An Automated System to Limit COVID-19 Using Facial Mask Detection in Smart City Network

Mohammad Marufur Rahman
Computer Science and Engineering
Khulna University of Engineering & Technology
Khulna-9203, Bangladesh
inkmarufnayem@gmail.com

Md. Motaleb Hossen Manik
Computer Science and Engineering
Khulna University of Engineering & Technology
Khulna-9203, Bangladesh
mkmanik557@gmail.com

Md. Milon Islam
Computer Science and Engineering
Khulna University of Engineering & Technology
Khulna-9203, Bangladesh
milonislam@cse.kuet.ac.bd

Saifuddin Mahmud
Advanced Telerobotics Research Lab
Computer Science
Kent State University
Kent, Ohio, USA
smahmud2@kent.edu

Jong-Hoon Kim
Advanced Telerobotics Research Lab
Computer Science
Kent State University
Kent, Ohio, USA
jkim72@kent.edu

***Abstract*— COVID-19 pandemic caused by novel coronavirus is continuously spreading until now all over the world. The impact of COVID-19 has been fallen on almost all sectors of development. The healthcare system is going through a crisis. Many precautionary measures have been taken to reduce the spread of this disease where wearing a mask is one of them. In this paper, we propose a system that restrict the growth of COVID-19 by finding out people who are not wearing any facial mask in a smart city network where all the public places are monitored with Closed-Circuit Television (CCTV) cameras. While a person without a mask is detected, the corresponding authority is informed through the city network. A deep learning architecture is trained on a dataset that consists of images of people with and without masks collected from various sources. The trained architecture achieved 98.7% accuracy on distinguishing people with and without a facial mask for previously unseen test data. It is hoped that our study would be a useful tool to reduce the spread of this communicable disease for many countries in the world.**

Keywords—Facial Mask Detection, COVID-19, Deep Learning, Convolutional Neural Network, Smart City.

I. Introduction

A new strain which has not previously been identified in humans is novel coronavirus (nCoV). Coronaviruses (CoV) are a wide group of viruses which cause illness that range from colds to deadly infections like Middle East Respiratory Syndrome (MERS) and Severe Acute Respiratory Syndrome (SARS) [1]. The first infected patient of coronavirus has been found in December 2019. From that period, COVID-19 has become a pandemic all over the world [2]. People all over the world are facing challenging situations due to this pandemic. Every day a large number of people are being infected and died. At the time of writing this paper, almost 16,207,130 infected cases have been confirmed where 648,513 are death [3]. This number is increasing day by day. Fever, dry cough, tiredness, diarrhea, loss of taste, and smell are the major symptoms of coronavirus which is declared by the World Health Organization (WHO) [4]. Many precautionary measures have been taken to fight against coronavirus. Among them cleaning hands, maintaining a safe distance, wearing a mask, refraining from touching eyes, nose, and mouth are the main, where wearing a mask is the simplest one. COVID-19 is a disease that spread from human to human which can be controlled by ensuring proper use of a facial mask. The spread of COVID-19 can be limited if people strictly maintain social distancing and use a facial mask. Very sadly, people are not obeying these rules properly which is speeding the spread of this virus. Detecting the people not obeying the rules and informing the corresponding authorities can be a solution in reducing the spread of coronavirus.

A face mask detection is a technique to find out whether someone is wearing a mask or not. It is similar to detect any object from a scene. Many systems have been introduced for object detection. Deep learning techniques are highly used in medical applications [5], [6]. Recently, deep learning architectures [7] have shown a remarkable role in object detection. These architectures can be incorporated in detecting the mask on a face. Moreover, a smart city [8] means an urban area that consists of many IoT sensors to collect data. These collected data are then used to perform different operations across the city. This includes monitoring traffic, utilities, water supply network, and many more. Recently, the growth of COVID-19 can be reduced by detecting the facial mask in a smart city network.

This paper aims at designing a system to find out whether a person is using a mask or not and informing the corresponding authority in a smart city network. Firstly, CCTV cameras are used to capture real-time video footage of different public places in the city. From that video footage, facial images are extracted and these images are used to identify the mask on the face. The learning algorithm Convolutional Neural Network (CNN) is used for feature extraction from the images then these features are learned by multiple hidden layers. Whenever the architecture identifies people without face mask this information is transferred through the city network to the corresponding authority to take necessary actions. The proposed system appraised promising output on data collected from different sources. We also represented a system that can ensure proper enforcement of the law on people who are not following basic health guidelines in this pandemic situation.

The remainder of the paper is arranged accordingly. The most recent works for facial mask detection is described in

Section II. In Section III, the proposed methodology for developing the whole system is described. Section IV analyses the results obtained from the developed system. The conclusion is drawn in Section V. Lastly, the limitations with potential further works are depicted in Section VI.

II. Related Works

In the meantime, many systems have been developed for COVID-19 in smart city networks. BlueDot and HealthMap services have been introduced in [9]. BlueDot method was first used to mark the cluster of unusual pneumonia in Wuhan which finally detected the disease as a pandemic. It also predicted that the virus would spread from Wuhan to Bangkok, Taipei, Singapore, Tokyo and Hong Kong. HealthMap service, based on San Francisco, spotted the patients with a cough which is the initial sign of COVID-19, using Artificial Intelligence (AI) and big data. A study on using facemask to restrict the growth of COVID-19 is introduced in [10]. The study indicated that the masks that are adequately fit, effectively interrupt the spread of droplets expelled when coughing or sneezing. Masks that are not perfectly fitted, also capable of retaining airborne particles and viruses. Allam and Jones [11] proposed a framework on smart city networks focusing on how data sharing should be performed during the outbreak of COVID-19. The proposed system discussed the prospects of Urban Health Data regarding the safety issues of the economy and national security. In the system, the data is collected from various points of the city using sensors, trackers, and from laboratories.

A face mask detecting model named RetinaFaceMask combining with a cross-class object removal algorithm is proposed by Jiang et al. [12]. The developed model includes one stage detector consisting feature pyramid network that results in slightly higher precision and recall than the baseline result. For reducing the shortage of datasets, they have applied transfer learning, a well-known deep learning technique. Gupta et al. [13] proposed a model to enforce the social distance using smart city and Intelligent Transportation System (ITS) during COVID-19 pandemic. Their model described the deploying sensors in different places of the city to monitor the real-time movement of objects and offered a data-sharing platform. A noticeable contribution of a smart city in controlling the spread of coronavirus in South Korea is explained by Won Sonn and Lee [14]. A time-space cartographer speeded up the contact tracking in the city including patient movement, purchase history, cell phone usages, and cell phone location. Real-time monitoring has been carried out on CCTV cameras in the hallways of residential buildings.

Singh et al. [15] put their focus on how IoT can fight against COVID-19. The developed system emphasizes on inter-connected devices or operations to track the patients along with wary cases. A well-informed group using inter-connected devices is formed to identify the clusters significantly. A remarkable pandemic control model without lockdown in a smart city has been outlined by Sonn et al. [16]. The patients have been interviewed and their past movement has been monitored. They have claimed that some patients tried to conceal about their past mobility but real-time tracking system found the exact information. Jaiswal et al. [17] proposed a way to minimize the risk during COVID-19. Their proposed model used the position of technology to track infected people. Drones and Robot technologies have been applied as medical personnel for providing adequate services to infected people. The development of smart cities under COVID-19 and controlling the pandemic in China has been reviewed by Wang et al. [18]. The continuous supply of essential materials and contactless logistic distribution of systems to society made the way to reduce the spread of coronavirus. ITS and real-time map reflection methods have been used to block the movement of vehicles during the pandemic. In addition, driverless vehicles have been used to monitor the scenarios across the city.

III. Methodology

We proposed an automated smart framework for screening persons who are not using a face mask in this paper. In the smart city, all public places are monitored by CCTV cameras. The cameras are used to capture images from public places; then these images are feed into a system that identifies if any person without face mask appears in the image. If any person without a face mask is detected then this information is sent to the proper authority to take necessary actions. The block diagram of the developed framework is depicted in Fig. 1. All the blocks of the developed system are described as follows.

A. Image Preprocessing

The images captured by the CCTV cameras required preprocessing before going to the next step. In the preprocessing step, the image is transformed into a grayscale image because the RGB color image contains so much redundant information that is not necessary for face mask detection. RGB color image stored 24 bit for each pixel of the image. On the other hand, the grayscale image stored 8 bit for each pixel and it contained sufficient information for classification. Then, we reshaped the images into (64×64) shape to maintain uniformity of the input images to the architecture. Then, the images are normalized and after normalization, the value of a pixel resides in the range from 0 to 1. Normalization helped the learning algorithm to learn faster and captured necessary features from the images.

B. Deep Learning Architecture

The deep learning architecture learns various important nonlinear features from the given samples. Then, this learned architecture is used to predict previously unseen samples. To train our deep learning architecture, we collected images from different sources. The architecture of the learning technique highly depends on CNN. All the aspects of deep learning architecture are described below.

i) Dataset Collection: Data from two different sources [19], [20] are collected for training and testing the model. We collected a total of 858 images of people with masks and 681 images of people without a mask. For training purposes, 80% images of each class are used and the rest of the images are utilized for testing purposes. Fig. 2 shows some of the images of two different classes.

ii) Architecture Development: The learning model is based on CNN which is very useful for pattern recognition from images [21]. The network comprises an input layer, several hidden layers and an output layer. The hidden layers consist of multiple convolution layers that learn suitable filters for important feature extraction from the given samples. The features extracted by CNN are used by multiple dense neural networks for classification purposes. The architecture of the developed network is illustrated in Table I. The architecture contains three pairs of convolution layers each followed by

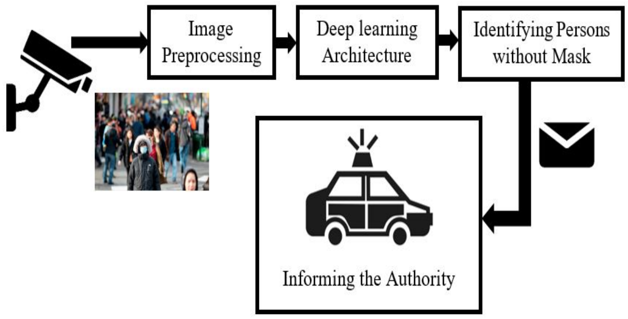

Fig. 1. Block diagram of the proposed system.

People without mask People with mask

Fig. 2. Sample images from the used dataset.

TABLE I. THE ARCHITECTURE OF THE DEEP LEARNING NETWORK

Layer	Type	Kernel	Kernel Size	Output Size
1	Convolution2D	32	(3×3)	(62×62×32)
2	Convolution2D	32	(3×3)	(60×60×32)
3	MaxPooling2D	-	(2×2)	(30×30×32)
4	Convolution2D	32	(3×3)	(28×28×32)
5	Convolution2D	32	(3×3)	(26×26×32)
6	MaxPooling2D	-	(2×2)	(13×13×32)
7	Convolution2D	32	(3×3)	(11×11×32)
8	Convolution2D	32	(3×3)	(9×9×32)
9	MaxPooling2D	-	(2×2)	(4×4×32)
10	Flatten	-	-	512
11	Dense	-	-	100
12	Dropout	-	-	100
13	Dense	-	-	30
14	Dropout	-	-	30
15	Dense	-	-	10
16	Dropout	-	-	10
17	Dense	-	-	2

one max pooling layer. This layer decreases the spatial size of the representation and thereby reduces the number of parameters. As a result, the computation is simplified for the network. Then, a flatten layer reshapes the information into a vector to feed into the dense network. Three pairs of dense and dropout layers learn parameters for classification. The dense layer comprises a series of neurons each of them learn nonlinear features. The dropout layer prevents the network from overfitting by dropping out units. Finally, a dense layer containing two neurons distinguishes the classes.

iii) Screening and Informing the Authority: The main goal of our proposed system is screening persons who are not following guidelines of using a facial mask. The learning architecture identifies whether any input image contains persons without a face mask. If such a person is detected, then this information is sent to the proper authority. The GPS location of the CCTV camera captured the person without a mask along with the image and the exact time is sent via SMS to the corresponding authority. They would come to the locality where the person without a face mask was detected and took necessary actions. If proper actions are taken, then people might not come in public places without a facial mask that would help greatly to limit the growth of COVID-19.

IV. RESULT ANALYSIS

By preserving a reasonable proportion of different classes, the dataset is partitioned into training and testing set. The dataset comprises of 1539 samples in total where 80% is used in training phase and 20% is used in testing phase. The training and testing dataset contains 1231 and 308 images respectively. The developed architecture is trained for 100 epochs since further training results cause overfitting on the training data. Overfitting occurs when a model learns the unwanted patterns of the training samples. Hence, training accuracy increases but test accuracy decreases. Fig. 3 and Fig. 4 show the graphical view of accuracy and loss respectively. The trained model showed 98.7% accuracy and AUC of 0.985 on the unseen test data.

In Fig. 3, the accuracy curve of training and testing is shown for about 100 epochs. From Fig. 3, it is realized that the training and testing accuracy are almost identical. This means the model has a decent generalization ability for previously unseen data and it does not cause overfitting of the training data. In Fig. 4, loss curves of training and testing phases are shown. Here, it is evident that the training loss is decreasing over increasing epochs. The testing loss is lower than training loss for about 30 epochs but after that, it started increasing w means the confidence of prediction started decreasing. The testing loss fluctuates between an acceptable range and it falls about at 98^{th} epoch.

Table II represents the confusion matrix of the testing phase. The developed architecture misclassifies only 04 samples out of 308 samples. It classifies 01 sample as with mask while it is in without mask class and classifies 03 samples as without mask while these were in with mask class. The main aim of the system is to identify samples within without mask class and this architecture misclassified only 01 sample of this class that shows the reliability of the developed system.

Fig. 5 depicts the receiver operating characteristic (ROC) curve of the proposed framework. This illustrates the prediction ability of the classifier at different thresholds. Two parameters are plotted in the ROC curve; one is the true positive rate (TPR) and other is the false positive rate (FPR) measured using (1) and (2) respectively. TPR and FPR are calculated for different threshold and these values are plotted as ROC curve. The area under the ROC curve (AUC) measures the performance of the binary classifier for all possible thresholds. The value of AUC ranges from 0 to 1. When a model predicts 100% correct its AUC is 1 and when it predicts 100% wrong then its AUC is 0. The AUC achieved form our classifier is 0.985 that points towards a decent classifier.

$$\textit{True Positive Rate} = \frac{\textit{True Positive}}{\textit{True Positive} + \textit{False Negative}} \quad (1)$$

$$\textit{False Positive Rate} = \frac{\textit{False Positive}}{\textit{True Negative} + \textit{False Positive}} \quad (2)$$

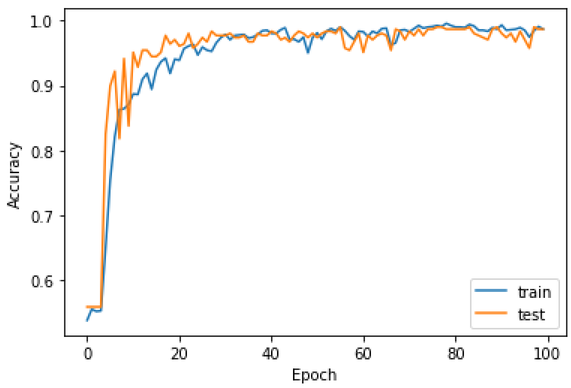

Fig. 3. Accuracy of the developed system for training and testing phase.

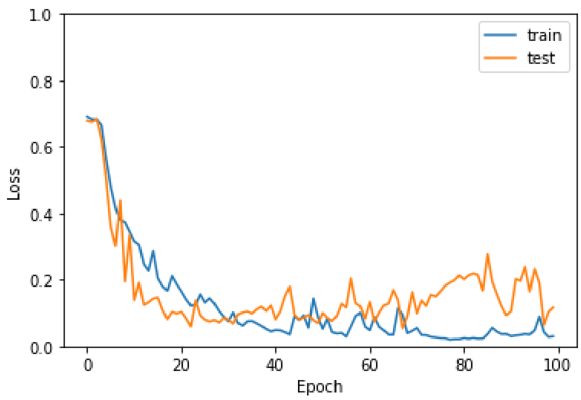

Fig. 4. Loss of the developed system for training and testing phase.

