Deep Learning - Based COVID-19 Safety Monitoring

Steffy Francis
DepartmentOf Computer Science & Engg.
I.E.S College Of Engineering , Chittillappilly,
Thrissur, India.

**Jensy V*
Assistant Professor in Computer Science & Eng.,
I.E.S College Of Engineering Chitillapilly,
Thrissur, India

Abstract

As of June 2, 2021, the fatal coronavirus illness 2019 (COVID-19) has spread to over 180 nations, resulting in 28,175,044 confirmed cases and 331,895 deaths in India alone. The population's vulnerability is heightened by the lack of effective treatment drugs and immunity to COVID-19. The population's vulnerability is heightened by the lack of effective treatment drugs and immunity to COVID-19. Because there are no effective vaccines or medications available, the only viable way to combat the pandemic is to follow covid guidelines, such as social distancing and wearing masks etc. In order to automate the task of monitoring social-distancing, mask, and age of a person using surveillance footage, this study provides a deep learning-based system. The suggested framework employs the YOLO v3 object detection model to distinguish persons from the background using bounding boxes and assigned IDs.It was discovered that it had improved accuracy for all of the input videos that were evaluated.

Keywords: *Covid-19, Social-distancing, Person detection, Masks,Age detection, YOLO*

I. INTRODUCTION

When the novel coronavirus (Covid-19) pandemic breaks out, the population is concerned about the virus's spread if there is no viable cure. Due to an upsurge in the number of cases reported around the world, the World Health Organization (WHO) has designated Covid-19 a pandemic.In order to contain the pandemic, many governments have imposed a lockdown, in which citizens are required to stay at home during this critical period. The Centers for Disease Control and Prevention (CDC) and other public health organisations had to make it apparent that avoiding close contact with other individuals is the most effective strategy to halt the spread of Covid-19.Citizens all across the world are exercising physical distancing,wearing masks etc, in order to flatten the curve on the Covid-19 pandemic.

During the quarantine period, group activities and congregations such as travel, meetings, gatherings, workshops, and prayers were prohibited in order to achieve social distance. People are urged to manage and conduct events as much as possible via phone and email in order to reduce face-to-face contact.To help stop the virus from spreading further, people are being encouraged to practise good hygiene, such as washing their hands regularly, wearing masks, and avoiding close contact with sick people. However, there is a difference between knowing what to do to decrease virus spread and actually doing it.

To lessen the economic burden of the pandemic, various countries have allowed a restricted number of economic activities to restart after the number of new Covid-10 cases has fallen below a particular threshold. Concerns about worker safety have surfaced in the new post-Covid-19 climate as these countries cautiously recommence their economic activity.People should avoid any person-to-person contact, such as shaking hands, and keep a distance of at least 1 metre from each other while wearing masks to limit the risk of infection.

We realised this need and created a model that is particularly well adapted to detecting specific breaches in real time.Our model's first application is to recognise people's faces in order to assess their age and whether or not they're wearing an approved mask.The second application is to evaluate whether or not social distance is being maintained between two individuals in the most efficient, accurate, and easy way possible, requiring the least amount of effort from supervisory authorities.

II. RELATED WORK

This section showcases some of the relevant deep learning-based human detection research. Deep learning is used in a large number of recent research on object categorization and detection, which is also explored. The current state-of-the-art review focuses mostly on current machine learning-based object identification research.

In the computer vision job of categorization and localisation of its form in video footage, human detection may be regarded as object detection. Deep learning has established a research trend in artificial intelligence's multi-class object recognition and detection, achieving exceptional results on difficult datasets.

Human descriptors, machine learning techniques, occlusion, and real-time detection are all included in the survey. Techniques based on deep convolutional neural networks (CNN) have been demonstrated to outperform others on a variety of image recognition benchmarks [5].

In terms of accuracy and speed, the present state-of-the-art object detectors using deep learning have their advantages and disadvantages. Within the picture, the item may have multiple spatial positions and aspect ratios. As a result, real-time object identification algorithms based on the CNN model, such as R-CNN [2] and YOLO [8], have been developed to recognise multi-classes in distinct regions in pictures.In terms of both speed and accuracy, YOLO (You Only Look Once) is the most popular deep CNN-based object identification algorithm.

We describe a computer vision system for recognising humans using a camera positioned at the roadside or at a workplace, based on the approach presented in [6]. The persons strolling in a certain location are covered by the camera's field of vision.These existing deep CNN approaches can recognise the number of persons in an image or video with bounding boxes, while the YOLO approach was used to identify the video stream recorded by the camera. The programme will emphasise if there is adequate social distance between persons in the video by measuring the Euclidean distance between them.

III. METHODOLOGY

Deep learning has ushered in the most advanced methodologies for a wide range of jobs and obstacles, including medical diagnosis, machine translation, speech recognition, and much more. Object categorization, detection, segmentation, tracking, and identification are at the heart of the majority of these activities.Convolution neural network (CNN)-based designs have demonstrated considerable performance increases in recent years, leading to high-quality object identification.

In this paper, a deep learning-based approach is developed that uses object detection and tracking models to help in the social distancing remedy for COVID-19 patients that are escalating.To strike a compromise between speed and precision, YOLO v3 and Deepsort are used as object detection and tracking techniques, with bounding boxes encircling each discovered object.These bounding boxes are then used to compute the pairwise L2 norm with a computationally efficient vectorized representation in order to detect groups of persons who are not following the social distancing order and are wearing masks.

1. Dataset

For testing the present technique, three datasets were employed. Dataset 1 contains photographs of individuals wearing face masks (Fig. 1),

Fig 1. Photos of people wearing masks

Dataset 2 contains photographs of individuals who do not wear face masks (Fig. 2),

Fig 2. Photos of people without masks

and Dataset 3 has photographs of individuals of various ages.Mostly Contains a lot of front face poses with only one face in the picture.

Fig 3. Photos of people in different age group

2. *Age Recognition*

I used a dataset of 8800 files in this example. It has 8800 photos of male and female faces ranging in age from 0 to 106 years old. Each photograph contains labels that indicate the age and gender of the subject. Males are assigned a number of 0 and females are assigned a number of 1.The ‘images' list contains all 8800 pictures, each of which has a different size (48 x 48 x 3). The respective ages are listed in the ‘ages' list, while the matching genders are listed in the ‘genders' list.

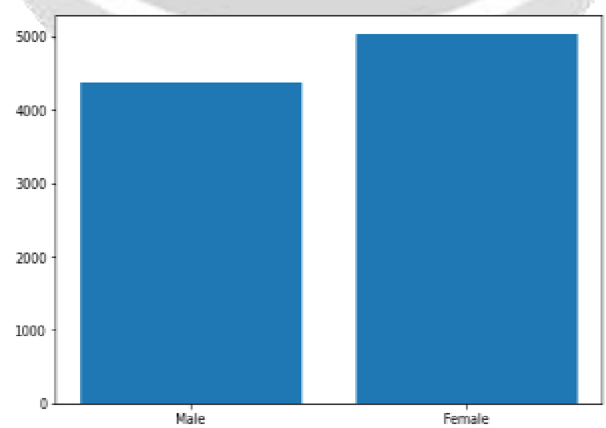

Fig 4. distribution of gender

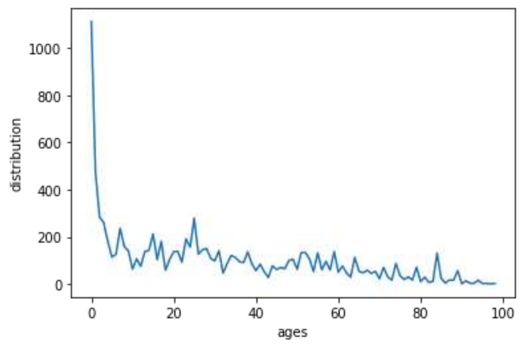

Fig 4. variation of samples of different ages

The first bar graph depicts the gender distribution. It appears to be well-balanced. The fluctuation of samples of various ages is seen in the second line graph. We can observe that the number of samples under 40 is significantly higher than the number of samples beyond 40. The train set distribution becomes skewed as a result of this.

As we've seen in this scenario, we'll need to use the same model to predict both age and gender. So, in order to produce the real labels for our training set, we'll have to go through several steps.The one-dimensional label vectors must be created.As a result, the ‘labels' list will have the following structure.

