Advances in Parallel Computing Algorithms, Tools and Paradigms
D.J. Hemanth et al. (Eds.)

doi:10.3233/APC220011

Effective Face Mask And Social Distance Detection with Alert System for Covid-19 Using YOLOv5 Model

Abirami T[a,1], Priakanth P[b]and Madhuvanthi T[b]
[a]*Associate Professor, Department of IT, Kongu Engineering College, Perundurai*
[b]*Department of CSE, Kongu Engineering College, Perundurai*

Abstract. Globally, numerous preventive measures were taken to treat the COVID-19 epidemic. Face masks and social distancing were two of the most crucial practices for limiting the spread of novel viruses. With YOLOv5 and a pre-trained framework, we present a novel method of complex mask detection. The primary objective is to detect complex different face masks at higher rates and obtain accuracy of about 94% to 99% on real-time video feeds. The proposed methodology also aims to implement a structure to detect social distance based on a YOLOv5 architecture for controlling, monitoring, accomplishing, and reducing the interaction of physical communication among people in the day-to-day environment. In order for the framework to be trained for the different crowd datasets from the top, it was trained for the human contrasts. Based on the pixel information and the violation threshold, the Euclidean distance between peoples is determined as soon as the people in the video are spotted. In the results, this social distance architecture is described as providing effective monitoring and alerting.

Keywords.yolov5, Facemask detection, social distance detection, yolov4, Confidence Score, mAP.

1. Introduction

In December 2019, a disease called COVID-19 emerged from the respiratory virus outbreak. It was severe, and millions of people died as a result. Also, it caused health problems for those who survived. The novel virus can spread from one person to another. The researchers found that the virus particles and droplets were released into the air if an infected person coughs, breathes or sneezes. Those droplets can clog and accumulate in the environment. [1]The primary cause of COVID-19 was recorded on Dec. 1, 2019, and a new coronavirus was later identified as SARS-CoV-2.

The affected person initially gets the infection in their lungs through their respiratory system, which recreates the affected cells and reduces their ability to breathe [2] . COVID 19 symptoms include dry cough, fever, discomfort in breathing, headache, and sometimes migraine problems. These also affect the muscle torment of the person affected, causing him to become invulnerable to violence. Immediately, physical distancing is one of the most crucial calls for humanity at this critical moment.

[1]Abirami T,Associate Professor, Associate Professor, Department of Information Technology, Kongu Engineering College, Perundurai; E-mail: abi.it@kongu.edu

The World Health Organization recommends keeping three feet away from anyone sneezing or coughing. [3] Monitoring the social distance in the gatherings, schools and universities are huge but with the help of surveillance camera available at most public places.

2. Related Works

(3)This article includes the detection of social distance monitoring using Yolo v3, calculating the distance between the people using Euclidian distance with a minimum threshold value in pixels. The result is an accuracy of 92%. [4]A face mask detection method is presented in this paper focusing on the medical masked face, which uses the YOLO version 2 model based on the Resnet-50 network, and it also introduces the you mean to detect the smallest number of anchor boxes to provide more accurate detection with high-performance results.A recent study by [5]showed an IoT UAV-based thermal sensors used to gather the dataset. It had an average precision of 99.5% utilizing different performance measurement techniques in real-time outdoor synopses.

(1) it demonstrates the blob segmentation of pixel images of humans to monitor the social distance between each other with CNN. So, this method in deep learning for social distance detection has a lower accuracy than 70%. The paper demonstrates the use of deep learning approaches for face mask recognition and face classification using clustering for the dataset.[6] this paper develops a more novel face detection system of principal component analysis (PCA) and convolution neural network (CNN) combined with a probabilistic Bayes classifier, SVM (support vector machine), and MLP (multilayer perceptron algorithm) to detect multi-face model for mask detection. [7]this paper describes the human detection model which measures the distance between humans using a sound wave sensor.[8] has demonstrated an artificial intelligence-based IoT with temperature monitoring, automatic sanitization and mask detection with an accuracy of 98%. In this method, each person should show their face to the camera. In order to get sanitized, the hand needs to be below the sanitizer. Then the temperature is checked with the help of a temperature sensor, if it is below the threshold value, it allows the person to enter.The work given in [17] using CNN for detecting abnormal activities through image analysis.

3. Methodology

We propose face mask detection and physical distance recognition using object recognition algorithms in instantaneous applications. In single-stage detection modules, YOLO is preferred to predict the probability of the class and a bounding box simultaneously for the complete input image. Another detection module is examining the entire image or video frames meant for the faces. Because it takes a long time to progress, and it also detects the faces incorrectly, Yolo is used to depict a region of interest(ROI) in the face detection module, (5) remains the fastest object detector to find a region of interest.[9] There are various versions such as YOLOV1, YOLOV2, YOLOV3, YOLOV4 and YOLOV5. The proposed work improves the predictive accuracy of human perception by Yolov5, especially for small scale objects. With its network structure, multi-scale object detection can be made more effective by adjusting it. Besides, it uses logistic rather than softmax independently for object classification. It is a single-

stage architecture network that predicts the probability of the class and an individual bounding box for the complete input image simultaneously. The architecture appropriates a convolution layer for feature extraction and fully connected layers for class prediction.

3.1 Dataset Characteristics

This paper proposes facemask detection based on the two datasets. The first is the Kaggle dataset which was published by Ashish Jangra (Face Mask Detection ~12K Images Dataset | Kaggle). This ~12K dataset consists of approximately ~6K are preprocessed from the Celebrity dataset, and another 6K images scrapped from the Google search. It consists of 5,883 images of with masks and 5,909 images of without masks. The second dataset is a custom dataset that is collected from the college database. It includes of ~26k images of database images.

3.2 Face Mask Detection

A model which directly predicts the object with bounding boxes for the image or video is called a single-stage detector. It is a faster and simpler architecture. The first step is data collection upon recognition of the mask. The final accuracy of the CNN model is impacted by trained data. So, a large set of data is taken, which could be directly fed into it. So, labels are generated by the labelling tool which allows the creation of labels for the images in XML format. The dataset is classified into three labels. The top label has who wears a facemask, the second label has who didn't wear a facemask and the third label has who didn't wear the mask properly.[10] In this proposed method, yolo along with Tensor is used for real-time face mask detection. A comparison of various yolo models is performed. Initially, augmentation of the image is performed, including the image flipping vertically and horizontally, 0.2 of a shearingfactor and zooming. By defining the upper and lower bounds, the learning rate is obtained through the learning rate(lr) finder by Keras. When the highest lr is obtained, there will be a reduction in the loss. Continuously the model is trained till it achieves a reasonable fit. Here, 100 epochs are optimally used. A threshold point is observed and accuracy and loss on training, test and validations are compared with regard to the epochs.

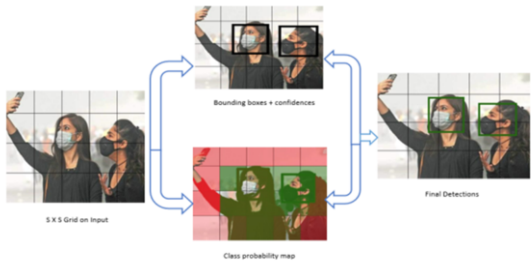

Figure 1. Working of Single stage Face Mask Detection

While training the batch size is given as 16 and 32, ‘prediction', 'recall' as the metrics, and ‘box’, ‘object’, 'class' as the loss function. Finally, a model is developed based on real-time video input. Essentially, the algorithm performs the same as the training

dataset, except for the initial conversion of video into frames for the batch size. **Figure 1.,**shows that the alert is created with bounding boxes, where red boxes denote No mask (NM), and Green denotes With Mask (WM). As part of the multi-stage face mask detector, the dataset is loaded. College database uses a number of CNN models including MobileNet, VGG16, and VGG19, and the Haar cascade face detector, Tensorflow, and Keras in this procedure. It has 70 percent for trains, 20 percent for tests, and 10 percent for validating the model's accuracy. Personal information, including name, contact information, email address, and academic information, is stored in the database. After detection, the database is compared with the faces. A message and email are sent if his/her face matches the one in the database. Furthermore, the output of the detector is automatically stored as an image for future reference.

3.3 Single Stage Social Distancing Detection

[11]The camera is positioned at a different angle to capture video, and the video frame is interpreted as if it were a perspective shot. Therefore, it is automatically converted from perspective view into top-down view, or birds-eye view. Within the first frame, seven points are selected, four represent the RoI, and the last three determine the limitation of horizontal and vertical distances.[3] and[12]In the RoI mapping, the four points are top to left, bottom to right, top to left and bottom to right for monitoring social distance from the bird's eye view. The four points are computed using the following method with X-Y coordinates. [13].

Figure 2. Detection of Social Distance maintaining from college database with alert boxes

[14]Additionally, the three dots correspond to the order in which the top region lies on the left side, the bottom region lies on the left side and the bottom region lies on the right side. RoI points can be seen as rectangle boxes in real-time video. Vertically and horizontally, these birds-eye views are categorized consistently.[15] By generating a rectangle box map, a transformation can be derived and applied to the whole perspective image. The bounding box for each person is applied using YOLO after the person has been detected.[16] Distances between two individuals can be calculated using Euclidean distances. This method provides an accurate measurement of distance. The proposed methodology assumes that individuals seen in the video are walking at a flat plane corresponding to the video frame. A bird's eye view is then constructed by selecting four points on the frame. Based on **Figure 2,** it can be clearly seen that the distance algorithm will compute the distance in pixels that are less than 300 or greater than 300. The bird's eye view is created after defining the preset distance based on the RoI, width, and height. In addition to detecting and estimating each person's position, the gap between two or more people is also scaled and measured.Based on the minimum distance, a violation or breach is indicated. I.e., if a distance between two people is less than the preset distance, then a red rectangle box is automatically drawn on them, where the red

box signifies the social distance violation. Whenever there is a violation of social distance, red appears, whereas green indicates no violation.

4. Experimental Results

For social distance, two different types of real-time video data have been analyzed. It also classifies and detects all the instances of the face flawlessly and more accurately, as is shown in the video. But other models of Yolo gave true classification but missed the detection of faces in the crowded video. At first, the precision of the custom model is considered. From the **Table 1.,**various versions of Yolo are used for detecting masks in real-time. Yolov5xl showed the highest precision and recall of all Yolo5 versions and also other Yolo algorithms. On steady-state after the 100 epochs, the mAp value is approximately 0.97; similarly, the precision value is approximately 0.96, and the recall value is approximately 0.92.

Table 1. Comparison of different yolo models and its version for college database

METRICS	VARIOUS YOLO MODELS					
	v5s	**v5m**	**v5l**	**v5xl**	**v4**	**v3**
DATASET2	Epoch = 100		Batch = 16			
Precision	0.965	0.916	0.935	**0.992**	0.824	0.691
Recall	**0.99**	0.957	0.971	**1.0**	0.873	0.717
mAp	0.965	0.972	0.965	**0.98**	0.861	0.74
Time(hrs)	1.448	3.834	5.536	7.110	0.949	9.371

On training the model, the loss functions gradually decrease, therefore the class loss is 0.003638, box loss and object loss are 0.2597. On validation, the class loss is 0.002348, the object loss is 0.01414, and the box loss is 0.01919. **Table 2**, shows the convolutional neural network models with training accuracy, validation accuracy, and validation loss with Adam as an optimizer. Among these models, MobileNet and InceptionV3 show high accuracy but false predictions are not avoidable in CNN models.

Table 2. Training and validation accuracy with its loss of CNN models with adam optimizer

MODEL	OPTIMIZER	Trainable Params	Train_acc	Val_loss	Val_acc
MobilenetV2	adam	164,226	0.9954	0.0253	0.9908
VGG19	adam	16,386	0.9850	0.0470	0.9762
VGG16	adam	8,914	0.8952	0.6742	0.9103
InceptionV3	adam	21,768,352	0.9641	0.0087	0.9944

5. Conclusion

Research and doctors have suggested numerous security systems for society in response to the rapid increase in transit of Corona virus. One of the most important things to consider is to protect themselves from this pandemic disease. At the entrance of the place a thermal Flix camera continuously measures the temperatures of the students and other people. Out of the various yolo models, physical distance scrutiny and face mask

identification using the Yolo methodology yield the best results using fast detection mechanisms. A system to alert people will also be included with this analysis so they can be alerted when there are too many people nearby and if they are not wearing a mask or if the mask is improperly worn. So, this proposed method is really useful, as it makes it easier to collect data from institutions and organizations, thereby helping organizations to identify those without masks. In addition, individual emails and alerts are sent out, so that those affected can take necessary precautions as a warning or restraint. Afterwards, the information about the person who doesn't follow the rules governed by the Health Department is stored. Following a thorough analysis of the person's report, it was decided what should be done.

References

[1] Ahamad AH, Zaini N, Latip MFA, editors. Person Detection for Social Distancing and Safety Violation Alert based on Segmented ROI. 2020 10th IEEE International Conference on Control System, Computing and Engineering (ICCSCE); 2020: IEEE.
[2] Degadwala S, Vyas D, Dave H, Mahajan A, editors. Visual Social Distance Alert System Using Computer Vision & Deep Learning. 2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA); 2020: IEEE.
[3] Ahmed I, Ahmad M, Rodrigues JJ, Jeon G, Din S. A deep learning-based social distance monitoring framework for COVID-19. Sustainable Cities and Society. 2021;65:102571.
[4] Loey M, Manogaran G, Taha MHN, Khalifa NEM. Fighting against COVID-19: A novel deep learning model based on YOLO-v2 with ResNet-50 for medical face mask detection. Sustainable cities and society. 2021;65:102600.
[5] Barnawi A, Chhikara P, Tekchandani R, Kumar N, Alzahrani B. Artificial intelligence-enabled Internet of Things-based system for COVID-19 screening using aerial thermal imaging. Future Generation Computer Systems. 2021.
[6] Cristani M, Del Bue A, Murino V, Setti F, Vinciarelli A. The visual social distancing problem. IEEE Access. 2020;8:126876-86.
[7] Garg D, Goel P, Pandya S, Ganatra A, Kotecha K, editors. A deep learning approach for face detection using YOLO. 2018 IEEE Punecon; 2018: IEEE.
[8] Shenvi DR, Shet K, editors. CNN Based COVID-19 Prevention System. 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS); 2021: IEEE.
[9] SHUKLA R, MAHAPATRA AK, PETER JSP. Social distancing tracker using yolo v5. Turkish Journal of Physiotherapy and Rehabilitation. 2021;32:2.
[10] MacIntyre CR, Wang Q. Physical distancing, face masks, and eye protection for prevention of COVID-19. The Lancet. 2020;395(10242):1950-1.
[11] Hou YC, Baharuddin MZ, Yussof S, Dzulkifly S, editors. Social distancing detection with deep learning model. 2020 8th International Conference on Information Technology and Multimedia (ICIMU); 2020: IEEE.
[12] Meivel S, Devi KI, Maheswari SU, Menaka JV. Real time data analysis of face mask detection and social distance measurement using Matlab. Materials Today: Proceedings. 2021.
[13] Ahmed I, Ahmad M, Jeon G. Social distance monitoring framework using deep learning architecture to control infection transmission of COVID-19 pandemic. Sustainable Cities and Society. 2021;69:102777.
[14] Redmon J, Farhadi A. Yolov3: An incremental improvement. arXiv preprint arXiv:180402767. 2018.
[15] Rahim A, Maqbool A, Rana T. Monitoring social distancing under various low light conditions with deep learning and a single motionless time of flight camera. Plos one. 2021;16(2):e0247440.
[16] Sathiyavathi, V., M. Jessey, K. Selvakumar, and L. SaiRamesh. "Smart Surveillance System for Abnormal Activity Detection Using CNN." Advances in Parallel Computing Technologies and Applications 40 (2021): IOS Press, 341-349.