TABLE II. THE CONFUSION MATRIX OF THE DEVELOPED SYSTEM

		Predicted Class	
		Without Mask	With Mask
True Class	Without Mask	134	1
	With Mask	3	170

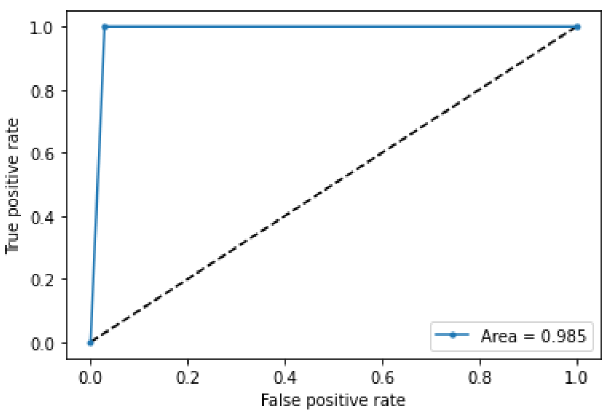

Fig. 5. ROC of the classification network.

V. CONCLUSION

This paper presents a system for a smart city to reduce the spread of coronavirus by informing the authority about the person who is not wearing a facial mask that is a precautionary measure of COVID-19. The motive of the work comes from the people disobeying the rules that are mandatory to stop the spread of coronavirus. The system contains a face mask detection architecture where a deep learning algorithm is used to detect the mask on the face. To train the model, labeled image data are used where the images were facial images with masks and without a mask. The proposed system detects a face mask with an accuracy of 98.7%. The decision of the classification network is transferred to the corresponding authority. The system proposed in this study will act as a valuable tool to strictly impose the use of a facial mask in public places for all people.

VI. LIMITATIONS AND FUTURE WORKS

The developed system faces difficulties in classifying faces covered by hands since it almost looks like the person wearing a mask. While any person without a face mask is traveling on any vehicle, the system cannot locate that person correctly. For a very densely populated area, distinguishing the face of each person is very difficult. For this type of scenario, identifying people without face mask would be very difficult for our proposed system. In order to get the best result out of this system, the city must have a large number of CCTV cameras to monitor the whole city as well as dedicated manpower to enforce proper laws on the violators. Since the information about the violator is sent via SMS, the system fails when there is a problem in the network.

The proposed system mainly detects the face mask and informs the corresponding authority with the location of a person not wearing a mask. Based on this, the authority has to send their personnel to find out the person and take necessary actions. But this manual scenario can be automated by using drones and robot technology [22], [23] to take action instantly. Furthermore, people near to the person not wearing a mask may be alerted by an alarm signal on that location, and displaying the violators face in a LED screen to maintain a safe distance from the person would be a further study.

REFERENCES

[1] WHO EMRO | About COVID-19 | COVID-19 | Health topics. [Online]. Available: *http://www.emro.who.int/health-topics/corona-virus/about-covid-19.html*, accessed on: Jul. 26, 2020.

[2] H. Lau *et al.*, "Internationally lost COVID-19 cases," *J. Microbiol. Immunol. Infect.*, vol. 53, no. 3, pp. 454–458, 2020.

[3] Worldometer, "Coronavirus Cases,". [Online]. Available: *https://www.worldometers.info/coronavirus*, accessed on: Jul. 26, 2020.

[4] L. Li *et al.*, "COVID-19 patients' clinical characteristics, discharge rate, and fatality rate of meta-analysis," *J. Med. Virol.*, vol. 92, no. 6, pp. 577–583, Jun. 2020.

[5] M. Z. Islam, M. M. Islam, and A. Asraf, "A Combined Deep CNN-LSTM Network for the Detection of Novel Coronavirus (COVID-19) Using X-ray Images," *Informatics in Medicine Unlocked*, vol. 20, pp. 100412, Aug. 2020.

[6] L. J. Muhammad, M. M. Islam, S. S. Usman, and S. I. Ayon, "Predictive Data Mining Models for Novel Coronavirus (COVID-19) Infected Patients' Recovery," *SN Comput. Sci.*, vol. 1, no. 4, p. 206, Jun. 2020.

[7] L. Liu *et al.*, "Deep Learning for Generic Object Detection: A Survey," *Int. J. Comput. Vis.*, vol. 128, no. 2, pp. 261–318, Sep. 2018.

[8] L. Calavia, C. Baladrón, J. M. Aguiar, B. Carro, and A. Sánchez-Esguevillas, "A Semantic Autonomous Video Surveillance System for Dense Camera Networks in Smart Cities," *Sensors*, vol. 12, no. 8, pp. 10407–10429, Aug. 2012.

[9] G. Halegoua, "Smart City Technologies," *Smart Cities*, 2020, doi: 10.7551/mitpress/11426.003.0005.

[10] L. P. Garcia, "Uso de máscara facial para limitar a transmissão da COVID-19," *Epidemiol. e Serv. saude Rev. do Sist. Unico Saude do Bras.*, vol. 29, no. 2, p. e2020023, 2020.

[11] Z. Allam and D. S. Jones, "On the Coronavirus (COVID-19) Outbreak and the Smart City Network: Universal Data Sharing Standards Coupled with Artificial Intelligence (AI) to Benefit Urban Health Monitoring and Management," *Healthcare*, vol. 8, no. 1, p. 46, 2020.

[12] M. Jiang, X. Fan, and H. Yan, "RetinaMask: A Face Mask detector," 2020. [Online]. Available: *http://arxiv.org/abs/2005.03950*.

[13] M. Gupta, M. Abdelsalam, and S. Mittal, "Enabling and Enforcing Social Distancing Measures using Smart City and ITS Infrastructures: A COVID-19 Use Case," 2020. [Online]. Available: *https://arxiv.org/abs/2004.09246.*

[14] J. Won Sonn and J. K. Lee, "The smart city as time-space cartographer in COVID-19 control: the South Korean strategy and democratic control of surveillance technology," *Eurasian Geogr. Econ.*, pp. 1–11, May. 2020.

[15] R. P. Singh, M. Javaid, A. Haleem, and R. Suman, "Internet of things (IoT) applications to fight against COVID-19 pandemic," *Diabetes Metab. Syndr. Clin. Res. Rev.*, vol. 14, no. 4, pp. 521–524, Jul. 2020.

[16] J. W. Sonn, M. Kang, and Y. Choi, "Smart city technologies for pandemic control without lockdown," *Int. J. Urban Sci.*, vol. 24, no. 2, pp. 149–151, 2020.

[17] R. Jaiswal, A. Agarwal, and R. NEGI, "Smart Solution for Reducing the COVID-19 Risk using Smart City Technology," *IET Smart Cities*, vol. 2, pp. 82–88, 2020.

[18] X. Wang, X. Le, and Q. Lu, "Analysis of China's Smart City Upgrade and Smart Logistics Development under the COVID-19 Epidemic," *J. Phys. Conf. Ser.*, vol. 1570, p. 012066, 2020.

[19] Face Mask Detection | Kaggle. [Online]. Available: *https://www.kaggle.com/andrewmvd/face-mask-detection*, accessed on: Jul. 27, 2020.

[20] GitHub-prajnasb/observations. [Online]. Available: *https://github.com/prajnasb/observations*, accessed on: Jul. 27, 2020.

[21] A. Khan, A. Sohail, U. Zahoora, and A. S. Qureshi, "A Survey of the Recent Architectures of Deep Convolutional Neural Networks," *Artif. Intell. Rev.*, Jan. 2019.

[22] I. S. Cardenas *et al.*, "Telesuit: design and implementation of an immersive user-centric telepresence control suit," in *Proceedings of the 23rd International Symposium on Wearable Computers - ISWC '19*, New York, NY, USA, 2019, pp. 261–266.

[23] D. Y. Kim, I. S. Cardenas, and J.-H. Kim, "Engage/Disengage: Control Triggers for Immersive Telepresence Robots," in *Proceedings of the 5th International Conference on Human Agent Interaction*, New York, NY, USA, 2017, pp. 495–499.