[[[age(1)],[gender(1)]],
[[age(2)],[gender(2)]],…………
[[age(n)],[gender(n)]]]

The labels and pictures list are then converted to NumPy arrays, the photos are normalised, and the training and test data splits are created.

3. *Identifying the face mask*

This research examines based on the deep convolutional neural network YOLOv3 of the single-step target detection in order to increase the accuracy of target detection.The YOLOv3 network model was trained, and the model with the highest detection accuracy was chosen by comparing mAP evaluation indexes and used for mask detection.

Face mask Net's mission is to identify if persons in public settings are wearing masks. On the basis of YOLOv3, the loss function in the original network structure was modified, i.e. IoU was replaced with DIoU, in order to increase detection accuracy.Second, the NMS function in the yolo module was replaced with DIou-NMS in order to increase the network's accuracy even further.

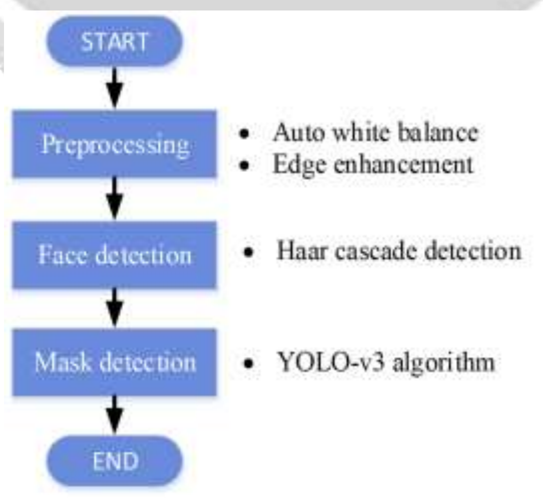

Fig 5. Algorithm for face mask detection

FPN was introduced by YOLOv3. The goal is to recognise things on three distinct scales by merging the shallow layer's high resolution characteristics with the high semantic information's high level characteristics. The dataset is clustered before network training, and K-means is used, based on the target criteria discussed in this study.The algorithm can retrieve the anchor value with great precision, allowing more precise boxes to be assigned to the target on the wider feature map. The size of anchors was determined using K-means clustering, and nine anchor point frames were selected: (11,14), (18,23), (27,33), (39,50), (57,72), (80,106), (118,154), (182,235). (305,388).As a result, each cell in each scale should anticipate three border boxes using three anchor points.

4. Social Distancing

This technique for detecting social distancing was created to identify the safety distance between individuals in public areas.In this study, the deep CNN approach and computer vision techniques are used. Initially, the person in the video frame was detected using an open-source object detection network based on the YOLOv3 method.Only the person class was utilised as a consequence of the detection, and other object types were disregarded in this application. As a result, the bounding box that best matches each identified person may be created in the image, and this data will be utilised to calculate distance.

The position of the bounding box for each individual (x, y, w, h) in the perspective view is recognised and translated into a top-down view in the distance measurement stage. The bottom-center point of the bounding box is used to estimate each person's position in the top-down view.From the top-down perspective, the distance between each person pair may be calculated, and the distances are scaled by the scaling factor obtained from camera view calibration. Given the positions of two pedestrians in a picture (x1, y1) and (x2, y2), the distance between them, d, may be calculated as follows:

$$d = \sqrt{(x2 - x2)2 + (y2 - y1)2}$$

The pair whose distance is less than the minimum permissible distance, t, is highlighted in red, while the remainder is highlighted in green. A red line is drawn between the pair of people whose distance is less than the pre-determined threshold. The colour threshold operation of the bounding box, c, may be defined as:

C = red if $d<t$
C = green if $d>=t$

IV. EXPERIMENTAL RESULTS

The proposed framework generates a processed frame with the identified people confined in bounding boxes, as well as a statistical analysis that displays the total number of social groups using the same colour encoding and a violation index term calculated as the ratio of the number of people to the number of groups.

