



Face mask detection in COVID-19: a strategic review

Vibhuti¹ · Neeru Jindal² · Harpreet Singh³ · Prashant Singh Rana³

Received: 14 May 2021 / Revised: 12 January 2022 / Accepted: 27 March 2022 /
Published online: 5 May 2022

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract

With the outbreak of the Coronavirus Disease in 2019, life seemed to be had come to a standstill. To combat the transmission of the virus, World Health Organization (WHO) announced wearing of face mask as an imperative way to limit the spread of the virus. However, manually ensuring whether people are wearing face masks or not in a public area is a cumbersome task. The exigency of monitoring people wearing face masks necessitated building an automatic system. Currently, distinct methods using machine learning and deep learning can be used effectively. In this paper, all the essential requirements for such a model have been reviewed. The need and the structural outline of the proposed model have been discussed extensively, followed by a comprehensive study of various available techniques and their respective comparative performance analysis. Further, the pros and cons of each method have been analyzed in depth. Subsequently, sources to multiple datasets are mentioned. The several software needed for the implementation are also discussed. And discussions have been organized on the various use cases, limitations, and observations for the system, and the conclusion of this paper with several directions for future research.

Keywords COVID-19 · Face mask detection · Object detection · Classification

✉ Vibhuti
vvibhuti_be17@thapar.edu

Neeru Jindal
neeru.jindal@thapar.edu

Harpreet Singh
harpreet.s@thapar.edu

Prashant Singh Rana
prashant.singh@thapar.edu

¹ Research Scholar, ECED, Thapar Institute of Engineering and Technology, Patiala, Punjab, India

² Faculty, ECED, Thapar Institute of Engineering and Technology, Patiala, Punjab, India

³ Faculty, CSED, Thapar Institute of Engineering and Technology, Patiala, Punjab, India

1 Introduction

The spread of Coronavirus disease 2019, commonly known as COVID-19, is a significant concern for everyone worldwide. It is a contagious disease that has affected human life globally [108, 117]. The health specialists suggest that the virus might transmit by direct or indirect contact with the infected person [119], hence measures like compulsory wearing of face masks [40], as illustrated in Fig. 1, have been strictly put into effect by medical bodies. Numerous studies advise putting face masks on even if a person is not feeling sick. It is not the first time, during COVID-19, that wearing face masks has been stressed to combat the transmission. It is a practice that can be dated back to the 1910–11 Manchurian epidemic in China [60]. Various pandemics of history have been survived by wearing face masks. Besides, it is well proven by various studies that not just wearing face masks instead wearing them properly limits the transmission of the virus to quite an extent. The observation that greater the proportion of population wearing face masks in a country, the lesser the cases of COVID-19 in the nation has created the need for an automated face mask detector.

Further, the coronavirus pandemic has necessitated the scientific contribution across the globe to help in battling the pandemic. Leveraging the contemporary technical advancements. Numerous solutions to prevent the transmission of the virus have been formulated. As observed in [71], the authors have put forward an updated mask detection architecture working with noteworthy efficiency of 97%. In [5], the spotting of face masks involved PyTorch, with results being 97% accurate. Further, [95] proposed the detection of several kinds of masks using ultramodern method, and also, the output was obtained after applying the model in real-time. CNN based detectors have been used on custom collected face mask datasets in [21]. Another study was performed to formulate an application that inspects people wearing face masks in public areas [31]. Additionally, the already existing dataset was enriched by including more images in [79]. The proposed work used the Faster R-CNN model to implement the task and achieved an accuracy of 99.8%. In [30], the authors have put forward a system of verifying the correct position of the face mask of an individual, while [72] includes discussions on the various technological methods available to deal with the virus.

With the advancements in technology that the world has been witnessing, there are various available techniques [7, 48, 74, 76, 113] that could prove valuable to society if used

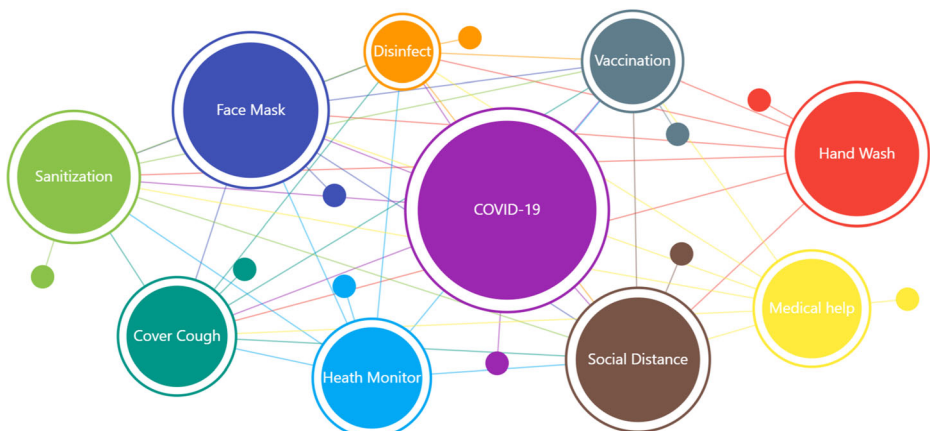


Fig. 1 Precautions to avoid COVID-19

effectively. A real-time system which could itself classify, seeing a person, in two categories [77]:

- 1) A person wearing a face mask
- 2) A person not wearing a face mask

could be useful in recent times. Such systems could find applications in public areas like hospitals, airports, malls, etc. One of the methods to make the detector is by first detecting the faces in real-time. And, after detecting the faces from the webcam stream, saving the frames containing the faces and next applying a classifier. The numerous algorithms that could be used for categorization have been discussed in the subsequent section. Another way that could be opted to execute the same is by using an object detection model. Following are the contributions in view of the current state-of-the-art.

- Although several precautions are recommended to get safe from covid-19, still face masking, and social distancing are significant factors. So, it was necessary to propose many face masking techniques under one umbrella for the research community.
- Pertaining to the need of the current time, the proposed work reviews several studies conducted in the field of face mask detection. The strong suit of plenty of publications has been discussed on face masking, which is still missing in terms of observations, future trends, a vast number of references, current trends, etc.
- Performance parameters of several algorithms are compared, and discussions on them are presented to increase the efficacy of the review paper.

1.1 Motivation and trends in recent years

With time, the surge in COVID-19 cases urged people to be cautious, alert, and take all safety measures possible. In situations such as this, where a mere sneeze could be harmful to many people, safety remains the priority. To ensure the well-being of all humans, a system that could itself monitor if a face mask is on or not is necessitated. It would not only secure a being rather fellows in the vicinity as well. Having access to the ultra-modern technological methods, implementing such a system could be a boon to society.

After analyzing the problem statement, numerous studies performed on the same were scrutinized to commence the research. Then, the content relevant to the issue was filtered, and a depth understanding of the topic was attained. Further, several existing datasets were explored, along with the techniques available. The literature survey of the available methods was conducted, followed by a comparison of the different algorithms. Further, the software was explored and thereby applications. Eventually, the future scope was inspected as shown in Fig. 2.

Initially, around 180 papers were identified belonging to varied publications like Springer. Later, the collected documents were checked for duplication and removed, if any. Then, the articles were screened for their eligibility in context with relevance to the problem statement and thereby, leaving just about 140 papers. Further, the papers were assessed for quality, bringing down the count to 130. Besides, around 100 papers were analyzed for understanding the various techniques available, including

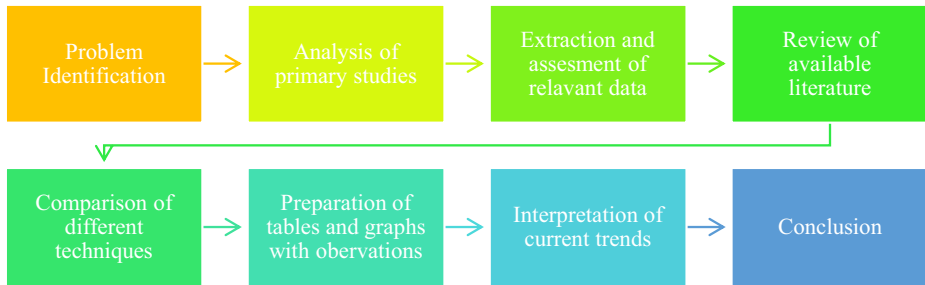


Fig. 2 Methodology used

state-of-the-art. Few more publications were investigated to gather knowledge about the available datasets.

The paper's organisation is as follows: Section 2 deals with the general flow chart of the face mask detector. Section 3 discusses the various techniques that could be used to implement a face mask detector, while Section 4 reviews some of the real-time methods. Section 5 analyses the trends of techniques in the last two decades along with the advantages and challenges of the techniques discussed in Section 3. In Section 6, the URLs for multiple online available datasets are mentioned. Section 7 suggests several useful software that could be used to carry out the process, followed by Section 8 that states the use cases, drawbacks and the observations made for the process. Section 9 provides conclusions of the study along with future directions.

Figure 3 illustrates the number of publications in face mask detectors in the last two decades. Owing to COVID-19, such detectors have gained to be a hot topic of study in 2020 among researchers.

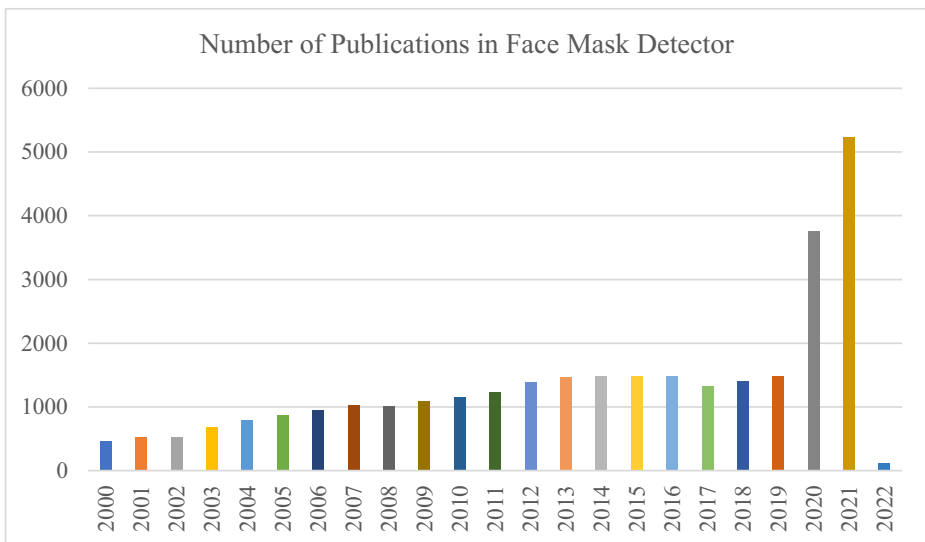


Fig. 3 The number of publications in face mask detector from the year 2000 to 2022(The year 2022 includes data till January 11) as taken from Semantic Scholar using words “Face Mask Detector”

2 General flow chart

The implementation of the face mask detector system could be executed in two phases, as shown in Fig. 4.

The first phase is the training phase. This stage is initiated with the collection of the dataset. One of the most crucial steps is to have a good quantity and quality of data [1]. One can prepare the dataset or use already existing datasets from the various available sources. If preparing yourself, the size of data could be increased by using techniques like data augmentation. Also, the data has to be cleaned before use because it plays a significant role in building a model. Various Steps involved in data cleaning are shown in Fig. 5. After obtaining a good quality dataset, the model is selected under the system’s demands and trained on the chosen dataset. Multiple techniques could be used to accomplish the target.

By acquiring the most suitable trained model, the first phase comes to an end. In the subsequent step, the frames from the live video feed or the images are used as input to the trained model. The live video feed could be obtained using a mobile phone, a camera, or a surveillance camera and hence could vary in format, i.e., H.265, H.264, etc.

There are several cases where the video frame cannot capture the images as desired. There is a possibility of the video recorded being blurred or having noise, etc. In scenarios like these,

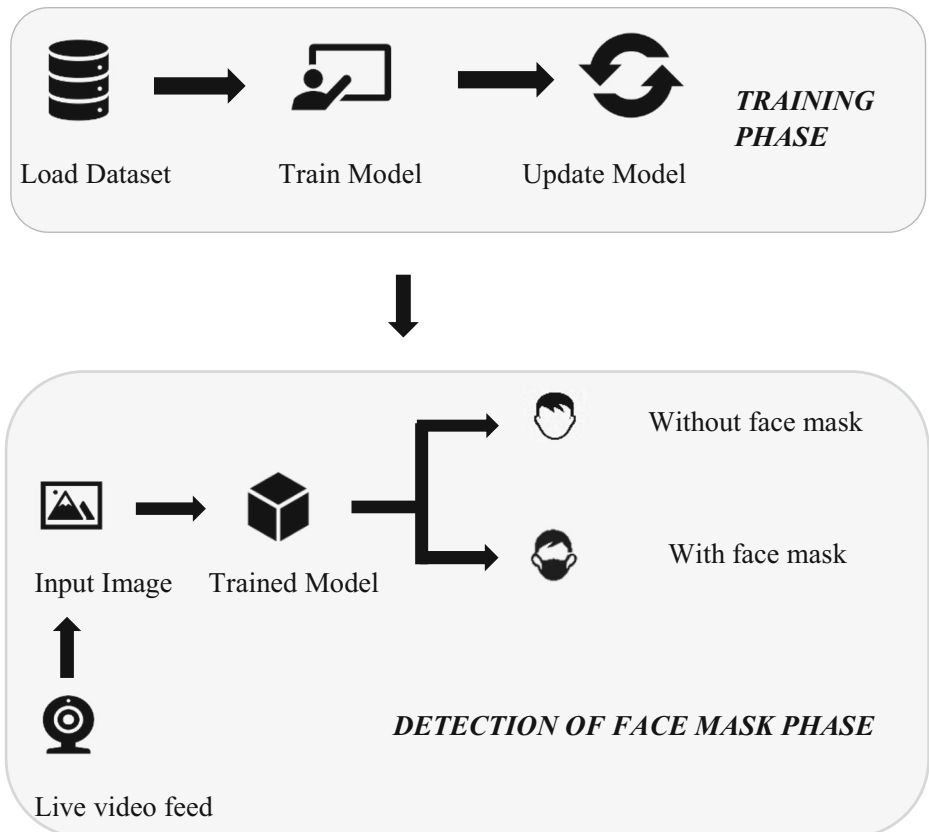


Fig. 4 The proposed flow diagram for face mask detection system



Fig. 5 Steps involved in data cleaning

image pre-processing comes to the rescue. Further, there are several methods in OpenCV that could be used to enhance the quality of the image. For instance, blurriness could be reduced using the filter2D function of OpenCV, which enhances the sharpness of the picture. Also, image denoising techniques of the same library are helpful to deal with noisy images. Various transforms or histograms could be used for the same. Additionally, object tracking could also be considered to detect faces. Though these are the ways to deal with the discrepancies, the target should be to capture good quality videos (Fig. 6).

3 Face mask detection techniques

Some of the several techniques used in face mask detection are discussed below (Fig. 7):

3.1 Object detection

Deep Learning techniques have managed to pick up steam currently because of their ability to train vast data with high accuracy [102]. These state-of-art methods prioritise accuracy in some cases whereas speed in others. In place of the advantages of deep learning techniques in a real-time application, this section discusses object detection using the deep learning approach [19, 29, 42, 46, 109, 114].

At the hands of Computer Vision, Object Detection works to identify and locate objects of certain classes in images and videos. This is imitated in Fig. 8. Besides, this technique uses bounding boxes to localize the things in the input image. This can also enumerate the number

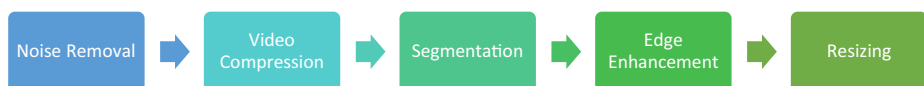


Fig. 6 Video pre-processing techniques

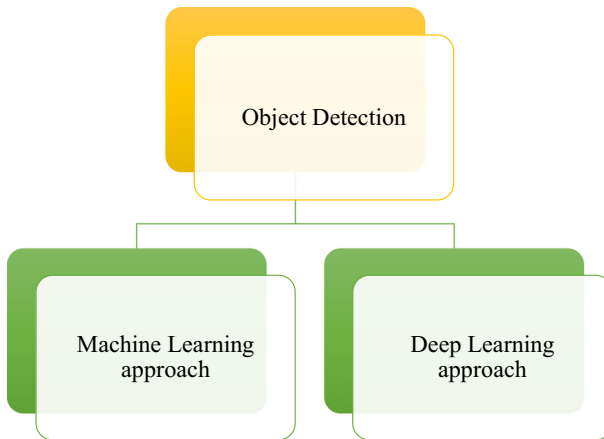


Fig. 7 Approach for object detection methods

of objects in the given image. Various object detection algorithms are available lately [37, 41, 121]. They are categorized into [92].

- Two-Shot Detection
- Single-Shot Detection

3.1.1 Two shot detectors

This model achieves the target in two steps: Region proposal followed by classification of those regions and refinement of location prediction. Various models for this category are:

- **Faster Region-Based Convolutional Neural Network**

It is the improvised model of earlier proposed R-CNN [91] and fast R-CNN. It comes with better region-based CNN architecture [25]. Moreover, it is one of the extensively employed advanced algorithm with the R-CNN backbone. Compared to earlier models, it replaces the selective search algorithm used to identify RoI. The detailed diagram explaining the same is shown in Fig. 9. Additionally, when accuracy is of concern, this algorithm is given preference. In [82], the author performs company logo detection using the mentioned technique. Also, in [22], this algorithm is used to identify the stages in malaria-infected blood. In [39], the author



Fig. 8 Object detection segments

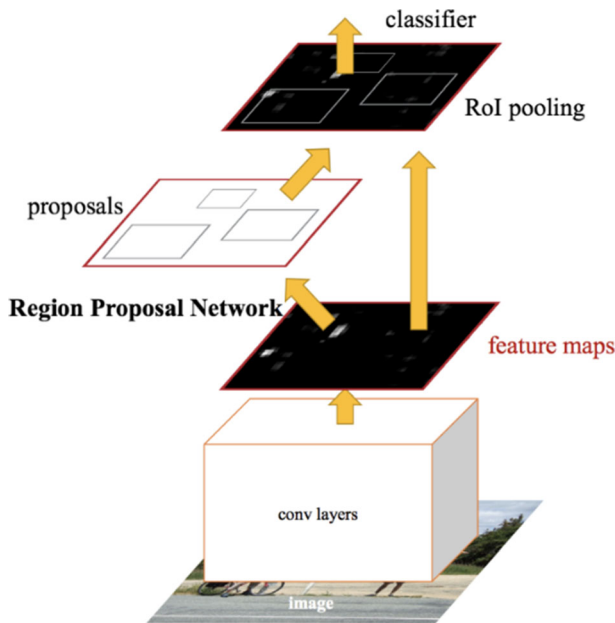


Fig. 9 Faster R-CNN [33]

uses this state-of-art model to monitor people wearing face masks in public areas. Furthermore, several researchers [6, 14, 63, 87, 94, 103, 115] have taken leverage of this method.

- **Region-Based Fully Convolutional Network**

It is a two-shot architecture that is developed, taking inspiration from Faster-RCNN. Unlike Faster R-CNN, all the composite work is finished before ROI pooling, which is applied on score maps. All regional proposals utilize the same score maps to perform average voting. Also, all the layers are convolutional and computed on the image. It can be taken as a hybrid model of one-shot and two-shot models. The architecture is shown in Fig. 10. Besides, the related works are talked over in [15, 54, 106] closely.

3.1.2 Single-shot detectors

They are usually used when speed is a priority to implement a study. This is because of their method to predict the boundary boxes and the classes, which does not involve a dedicated step for the proposal of bounding boxes and utilizes a single deep neural network. Therefore, they find numerous applications in real-time detections.

- **You Only Look Once**

Unlike selecting an image in parts, the algorithm performs categorization in a single pass. The input image is made to pass through multiple layers of the network which eventually produces a prediction as an output [62]. Moreover, Yolov3 makes use of DarkNet-53 to detect features. DarkNet-53 is a 53 layers CNN trained on ImageNet. It even uses Residual networks, which

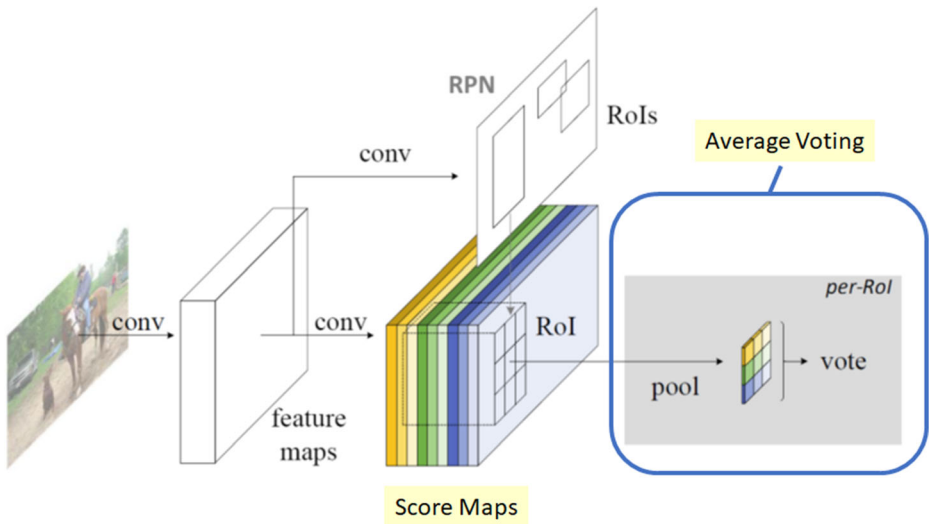


Fig. 10 R-FCN [34]

skip connections [80]. Besides, anchor boxes are used as a pre-trained landmark by the bounding boxes to provide the detected object location. Again, it predicts the class probabilities for each grid cell. In this model, the Non-max Suppression algorithm finds usage to eliminate anchor boxes that are not required. The bounding boxes are discarded using IoU (Intersection over Union) (Fig. 11)

Further, YOLO has gained attention because of its speed [58]. Moreover, its excellence in learning even on the generalized images of the objects and making predictions with high

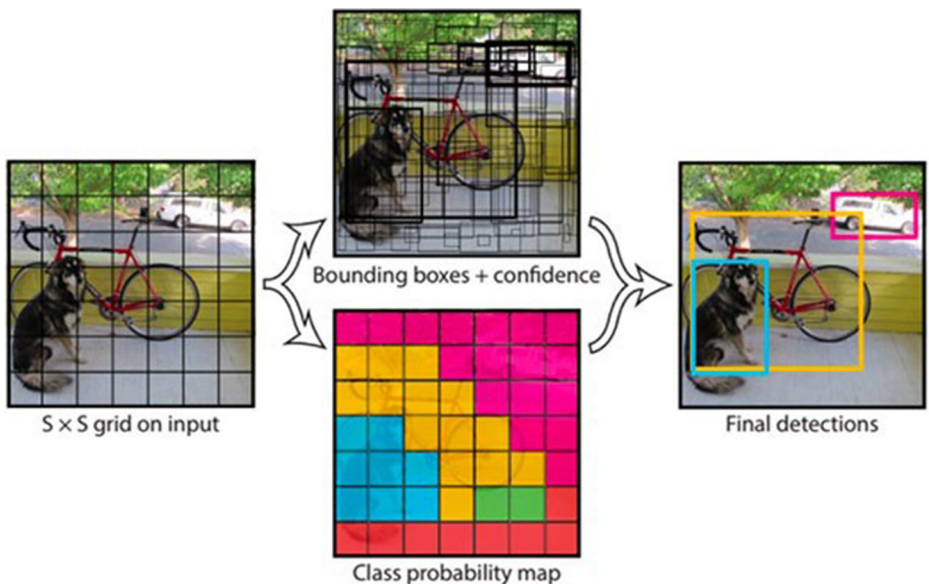


Fig. 11 Working of Yolo [32]

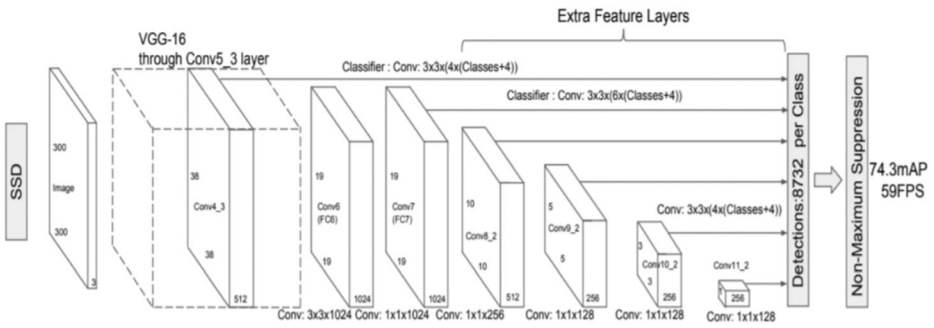


Fig. 12 Single Shot Multibox Architecture [35]

accuracy aids it is outperforming other fellow models. In [85], the author has enhanced the traditional Yolov4 series to propose a novel detector. Likewise, in [11], this state-of-art technique has been implemented to improve the performance of mask detectors. Also, a similar approach is elucidated in various [2, 8, 38, 43, 49, 52, 55, 57, 83, 88, 98] compositions.

• **Single shot multibox detector**

It uses VGG-16 as its backbone architecture, discarding the fully connected layers [12]. The model can be set up in two components, i.e., extraction of feature maps, followed by application of convolution filter in order to detect objects. It works by matching objects with default boxes of distinct aspects. Whenever any box meets the set minimum threshold value of IoU, a match becomes considerable. Besides, after approximation, each feature map location is scaled, and the predictions by the model are made by feature maps to consider objects of multiple sizes as shown in Fig. 12.

In [65], real-time face mask detection is discussed with changes in architecture used. [68] provides a way to execute the algorithm. Also, [81] talks about the model used in detecting objects for the blinds. Further, a different approach is used in [23] for object detection. In [17], an improvised way of detecting face masks using SSD has been executed. The authors have improved the algorithm by using inverse convolution and feature fusion. While [53] brings up a similar technique for executing their study.

It can be observed from Table 1 that single shot detectors, including YOLO and SSD, have higher inference speed owing to faster localization and categorization followed by Faster R-CNN. Additionally, the algorithm to be used is chosen depending on the requirement of the problem. Generally, Faster R-CNN, because of the detection speed, is employed when the results are not to be obtained in real-time, whereas YOLO is the choice of practitioners when

Table 1 Comparison table of state-of-the-art detection models [78]

Model	Accuracy	FPS
Faster R-CNN	70.4	17
R-FCN	77.6	6
YOLOv3	78.6	91
SSD	78.5	59

working with live data feed. Also, SSD maintains a balance between speed and detection effectiveness.

3.1.3 Feature extraction

Extraction of features is a way to get rid of unnecessary information from the data, thereby reducing the computational cost and still having imperative and relevant data reserved. Also, the reduced data helps increase the model's learning rate. Moreover, real-time face mask detection leverages machine learning and deep learning techniques for feature extraction. In deep learning, neural networks themselves facilitate extracting features without human intervention. The input data is passed to the feature extraction network, with different backbone architectures, including MobileNetv2 and Xception [71]. Subsequently, the result is forwarded to the classifier network categorizing a person with or without a mask. On the other hand, algorithms, like histogram of oriented gradients (HOG) and Principal Component Analysis (PCA), could be utilized to obtain features in the machine learning model [29, 71]. Additionally, features could be extracted manually by incorporating the methods mentioned in Fig. 13.

3.2 Other techniques

Diversely, another path that could be taken to execute the study is by considering the problem in two sections. The problem statement, here face mask detector, could be constructed by first performing face detection on the frames coming from the video feed and later giving the frames with faces as an input to the classifier, which hence furnishes the desired output, i.e., faces with or without masks (Fig. 14).

Elaborating on above-mentioned points, FACE DETECTION is a technical advancement in the contemporary world where human faces could be detected in an image. The location of the face is marked using bounding boxes. Also, numerous aspects are to be considered to

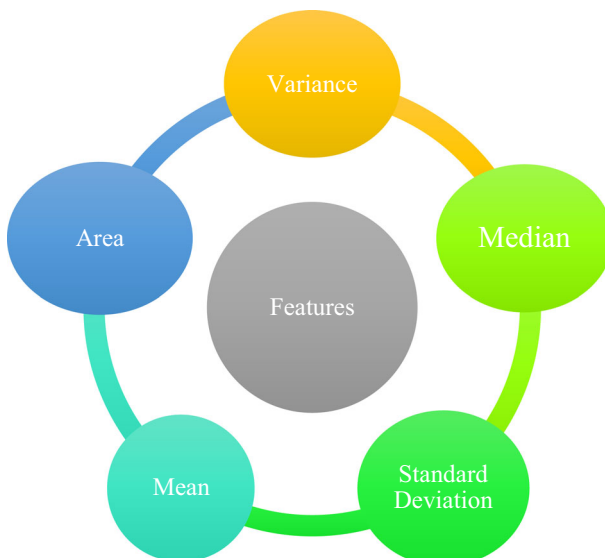


Fig. 13 Features extraction techniques in face mask detection [12, 15, 42, 93]



Fig. 14 The Face Mask Detection could be implemented by first performing face detection followed by face mask classification on an individual

perform successful detection [51]. Due to the advantages of neural networks, even they are used in detection [104]. The innovation is in use in various applications. Some of the different methods to perform the same are listed below:

- **Dlib**

Dlib performs face detection using deep learning through Convolutional Neural Networks. It performs better than HOG based method even on the faces at odd angles. A delicate implementation of the library is well illustrated in [86, 111].

- **Multi-task Cascaded Convolutional Neural Network**

A CNN-based proposed works in three different stages to detect and localize faces and vital facial points. [120]. Besides, [110] conducted facial recognition using MTCNN. In [28], the real-time application of detecting people with or without face masks using the mentioned method is illustrated. Likewise, a detailed study is executed in [50].

- **RetinaFace**

It is a single-stage detector that works on pixel-wise face localization and simultaneously predicts face box, face score, and facial key points. An elaborate discussion is presented in multiple pieces of research [16, 26, 69].

3.2.1 Performance analysis

From the analysis in Fig. 15, it can be observed that all the algorithms perform efficiently on images. However, some studies maintain the poor performance of dlib in scenarios with a lot of faces in it. While analyzing the performance of the different methods, the quality of the image should be considered. Also, the model's accuracy varies with the angle of the face in an image, as studied in [64].

Although the effectiveness of the architecture can be influenced by the size and the quality of the dataset, there are precisely defined parameters used to assess the classification outcomes. Precision and recall are the evaluation metrics to check the performance of the model. Additionally, precision is taken to be the measure of correct positive identifications while recall represents the proportion of correctly classified actual positives. The closer the value of precision and recall is to 1, the more accurate is the used backbone network. From Table 2, Dlib based on ResNet50 has the precision and recall value closest to 1, in comparison to other algorithms, thereby conducive to an effective model.

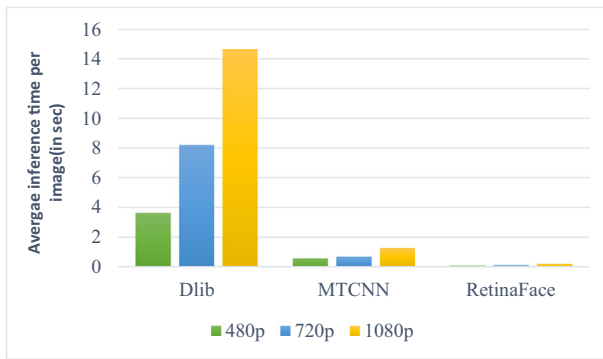


Fig. 15 Face detection speed analysis [64]

After successfully performing face detection, the next step to classify the faces detected is carried off. CLASSIFICATION is considered supervised learning in machine learning [90], which specifies the class label to which the input data belongs. The methods that can be used to perform the same are considered below.

- **Convolutional Neural Network**

In deep learning, a CNN model is usually fed with an image as an input which is then made to pass through multiple layers [3]. To begin with, the input is made to pass through convolutional layers with kernels in succession, followed by a pooling layer. This layer then reduces the number of learning parameters and hence computations by turning down the size of feature maps. It is afterwards carried through fully connected layers, which at the end apply a softmax function that predicts the probabilistic values for each class. The class having the maximum value is then taken to be the class to which the object belongs.

CNN can make use of varied backbone architectures to achieve the task. In [13], the VGG-16 architecture of CNN is discussed. Further, a real-time face mask detector which could be helpful in times like those of COVID-19, is demonstrated in [27]. Besides, [75, 93, 97, 99, 122] analyses the usage of the technique.

- **Support Vector Machines**

It is a method leading to the division of the input data into different classes by making boundaries using hyper-planes. When working on multi-class data, each class is considered to have its binary classifier. [59] describes and exhibits how SVM is used for image classification. It uses SVM on

Table 2 Detection accuracy comparison of algorithms

Model	Face Detection		Mask Detection	
	Precision (%)	Recall (%)	Precision (%)	Recall (%)
Dlib based on ResNet50 [84]	99.20	99.0	98.92	98.24
MTCNN based on MobileNet [44]	94.50	86.38	84.39	80.92
RetinaFace based on MobileNet [44]	83.0	95.60	82.30	89.10
RetinaFace based on ResNet [84]	91.9	96.3	93.4	94.5

several datasets and later even compares the performances on each dataset and with multiple other classifiers. Also, discussion about similar aspects is done in [45, 123].

- **Decision Trees**

It is among the most useful algorithms that are availed to deal with classification problems. It is a flow-chart-like structure where each internal node tests on a feature, and the branch represents the test result while the leaf node represents the decision, i.e., class label [24]. In [73], decision trees and their specific algorithms are reviewed in depth. Correspondingly, [100] talks about work in the same domain.

- **Ensemble**

This type of learning produces an optimal predictive model because it combines several other models. The model works either by bagging or by bootstrap aggregation. [4] reviews about the available hybrid and Ensemble methods in detail. Besides, an assessment of the process is described in [20].

The accuracy comparison chart, as shown in Fig. 16, analyses the result of several algorithms obtained on the Simulated Masked Face Dataset (SMFD) as studied in [67, 75]. Although it can be observed that SVM has achieved the highest possible accuracy, it cannot be neglected that the other components, like the selection of hyperparameters, play a crucial role while deciding the feasibility of an algorithm. The amalgamation of architecture, dataset, pre-processing, and requirement of the problem statement result in selecting the technique to be used.

4 Analysis of real-time techniques

The comparison of different contemporary real timedetection techniques has been shown in Table 3.

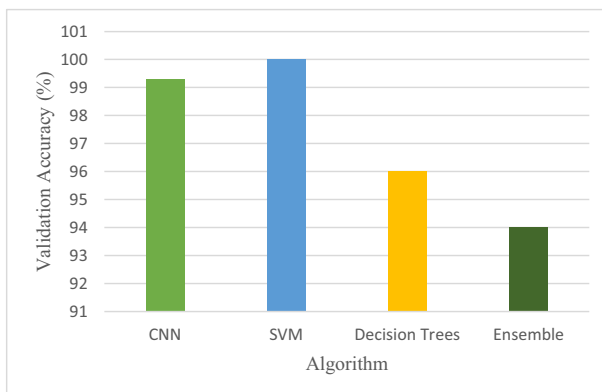


Fig. 16 Comparison of various classification algorithms

Table 3 Detection accuracy comparison of algorithms

Real-time technique	Algorithm	Runtime	Theoretical Significance
Face Mask Detection, Su et al. [92]	YOLOv3	14.62 fps	A novel fusion transfer learning in the amalgamation of YOLOv3 has been proposed for face mask detection, with EfficientNet being the backbone architecture for feature extraction.
Real-time Face Mask Detection, Gadge et al. [62]	YOLOv4	49.5 fps	A face mask detector has been fabricated using deep learning model which works with real-time streams. The training is performed over several thousand epochs to achieve notable results in real-time.
Real-Time AI based Face Mask Detection, Teboulbi et al. [97]	Neural Networks	–	Multiple neural network architectures, after evaluation, have been used with Raspberry Pi and webcam to enforce effective real-time methods for face mask detection in public places.
Real-Time Face Mask Detection in Video, Ding et al. [19]	YOLOv5	52 fps	Deep Learning techniques are established to be producing improved detection results in real-time. Also, the availability of enriched dataset assists in improving the efficiency of the model.
Personal Protective Equipment Detection, Nath et al. [66]	YOLOv3	11 fps	A deep learning model built to verify whether construction workers wear personal protective equipment, including hard hats, vests, etc., in real-time. Also, different approaches have been compared to execute the task.
Front Vehicle Detection, Cao et al. [10]	SSD	66.6 fps	A vehicle detection model is proposed for intelligent cars having an improvised SSD architecture. Also, the system has shown to be robust in excessive traffic.
Illegal Parking Detection, Tang et al. [96]	SSD	40 fps	A deep learning-based model with an improved SSD has been proposed to detect illegal vehicle parking. The system has achieved a precision of 97.3%.
Vehicle Type Recognition, Kim et al. [47]	Faster R-CNN YOLOv4 SSD	36.32 fps 82.1 fps 105.14 fps	The study aims to analyze several techniques for vehicle type recognition in real-time. The analysis has maintained the performance of YOLOv4 to be the best.
Social Distancing monitoring, Jindal et al. [61]	YOLOv3	21fps	The proposed study uses YOLOv3 to implement a bird-eye view based social distancing monitoring system. Also, by showcasing red boxes, an alert is shown if social distancing is violated.

5 Face mask detection techniques analysis

In terms of the approach being used for the execution of tasks, Fig. 17 demonstrates the growth of each method since the year 2011. It can be observed that deep learning has gained much attention freshly. Also, the data has been collected using Semantic Scholar.

Figure 18 shows the comparative percentage usage of reviewed techniques in articles available on different e-sources from 2000 till 2021(April). The articles in Fig. 17a have been selected from Semantic Scholar using the keywords “technique name” + Face Mask Detection. Further, Fig. 17b depicts the articles chosen from Semantic Scholar using “technique

Table 4 List of advantages and disadvantages of some of object detection algorithms (deep learning approach)

Name	Faster R-CNN [91]	R-FCN [15]	YOLO [62]	SSD [56]
Advantages	Performs well on small object detection because of the presence of nine anchors in a single grid	The speed is faster in comparison to that of other region-based CNN	Objects are detected efficiently and with great speed, hence finding application in real-time	Faster in object localization than Faster R-CNN because of one step architecture
Challenges	The training time is more and, despite efficiency, fails to perform real-time detection because of two-step architecture	The mAP of R-FCN is appreciable but lesser than R-CNN	Difficulty in detecting small objects and close objects	More training data is required and relatively less accuracy

Table 5 List of advantages and disadvantages of various face detection methods

Name	Dlib [9]	MTCNN [118]	RetinaFace [112]
Advantages	Along with an easy training process, it can work well with different face orientations and that too at a fast speed	Efficient detection even when faces are not aligned initially because of jointly performing face detection and alignment	Can generate an accurate rectangular face bounding box together with five points facial landmark
Challenges	Fails in detecting small faces as it is trained for a minimum 80×80 face size	It is slow as compared to other models and also does not perform well with small faces	It tends to fail when input images contain a large face

Table 6 List of advantages and disadvantages of various classifiers

Name	CNN [70]	SVM [89]	Decision Trees [24]	Ensemble [4, 100]
Advantages	High accuracy while performing image recognition tasks and automatically detects crucial features without any intervention	Requires few input data in addition to being highly efficient in high-dimensional spaces. Also, it is memory efficient	In comparison to other models, has fewer requirements for data pre-processing and performs well for categorical data	Capable of making better predictions and achieving preferable performance than any other single model.
Challenges	Many input data required along with difficulty in classifying images that are variant to input data	Poor performance in overlapping target classes and, in cases where several features of data points exceed the number of training data samples	Often higher training time and a small change in data leads to huge structure change which results in instability	Usually computationally expensive which results in learning time and memory constraints

Table 7 List of different datasets available on online platforms for the study

NAME	DESCRIPTION	SOURCE(URLs)
YOLO Medical Mask Dataset Dataset by Prajna Bhandari at PyImagesearch	Contains ~631 images of people with medical mask Contains ~1376 images including 690 with mask images and 686 without a mask	https://www.kaggle.com/googryolo-medical-mask-dataset https://github.com/prajnasb/observations
Real-World Masked Face Dataset (RMFD)	Contains ~5000 masked face images of 525 people and ~90,000 normal face	https://github.com/X-zhangyang/Real-World-Masked-Face-Dataset
LFW- Simulated Masked face	The training dataset contains ~13,027 masked faces images of 5713 people, and the testing dataset contains ~70 masked faces of 48 people	https://www.kaggle.com/muhammeddalkran/lfw-simulated-masked-face-dataset
Face Mask Detection Dataset	Contains ~3725 images of faces with masks and ~ 3828 images of faces without masks	https://www.kaggle.com/omkargurav/face-mask-dataset

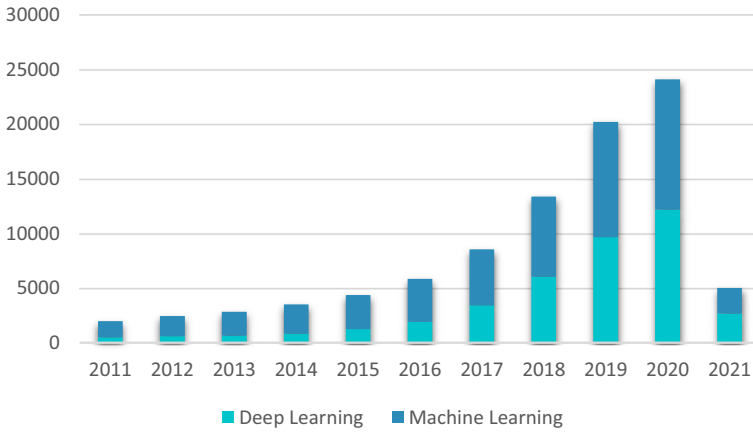


Fig. 17 Deep Learning and Machine Learning usage trends over the years 2011–2021(till April’2021)

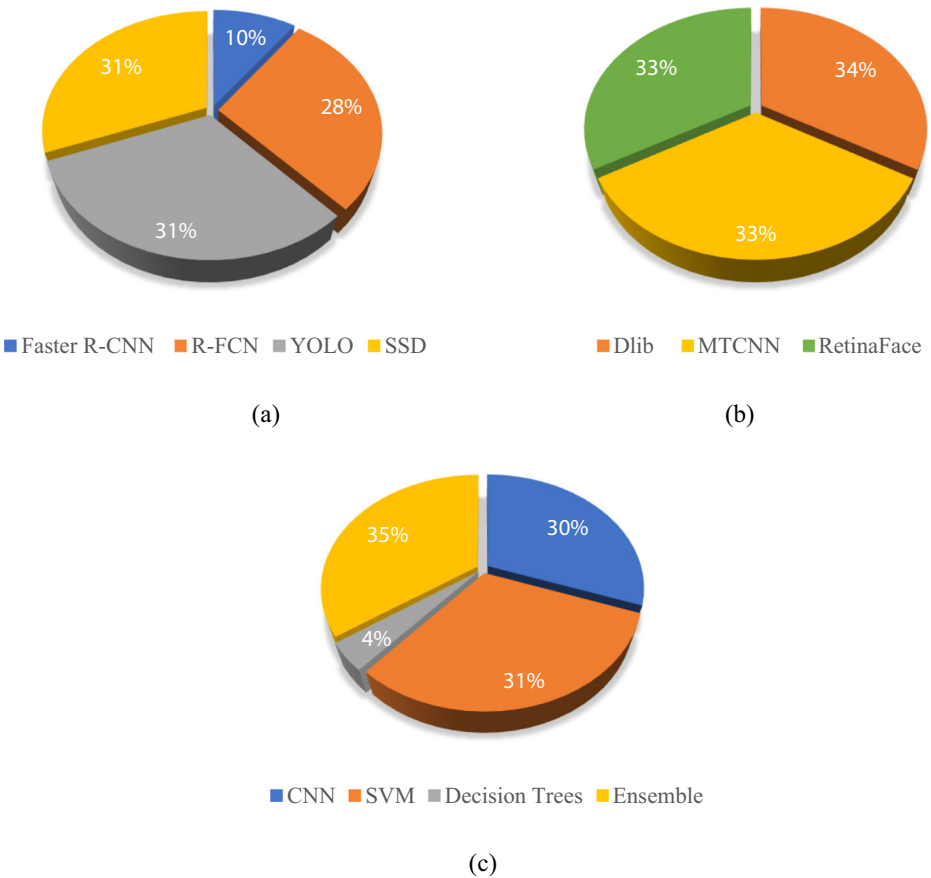


Fig. 18 Records of the different (a) Object detection (b) face detection (c) classification techniques analyzed over the year 2000–2021(April) using Semantic Scholar

name” + Face Detection. Whereas for Fig. 17c, keywords “technique name” + Classification were utilized.

5.1 Popular techniques with advantages and challenges

A single algorithm cannot suffice for all the needs. The choice of the algorithm relies on many factors. The specific parameters that rule the decision-making include the size of training data, speed, accuracy, training time, number of features, etc. None of the models can be declared best among the counters, but a comparison can be put together to help in the choosing process [101, 105] (Tables 4, 5 and 6).

6 Dataset

It is a collection of instances used to train models for learning. It can either be created by scraping from the internet or accessing various online websites [107]. Few of the sources that are currently available on different sources are shared in this article (Table 7).

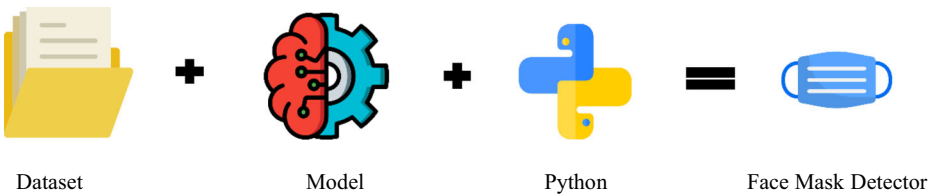


Fig. 19 Requirements for face mask detection system [36]

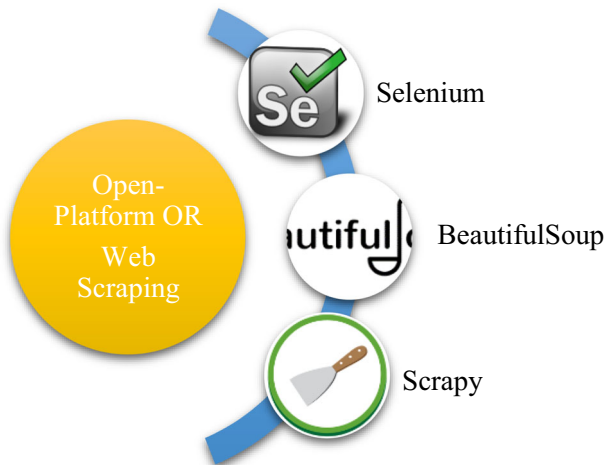
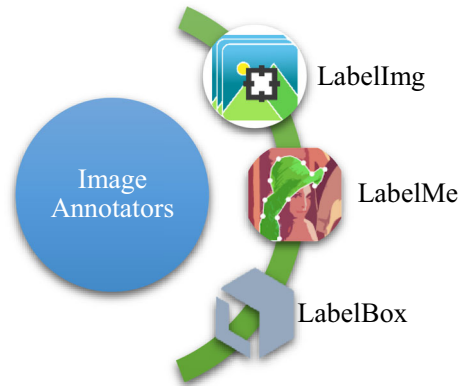


Fig. 20 List of tools used in data collection [36]

Fig. 21 List of tools used in Image Annotation [36] *The dataset could be enriched by deploying techniques like data augmentation



7 Several supporting software

These days, there is a plethora of programming languages, programming tools, libraries, and frameworks to choose from while working on a project. Further, there are no stringent rules to choose from numerous such sources. Nonetheless, the article lists specific tools that could be useful in a study.

The basic requirements for completing a face mask detector project are illustrated below (Fig. 19)

7.1 Dataset

A rich and relevant dataset can be accessed using the below-mentioned methods:

7.1.1 Data collection

It involves accumulation of content pertinent to the problem situation at hand. It is usually performed in accordance with the task to be executed. There are various methods available that

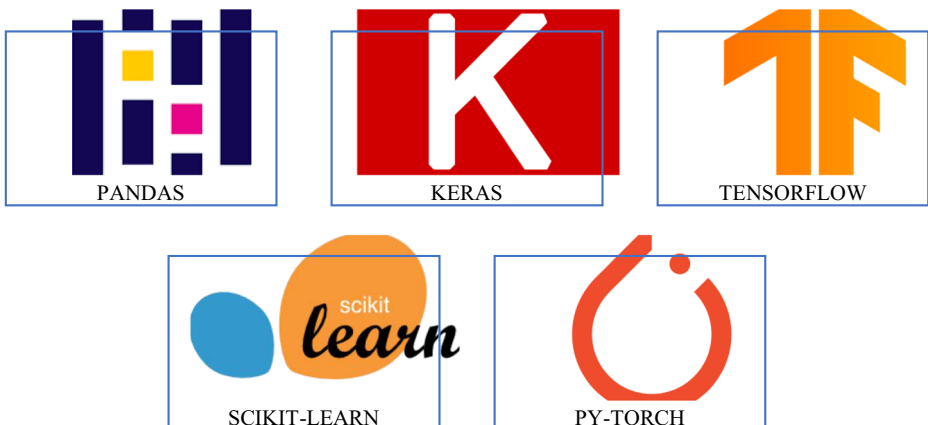


Fig. 22 List of useful Python libraries [18]

could be used to prepare one's dataset. Some of the tools that could be used for the purpose are shown below (Fig. 20).

7.1.2 Annotating image

One of the essential steps while dealing with the image dataset is to annotate it. It refers to labeling images to be later utilized in the machine learning model. Lately, various approaches are viable to execute the same. Some of them are (Fig. 21):

7.2 Model

Below are libraries and frameworks typical of the different implementation techniques mentioned above. They can be installed under the demand of the action, and the model used. To add, the desired file can be imported from the concerned library (Fig. 22).

Since, while working with model creation, open-source libraries and frameworks play a significant role. Figure 23 reviews the ranking of the numerous libraries consistent with the GitHub star count as reported by the official documentation of the respective library on PyPi till April 2021. The assessment could be helpful for the uninitiated to begin working with such user-friendly libraries.

7.3 Python

Some of the other useful open source libraries that can be amalgamated with the essential packages are talked through in this section (Fig. 24).

8 Applications, limitations, and observations

Certain areas where face mask detection can be effectively employed are discussed below.

- **Transit hubs**

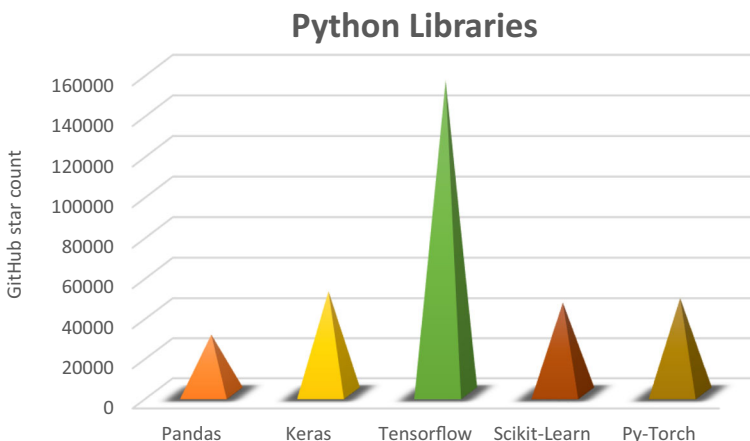


Fig. 23 Popularity of useful python libraries based on statistics of GitHub stars (till April' 2021)



Fig. 24 List of other supporting libraries [18]

At places like airports, railway stations, etc., face mask detectors, integrated with security cameras, can be implemented to keep a check on travellers wearing face masks or not. The passenger's face could be detected throughout the premises, and the authorities could be informed immediately if any violation is detected.

- **Workplaces**

A mechanism to observe if an employee has worn a face mask or not could be incorporated in an office. A warning message could be sent to people who are not following the safety precautions. Also, a daily record of people not complying with the regulations could be maintained.

- **Healthcare centres**

In various healthcare organizations and hospitals, a face mask detection system could track health workers wearing face masks during their shifts. Besides, it could be helpful in alerting the visitors entering the site without face masks. The officials could be immediately informed in case of defiance.

- **Surveillance systems**

Utilizing face mask detection systems unified with surveillance cameras can help strictly track people wearing face masks or not in public areas.

8.1 Limitations

Although the system performs efficiently in real-time, it faces the following challenges.

- Although different network architecture performs better in mask detection tasks, the model suffers limitations due to large dataset performance [65].
- The irregularities in images, like those with insufficient light and side angle, need proper attention [116].
- Also, another major challenge is to achieve high accuracy in the least possible time [97].
- Additionally, the video analysis has difficulties, including motion blur, transitioning between frames, etc. [64].

8.2 Observations

- Although two-stage detectors excel in accuracy, one-stage detectors outperform them when used for real-time requirements. Hence, for real-time video feed detection use of algorithms like YOLO, SSD is appreciable.
- Since training a deep neural network is expensive as it involves high computational complexity, transfer learning, i.e., utilizing pre-trained models like MobileNet, VGG-16, etc., is recommended.
- Owing to the exceptional results that deep learning models produce, they have become the choice of various practitioners. Though they perform efficiently with high accuracy, applying disparate backbone architectures with different hyperparameters could result in even better accuracy.
- Also, poor images, like insufficient light, side angle, etc., in the dataset have affected the performance of the model. Hence, the dataset's quality could be improved further for future use.
- Though there have been many studies and research work dedicated to COVID-19 these days, there is still a scope for a lot more analysis that could be done in the healthcare domain.

After reviewing many studies, it can be inferred that despite the variety of techniques being available to implement the model, one-stage object detectors are the preferred choice for real-time requirements. The accuracy with which it works in real-time makes the application possible. Also, because of the computational costs, drawbacks could be dealt with by altering the architectures, hyperparameters, input size, etc.

9 Conclusion and future directions

To deal with the pandemic more effectively, developing central systems capable of automatically detecting whether a person is wearing a face mask or not has become an engaging topic for people working in this sphere. A countless number of researches have been initiated lately in this domain. However, this paper aims to provide a detailed review of the various ways that could be opted for executing such an advanced system. After inspecting all the implementation techniques, it could be safely stated that deep learning has become popular among researchers in recent times. The efficiency of the approach makes it suitable for use in such tasks. Additionally, despite many datasets being available, the RMFD dataset is widely used. If used constructively, the deployment of the model could be beneficial in public areas. The proposed

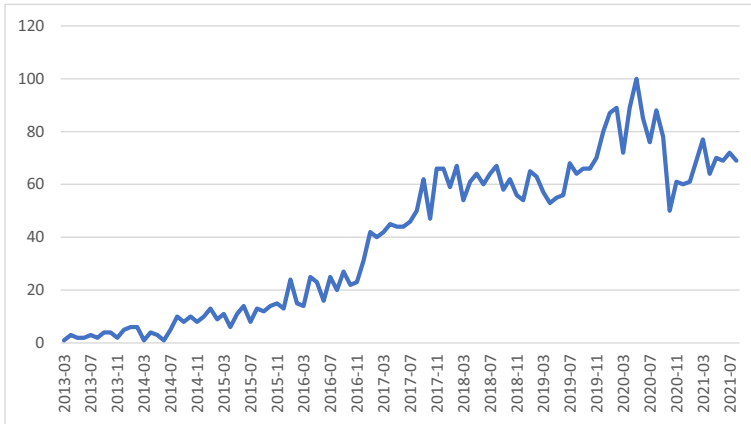


Fig. 25 Upsurge of deep learning from March 2013 to August 2021 (Created by Google Trends)

system could be upgraded for future works by integrating them with automated thermal detection systems. Also, a check on whether social distancing is being practised in crowded areas could be an add-on to the system. A feature of facial landmark detection could be added for biometric purposes. Moreover, owing to the versatility of the state-of-art techniques, their architectures could be enhanced to achieve better results at a faster speed. As shown in Fig. 25, there has been an upsurge in the usage of deep learning methods. Taking advantage of the enormous utility of these methods, various future studies could be executed in this domain. The quality of datasets could be improved by removing images with insufficient light. Nonetheless, the system could be integrated with a model to check if sufficient physical distance is being maintained between people. It could also be blended with a design that detects the mask type of a person. Besides, new feature extraction techniques could be explored using machine learning algorithms.

Code availability Not applicable.

Data Availability Not applicable.

Declaration

Conflict of interest Not applicable.

References

1. Abdallah ZS, Du L, Webb GI (2017) Data preparation. In: Sammut C, Webb GI (eds) Encyclopedia of machine learning and data mining. Springer
2. Adarsh P, Rathi P, Kumar M (2020) YOLO v3-tiny: object detection and recognition using one stage improved model. In: 2020 6th international conference on advanced computing and communication systems (ICACCS). <https://doi.org/10.1109/ICACCS48705.2020.9074315>

3. Albawi S, Mohammed TA, Al-Zawi S (2017) Understanding of a convolutional neural network. In: 2017 international conference on engineering and technology (ICET). <https://doi.org/10.1109/ICEngTechnol.2017.8308186>
4. Ardabili S, Mosavi A, Várkonyi-Kóczy AR (2020) Advances in machine learning modeling reviewing hybrid and ensemble methods. In: Várkonyi-Kóczy A (ed) Engineering for sustainable future. INTER-ACADEMIA 2019. Lecture notes in networks and systems, vol 101. Springer. https://doi.org/10.1007/978-3-030-36841-8_21
5. Basha Z, Pravallika BNL, Shankar EB (2021) An efficient face mask detector with PyTorch and deep learning. EAI Endorsed Trans. Pervasive Health and Technol. <https://doi.org/10.4108/cai.8-1-2021.167843>
6. Benjdira B, Khursheed T, Koubaa A, Ammar A, Ouni K (2019) Car detection using unmanned aerial vehicles: comparison between faster R-CNN and YOLOv3. In: 2019 1st international conference on unmanned vehicle systems-Oman (UVS). <https://doi.org/10.1109/UVS.2019.8658300>
7. Bhadani AK, Sinha A (2020) A facemask detector using machine learning and image processing techniques. Engineering Science and Technology an International Journal
8. Bhambani K, Jain T, Sultanpure KA (2020) Real-time face mask and social distancing violation detection system using YOLO. In: 2020 IEEE Bangalore humanitarian technology conference (B-HTC). <https://doi.org/10.1109/B-HTC50970.2020.9297902>
9. Boyko N, Basytiuk O, Shakhovska N (2018) Performance evaluation and comparison of software for face recognition, based on Dlib and OpenCV library. In: 2018 IEEE second international conference on Data Stream Mining & Processing (DSMP). <https://doi.org/10.1109/DSMP.2018.8478556>
10. Cao J, Song C, Song S, Peng S, Wang D, Shao Y, Xiao F (2020) Front vehicle detection algorithm for smart Car based on improved SSD model. Sensors 20(16):4646. <https://doi.org/10.3390/s20164646>
11. Cao Z, Shao M, Xu L, Mu S, Qu H (2021) MaskHunter: real-time object detection of face masks during the COVID-19 pandemic. IET Image Process 14(16):4359–4367. <https://doi.org/10.1049/iet-ipr.2020.1119>
12. Chandan G, Jain A, Jain H, Mohana (2018) Real time object detection and tracking using deep learning and OpenCV. In: 2018 international conference on inventive research in computing applications (ICIRCA). <https://doi.org/10.1109/ICIRCA.2018.8597266>
13. Chavda C, Dsouza J, Badgujar S, Damani A (2020) Multi-stage CNN architecture for face mask detection. arXiv:2009.07627
14. Chen Y, Li W, Sakaridis C, Dai D, Gool LV (2018) Domain adaptive faster R-CNN for object detection in the wild. In: 2018 IEEE/CVF conference on computer vision and pattern recognition, pp 3339–3348. <https://doi.org/10.1109/CVPR.2018.00352>
15. Dai J, Li Y, He K, Sun J (2016) R-FCN: object detection via region-based fully convolutional networks. arXiv:1605.06409
16. Deng J, Guo J, Ververas E, Kotsia I, Zafeiriou S (2020) RetinaFace: single-shot multi-level face localisation in the wild. In: 2020 IEEE/CVF conference on computer vision and pattern recognition (CVPR). <https://doi.org/10.1109/CVPR42600.2020.00525>
17. Deng H, Zhang J, Chen L, Cai M (2021) Improved mask wearing detection algorithm for SSD. Phys: Conf Ser 1757(1):012140. <https://doi.org/10.1088/1742-6596/1757/1/012140>
18. Dewani R (2020) Top 13 Python libraries every data science aspirant must know! (and their resources). Analytics Vidya <https://www.analyticsvidhya.com/blog/2020/11/top-13-python-libraries-every-data-science-aspirant-must-know-and-their-resources/>. Accessed 18 November 2020
19. Ding Y, Li Z, Yastremsky D (2021) Real-time face mask detection in video data. arXiv:2105.01816
20. Dong X, Yu Z, Cao W, Shi Y, Ma Q (2020) A survey on ensemble learning. Front Comput Sci 14:241–258. <https://doi.org/10.1007/s11704-019-8208-z>
21. Draughon GTS, Sun P, Lynch JP (2020) Implementation of a computer vision framework for tracking and visualizing face mask usage in urban environments. In: Proceedings of IEEE international smart cities conference (ISC2), pp 1–8. <https://doi.org/10.1109/ISC251055.2020.9239012>
22. Eggert C, Brehm S, Winschel A, Zecha D, Lienhart R (2017) A closer look: small object detection in faster R-CNN. In: 2017 IEEE international conference on multimedia and expo (ICME). <https://doi.org/10.1109/ICME.2017.8019550>
23. Fu CY, Liu W, Ranga A, Tyagi A, Berg A (2017) DSSD: Deconvolutional single shot detector. Computer vision and pattern recognition. arXiv:1701.06659
24. Ghasemi E, Gholizadeh H, Adoko AC (2020) Evaluation of rockburst occurrence and intensity in underground structures using decision tree approach. Eng Comput 36(2):213–225. <https://doi.org/10.1007/s00366-018-00695-9>
25. Girshick R (2015) Fast R-CNN. In: 2015 IEEE International Conference on Computer Vision (ICCV). <https://doi.org/10.1109/ICCV.2015.169>

26. Guo X, Nie J (2020) Face recognition system for complex surveillance scenarios. *J Phys Conf Ser*. <https://doi.org/10.1088/1742-6596/1544/1/012146>
27. Guo G, Wang H, Yan Y, Zheng J, Li B (2018) A fast face detection method via convolutional neural network. *Neurocomputing* 395:128–137. <https://doi.org/10.1016/j.neucom.2018.02.110>
28. Gupta V, Rajput R (2021) Face mask detection using MTCNN and MobileNetV2. *International Research Journal Of Engineering and Technology (IRJET)* 8(3)
29. Gupta S, Sreenivasu SVN, Chouhan K, Shrivastava A, Sahu B, Potdar RM (2021) Novel face mask detection technique using machine learning to control COVID-19 pandemic. *Materials Today: Proceedings*. <https://doi.org/10.1016/j.matpr.2021.07.368>
30. Hammoudi K, Cabani A, Benhabiles H, Melkemi M (2020) Validating the correct wearing of protection mask by taking a selfie: design of a mobile application "CheckYourMask" to limit the spread of COVID-19. *CMES Comput Model Eng Sci* 3(124):1049–1059. <https://doi.org/10.36227/techrxiv.12355970>
31. Henderi Rafika AS, Warnars HLH, Saputra MA (2020) An application of mask detector for prevent Covid-19 in public services area. *J Phys Conf Ser*. <https://doi.org/10.1088/1742-6596/1641/1/012063>
32. <https://images.app.goo.gl/6UCnpsfXmUSgaic6>
33. <https://images.app.goo.gl/b3SnApDsmFxBL7g6>
34. <https://images.app.goo.gl/gw6qLXsrLsNJ3KdMA>
35. <https://images.app.goo.gl/wY5Hbud3o8QT9D2R8>
36. <https://www.google.com/imghp?hl=EN>
37. Huang J, Rathod V, Sun C, Zhu M, Korattikara A, Fathi A, Fischer I, Wojna Z, Song Y, Guadarrama S, Murphy K (2017) Speed/accuracy trade-offs for modern convolutional object detectors. In: 2017 IEEE conference on computer vision and pattern recognition (CVPR)
38. Huang R, Pedoem J, Chen C (2018) YOLO-LITE: a real-time object detection algorithm optimized for non-GPU computers. In: 2018 IEEE international conference on big data (big data), Seattle, WA, USA, pp 2503–2510. <https://doi.org/10.1109/BigData.2018.8621865>
39. Hung J, Goodman A, Lopes S, Rangel G, Ravel D, Costa F, Duraisingh M, Marti M, Carpenter A (2018) Applying faster R-CNN for object detection on malaria images. arXiv:1804.09548
40. Interim Infection Prevention and Control Recommendations for Patients with Suspected or Confirmed Coronavirus Disease 2019 (COVID-19) in Healthcare Settings (2021). <https://www.cdc.gov/coronavirus/2019-ncov/hcp/infection-control-recommendations.html>. Accessed 10 February 2021
41. Jha S, Seo C, Yang E, Josh GP (2021) Real time object detection and tracking system for video surveillance system. *Multimed Tools Appl* 80:3981–3996. <https://doi.org/10.1007/s11042-020-09749-x>
42. Jiang X, Hadid A, Pang Y, Granger E, Xiaoyi F (2019) Deep learning in object detection and recognition. <https://doi.org/10.1007/978-981-10-5152-4>
43. Jiang X, Gao T, Zhu Z, Zhao Y (2021) Real-time face mask detection method based on YOLOv3. *Electronics* 10(7):837. <https://doi.org/10.3390/electronics10070837>
44. Joshi AS, Joshi SS, Kanahasabai G, Kapil R, Gupta S (2020) Deep learning framework to detect face masks from video footage. In: 2020 12th international conference on computational intelligence and communication networks (CICN), pp 435–440. <https://doi.org/10.1109/CICN49253.2020.9242625>
45. Kar NB, Babu KS, Sangaiah AK, Bakshait S (2019) Face expression recognition system based on ripplelet transform type II and least square SVM. *Multimed Tools Appl* 78:4789–4812. <https://doi.org/10.1007/s11042-017-5485-0>
46. Kim DY, Park JH, Lee Y, Kim S (2020) Network virtualization for real-time processing of object detection using deep learning. *Multimed Tools Appl*. <https://doi.org/10.1007/s11042-020-09603-0>
47. Kim JA, Sung JY, Park SH (2020) Comparison of faster-RCNN, YOLO, and SSD for real-time vehicle type recognition. In: 2020 IEEE international conference on consumer electronics - Asia (ICCE-Asia), pp 1–4. <https://doi.org/10.1109/ICCE-Asia49877.2020.9277040>
48. Kong X, Wang K, Wang S, Wang X, Jiang X, Guo Y, Shen G, Chen X, Ni Q (2021) Real-time mask identification for COVID-19: an edge computing-based deep learning framework. *IEEE Internet Things J*. <https://doi.org/10.1109/JIOT.2021.3051844>
49. Kos M, Kristo M, Pobar M (2019) Human detection in thermal imaging using YOLO. In: Proceedings of the 2019 5th international conference on computer and technology applications. <https://doi.org/10.1145/3323933.3324076>
50. Ku H, Dong W (2020) Face recognition based on MTCNN and convolutional neural network. *Frontiers in Signal Processing*. <https://doi.org/10.22606/fsp.2020.41006>
51. Kumar A, Kaur A, Kumar M (2019) Face detection techniques: a review. *Artif Intell Rev* 52:927–948. <https://doi.org/10.1007/s10462-018-9650-2>
52. Lan W, Dang J, Wang Y, Wang S (2018) Pedestrian detection based on YOLO network model. In: 2018 IEEE international conference on mechatronics and automation (ICMA). <https://doi.org/10.1109/ICMA.2018.8484698>

53. Li Z, Zhou F (2017) FSSD: feature fusion single shot multibox detector. arXiv:1712.00960
54. Lin J, Jiang F, Shen R (2018) Hand-raising gesture detection in real classroom. In: 2018 IEEE international conference on acoustics, speech and signal processing (ICASSP). <https://doi.org/10.1109/ICASSP.2018.8461733>
55. Liu R, Ren Z (2020) Application of Yolo on mask detection task. In: 2021 IEEE 13th international conference on computer research and development (ICCRD). <https://doi.org/10.1109/ICCRD51685.2021.9386366>
56. Liu W, Anguelov D, Szegedy C, Reed S, Fu CY, Berg AC (2016) SSD: single shot MultiBox detector. In: Leibe B, Matas J, Sebe N, Welling M (eds) Computer vision – ECCV 2016. ECCV 2016. Lecture notes in computer science, vol 9905. Springer. https://doi.org/10.1007/978-3-319-46448-0_2
57. Liu G, Nouaze JC, Mbouembe PLT, Kim JHK (2020) YOLO-tomato: a robust algorithm for tomato detection based on YOLOv3. *Sensors* 20(7):2145. <https://doi.org/10.3390/s20072145>
58. Loey M, Manogaran G, Taha MHN, Khalifa NEM (2021) Fighting against COVID-19: a novel deep learning model based on YOLO-v2 with ResNet-50 for medical face mask detection. *Sustain Cities Soc* 65:102600. <https://doi.org/10.1016/j.scs.2020.102600>
59. Loey M, Manogaran G, Taha MHN, Khalifa NEM (2021) A hybrid deep transfer learning model with machine learning methods for face mask detection in the era of the COVID-19 pandemic. *Measurement* 167:108288. <https://doi.org/10.1016/j.measurement.2020.108288>
60. Lynteris C (2018) Plague masks: the visual emergence of anti-epidemic personal protection equipment. *Med Anthropol* 37(6):442–257. <https://doi.org/10.1080/01459740.2017.1423072>
61. Magoo R, Singh H, Jindal N, Hooda N, Rana PS (2021) Deep learning-based bird eye view social distancing monitoring using surveillance video for curbing the COVID-19 spread. *Neural Comput & Applic.* <https://doi.org/10.1007/s00521-021-06201-5>
62. Mahurkar RR, Gadge NG (2021) Real-time Covid-19 face mask detection with YOLOv4. 2021 second international conference on electronics and sustainable communication systems (ICESC). <https://doi.org/10.1109/ICESC51422.2021.9533008>
63. Meng R, Rice SG, Wang J, Sun X (2018) A fusion steganographic algorithm based on faster R-CNN. *Computers, Materials, and Continua* 55(1):1–16. <https://doi.org/10.3970/cmc.2018.055.001>
64. Militante SV, Dionisio NV (2020) Real-time facemask recognition with alarm system using deep learning. In: 2020 11th IEEE control and system graduate research colloquium (ICSGRC). <https://doi.org/10.1109/ICSGRC49013.2020.9232610>
65. Nagrath P, Jain R, Madan A, Arora R, Kataria P, Hemanth J (2021) SSDMNV2: a real time DNN-based face mask detection system using single shot multibox detector and MobileNetV2. *Sustain Cities Soc* 66: 102692. <https://doi.org/10.1016/j.scs.2020.102692>
66. Nath ND, Behzadan AH, Paal SG (2020) Deep learning for site safety: real-time detection of personal protective equipment. *Autom Constr* 112. <https://doi.org/10.1016/j.autcon.2020.103085>
67. Negi A, Kumar K, Chauhan P, Rajput RS (2021) Deep neural architecture for face mask detection on simulated masked face dataset against Covid-19 pandemic. In: 2021 international conference on computing, communication, and intelligent systems (ICCCIS), pp 595–600. <https://doi.org/10.1109/ICCCIS51004.2021.9397196>
68. Ning C, Zhou H, Song Y, Tang J (2017) Inception single shot MultiBox detector for object detection. In: 2017 IEEE International Conference on Multimedia & Expo Workshops (ICMEW). <https://doi.org/10.1109/ICMEW.2017.8026312>
69. Oh C, Park G, Kim S, Kim D, Yi Y (2020) Towards real-time CNN inference from a video stream on a Mobile GPU (WiP Paper). In: The 21st ACM SIGPLAN/SIGBED conference on languages, compilers, and tools for embedded systems. <https://doi.org/10.1145/3372799.3394366>
70. O'Shea K, Nash R (2015) An introduction to convolutional neural networks. arXiv:1511.08458
71. Oumina A, Makhfi NE, Hamdi M (2020) Control the COVID-19 pandemic: face mask detection using transfer learning. In: IEEE international conference on electronics, control, optimization and computer science (ICECOCs), pp 1–5. <https://doi.org/10.1109/ICECOCs50124.2020.9314511>
72. Pandiyan P (2020) Social distance monitoring and face mask detection using deep neural network.
73. Patel HH, Prajapati P (2018) Study and analysis of decision tree based classification algorithms. *Int J Comput Sci Eng* 6(10):74–78. <https://doi.org/10.26438/ijcse/v6i10.7478>
74. Pooja S, Preeti S (2021) Face mask detection using AI. In: Khosla PK, Mittal M, Sharma D, Goyal LM (eds) Predictive and preventive measures for Covid-19 pandemic. Algorithms for intelligent systems. Springer, pp 293–305. https://doi.org/10.1007/978-981-33-4236-1_16
75. Qi R, Jia RS, Mao QC, Sun HM, Zuo LQ (2019) Face detection method based on cascaded convolutional networks. *IEEE Access* 7:110740–110748. <https://doi.org/10.1109/ACCESS.2019.2934563>
76. Qin B, Li D (2020) Identifying facemask-wearing condition using image super-resolution with classification network to prevent COVID-19. *Sensors* 20(18):5236. <https://doi.org/10.3390/s20185236>

77. Rahman MM, Manik MH, Islam MM, Mahmud S, Kim JH (2020) An automated system to limit COVID-19 using facial mask detection in Smart City network. In: 2020 IEEE international IoT, electronics and mechatronics conference (IEMTRONICS), pp 1–5. <https://doi.org/10.1109/IEMTRONICSS51293.2020.9216386>
78. Ramos L, Morales B (2020) SwiftFace: real-time face detection. arXiv:2009.13743
79. Razavi M, Alikhani H, Janfaza V, Sadeghi B, Alikhani E (2021) An automatic system to monitor the physical distance and face mask wearing of construction workers in COVID-19 pandemic. arXiv: 2101.01373
80. Redmon J, Divvala S, Girshick R, Farhadi A (2016) You only look once: unified, real-time object detection. In: 2016 IEEE conference on computer vision and pattern recognition (CVPR). <https://doi.org/10.1109/CVPR.2016.91>
81. Reka S, Grover A, Arora A, Chugh R (2019) Real time multi object detection for blind using single shot multibox detector. Wirel Pers Commun. <https://doi.org/10.1007/s11277-019-06294-1>
82. Ren S, He K, Girshick R, Sun J (2015) Faster R-CNN: towards real-time object detection with region proposal networks. IEEE Trans Pattern Anal Mach Intell 39(6):1137–1149. <https://doi.org/10.1109/TPAMI.2016.2577031>
83. Rodriguez CR, Luque D, Rosa CL, Esenarro D, Pandey B (2020) Deep learning applied to capacity control in commercial establishments in times of COVID-19. In: 2020 12th international conference on computational intelligence and communication networks (CICN). <https://doi.org/10.1109/CICN49253.2020.9242584>
84. Sethi S, Kathuria M, Kaushik T (2021) A real-time integrated face mask detector to curtail speed of coronavirus. Comput Model Eng Sci 127(2):389–409. <https://doi.org/10.32604/cmescs.2021.014478>
85. Shafiee MJ, Chywl B, Li F, Wong A (2017) Fast YOLO: a fast you only look once system for real-time embedded object detection in video. Journal of Computational Vision and Imaging Systems 3(1). <https://doi.org/10.15353/vsnl.v3i1.171>
86. Sharma S, Shanmugasundaram K, Ramasamy SK (2016) FAREC — CNN based efficient face recognition technique using Dlib. In: 2016 international conference on advanced communication control and computing technologies (ICACCCT). <https://doi.org/10.1109/ICACCCT.2016.7831628>
87. Singh S, Ahuja U, Kumar M, Kumar K, Sachdeva S (2021) Face mask detection using YOLOv3 and faster R-CNN models: COVID-19 environment. Multimed Tools Appl 80:19753–19768. <https://doi.org/10.1007/s11042-021-10711-8>
88. Somaldo P, Ferdiansyah FA, Jati G, Jatmiko W (2020) Developing smart COVID-19 social distancing surveillance drone using YOLO implemented in robot operating system simulation environment. In: 2020 IEEE 8th R10 humanitarian technology conference (R10-HTC). <https://doi.org/10.1109/R10-HTC49770.2020.9357040>
89. Somvanshi M, Chavan P, Tambade S, Shinde SV (2016) A review of machine learning techniques using decision tree and support vector machine. In: 2016 international conference on computing communication control and automation (ICCUBEA). <https://doi.org/10.1109/ICCUBEA.2016.7860040>
90. Soofi AA, Awan A (2017) Classification techniques in machine learning: application and issues. J Basic Appl Sc. <https://doi.org/10.6000/1927-5129.2017.13.76>
91. Soviany P, Ionescu RT (2018) Optimizing the trade-off between single-stage and two-stage deep object detectors using image difficulty prediction. In: 2018 20th international symposium on symbolic and numeric algorithms for scientific computing (SYNASC). <https://doi.org/10.1109/SYNASC.2018.00041>
92. Su X, Gao M, Ren J, Li Y, Liu X (2021) Face mask detection and classification via deep transfer learning. Multimed Tools Appl. <https://doi.org/10.1007/s11042-021-11772-5>
93. Sufian A, Ghosh A, Sultana F, Chakrabarti A, De D (2020) Fundamental concepts of convolutional neural network. In: Balas V, Kumar R, Srivastava R (eds) Recent trends and advances in artificial intelligence and internet of things. Intelligent systems reference library, vol 172. Springer. https://doi.org/10.1007/978-3-030-32644-9_36
94. Sun X, Wu P, Hoi SCH (2017) Face detection using deep learning: an improved faster RCNN approach. Neurocomputing 299:42–50. <https://doi.org/10.1016/j.neucom.2018.03.030>
95. