2020 IEEE Pune Section International Conference (PuneCon)
Vishwakarma Institute of Technology, Pune India. Dec 16-18, 2020

A Fast, Automatic Risk Detector for COVID-19

Bhushan Bhagwan Gawde
Dept. of Computer Engineering and IT
Veermata Jijabai Technological Institute
Mumbai, India.
bbgawde_b14@ce.vjti.ac.in

2020 IEEE Pune Section International Conference (PuneCon) | 978-1-7281-9600-8/20/$31.00 ©2020 IEEE | DOI: 10.1109/PuneCon50868.2020.9362389

***Abstract*— With its lethal spread to more than 200 countries, COVID-19 has brought a global crisis, affecting more than 3 crore people across the world. Viruses don't have a cure, and this makes the population vulnerable and heavily rely on preventing the infection. Hence, following the rules of social distancing and wearing a face mask are two very essential approaches to fight against this pandemic. Motivated by this notion, this work proposes a deep learning-based framework for automating the detection of risk due to COVID-19. The proposed framework utilizes YOLOv3 object detector to detect whether a person has worn a mask. In case of absence of mask, to categorize the level of risk, the person's age category is estimated, and the result of the risk detector is displayed on the image with a bounding box. In case of multiple boxes, the framework also calculates the distance between them to check whether the rules of social distancing are being followed. The result of the YOLOv3 model is compared with popular state-of-the-art model, Faster Region-based Convolutional Neural Network. From the experimental analysis, it is concluded that YOLOv3 object detector displays best results with respect to the trade-off between speed and accuracy.**

Keywords*— *COVID-19, Object Detection, Social Distancing, Age Classification, Object Classification, Face mask Detection.

I. Introduction

COVID-19 belongs to a family of infectious diseases caused by the coronavirus. Initially, the novel coronavirus (SARS-CoV-2) was reported in Wuhan, China, during late December 2019. As of September 21, this novel coronavirus has spread to about 212 countries and territories across the world, with 74,72,248 active cases and 9,69,252 people succumbed to death [1]. All around the world, several healthcare organizations, distinguished scientists, and many other medical experts have been trying to develop vaccines for this deadly virus. The University of Oxford along with AstraZeneca, a multinational pharmaceutical company has been developing a potential vaccine for this virus. The vaccine is currently in the stage of human trials and is showing promising results as per the reports [3]. But still, no one has been successful in inventing a proper vaccine which could be eligible to be sent for mass production and distribution. The situation all around the world has been deteriorating and hence, all of us need to look for ways to control the spread of this deadly virus. Respiratory diseases are transmissible diseases where the mode of transmission of the virus is one of the most critical factors to be considered for stopping its spread in the community. Viruses do not have a cure and the infected person's immune system should generate anti-bodies in order to fight the virus. Generally, the immune system of young people is better than that of aged ones. As per the reports [4], this virus has proved to be more fatal for aged people (people with age greater than 45 years). The solution to this problem is two-fold; following social distancing and always wearing a face mask in public. For controlling the spread, almost all the affected countries had implemented a lockdown but it was very difficult to follow a complete lockdown because essential services like the supply and purchase of groceries and medicines could not be stopped. Hence, many governments have mandated the use of a face mask for the people in public places like grocery shops and pharmacies while strictly following social distancing norms. According to WHO, people should be at least 6 feet apart from each other [5]. Social distancing aims to minimize or interrupt the transmission of the virus by avoiding the physical contact between infected and healthy people. Whenever an infected person sneezes or coughs, the virus can spread through the droplets of that person to the people close to him/her. A recent study has indicated that infected people with mild or no symptoms can transmit the virus to someone who has a weaker immune system and can prove to be fatal [6]. Hence, not just social distancing, but wearing a mask in public is also very important for stopping the transmission of the virus. This project is an attempt to automatically detect how much risk a person is at, due to COVID-19. The contribution of this project can be summarized by the following points:

1) *Given a video or a list of images, the proposed deep learning framework will check whether a person has worn a face mask or not.*
2) *If the person hasn't worn a mask, the age of the person will be estimated and he/she will be notified of the severity of the risk due to COVID-19, depending upon his/her age.*
3) *When there are multiple people in the video, this framework will also calculate the distance between them to check whether the people are following the rules of social distancing or violating them.*

This paper also compares the performance of the popular state-of-the-art object detection networks in the task of face mask detection. The paper structure is as follows: Section II will discuss about the background study related to the importance of face mask and social distancing during the coronavirus pandemic and will also talk briefly about the known object detection and age classification methodologies. Section III will introduce the deep learning-based framework proposed to automatically estimate the risk due to COVID-19. Section IV will discuss the experiments and the corresponding results. Lastly, Section V and Section VI will discuss the future scope of this project and the conclusion respectively.

II. Background Study and Related Work

Initially, when the virus was spreading very rapidly in different countries across the world, in order to control the spread, many countries had implemented a lockdown. In countries like Italy and Spain, which were worst hit in the initial stages of the spread of this virus, a complete lockdown was implemented with the belief that the virus will cease to spread. In India, the lockdown was implemented in different stages. But this lockdown caused many industries to shut down and lose their businesses which indirectly led to the economy of the country to suffer. We all have realized, and the World Health Organization too, has mentioned that the coronavirus will never be completely out of our lives and we have to learn to live with it [7]. Now, as the pandemic continues, different parts of the countries across the world are being unlocked slowly in order to bring back the economy where it was before the lockdown. But now, the situation has completely changed and it has become mandatory for businesses, that the employees strictly follow social distancing, and continue to wear face masks when they step out of their houses to help prevent the spread of the virus. During the initial stages, when the virus started spreading, the usage of masks was not recommended because the transmission dynamics of the novel coronavirus were unknown to many people. There were a lot of questions arising throughout the world regarding the usage of masks. The experts weren't aware of the extent to which asymptomatic people affected with COVID-19 could spread the virus unknowingly to others. Ferretti et al. [8] aimed to develop a mathematical model for estimating the contribution from different types of transmission routes. According to the estimations made by the model, the contribution to the basic reproductive number (R0) includes 46% from presymptomatic individuals, 38% from symptomatic individuals, and 10% from asymptomatic individuals. In order to unearth the reason behind the rapid dissemination of coronavirus, Li et al. [9] proposed a model for quantifying the fraction of undocumented but infectious cases, since this is a very critical characteristic that defines the potential of a pandemic. The authors developed a mathematical model which not only took into consideration the number of reported cases, but also the number of unreported infectious cases. The spatiotemporal dynamics of the infections were simulated depending on the number of people moving to and from various cities in China. They used a networked metapopulation model, the number of reported infections within China, and mobility data of people to infer epidemiological characteristics associated with SARS-CoV-2. According to the studies of Yan et al. [10], not just while sneezing and coughing, the patients infected with influenza shed infectious virus particles while normal breathing as well. In a country like India, China, etc. where the population is very huge, the usage of masks by individuals would be much more important than in other countries. This will help in retaining the contagious droplets, and aerosols that can infect others and contaminate surfaces. According to the studies of Tracht et al. [11], the mathematical model of an influenza pandemic suggested that substantial number of cases might still be prevented even if masks are only 20% effective at reducing transmission. Not just theoretically, it has also been practically proven that masks are effective in controlling the spread of the virus. Healthcare workers usually wear surgical masks to protect themselves from patients with respiratory infections who have viruses like influenza and common cold virus in their exhaled breath. N-95 mask which is a medical-grade mask is said to provide the best protection against the virus [12]. The use of face masks is two-fold. It not only helps to reduce the inhalation of aerosols released from the infectious individual's nostrils, but also helps to block the virus particles exhaled by an infectious individual. Also, it serves as a successful barrier when a person coughs or sneezes. According to the research work in [13], the normally exhaled breath of a person is made up of tiny water droplets and fluid from the lining of the lungs. This lining consists of dissolved viruses and bacteria, along with metabolites and other compounds. When a face mask is worn, these particles get collected on the inside of the mask, and a layer is formed which attracts more particles. It was found that these masks capture a wide variety of proteins that are even smaller in size than the coronavirus particles. Hence, the usage of masks during this COVID-19 pandemic is not only our way of safeguarding ourselves from the virus but also our responsibility towards safeguarding others around us.

Similarly, following the norms of social distancing is also very important to stop the spread of this infectious disease. Countries like China, Italy, Spain etc. where the number of COVID-19 cases were skyrocketing in initial stages, started following the approach of social distancing and after a month or two, the daily rise in confirmed cases of coronavirus significantly decreased in these countries [15]. The benefits of the practice of social distancing in controlling COVID-19 is clearly visible from this fact. Social distancing helps to control the spread of the virus by keeping people apart from each other which makes it difficult for the virus to propagate towards the nasal tract of people. Prem et al. [16] along with his fellow authors, aimed to use SEIR models to simulate the trajectory of the coronavirus outbreak. On 16th of April 2020, Landing AI, under the leadership of one of the most infamous personalities in the world of AI, Dr. Andrew Ng, announced that they will be developing an AI tool [17] which will easily monitor whether safe physical distancing is being followed in the workplace. Since the novel coronavirus pandemic began, many countries have been trying to implement technology-based solutions to contain the spread of the virus. Many countries have developed real-time systems for monitoring social distancing. Countries like India and South Korea, are utilizing GPS to track the movements of the suspected infectious people to monitor any possibility of the healthy people getting exposed to them. In India, the government has made it mandatory for people to install the "Arogya Setu" Application [18]. This application works with the help of GPS and Bluetooth to locate the COVID-19 patients in the vicinity area of an individual. It also lets a person know his/her health status i.e. Safe or Not safe depending on the people whom he/she has been in contact with. Some law enforcement departments have been using surveillance cameras to detect mass gatherings of people [20]. Such manual intervention in these critical times might help control the situation, but it also brings a unique set of challenges to the government bodies.

Human detection using a surveillance system is an established area of research. It uses manual methods for identifying unusual activities and hence, it has limited capabilities. Human detection is a very ambitious goal, the reason being various constraints like low-resolution video, varying pose, background complexities and limited machine capabilities [21]. Detecting an object which is in motion, is a very challenging task. It incorporates two stages: object detection and object classification [22]. The primitive approaches used for detecting moving objects in a video are background subtraction, spatio-temporal filtering and optical flow. In the background subtraction method, the pixel-level or block-level differences between the current frame and a

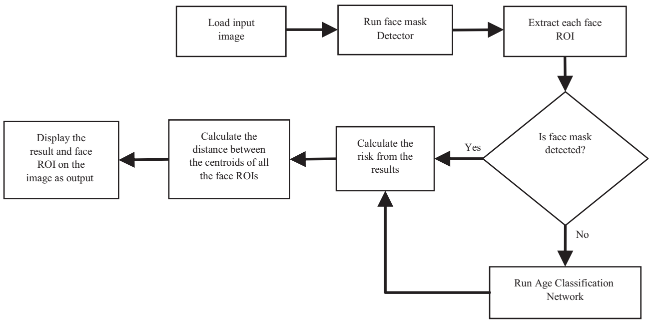


Fig. 1. Flowchart of the Proposed Methodology

background frame is computed. Adaptive Gaussian mixture, temporal differencing, hierarchical background models, and non-parametric background are the most popular approaches of background subtraction [23], [24]. In the optical flow-based object detection technique, flow vectors associated with the motion of the object are calculated over a time interval in order to identify which regions of the images are in motion for a given sequence of images [27]. The major downfall of the optical flow-based techniques is the computational complexities and also, their sensitivity to noise, colour, lighting, etc. Aslani et al. [26] aimed to use a spatio-temporal filtering approach in which the motion parameters are identified by using 3D features of the moving person in the particular image sequence. These methods are advantageous due to their lower computational complexity but show limited performance because of noise and changing background patterns. There have been recent advancements in the domain of AI wherein intelligent systems can be used to detect and capture human activities. Human detection is a specialized task of a more generalized task known as object detection. In the last decade, convolutional neural networks (CNNs) which use a region-proposal based approach for object detection have been developed. They not only generate the objectness score and the class labels but also generate the bounding boxes around the object of interest for visualization. Some of the networks employing region-based detection techniques are RCNN [29], Fast RCNN [30], and Faster RCNN [31]. These approaches do not employ single-stage detection approaches. Instead, they are two-stage networks, and hence, although these methods are efficient, they suffer in terms of larger training time requirements. As opposed to these models, YOLO family of models considers a regression-based method to interpret the class probabilities and the bounding box coordinates. There are many single stage detector models like YOLOv1 [32], YOLOv2 [33], SSD [34], YOLOv3 [35] etc. This project employs YOLOv3 for the task of face mask detection. The network divides the image into several grids and each of these grids is responsible for representing bounding boxes along with the class probability scores. This approach offers excellent improvements in terms of speed while they suffer in terms of accuracy as compared to region-proposal based approaches. The YOLO model is trained to detect various objects present in a given image in a single stage. This model has been efficiently incorporated in this project work for the detection of the presence of mask on a person's face.

For the task of age classification, many research findings have been reported in the last few years. The task of automatic age classification using Machine Learning methods is not new. There were numerous research attempts wherein researchers developed different algorithms like Haar Cascade Classifier, Local Binary Patterns Histogram Face recognizer, KNN classifier for finding a face in the given input image. This process was supplemented with techniques for predicting the age group. Zhang et al. [36] focused on using a multi-task extension of a variant of the Gaussian Process called Warped Gaussian Process for estimating the age of a person. This method considers the age as a regression function which is implicitly defined by the kernel functions. Han et al. [37] proposed a hierarchical approach for automatic age estimation from facial images. The authors used a binary decision tree based on SVM also called SVM-BDT for this task. Following this, Almeida et al. [38] developed a biometric system for automatic age detection of people, that uses SVM for the final classification stage. In another research, Tal et al. [39], [40] proposed the use of a deep convolutional neural network for the task of automated age and gender prediction from images. Their network is a simple network with 3 convolutional and 3 fully-connected layers. The age classification network used in this project work is a similar network inspired by the research work of Tal et al.

III. Proposed Framework

In this paper, a deep learning-based framework for automatically determining how much risk a person is at due to COVID-19, has been presented. It utilizes object detection and image classification approach for determining the same. Figure 1 shows the proposed methodology of this framework. A three-step approach has been followed.

A. Face mask Detection Network

Firstly, a face mask detection network is used for predicting whether the person is wearing a face mask or not. For this, a two-class object detection network is trained. The two classes are as follows: "with_mask" for the people who are wearing a mask and "without_mask" for those who are not wearing a mask. For object detection backbone, experiments were tried with YOLOv2, YOLOv3 and Faster R-CNN. Of these experiments, YOLOv3 seemed to be the best considering the trade-off between speed and accuracy. When the given image is passed through the network, the final class prediction will tell whether the person in the image is wearing a mask or not. If the person is wearing a mask, then he/she is safe. If the person is not wearing a mask, the detected ROI for the face will be further sent to the age classification network for classifying the age of the person. Since the number of output class labels (C) in this detector was 2, and the number of bounding boxes per cell (B) was 3, the size of the final detection kernel of YOLOv3 was 1 x 1 x 21.

B. Age Classification Network

As we know, fatalities due to COVID-19 are generally directly proportional to age groups i.e. higher age groups are at more risk. This assumption is based on the facts provided by the worldometer website and the research study done by various research institutes. There might be a few exceptions to this general assumption wherein we have seen and read about many aged people recovering from COVID -19. But, since the immune system of a person deteriorates after a certain age, this assumption has been considered for this project work. For age group prediction, the network architecture similar to the one proposed by Hassner et al. [39], [40] was used. This network has AlexNet-like architecture. The network comprises of three convolutional layers and three fully-connected layers. For age classification, eight age brackets were considered. They are as follows: 0-2, 4-6, 8-13, 15-20, 25-32, 38-43, 48-53, 60-100. The output of the last FC layer was fed to a soft-max layer that assigns a probability for each of the age bracket classes. This final output age bracket prediction is made by considering the class giving the maximum probability as the output for any given test image. There are four levels of risk depending on the age groups. They are low risk, medium risk, high risk and very high risk. The age groups 0-2 and 4-6 are considered to be at high risk and medium risk respectively since new born babies and small children do not have a properly matured immune system. The age groups 8-13, 15-20, and 25-32 are considered to be at low risk because during this age, the immune system of a person gets matured enough to fight with the foreign virus particles entering the body. The age groups 38-43 and 48-53 are considered to be at medium risk and high risk respectively since after the age of 35/36, the bodies of people become less immune to tackling the virus as years go by. Finally, the age group comprising of people with age > 60 are considered to be at a very high risk as during this period, the immune system starts deteriorating quickly.

TABLE I. PERFORMANCE COMPARISON OF DIFFERENT OBJECT DETECTION MODELS.

Detection Model	mAP	FPS
Faster RCNN	94.43	3
YOLO v2	75.61	34
YOLO v3	**80.56**	**23**

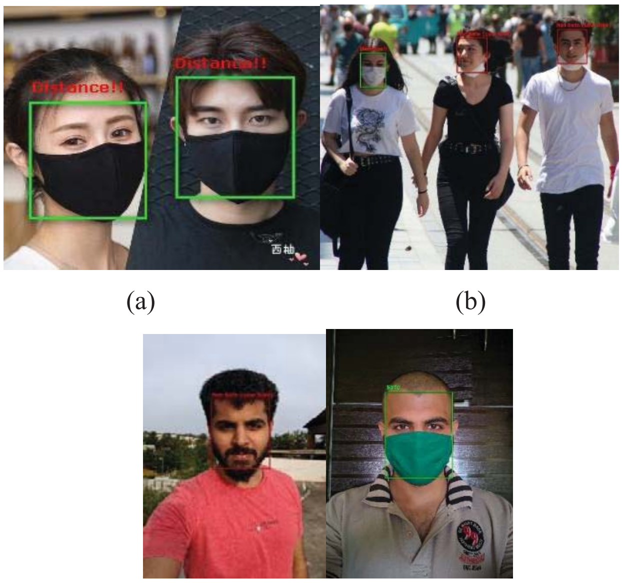
C. Social Distancing Calculator

When multiple bounding boxes are detected for a single image, the distances between each of the boxes is calculated. Consider *h* is the height of the bounding box, *w* is width of the bounding box, *x* and *y* are the co-ordinates of the centroid of the bounding box. The depth *d* of the box from the camera can be calculated by the formula proposed by Paul et. al. [41].

$$d = \left(\frac{2 * 3.14 * 180}{w + h * 360} * 1000 + 3 \right) \quad (1)$$

Now, each of the box can be mapped to a 3-D plane which has the co-ordinates *(x, y, d).* For the set of detected bounding boxes, the *L2* norm is calculated between each pair of boxes. The closeness threshold can be changed according to the requirement (Default value = 120 pixels). For all the boxes which do not satisfy this closeness condition, a message “Distance!!” is shown on the output image along with the estimated level of risk.

Datasets. For training and testing the object detection network, the Real-world Masked Face Dataset provided by X-Zhangyang [42] was used. This dataset comprises of two parts, Real-world masked face recognition dataset and Simulated masked face recognition dataset. The Real-world masked face recognition dataset consists of 90000 normal faces and 5000 masked faces of 525 people while the simulated dataset consists of 5,00,000 masked faces of 10000 subjects. Of these, 1000 normal faces and 1000 masked faces were selected. For labelling these images, the Computer Vision Annotation Tool (CVAT) [43] was used since this



(a) (b)

(c) (d)

Fig. 2. a) Face mask detected but social distancing not followed. b) Multi-person scenario with 2 people without face mask (unsafe with low risk) and one person wearing face mask, but social distancing not followed. c) Not wearing face mask so person is unsafe. Age classified as 25-32, hence, risk categorized as low. d) Face mask detected so person is safe.

tool directly exports the annotations in YOLO format. Additionally, the face mask dataset provided by Lorenzo et al. [44] on Kaggle was used. For the age classification network, the Adience benchmark dataset [45] was used for training and testing. The main reason why this dataset was used is that it captures all the variations in appearance, pose, lighting, noise, etc. This helps the model to get trained on images of a variety of people in various situations. Then, the age classification network was fine-tuned on 1000 images out of those containing normal faces used to train face mask detection network.

IV. EXPERIMENTS AND RESULTS

For training the face mask detection network, the input images were resized to the fixed resolution of 320 x 320. The training, validation, and test sets were split as 70%, 20%, and 10% of the selected dataset. The weights for all the layers except the last fully connected layer were initialized with the weights of the YOLOv3 model pre-trained on the COCO dataset. Then, the model was trained for 8 epochs using a batch size of 32 on the Nvidia GTX 1050 GPU. After experimenting with a variety of optimizers, Adam optimizer was selected to be used during training since it proved to be faster, and more stable than the other optimizers. A learning rate of 1e-4 was kept during training. Around the 7th epoch, the model achieved an mAP of 78.3. During inference, the model ran at 23 FPS. The non-maximum suppression threshold selected was 0.55. Table 1. shows the performance evaluation of the different object detectors used for the task of face mask detection. When the face mask detection network predicts "without_mask" for the face of a particular person, the face ROI box is stretched to the fixed size of 224 x 224 and is sent through the age classification network for predicting the age category of the person. As mentioned earlier, this network was trained using the Adience benchmark dataset. In order to avoid the overfitting, three dropout layers with a ratio of 0.5 were added; 1 after the 3rd convolution layer and 2 after the first 2 FC layers. The learning rate was kept as 1e-4. The validation accuracy of this network was 79%.

TABLE II. SPEED COMPARISON BETWEEN DIFFERENT NETWORK ARCHITECTURES

Architecture	FPS
Age Classification Network	20
Faster RCNN	3
Faster RCNN + Age Classification Network	2.5
YOLO v2	34
YOLO v2 + Age Classification Network	12
YOLO v3	23
YOLO v3 + Age Classification Network	**10**

During inference, the age classification network took around 50ms per image i.e. the network ran at 20fps. Finally, when both the networks ran one after the other, the final speed for the end to end system was around 10 fps and the mAP achieved by YOLOv3 model was 80.56. This experiment was repeated wherein the backbone of the face mask detection network was changed to YOLOv2 and Faster RCNN. Table 2. shows the total speed (in fps) of the individual stages and combined end to end system using different network architectures. With the face mask detector backbone as YOLOv2, the network performs the fastest (~ 12 fps) with a significant drop in mAP than with YOLOv3. With Faster RCNN as the backbone, the performance goes down to 2.5 fps but the mAP increased to around 94.43 which was highest among the experiments. When multiple bounding boxes are detected, this framework also calculates the distance between the bounding boxes to check whether the norms of social distancing are being followed. Figure 2 shows some sample results generated using this framework. This framework does not require manual intervention and hence, it makes the risk detection and estimation process completely automatic.

V. FUTURE SCOPE AND CHALLENGES

There is surely going to be a lot of difference between the world before coronavirus and the world after coronavirus and we need to learn to live with it. For preventing the infection due to the coronavirus, in case of absence of a vaccine, the two methods, i.e. following social distancing and using a face mask are the ones on which everyone would have to rely upon. In such a situation, this framework would be definitely beneficial for monitoring the risk due to COVID-19. Since contagious viruses generally spread through direct contact between people, and enter the body through nasal passages, this framework would also be beneficial to estimate the risk due to those types of contagious virus whose transmission dynamics are similar to the novel coronavirus. Since this project is aimed at estimating the risk due to COVID-19, it needs to be highly accurate. This detector works very fast but is still not a real-time detector. The current model faces a trade-off between accuracy and speed which is common in almost all of the deep learning models known. Experiments can be performed with the hyperparameters, and the architecture of the face mask detector and the age classifier so that the models can run in a more optimized way while maintaining or possibly improving their accuracy. This project work is based on the assumption that very small babies, and aged people are at a very high risk of getting affected due to COVID-19 than others. But even if this assumption is true for most of the cases, a few outliers do exist. Also, we do not consider those cases where people have an underlying disease related to their respiratory system. In such cases, a person would be at very high risk even if he/she falls in the age bracket 25-32 years. So, the assumption made for the project turns out to be wrong in these outlier cases. A new heuristic which takes into consideration the existing illnesses of the people can be added in this framework, which will help this framework to be more accurate and exhaustive.

VI. CONCLUSION

This research work proposes an automatic, efficient and fast deep learning-based framework to automate the process of predicting the amount of risk a person is at due to COVID-19. This is a two-stage framework that consists of 2 different deep neural networks; one for face mask detection, and other for age classification. In order to choose the architecture for the face mask detection network, a comparative study was done wherein the object detectors like Faster RCNN, YOLOv2, and YOLOv3 were used. Of all these models, YOLOv3 turned out to have a very efficient performance with a proper balance between FPS and mAP scores. For every input image, the face mask detector identifies the face of each individual in the frame and classifies whether he/she has worn a face mask. If the person has worn a mask, he/she is safe. If not, the face ROI is sent to the age classification network which predicts the age category of the person. So, from the predicted age, the output will give an estimate of how much risk a person is at, due to COVID-19. In case of multiple people, the distance between them is calculated, and notified in output in case of violation of social distancing norms.

REFERENCES

[1] Worldometer, "COVID-19 Coronavirus Pandemic", 2020. [Online: accessed July 14, 2020]

[2] World Health Organization "Coronavirus Disease (COVID-19 Pandemic)", 2020. [Online: accessed July 02, 2020]

[3] DNA Web Team, "Coronavirus Vaccine updates: AstraZeneca-Oxford University's COVID-19 vaccine could roll out by year-end", 2020. [Online: accessed July 21, 2020]

[4] Worldometer, "Age, Sex, Existing Conditions of COVID-19 Cases and Deaths", 2020. [Online: accessed July 21, 2020]

[5] World Health Organization, "COVID-19 : Physical Distancing", 2020. [Online: accessed July 21, 2020]

[6] World Health Organization, "Transmission of SARS-CoV-2: implications for infection prevention precautions", 2020. [Online: accessed July 21, 2020]

[7] World Health Organization, "Coronavirus may never go away, WHO warns", May 9, 2020. [Online: accessed July 19, 2020]

[8] L. Ferretti, C. Wyamant, M. Kendall, et al., "Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing", Science, Vol. 368, Issue 6491, May 8, 2020.

[9] R. Li, S. Pei, B. Chen, et al., "Substantial undocumented infection facilitates the rapid dissem-ination of novel coronavirus (SARS-CoV-2)", Science, Vol. 368, Issue 6490, May 01, 2020.

[10] J. Yan, M. Grantham, J. Pantelic, et al., EMIT Consortium. "Infectious virus in exhaled breath of symptomatic seasonal influenza cases from a college community." Proceedings of the National Academy of Science of the USA, Vol. 115, January 30, 2018.

[11] M. Tracht, S. Y. Del Valle, and J. M. Hyman, "Mathematical modeling of the effectiveness of face masks in reducing the spread of novel influenza A (H1N1)", PLOSOne, February 10, 2010.

[12] M. P. Weekes, N. J. Matheson, "Covid-19: Should the public wear face masks", 2020. [Online: access July 09, 2020]

[13] M. AG Wallace, J. D. Pliel and M. C. Madden, "Identifying organic compounds in exhaled breath aerosol: Non-invasive sampling from

respirator surfaces and disposable hospital masks”, ScienceDirect, Vol.137, 105444, ISSN 0021-8502, November 2019.

[14] X. Liu and S. Zhang, “COVID-19 : Face masks and human-to-human transmission”, Wiley Online Library, Volume 14, Issue 4, pp. 472-473, March 29, 2020.

[15] K. Langton, “China lockdown: How long was China on lockdown?”, Express Co.uk, May 30, 2020. [Online: accessed May 30, 2020]

[16] K. Prem, Y. Liu, S. Flasche, A. J. Kucharski, et al., “The effect of control strategies to reduce social mixing on out-comes of the covid19 epidemic in wuhan, china: a modelling study,” The Lancet Public Health, Vol. 5, Issue 5, pp. 261-270, March 25, 2020.

[17] Landing AI, “Landing AI creates an AI tool to help customers monitor social distancing in the work-place”, Landing AI, April 16, 2020. [Online: accessed April 21, 2020].

[18] Government of India website, “Aarogya Setu Mobile App”, May 2020.

[19] M. Robakowska, A. Tyranska-Fobke, J. Nowak, et al., “The use of drones during mass events,” Disaster and Emergency Medicine Journal, pp. 129-134, October 2017.

[20] N. S. Punn and S. Agarwal, “Crowd analysis for congestion control early warning system on foot over bridge” , 2019 Twelfth International Conference on Contemporary Computing (IC3), pp. 1-6, September 19, 2019.

[21] M. Paul, S, Haque, and S. Chakraborty, “Human Detection in surveillance videos and its applications – A review” EURASIP Journal on Advances in Signal Processing, ResearchGate, Novermber, 2013

[22] S. Varma and M. Sreeraj, “Object Detection and Classification in surveillance system”, 2013 IEEE Recent Advances in Intelligent Computational Systems, IEEE, pp. 299-303, February 2014.

[23] S. Brutzer, B. Hoferlin, and G. Heidemann, “Evaluation of background subtraction techniques for video surveillance” in CVPR 2011, pp 1937 – 1944, June 25, 2011, IEEE.

[24] M. Piccardi, “Background subtraction techniques: a review”, 2004 IEEE International Conference on Systems, Man and Cybernetics IEEE, 2004, pp. 3099-3104, Vol. 4, October 2004.

[25] P. Dollar, V. Rabaud, G. Cottrell, and S. Belongie, “Behavior recognition via sparse spatio-temporal features”, 2005 IEEE International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance, pp. 65-72, January 10, 2006.

[26] S. Aslani and H. Mahdavi-Nasab, “Optical flow based moving object detection and tracking for traffic surveillance”, International Journal of Electrical and Computer Engineering, 2013, Vol. 7, pp. 1252 – 1256, 2013.

[27] A. Agarwal, S. Gupta, and D. K. Singh, “Review of optical flow technique for moving object detection”, 2nd International Conference on Contemporary, 2016, pp. 409-413.

[28] Z. Zhao, P. Zheng, S. Xu, and X. Wu, “Object Detection with Deep Learning: A Review”, IEEE Transactions on Neural Networks and Learning Systems, Vol. 30, pp. 3212-3232, November 2019.

[29] R. Girshick, J. Donahue, T. Darrell, and J. Malik, “Rich feature hierarchies for accurate object detection and semantic segmentation”, Proceedings of the IEEE conference on Computer Vision and Pattern Recognition, pp. 580-587, arXiv:1311.2524, CVPR 2014.

[30] R. Girshick, “Fast R-CNN”, Proceedings of the IEEE conference on Computer Vision and Pattern Recognition, pp. 1440-1448, arXiv:1504.08083, ICCV 2015

[31] S. Ren, K. He, R. Girshick, and J. Sun, “Faster R-CNN: Towards real-time object detection with Region proposal networks”, Proceedings of the IEEE conference on Computer Vision and Pattern Recognition, 2016

[32] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You Only Look Once: Unified Real-time object detection”, Proceedings of the IEEE conference on computer vision and pattern recognition, (CVPR), pp. 779-788, 2016.

[33] J. Redmon, and A. Farhadi, “Yolo9000: better, faster, stronger”, Proceedings of the IEEE conference on computer vision and pattern recognition, (CVPR) 2017, PP. 6517-6525.

[34] W. Liu, D. Anguelov, D. Erhan, et al., “SSD: Single shot multibox detection”, European Conference on Computer Vision, 2016.

[35] A. Farhadi and J. Redmon, “Yolov3: An incremental improvement”, Proceedings of the IEEE conference on computer vision and pattern recognition, (CVPR) 2018

[36] Y. Zhang and D. Yeung, “Multi-task warped Gaussian process for personalized age detection”, 2010 IEEE Computer Society conference on Computer Vision and Pattern Recognition, 2010, pp. 2622-2629.

[37] H. Han, C. Otto, and A. K. Jain, “Age Estimation from face images: Human vs Machine performance”, 6th IAPR International Conference on Biometrics (ICB), pp. 1-8, 2013, IEEE.

[38] V. Almeida, M. K. Dutta, C. Traviesco, A. Singh, and J. Alonso, “Automatic Age detection from facial images”, 2016 2nd International Conference on Communication control and Intelligent Systems (CCIS), 2016, pp. 110-114, IEEE.

[39] T. Hassner and G. Levi, “Age and Gender Classification using Convolutional Neural Networks”, CVPR 2015.

[40] T. Hassner, Github repository, “Age and Gender Classification using Convolutional Neural Networks”, 2015. [Online: accessed April 27, 2020]

[41] Paul Pias, “Object Detection and Distance Measurement”, Github repository, 2020. [Online: accessed April 1, 2020]

[42] X. Zhangyang, “Real World Masked Face Dataset”, 2020. [Online: accessed April 12, 2020].

[43] Computer Vision Annotation tool, https://github.com/opencv/cvat [Online: accessed April 22, 2020]

[44] Alexandra Lorenzo, “Mask Detection at YOLO format”, 2020, [Online: accessed April 12, 2020]

[45] The OUI-Adience Face Image Project, “Adience collection of unfiltered faces for gender and age classification”, 2014. [Online: accessed April 4, 2020]