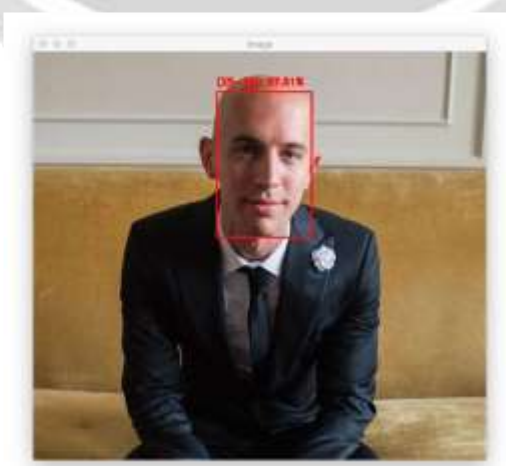

Fig 6. Age prediction result

The CNN system was trained on the image set which consists of color images of faces. This process determines the CNN parameters that were used to predict the apparent age of subjects' faces within the testing image set.

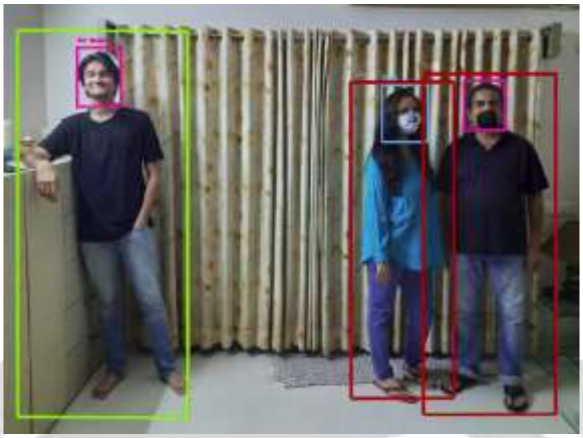

Fig 7. Result of detecting Face mask and social distancing

The violation indexes in the frames depicted in Fig are 3, 2, 2, and 2.33. For future examination, the frames containing identified violations are saved with a timestamp.

Object detection models are compared in terms of performance

Model	TT(in sec.)	NoI	mAP	TL
Faster RCNN	9651	1213	0.969	0.02
SSD	2124	1200	0.691	0.22
YOLO v3	**5659**	**7560**	**0.846**	**0.87**

V. CONCLUSION

As a result, we've developed a fully integrated real-time face mask, age, and social distance violation detection system, which employs YOLO v3 for object detection. Faces, both masked and unmasked, as well as full individuals, are all identified at the same time.

The relative distance between two persons is thus determined using optical principles utilising the coordinates provided by the detection of the class person. We found that the model produces fairly accurate results over a wide field of view, which is an important criterion for use in public settings, after extensive testing.Due to its high FPS and exceptional accuracy, this low weight model is easy to calibrate and may be utilised in real time without the necessity of time-consuming calculations or picture warping.

REFERENCES

1. Afiq Harith Ahamad, Norliza Zaini, Mohd Fuad Abdul Latip ,” Person Detection for Safety Violation Alert based on Segmented ROI”, 2020 10th IEEE International Conference on Control System, Computing and Engineering (ICCSCE2020), 21–22 August 2020, Penang, Malaysia.
2. D.T. Nguyen, W. Li, P.O. Ogunbona, “Human detection from images and videos: A survey”, Pattern Recognition, 51:148-75, 2016.
3. Yew Cheong Hou, Mohd Zafri Baharuddin, Salman Yussof, Sumayyah Dzulkifly, “Social Distancing Detection with Deep Learning Model”, 2020 8th International Conference on Information Technology and Multimedia (ICIMU)
4. R.Girshick,J.Donahue,T.Darrell,J.Malik."Richfeaturehierarchies for accurate object detection and semantic segmentation." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 580-587. 2014.
5. Kruti Goyal , Kartikey Agarwal , Rishi Kumar “Face Detection and Tracking ”,International Conference on Electronics, Communication and Aerospace Technology ICECA 2017.
6. J. Redmon, S. Divvala, R. Girshick, A. Farhadi, “You only look once: Unified, real-time object detection”, In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 779-788. 2016.
7. C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, Z. Wojna, “Rethinking the inception architecture for computer vision”, In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 2818- 2826, 2016.
8. Y. H. Kwon and V. Lobo, “Age Classification from Facial Images,” in IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1999, vol. 74, no. 1, pp. 1–21