IJRASET

International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL
FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

www.ijraset.com

Call: ⓒ08813907089 | E-mail ID: ijraset@gmail.com

Covid-19 Monitoring System Using Social Distancing Face Mask Detection and Classification on Live Video

Prof. Simran Pal R[1], Aaliya Syedi[3], Aditya C Kumar[5], Amrutha D[4], Akriti Raj[2]

[1, 2, 3, 4, 5]*Department of Computer Science HKBK College of Engineering, Bangalore, India*

Abstract: The spread of the Coronavirus has prompted individuals to remain indoors and adhere to COVID- appropriate practices, which include social distance, the use of face masks, hand sanitizers, and other measures to protect themselves against infection. In heavily populated locations with limited resources, it is impossible to physically supervise compliance with these standards. As a result, an automated, lightweight, and powerful video monitoring system is required to make the process more efficient. This paper proposes an extensive and productive solution for performing person detection, social distance detection, face mask detection, and face mask classification using object detection, clustering, and Convolution Neural Networks (CNN) On video datasets, in addition to YOLOv3, density-based spatial clustering of applications with noise (DBSCAN), Dual Shot Face Detector (DSFD), and MobileNetV2 based binary classifier other techniques have also been used to achieve the predicted outcomes. This study also provides parallels for numerous face mask detection and classification models.
Keywords: COVID-19, DBSCAN, DSFD, YOLOv3, MobileNetV2, CNN, clustering, Social Distance, Face Masks

I. INTRODUCTION

Coronaviruses are related RNA viruses that cause diseases in mammals and birds, including encephalitis in humans. They can cause respiratory tract infections that can be mild to fatal in humans and birds. Mild infections such as the common cold may be contracted by humans, whereas more fatal strains can cause SARS, MERS, and COVID-19, which is generating an ongoing pandemic. There have been 324,565,830 confirmed cases of COVID-19 worldwide as of January 2022, with 5,549,239 deaths. COViD-19 is believed to spread mostly through direct contact between people. The following are a few possibilities on how this may occur:

1) *Aerosols and Droplets:* Approximately 80% of all diseases are spread via this route. Infected droplets or microscopic particles called aerosols are released from the nose or mouth of an infected individual when he or she coughs, sneezes, or speaks. Inhalation is possible for anybody within six feet of the source.
2) *Airborne Transmission:* According to recent findings, the virus may survive for up to three hours in the open air. If someone who has it exhales and you inhale, it might enter your lungs
3) *Transmission on the Surface: Contacting* surfaces contaminated by someone else's cough or sneeze is another less common route to get the virus.

Some Anti-covid 19 precautions to prevent the spread of covid-19 are Social distancing, use of face masks, washing hands frequently, regular sanitization, etc. Social distancing is a term used in public health to refer to a collection of non- pharmaceutical treatments or procedures aimed at preventing the transmission of infectious illness by keeping physical space between individuals and minimizing the number of times people come into close contact. In both laboratory and clinical settings, the bulk of data suggests that wearing a mask lowers transmissibility per interaction by limiting the transfer of infectious respiratory particles

II. RELATED WORKS.

Rujula Singh R et al. [1] Offer a model that focuses on the utilization of YOLOV3, DBSCAN, and DSFD. Although the system obtained a high degree of accuracy, more efficient techniques can be applied to improve efficiency. The downside of this approach was that it evaluated prerecorded datasets rather than live video to simplify deployment and real-time use. Gupta Savyasachi et al. [2] Propose a strategy that utilizes deep neural networks to recognize individuals and a centroid tracking method to track the people throughout the film. Although the resulting technique was effective when tested on CVFD and CPID, significant barriers obscuring The visual field of the camera systems may hinder the tracking of persons, resulting in difficulties in appropriately assessing social distance.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue III Mar 2022- Available at www.ijraset.com

