for mask and one for without mask. The activation function is "sigmoid", which is suitable for binary classification. Next, the Keras Model function combines the MobileNet and all the layers to form the mask model. After the mask model is created, Binary Cross-Entropy is utilized as the loss function, which is suitable for image classification. Finally, the mask model is trained on the dataset using Keras fit function [45]. The hyperparameters for model training are mentioned in Table 2.

Table2. Hyperparameters for model training

Learning rate	0.0001
Batch-size	32
Epoch	20

Table 2 presents the training parameters that are selected after experimentation. These parameters provide better testing results for the classification of facial masks.

3.8 DETECTION

The masked model is trained on a dataset that contains only one face in every image. That is, the mask model can only recognize masks from images that contain only one face. Therefore, the mask model requires a face detector to identify the face from the input file and input the mask model for mask recognition. Here we import a face detector that can detect faces from videos and images. Next, the combination of the face detector and mask model becomes a full-face mask detector. The facemask detector can detect facemasks from three file types: image, video, and live camera streaming, and the detector uses the OpenCV library to process input and output files [47].

3.8.1. K-FOLD CROSS-VALIDATION

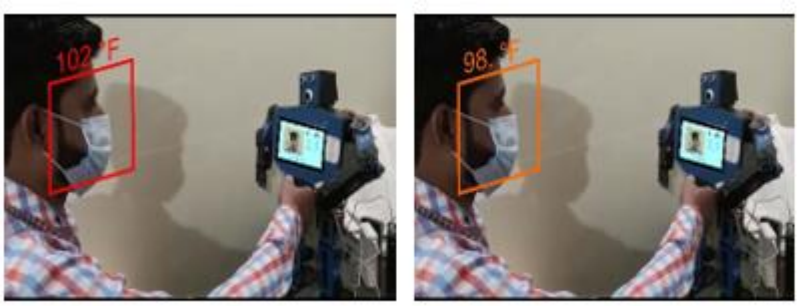
Since splitting dataset has randomness, if we use the normal way to evaluate our model, that is, using the function 'train-test-split', the result may not be convincing. This is because when we just evaluate our model once, some important features in the validation set will not be used, but when we evaluate it twice or more, the results may have big differences among each evaluation. Therefore, we choose the K-Fold Cross Validation to evaluate our model, which can help us obtain a more accurate result. The algorithm steps of K-Fold Cross Validation are shown below [48]. Divided the dataset equally into k parts. Use the first portion as the training set, and the rest as the test set. Train the model and determine its loss and accuracy. Repeat step two and three a total of k times, using a different section each time as the test set. To evaluate the model, use the average loss and accuracy rat. To implement this algorithm, we used the function "KFold" in the sklearn API [48].

3.8.2. VACCINATION STATUS

Each person has to submit his vaccination certificate then the administrator updates the persons who have submitted the vaccination certificate, from non-vaccinated to fully vaccinated or partially vaccinated according to their certificate. When a person comes in front of the robot, first it matches the person with the photos in the database and checks the vaccination record. If the person is vaccinated then move to the next step of checking the temperature [49].

3.8.3. CHECK TEMPERATURE

To check the temperature of a person we have used the MLX90614 temperature sensor as shown in Fig. 14. The MLX90614 series components are general infrared measurement device components. Its benefits include high precision, non-contact, low cost, and small size. Contact temperature measurement can only take the temperature after the object is measured and the temperature sensor has reached thermal equilibrium, so the reaction time is long and easily influenced by ambient temperature; whereas infrared temperature measurement determines the temperature of the human based on the infrared radiation energy of the human to be measured [49].

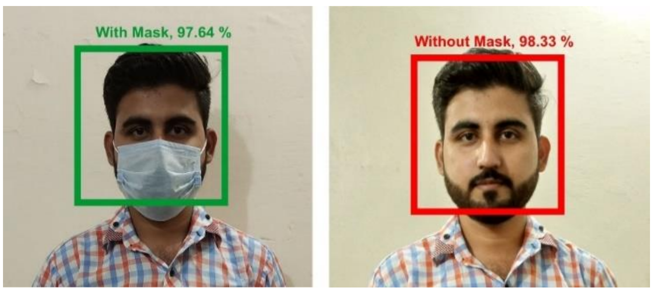


(a) (b)

Fig.14. Temperature Checking (a)abnormal (b) normal

4. RESULTS AND DISCUSSION

The proposed model performance is computed in terms of accuracy. It is an essential evaluation measure that determines how accurate our system is working as shown in Fig. 15.



(a) (b)

Fig. 15. Proposed Model Results in Terms of Accuracy (a) with Mask (b) without the Mask

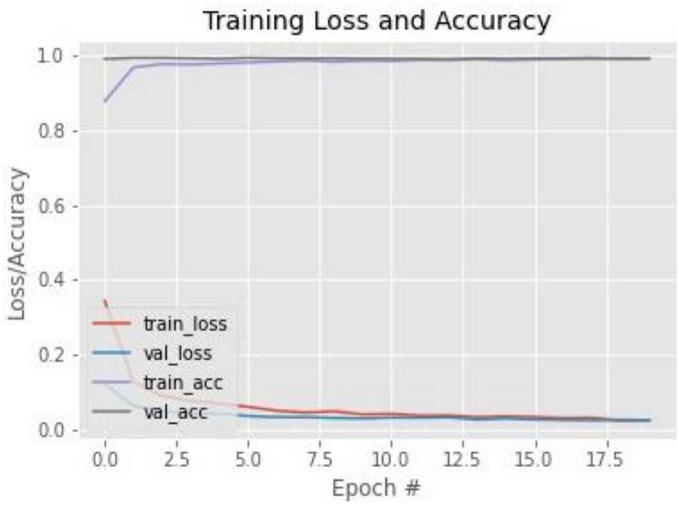


Fig. 16. Proposed Model Training with Loss Rate

The quantitative results of the proposed method are mentioned in Table 3.

Table 3. Quantitative Results of the Proposed Model

Classes	Precision	Recall	Macro average	Weighted average	F1-score
With mask	0.99	0.99	0.99	0.99	0.99
Without mask	0.99	0.99	0.99	0.99	0.99

Table 3 depicts the proposed method results, in which we achieved a 0.99 prediction rate. The proposed method results are compared to the existing methods as shown in Table 4.

Table 4. Comparison of the Proposed Method Results

Ref#	Year	Error rate
[50]	**2022**	**1.1**
[51]		**0.02**
[52]		**0.13**
Proposed Model		**0.01**

Table 4 present the comparison of the results, in which a deep learning classifier is employed for the detection of face masks with a 1.1 error rate [50]. The eleven-layer CNN model is designed for the classification of face/non-face mask images that gives an error rate of 0.02 [51]. The YOLOv4 detector is employed for the detection of the facial mask with a 0.13 error rate [52].

As compared to existing research, in this article HOG features are used with SVM for person recognition. The facial mask is detected through facemask detectors known as Mobilenet-v2. The model is trained on 20 epochs with selected hyperparameters which provides a 0.01 error rate.

5. CONCLUSION

The proposed security robot system helps to control the face-to-face interaction of security guards with people. Instead of a security guard here, the robot checks the person who belongs to the organization initially then moves to the vaccination step and check-in the record of that person if the person is vaccinated then proceeds to the mask step in which it checks person must wear the mask. If a person fulfills this condition then it checks the temperature of a person and if the temperature is normal then the person is permitted to enter the organization and the security robot system opens the gate and speaks that ‘you can enter in organization’. If any condition is not fulfilled the person is not allowed to enter the organization. In case of any issue, the person could directly talk to the head of the organization by live streaming through the camera. The proposed system achieves 99% precision, 99% Recall, 99% Macro Average, 99% Weighted Average and 99% F1-Score in both of the classes (with mask and without mask). The achieved results are far better compared to the existing methods. In the future, many more functionalities could be added to this such as robot size could be increased for future use and some enhancements could be done to increase picture or video quality. The dataset could be increased for more accurate results. Another thing that could be improved is that if someone wears a scarf or any cloth on his nose and mouth then we could adjust the model according to it. In the future, another functionality of self-defense could be added that if someone tries to touch the robot, the robot could attack for his safety. Solar panels could also be added instead of batteries so that we don't have to charge and don't have to face electricity issues. We believe that this system will help in controlling the spread of viruses too fast and it also provides security.

REFERENCES

[1] L. Yang, C. Fu, Y. Li, and L. Su, "Survey and study on intelligent monitoring and health management for large civil structure," *International Journal of Intelligent Robotics and Applications,* vol. 3, no. 3, pp. 239-254, 2019.

[2] K. W. J. I. T. Bowyer, and s. magazine, "Face recognition technology: security versus privacy," vol. 23, no. 1, pp. 9-19, 2004.

[3] A. Muiz Fayyaz, M. Kolivand, J. Alyami, S. Roy, and A. Rehman, "Computer Vision-Based Prognostic Modelling of COVID-19 from Medical Imaging," *Prognostic Models in Healthcare: AI and Statistical Approaches*, pp. 25-45: Springer, 2022.

[4] W. H. Organization, "COVID-19 transmission and protective measures," *World Health Organization. At https://www.who. int/westernpacific/emergencies/COVID-19/information/transmission-protective-measures*, 2020.

[5] R. Vaishya, M. Javaid, I. H. Khan, A. J. D. Haleem, M. S. C. Research, and Reviews, "Artificial Intelligence (AI) applications for COVID-19 pandemic," vol. 14, no. 4, pp. 337-339, 2020.

[6] J. Gee, P. Marquez, J. Su, G. M. Calvert, R. Liu, T. Myers, N. Nair, S. Martin, T. Clark, and L. Markowitz, "First month of COVID-19 vaccine safety monitoring—United States, December 14, 2020–January 13, 2021," *Morbidity and mortality weekly report,* vol. 70, no. 8, pp. 283, 2021.

[7] S. V. Militante, and N. V. Dionisio, "Real-time facemask recognition with alarm system using deep learning." pp. 106-110.

[8] A. M. Fayyaz, K. A. Al-Dhlan, S. U. Rehman, M. Raza, W. Mehmood, M. Shafiq, and J.-G. Choi, "Leaf Blights Detection and Classification in Large Scale Applications," *INTELLIGENT AUTOMATION AND SOFT COMPUTING,* vol. 31, no. 1, pp. 507-522, 2022.

[9] P. J. J. o. i. Brey, communication, and e. i. society, "Ethical aspects of facial recognition systems in public places," 2004.

[10] M. A. Shereen, S. Khan, A. Kazmi, N. Bashir, and R. J. J. o. a. r. Siddique, "COVID-19 infection: Emergence, transmission, and characteristics of human coronaviruses," vol. 24, pp. 91-98, 2020.

[11] P. Li, Y. Xu, Y. Wei, Y. J. I. T. o. P. A. Yang, and M. Intelligence, "Self-correction for human parsing," 2020.

[12] Y. Sun, B. Xue, M. Zhang, G. G. Yen, and J. J. I. t. o. c. Lv, "Automatically designing CNN architectures using the genetic algorithm for image classification," vol. 50, no. 9, pp. 3840-3854, 2020.

[13] D. K. Yadav, "Efficient method for moving object detection in cluttered background using Gaussian Mixture Model." pp. 943-948.

[14] W. Wójcik, K. Gromaszek, M. J. F. R.-S. C. Junisbekov, Subspace Projection, and E. Methods, "Face recognition: Issues, methods and alternative applications," pp. 7-28, 2016.

[15] P. Viola, and M. J. J. I. j. o. c. v. Jones, "Robust real-time face detection," vol. 57, no. 2, pp. 137-154, 2004.

[16] J. Mahajan, and A. Paithane, "Face detection on distorted images by using quality HOG features." pp. 439-444.

[17] Y.-L. Tian, T. Kanade, and J. F. Cohn, "Facial expression analysis," *Handbook of face recognition*, pp. 247-275: Springer, 2005.

[18] A. Liang, W. Liu, L. Li, M. R. Farid, and V. Le, "Accurate facial landmarks detection for frontal faces with extended tree-structured models." pp. 538-543.

[19] E. W. J. h. m. w. c. Weisstein, "Affine transformation," 2004.

[20] G. L. Wells, B. J. M. Hryciw, and Cognition, "Memory for faces: Encoding and retrieval operations," vol. 12, no. 4, pp. 338-344, 1984.

[21] A. Dosovitskiy, J. Tobias Springenberg, and T. Brox, "Learning to generate chairs with convolutional neural networks." pp. 1538-1546.

[22] H. A. Rowley, S. Baluja, T. J. I. T. o. p. a. Kanade, and m. intelligence, "Neural network-based face detection," vol. 20, no. 1, pp. 23-38, 1998.

[23] "Jessica Li (https://www.kaggle.com/jessicali9530), Accessed by 7/9/2022."

[24] J. Amin, M. Sharif, M. Yasmin, and S. L. Fernandes, "A distinctive approach in brain tumor detection and classification using MRI," *Pattern Recognition Letters,* vol. 139, pp. 118-127, 2020.

[25] J. Amin, M. Sharif, M. Yasmin, H. Ali, and S. L. Fernandes, "A method for the detection and classification of diabetic retinopathy using structural predictors of bright lesions," *Journal of Computational Science,* vol. 19, pp. 153-164, 2017.

[26] M. I. Sharif, J. P. Li, J. Amin, and A. Sharif, "An improved framework for brain tumor analysis using MRI based on YOLOv2 and convolutional neural network," *Complex & Intelligent Systems*, pp. 1-14, 2021.

[27] T. Saba, A. S. Mohamed, M. El-Affendi, J. Amin, and M. Sharif, "Brain tumor detection using fusion of hand crafted and deep learning features," *Cognitive Systems Research,* vol. 59, pp. 221-230, 2020.

[28] J. Amin, M. Sharif, M. Raza, T. Saba, and M. A. Anjum, "Brain tumor detection using statistical and machine learning method," *Computer methods and programs in biomedicine,* vol. 177, pp. 69-79, 2019.

[29] J. Amin, M. Sharif, M. Raza, and M. Yasmin, "Detection of brain tumor based on features fusion and machine learning," *Journal of Ambient Intelligence and Humanized Computing*, pp. 1-17, 2018.

[30] J. Amin, M. Sharif, and M. Yasmin, "A review on recent developments for detection of diabetic retinopathy," *Scientifica,* vol. 2016, 2016.

[31] J. Amin, M. Sharif, N. Gul, M. Yasmin, and S. A. Shad, "Brain tumor classification based on DWT fusion of MRI sequences using convolutional neural network," *Pattern Recognition Letters,* vol. 129, pp. 115-122, 2020.

[32] M. Sharif, J. Amin, M. Raza, M. Yasmin, and S. C. Satapathy, "An integrated design of particle swarm optimization (PSO) with fusion of features for detection of brain tumor," *Pattern Recognition Letters,* vol. 129, pp. 150-157, 2020.

[33] J. Amin, M. Sharif, M. Yasmin, T. Saba, M. A. Anjum, and S. L. Fernandes, "A new approach for brain tumor segmentation and classification based on score level fusion using transfer learning," *Journal of medical systems,* vol. 43, no. 11, pp. 1-16, 2019.

[34] J. Amin, M. Sharif, M. Raza, T. Saba, R. Sial, and S. A. Shad, "Brain tumor detection: a long short-term memory (LSTM)-based learning model," *Neural Computing and Applications,* vol. 32, no. 20, pp. 15965-15973, 2020.

[35] J. Amin, M. Sharif, M. Raza, T. Saba, and A. Rehman, "Brain tumor classification: feature fusion." pp. 1-6.

[36] J. Amin, M. Sharif, M. Yasmin, T. Saba, and M. Raza, "Use of machine intelligence to conduct analysis of human brain data for detection of abnormalities in its cognitive functions," *Multimedia Tools and Applications,* vol. 79, no. 15, pp. 10955-10973, 2020.

[37] J. Amin, A. Sharif, N. Gul, M. A. Anjum, M. W. Nisar, F. Azam, and S. A. C. Bukhari, "Integrated design of deep features fusion for localization and classification of skin cancer," *Pattern Recognition Letters,* vol. 131, pp. 63-70, 2020.

[38] J. Amin, M. Sharif, N. Gul, M. Raza, M. A. Anjum, M. W. Nisar, and S. A. C. Bukhari, "Brain tumor detection by using stacked autoencoders in deep learning," *Journal of medical systems,* vol. 44, no. 2, pp. 1-12, 2020.

[39] M. Sharif, J. Amin, M. Raza, M. A. Anjum, H. Afzal, and S. A. Shad, "Brain tumor detection based on extreme learning," *Neural Computing and Applications*, pp. 1-13, 2020.

[40] J. Amin, M. Sharif, A. Rehman, M. Raza, and M. R. Mufti, "Diabetic retinopathy detection and classification using hybrid feature set," *Microscopy research and technique,* vol. 81, no. 9, pp. 990-996, 2018.

[41] J. Amin, M. Sharif, M. A. Anjum, M. Raza, and S. A. C. Bukhari, "Convolutional neural network with batch normalization for glioma and stroke lesion detection using MRI," *Cognitive Systems Research,* vol. 59, pp. 304-311, 2020.

[42] N. Muhammad, M. Sharif, J. Amin, R. Mehboob, S. A. Gilani, N. Bibi, H. Javed, and N. Ahmed, "Neurochemical Alterations in Sudden Unexplained Perinatal Deaths—A Review," *Frontiers in pediatrics,* vol. 6, pp. 6, 2018.

[43] M. Sharif, J. Amin, M. W. Nisar, M. A. Anjum, N. Muhammad, and S. A. Shad, "A unified patch based method for brain tumor detection using features fusion," *Cognitive Systems Research,* vol. 59, pp. 273-286, 2020.

[44] R. Iyer, P. Shashikant Ringe, R. Varadharajan Iyer, and K. J. A. I. J. R. E. T. Prabhulal Bhensdadiya, "Comparison of YOLOv3, YOLOv5s and MobileNet-SSD V2 for Real-Time Mask Detection," vol. 8, pp. 1156-1160, 2021.

[45] I. B. Venkateswarlu, J. Kakarla, and S. Prakash, "Face mask detection using mobilenet and global pooling block." pp. 1-5.

[46] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "Mobilenetv2: Inverted residuals and linear bottlenecks." pp. 4510-4520.

[47] M. Jiang, X. Fan, and H. J. a. p. a. Yan, "Retinamask: A face mask detector," 2020.

[48] T. J. S. Fushiki, and Computing, "Estimation of prediction error by using K-fold cross-validation," vol. 21, no. 2, pp. 137-146, 2011.

[49] S. Subandi, A. S. B. Nugroho, N. Nurkamilia, and A. A. J. J. Syahidi, "Body Temperature Measurement Tool for Early Detection of COVID-19 Based on Interactive Augmented Reality Technology and Sensor MLX90614: Framework and Prototyping," vol. 8, no. 5, pp. 141-148, 2021.

[50] T. A. Kumar, R. Rajmohan, M. Pavithra, S. A. Ajagbe, R. Hodhod, and T. Gaber, "Automatic face mask detection system in public transportation in smart cities using IoT and deep learning," *Electronics,* vol. 11, no. 6, pp. 904, 2022.

[51] H. Goyal, K. Sidana, C. Singh, A. Jain, and S. Jindal, "A real time face mask detection system using convolutional neural network," *Multimedia Tools and Applications,* vol. 81, no. 11, pp. 14999-15015, 2022.

[52] Z. Qin, Z. Guo, and Y. Lin, "An Implementation of Face Mask Detection System Based on YOLOv4 Architecture." pp. 207-213.