



Social Distancing and Face Mask Detection Using YOLO Object Detection Algorithm

Riddhiman Raguraman¹, T. S. Gautham Rajan¹, P. Subbulakshmi¹(✉),
L. K. Pavithra¹, and Srimadhaven Thirumurthy²

¹ School of Computer Science and Engineering, Vellore Institute of Technology, Chennai,
Tamil Nadu, India

{riddhiman.r2020, gauthamrajan.ts2020}@vitstudent.ac.in,
{subbulakshmi.p, pavithra.lk}@vit.ac.in

² Thomas J. Watson College of Engineering and Applied Science, Binghamton University, State
University of New York, Binghamton, NY, USA

Abstract. Due to the COVID-19 pandemic, there has been a huge impact world-wide. The transmission of COVID-19 can be prevented using preventive measures like social distancing and face masks. These measures could slow the spreading and prevent newer ones from occurring. Social distancing can be followed even by those with weaker immune systems or certain medical conditions. With the new normal into play, maintaining distance in social and wearing masks are likely to be followed for the next two years. This paper studies about maintaining distance in social and detection of masks using deep learning techniques. Several object detection models are used for detecting social distance. The inputs used are in the form of images and videos. With this system, the violations can be detected which will reduce the number of cases. In conclusion, the proposed system will be very efficient and can also be used to introduce newer preventive measures.

Keywords: COVID-19 · social distancing detection · wearing mask detection · CNN · YOLO · Faster R-CNN · SSD · AlexNet · Inception v3 · MobileNet · VGG

1 Introduction

Deep learning is a combination of artificial intelligence (AI) and machine learning (ML) that imitates the method of how humans gain knowledge. In recent years, more attention has been gained by deep learning in object detection and its applications. A research trend has been shown by deep learning in several classes of object detection and recognition in AI. Even on challenging datasets, outstanding performance has been achieved.

The processing of artificial intelligence (AI) is done by merging large data with intelligent algorithms in which the features and patterns will be learnt automatically. Most of the AI examples are profoundly dependent on deep learning and natural language processing. AI is a wide-ranging field of study that comprises of several theories and methodologies. AI is divided into three subfields: Machine Learning, Neural networks

and Deep Learning. By means of these technologies, computers could be trained for completing complex tasks with large datasets [30].”

Coronavirus disease (COVID-19) is a transmittable virus caused by the SARS CoV-2 virus [28]. The risk of being affected by COVID-19 is highest in crowded and poorly ventilated places. Outbreaks have been reported in crowded indoor spaces where people speak or sing loudly such as restaurants, choir rehearsals, fitness centres, nightclubs, offices and worship places [29].

Due to the COVID-19 pandemic, we have started taking precautionary measures such as social distancing and face masks. The aim of maintaining distance in social is to decrease the fast spread of the disease through breaking the chain of communicating and prevent additional ones from occurring (by WHO - World Health Organization). These measures would reduce the number of COVID-19 cases and also reduce the possibilities of being affected. One of the best ways of prevention or slowing down the transmission is to be well informed about the virus and the disease. By following social distancing by staying at least 1 metre apart from others, properly using a tight mask and regularly cleaning your hands [28]. Social distancing can be followed even by those with weak immune systems or certain medical conditions. Similarly, face masks can be worn in hazardous or polluted environment.

Manual checking of people following social distancing and people wearing face masks correctly is a long and tedious process. In order to monitor the violations with ease, the system can detect social distancing and facial mask using Deep Learning techniques. This paper also provides information on implementation of distancing from others and detecting the mask using YOLO v5 object detection model [3]. The sections are divided as follows: Sect. 2 provides information on the works which are related to the paper and similar researches carried out by other fellow researchers, Sect. 3 describes the various object detection that may be implemented for social distancing and detection of mask in detail. Section 4 provides information on how the implementation was carried out using YOLO v5 object detection model. Section 5 explains the outcomes obtained from the implementation. Finally, the conclusion and future scope of the paper is explained in Sect. 6.

2 Related Works

Since the announcement of the new COVID-19 virus, everyone was advised to wear face masks and follow social distancing. Hence, there has been several researchers, who were keen on finding a proper monitoring or detecting system for these new norms. Different types of object detection models and algorithms have been used.

There have been several applications for which the detection models have been proposed. Also, the type of inputs varies with each application. The inputs are in the form of images, videos and real-time videos.

Fan Zuo et al. in their paper, analysed several object detection models and found that YOLO v3 had better performance in comparison with others [1]. Similarly, X. Kong et al. in their paper, has compared YOLO v3 object detection model with Faster RCNN and Single-Shot detector (SSD) on the basis of Frame rate, Inference time and Mean Average Accuracy (mAP). The observations from the comparison were that the YOLO v3 was better and efficient than Faster RCNN and SSD [8].

Few researchers proposed new models and frameworks. This also includes combining two or more models and algorithms and forming a hybrid model. Shilpa Sethi et al. in their paper, has proposed a real time mask identification framework based on edge computing. This uses deep learning and includes video restoration, face detection and mask identification [6]. Similarly, Qin B et al. used a face mask wearing condition identification method by using image SRCNet [5]. SRCNet stands for super-resolution and classification networks.

Most of the distance calculations are calculated using the Euclidean distance formula. X. Kong et al. and A. Gad et al. both have used Euclidean distance formula in their papers [2, 9]. Similarly, bird-eye is used for calibration. X. Kong et al. and A. Gad et al. both have used bird-eye for calibration in their papers [2, 9].

In the recent years, CNN models for example VGG16, AlexNet, ResNet-50 and InceptionV3 have been trained to achieve exceptional results in object detection [6]. With the help of the neural network structure, deep learning in object detection has been capable of self-constructing the object description and learning advanced features that cannot be obtained from the dataset. This has been the success of deep learning [4].

Raghav Magoo et al. proposes a model which can be used in embedded systems for surveillance applications. This model is a combination of one-stage and two-stage detectors. This model achieves less inference time and higher accuracy. An experiment was conducted using ResNet-50, MobileNet and AlexNet. From the experiment, it was observed that the proposed model obtains an accuracy of 98.2% when implemented with ResNet-50 [6].

B. Wang et al. in their paper proposes a two-stage model to identify the state of wearing a mask using a combination of machine learning methods. The first stage detects the mask wearing region of the person. This is on the basis of a transfer model of Faster RCNN and InceptionV2 structure. The second stage verifies the real face masks with the help of extensive methodology. The implementation is done by training a double stage model [7].

L. Jiao et al. in their paper have reviewed the various deep learning-based object detection methods. All the different methods have been explained and analysed in detail. This paper also lists the various applications and recent trends of object detection. This survey has been very helpful in the study of various deep learning techniques [12].

M. R. Bhuiyan et al. in their paper proposes a deep learning-based face mask detection method using YOLO v3 model. The authors have stated that “Experimental results that show average loss is 0.0730 after training 4000 epochs. After training 4000 epochs mAP score is 0.96. This unique approach of face mask visualization system attained noticeable output which has 96% classification and detection accuracy” [13].

M. Cristani in their paper has introduced the VSD (Visual Social Distancing) problem. The VSD analyses human behaviour with the help of video cameras and imaging sensors. The use of VSD beyond social distancing applications are also discussed in this paper [14].

M. Qian et al., in their paper has discussed about the social distancing measures for COVID-19. The author provides several information on social distancing other strategies. The author concludes that “social distance measure is the most effective practice to prevent and control the disease” [15].

S. Saponara et al., proposes an AI based system for classifying distance in social. This AI system can track people, detect social distancing and monitor body temperature with the images obtained from the thermal cameras [16].

Shashi Yadav proposes a computer vision based real time automated monitoring for social distancing and face masks. This approach uses computer vision and MobileNet V2 architecture. The process includes gathering and pre handling the data, development, training of the model, testing the model and model implementation [17].

Dzisi and Dei in their paper has discussed about the adherence to distancing in social and wearing mask in public transportation during COVID-19 in different countries. The paper mainly focuses on the situation in Ghana. A study on the people of Ghana about their compliance to face masks and social distancing in public transportation has also been discussed [18].

X. Peng et al., in their paper proposes a tracking algorithm in realtime for detection of face. Two steps of this approach: detection of face using RetinaFace and Kalman filter as tracking algorithm. Even though the proposed algorithm was not trained using datasets, it showed amazing results [19].

T. Q. Vinh and N. T. N. Anh in their paper propose a real-time face mask detector. This system uses Haar cascade classifier for face detection and YOLO v3 algorithm for identifying the face. The final output shows that the accuracy can achieve till 90.1% [20].

A. Nowrin et al., in their paper have studied about the various face mask detection techniques with different datasets. This paper also provides information on both machine learning and deep learning based algorithms which are used for object detection [21].

M. Z. Khan et al, have proposed a hotspot zone detection using computer vision and deep learning techniques. Experiments on different object detection algorithms have been carried out. Using confusion matrix, 86.7% accuracy was achieved for the whole human object interaction system [22].

J. Zhang et al., in their paper proposes a novel face mask detection framework namely Context-Attention R-CNN. This paper also proposes a new practical dataset which covers several conditions. Based on experiments, this proposed framework achieves 84.1% mAP on the proposed dataset [23].

S. Srinivasan et al., in their paper proposes an effective solution for person, distancing in social and detecting the mask using object detection and binary classifier based on CNN. This paper also studies the various face detection models. The final system achieves an accuracy of 91.2% [24].

A. Rahman et al., in their paper have analysed the various COVID-19 diagnosis methods which uses deep learning algorithms. The paper also provides information about the security threats to medical related deep learning systems [25].

M. Sharma et al., in their paper proposes an intelligent system for social distance detection using OpenCV. The process includes three steps: object detection, distance calculation, violation visualization [26].

K. Bhambani et al., in their paper has focused on providing a better solution for social distancing and face mask detection using YOLO object detection model on real

time images and videos. The proposed model has obtained an accuracy of 94.75% mAP [27].

3 Relevant Methodologies

Several methodologies are available to be used for detection of social distance and wearing of mask. Each methodology has its own characteristics and differ from each other. A few of them are explained below.

3.1 Convolutional Neural Network (CNN)

CNN has evolved into a popular class of deep learning methods. It has also been drawing interest from a wide range of fields. CNN can automatically learn the three-dimensional orders of features and through using backpropagation algorithm. The CNN architecture can be described as numerous layers, particularly convolution layer, pooling layer and fully connected layers. It takes an image as input, assigns priority with learnable weights and biases to numerous object features and distinguishes between them (Fig. 1).

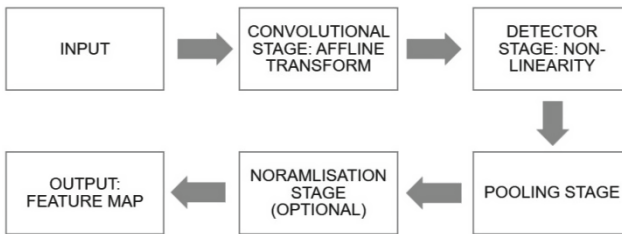


Fig. 1. Different stages in a typical CNN Layer.

The level of pre-processing required by CNN is significantly less in comparison with other classification algorithms. In comparison with its forerunners, CNN has the ability to identify key distinct features without any intervention of humans.

3.2 Object Detection

Object detection is associated with computer vision and image processing. It deals with feature detection of certain classes like humans, buildings or vehicles [12]. A traditional object detection algorithm can be divided into region selector, feature extractor and classifier [11]. Object detection comprises of both object classification and location regression [10] (Fig. 2).

The present deep learning-based object detectors are segregated into double stages namely, First-stage and second-stage detectors. Even though double-stage detectors obtain outstanding outcomes on various public datasets, they lag behind in low inference speed. On the other hand, one-stage detectors are fast and most preferred for several real-time object detection applications. On the whole, one-stage detectors have comparatively poor performance than two-stage detectors [10].

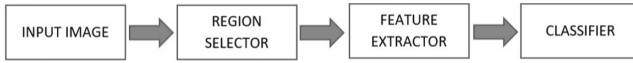


Fig. 2. Block diagram of a traditional object detection algorithm.

3.3 YOLO Object Detection Model

You Only Look Once (YOLO) is a popular object detection model used by research scholars around the world. YOLO is a CNN for executing the detection of objects in realtime. One of the advantages of YOLO is that it can be faster than other networks but still maintain the accuracy rate [29, 31] (Fig. 3).



Fig. 3. Block diagram of YOLO object detection model.

The YOLO model takes 45 fps (frames per second) to process images in real-time and the Fast YOLO model takes an extraordinary 155 fps to process images. The Fast YOLO model achieves twice the mAP of real-time conventional detection models. When generalising from natural photos to other domains like networks, the YOLO model outclasses existing detection models like R-CNN and DPM.

3.4 Faster R-CNN

Faster R-CNN is a combination of Fast R-CNN and RPN [6]. R-CNN represents Region-Based Convolutional Neural Network and RPN represents Region Proposal Network. To substitute the selection search algorithm, the Faster R-CNN uses RPN. This proposal generator is learned by using supervised learning methods. RPN is a fully convolutional network which uses a random sized images and generates various object detection proposal on each position of the feature map [10]. This allows cost free region proposals by integrating individual blocks of object detection in a single step. The individual blocks include feature extraction, proposal detection and bounding box regression [6].

3.5 Single-Shot Detector (SSD)

Initially, SSD identified for detecting objects using deep neural networks to solve computer vision problems. SSD will not reinitiate the features for the bounding box hypothesis [9]. RPN-based approaches like R-CNN require double stages. One stage for generating region proposal and another is for identifying object of each proposal. Therefore, two-shot approaches consume more time. In SSD, a single shot is enough to detect the various objects in an image. Hence, SSD is faster and time-efficient in comparison with the other two-shot RPN-based approaches [33].

3.6 AlexNet

AlexNet is a CNN architecture model which comprises a total of 8 layers along with weights. Among the 8 layers, the first 5 layers are convolutional layers and the remaining 3 layers are fully connected layers. The Rectified Linear Unit (ReLU) is used after every layer (both convolutional and fully connected) [34]. ReLU helps in preventing the computation growth required for operating the neural network. The next layer indiscriminately assigns inputs to zero with the occurrence rate of every stage. This helps to prevent overfitting [32]. The dropout method is used before or within the two fully connected layers. But on using this method, the time required for the network to converge increases [35].

3.7 Inception V3

Inception v3 is mainly used as assistance in image analysis and object detection. Inception v3 is the third version of the Inception CNN from Google. It comprises of 42 layers which slightly is higher than the v1 and v2 models. Inception v3 is an optimized edition of Inception v1. The main purpose of this version is to allow deeper networks, along with numerous parameters. In comparison with the parameters of AlexNet (60 million), Inception v3 has lesser parameters (25 million). In comparison with Inception v1, Inception v3 has a deeper network and higher efficiency. It is also less expensive.

3.8 MobileNet

MobileNet is a portable efficient CNN used in several applications. They are small in size, low latency and low powered models. Since it is light weighted, it has lesser parameters and higher classification accuracy. MobileNet comprises of depth-wise separable convolution layers which consists of a depth-wise convolution and a point wise convolution. Along with each depth-wise and point-wise convolution, a MobileNet includes a total of 28 layers. MobileNet also introduces two new universal hyperparameters: width and resolution multiplier. This permits developers to choose between latency and accuracy on the basis of the requirements.

3.9 Visual Geometry Group (VGG)

VGG or Visual Geometry Group is an advanced object detection and recognition model which supports up to 16 or 19 layers. As input, the VGG accepts 224×224 pixel RGB image. 3×3 convolutional layers are used. Also, a 1×1 convolutional filter is present. This acts as linear transformation for the input. The convolutional filter is then followed by a ReLU unit. The VGG is made up of three complete communicated layers. The initial layer comprises of 4095 channels each and the third layer comprises of 1000 channels. The hidden layers of VGG uses ReLU, as it is time efficient.

4 Implementation

Several papers with several methodologies and techniques of deep learning were studied. As per the study, it has been found that the social distancing and face mask can be detected efficiently by using CNN and YOLO object detection model. The latest version of the YOLO model, the YOLO v5 object detection model has been used.

Object detectors detect a corresponding label and a bounding box. In both the scenarios, bounding boxes are used for indication. All the codes were run in Google Colab. The inputs were mounted on Google Drive and imported into the Colab.

4.1 Face Mask Detection

Input is a custom dataset. The dataset consists of 50 images along with its corresponding labels to indicate the mask correctly. First step is to clone the repository from GitHub [36]. Then, all the required dependencies are installed. The inputs are mounted on Google Drive and imported into the Colab file (Fig. 4).

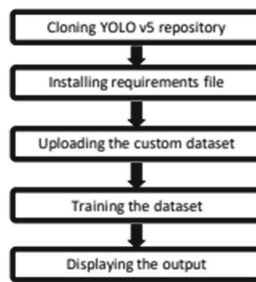


Fig. 4. Outline of Face Mask Detection

A configuration file is created to provide information about the dataset and its location to the model. The dataset path is provided for training and validation. Then the number of classes and name of the class are provided. The design of the model can be seen with all the layer informations. The model starts training and the results are logged. The models are trained for 50 epochs, tested and validated. PR and F1 curves are plotted. In addition to the custom data set, a video file is given as input, trained and validated (Fig. 5).



Fig. 5. Sample custom face mask dataset.

4.2 Social Distance Detection

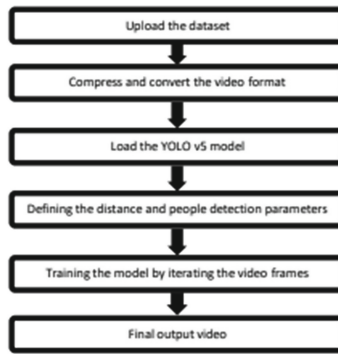


Fig. 6. Outline of Social Distancing Detection

The input is the processed video output of the face mask detection. This will use Euclidean distances. The ED between the centroids of all the regions projected will be computed (Fig. 6).

The Euclidean distance is expressed as

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

where,

(x_1, y_1) are the coordinates of one point.

(x_2, y_2) are the coordinates of the other point.

d is the distance between (x_1, y_1) and (x_2, y_2) .

All the required libraries are imported. The drive is mounted and input is copied to the Colab file. Pre-processing of the video like compression and conversion of the video format is carried out. This YOLO model uses PyTorch library. With the help of the specified distance formula, the YOLO model detects people and draws rectangle boxes surrounding them.

Bounding boxes will be used to indicate the person and lines will be used to indicate close distance contacts. The colour of the bounding boxes will switch from green to red, if the social distancing norms are violated.

To detect people in a video, frame by frame iteration is carried out and at the end of the process, the output file is saved.

5 Results

The results of the implementation are analysed and found to be satisfactory. From the PR and F1 curve graphs from the face mask detection, the efficiency of the trained model can be understood. A PR curve is plotted by taking the values of Precision and Recall on the y-axis and x-axis respectively. The F1 curve is plotted by taking the values of the harmonic mean of both precision and recall. The average precision rate is found to be 60%, the average recall rate is found to be 74% and average accuracy mAP was found to be 62% (Figs. 7 and 8).

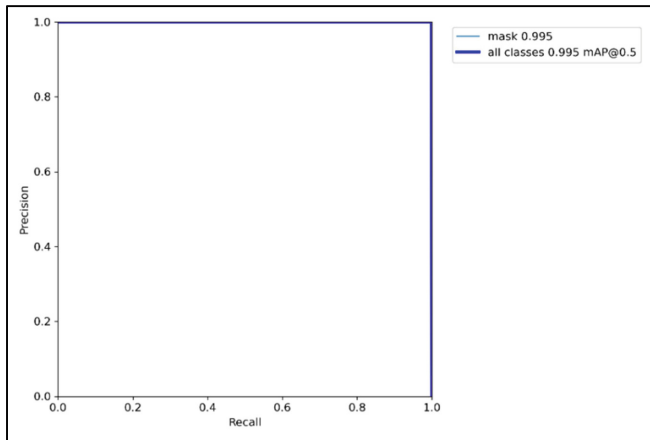


Fig. 7. Precision-Recall curve.

Figure 9 and 10 showcases the result outcome of the implementation. The red boxes detect the mask and the green box detects the social distance. Few limitations in the implementation are: uncontrolled background and illumination. The camera angle and lighting conditions affect the detection of the person. Even in a diverse environment, problems occur in the people detection.

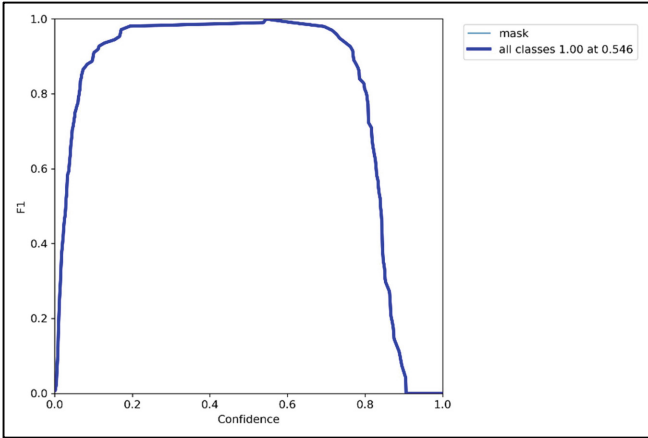


Fig. 8. F1 curve.



Fig. 9. Sample trained dataset with bounding boxes and mask prediction scores.



Fig. 10. A video frame of social distancing and face mask detection.

6 Conclusion and Future Scope

Even though the COVID-19 scenario has come to an end. New variants are still emerging around the world. To prevent further spread of the disease, everyone should follow the guidelines and norms advised by the WHO. The detection of social distancing and face mask is to provide an awareness to the people that the disease is still on the loose.

As future scope, the model can be trained to give comparatively better results. Real-time video can be taken as input and trained. To enhance the model further, temperature and cough detection can be added in addition to face mask and social distancing.

References

1. Zuo, F., et al.: Reference-free video-to-real distance approximation-based urban social distancing analytics amid COVID-19 pandemic. *J. Transp. Health* **21**. Science Direct (2021)
2. Gad, A., ElBary, G., Alkhedher, M., Ghazal, M.: Vision-based approach for automated social distance violators detection. In: 2020 International Conference on Innovation and Intelligence for Informatics, Computing and Technologies (3ICT). IEEE (2020)
3. Madane, S., Chitre, D.: Social distancing detection and analysis through computer vision. In: 2021 6th International Conference for Convergence in Technology (I2CT). IEEE (2021)
4. Hou, Y.C., et al.: Social distancing detection with deep learning model. In: 2020 8th International Conference on Information Technology and Multimedia (ICIMU). IEEE (2020)
5. Qin, B., Li, D.: Identifying facemask-wearing condition using image super resolution with classification network to prevent COVID-19. *Sensors (Basel)*. MDPI (2020)
6. Sethi, S., Kathuria, M., Kaushik, T.: Face mask detection using deep learning: an approach to reduce risk of Coronavirus spread. *J. Biomed. Inform.* **120**. Science Direct (2021)
7. Wang, B., Zhao, Y., Chen, C.L.P.: Hybrid transfer learning and broad learning system for wearing mask detection in the COVID-19 Era. *IEEE Trans. Instrum. Meas.* **70**, 1–12. IEEE (2021)
8. Kong, X., et al.: Real-time mask identification for COVID-19: an edge-computing based deep learning framework. *IEEE Internet Things J.* **8**(21), 15929–15938. IEEE (2021)
9. Magoo, R., et al.: Deep learning-based bird eye view social distancing monitoring using surveillance video for curbing the COVID-19 spread. *Neural Comput. Appl.* **33**, 1–8. Springer (2021)
10. Wu, X., et al.: Recent advances in deep learning for object detection. *Neurocomputing* **396**, 39–64. Science Direct (2020)
11. Xiao, Y., et al.: A review of object detection based on deep learning. *Multimed. Tools Appl.* **79**, 23729–23791. Springer (2020)
12. Jiao, L., et al.: A survey of deep learning-based object detection. *IEEE Access* **7**, 128837–128868. IEEE (2019)
13. Bhuiyan, M.R., et al.: A deep learning based assistive system to classify COVID-19 face mask for human safety with YOLOv3. In: 2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT). IEEE (2020)
14. Cristani, M., Bue, A.D., Murino, V., Setti, F., Vinciarelli, A.: The visual social distancing problem. *IEEE Access* **8**, 126876–126886. IEEE (2020)
15. Qian, M., Jiang, J.: COVID-19 and social distancing. *J. Public Health* **30**, 259–261. Springer (2022)
16. Saponara, S., Elhanashi, A., Gagliardi, A.: Implementing a real-time, AI-based, people detection and social distancing measuring system for Covid-19. *J. Real-Time Image Process.* **18**, 1937–1947. Springer (2021)

17. Yadav, S.: Deep learning based safe social distancing and face mask detection in public areas for COVID-19 safety guidelines adherence. *Int. J. Res. Appl. Sci. Eng. Technol.* **8**(VII), 1368–1375. Research Gate (2020)
18. Dzisi, E.K.J., Dei, O.A.: Adherence to social distancing and wearing of masks within public transportation during the COVID 19 pandemic. *Transp. Res. Interdiscip. Perspect.* **7**. Science Direct (2020)
19. Peng, X., Zhuang, H., Huang, G.-B., Li, H., Lin, Z.: Robust real-time face tracking for people wearing face masks. In: 2020 16th International Conference on Control, Automation, Robotics and Vision (ICARCV), pp. 779–783. IEEE (2020)
20. Vinh, T.Q., Anh, N.T.N.: Real-time face mask detector using YOLOv3 algorithm and haar cascade classifier. In: 2020 International Conference on Advanced Computing and Applications (ACOMP), pp. 146–149. IEEE (2020)
21. Nowrin, A., Afroz, S., Rahman, M.S., Mahmud, I., Cho, Y. -Z.: Comprehensive review on facemask detection techniques in the context of COVID-19. *IEEE Access* **9**, 106839–106864. IEEE (2021)
22. Khan, M.Z., Khan, M.U.G., Saba, T., Razzak, I., Rehman, A., Bahaj, S.A.: Hot spot zone detection to tackle COVID19 spread by fusing the traditional machine learning and deep learning approaches of computer vision. *IEEE Access* **9**, 100040–100049. IEEE (2021)
23. Zhang, J., Han, F., Chun, Y., Chen, W.: A novel detection framework about conditions of wearing face mask for helping control the spread of COVID-19. *IEEE Access* **9**, 42975–42984. IEEE (2021)
24. Srinivasan, S., Rujula Singh, R., Biradar, R.R., Revathi, S.: COVID-19 monitoring system using social distancing and face mask detection on surveillance video datasets. In: 2021 International Conference on Emerging Smart Computing and Informatics (ESCI), pp. 449–455. IEEE (2021)
25. Rahman, A., Hossain, M.S., Alrajeh, N.A., Alsolami, F.: Adversarial examples— security threats to COVID-19 deep learning systems in medical IoT devices. *IEEE Internet Things J.* **8**(12), 9603–9610. IEEE (2021)
26. Sharma, M.: Open-CV social distancing intelligent system. In: 2020 2nd International Conference on Advances in Computing, Communication Control and Networking (ICACCCN), pp. 972–975. IEEE (2020)
27. Bhambani, K., Jain, T., Sultanpure, K.A.: Real-time face mask and social distancing violation detection system using YOLO. In: 2020 IEEE Bangalore Humanitarian Technology Conference (B-HTC), pp. 1–6. IEEE (2020)
28. World Health Organization (WHO) – COVID-19. https://www.who.int/health-topics/coronavirus#tab=tab_1
29. World Health Organization (WHO) – Advice for the public: COVID-19. <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/advice-for-public>
30. Artificial Intelligence. What it is and why it matters. https://www.sas.com/en_us/insights/analytics/what-is-artificial-intelligence.html
31. YOLOv3: Real-Time Object Detection Algorithm. <https://viso.ai/deeplearning/yolov3-overview/>
32. Keras API reference. <https://keras.io/api/>
33. Tsang, S.-H.: Review: SSD — Single Shot Detector (Object Detection). <https://towardsdatascience.com/review-ssd-single-shot-detector-object-detection851a94607d11>

34. AlexNet – ImageNet Classification with Deep Convolutional Neural Networks. <https://neurohive.io/en/popular-networks/alexnet-imagenet-classification-with-deep-convolutional-neural-networks/>
35. Alake, R.: What AlexNet Brought To The World Of Deep Learning. <https://towardsdatascience.com/what-alexnet-brought-to-the-world-of-deep-learning46c7974b46fc>
36. GitHub for YOLO v5. <https://github.com/ultralytics/yolov5>