Shashi Yadav et al. developed a method for social distance and mask identification and suggested a deeplearning strategy based on (SSD) with MobileNet V2 and OpenCV. People having their hands over their faces or obstructed by items are categorized as masked when employing this method. This technique is not meant to address such situations. SSDs can recognize several things in a frame, but in this case, they can only detect one person at a time. There has been a lot of focus on social distancing monitoring and face mask detection in the publications. Additionally, even if both are employed, there is still room for development through the use of more accurate models. Inour study, we discuss the significance of inference time as an assessment metric, which is essential for the system's actual application but has not been discussed in previous publications. Unlike the previous publications, which all employ prerecorded datasets, we leverage real-time usabilitywith live video instead. A state-of-the-art object recognition model, YOLOv3, has been proposed in the study for the identification of individuals, followed by DBSCAN for calculating the lengths between people and performing clustering, which is superior to previous clustering techniques.

III. EXPERIMENTAL SETUP AND METHODS

A. Person Detection

YOLOv3 model was used for person detection. It consists offifty-three layers of Darknet-53 trained on ImageNet that actsas a strong feature extractor and a further fifty-three layers for detection giving a complete 106 bedded convolutional neural network. Anchor boxes with 3 scales are used: 13x13, 26x26,and 52x52. These three boxes are used to predict the presenceof a person. The output of this model when prediction may bea list of bounding boxes in conjunction with the confidence of the person category detected. Non-maximum suppression (NMS) is employed to unravel the difficulty of overlapping bounding boxes resulting in multiple detections for the identical object. This means only classes with more than 50%confidence are included and all those bounding boxes that have more than 30% overlap with another bounding box are discarded.

B. Social Distancing

DBSCAN algorithm is used to check if social distancing is maintained between the detected persons. It is an unsupervised learning algorithm that groups similar points. DBSCAN does not require the number of clusters to be set before training. It also ignores the outlier points duringthe formation of the clusters. Since social distancing is checked between a minimum of two individuals, the minimum needed points within the cluster is about to, and also the distance parameter was set to 200. Taking intothought person by person, if the distance between them is a smaller amount than the space parameter, then they're classified into a cluster. If an individual doesn't belong to anycluster, then they're categorized as safe and bound with orange boxes. individuals happiness to a cluster area unit denoted by red lines between every one of them, WHO area unit deemed too close, and certain with blue boxes.

C. Face Detection

For face detection, two pre-trained models: DSFD and RetinaNetMobileNetV1 in terms of accuracy and prediction time. RetinaNetMobileNetV1 is a lightweight single shot face detector developed for mobile readying of theface detection models. DSFD consists of 3 components: the primary shot detector, which consists of convolution layers, this feature enhances the module that generates additional options and therefore the second shot detector that incorporates these increased options and loss from the primary shot detector to grant the ultimate predictions. Since a second shot is used, the model performs far better than one shot detector however is sort of slow in prediction. There is acompromise between accuracy and prediction time. This can be because complicated models are comparatively correct however involve a great deal of computation and therefore take additional prediction time. In terms of accuracy, DSFD performs higher however encompasses an abundant higher detection time than RetinaNetMobileNetv1. Hence, a compromise on accuracy can't be created and therefore the DSFD model was chosen for Data Transformation. Here, all the categorical data were consolidated into an understandablenumerical format. The transactional dataset contains several data types with several ranges. Therefore, data transformation comprises data normalization.

D. Face Mask Detection

CNN binary image classification architecture is used to implement Face mask classification A variety of models weredeveloped to implement CNN for mask classification and their performance in terms of accuracy, precision, recall, andF1 scores were compared for sophistication zero (no mask) and class 1(masked). MobileNetV2 was chosen because of its performance in prediction time similarly to accuracy.MobileNetV2 provides the simplest set of performance values with associate accuracy of 93.2% on the check dataset and 95.6% on the training dataset. it's 2 varieties of blocks: Residual block with a stride of one and stride of two for curtailment.

There are 3 layers for each variety of blocks. theprimary layer is 1×1 convolution with ReLU6. The second layer is the depth-wise convolution. The third layer is another1×1 convolution however without any non-linearity. wherever C is the convolution layer and D is the depth-wise convolution layer. The output of MobileNetV2 is two- dimensional and passed to a 256 unit connected layer with a dropout regularization of four-hundredth followed by another sixty-four-unit-connected layer and one output for binary classification. The coaching and validation accuracy increase and also the training accuracy is often quite the validation accuracy. Thus, the model isn't overfitting. The coaching and validation loss decrease to some extent of stability with a little gap between the training and validation set, the low loss priceshowing the good match of the model-

IV. RESULTS

The accuracy and F1 score are calculated using the belowequations (1) and (2).

$$Accuracy = \frac{\text{True positives} + \text{True negatives}}{(\text{True positive} + \text{False positive}) + (\text{True negative} + \text{False negative})} \quad (1)$$

$$F1\ score = \frac{2 * P * R}{(P+R)} \quad (2)$$

where P is precision and R is the recall

The performance can be checked on manually labeled videos in table IV. F1 score, accuracy, and prediction time areportrayed to condense the model performance. The average accuracy of the system is 92.1% and the average F1 score is 91.29%. The average prediction time all together is 8.13 seconds for a 1-second video frame in which the time taken for person detection is 6.42 seconds.

V. FUTURE SCOPE

Even though the system performance is good in terms of accuracy along with the prediction time, subsequent improvement areas are recognized: First, the person detectionmodule consumes most of the time in video processing. A comparatively simpler person face detection algorithm can be put in place that takes lesser prediction time. Second, social distancing calculation and the face mask classification run separately, and thus correspondence can be used to execute them simultaneously. Third, there is a scarcity of datasets to be used for such a system and it is not diverse to work for allthe circumstances. For instance, the system may sometimes confuse the beard with the face mask due to insufficient negative examples with beards in it. When such datasets are at hand, a more influential model can be trained.

VI. CONCLUSION

This paper provides a productive solution to monitor social distancing practices in public areas where manual monitoring is challenging. Different modules have been developed for person detection, social distancing identification, face detection, and face mask categorization. This system performs satisfactorily well with an accuracy of 92.1% and an average F1 score of 91.29% on the labeled video datasets with an average prediction time of 8.13 seconds for a 1- second video frame in which the time taken for person detection is 6.42 seconds. It also provides data augmentation techniques to deal with the scarcity of datasets in the community.

REFERENCES

[1] R, R. S., Nayak, N., Srinivasan, S., & Biradar, R. (2021). COVID-19 Monitoring System using Social Distancing and Face Mask Detection on Surveillance video datasets. 2021 International Conference on Emerging Smart Computing and Informatics, ESCI 2021, 449–455. https://doi.org/10.1109/ESCI50559.2021.9396783

[2] Savyasachi Gupta, Rudraksha Kapily, Goutham Kanahasabaiz, Shreyas Srinivas Joshi, Aniruddha Srinivas Joshi SD-Measure: A Social Distancing Detector | IEEE Conference Publication IEEE Xplore. (n.d.), https://ieeexplore.ieee.org/document/9242628

[3] Yadav, S., & Kalam, A. P. J. A. (2020). Deep Learning-based Safe Social Distancing and Face Mask Detection in Public Areas for COVID-19 Safety Guidelines Adherence. 8. https://doi.org/10.22214/ijraset.2020.30560

[4] Bhambani, K., Jain, T., & Sultanpure, K. A. (2020). Real-Time Face Mask and Social Distancing Violation Detection System using YOLO. Proceedings of B-HTC 2020 - 1st IEEE Bangalore Humanitarian Technology Conference. https://doi.org/10.1109/BHTC50970.2020.9297902

[5] Sanjaya, S. A., & Rakhmawan, S. A. (2020). Face Mask Detection Using MobileNetV2 in the Era of COVID-19 Pandemic. 2020 International Conference on Data Analytics for Business and Industry: Way towards a Sustainable Economy ICDABI 2020. https://doi.org/10.1109/ICDABI51230.2020.9325631

[6] Suresh, K., Palangappa, M. B., & Bhuvan, S. (2021). Face Mask Detection by using Optimistic Convolutional Neural Network. Proceedings of the 6th International Conference on Inventive Computation Technologies, ICICT 2021, 1084–1089. https://doi.org/10.1109/ICICT50816.2021.9358653

INTERNATIONAL JOURNAL
FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY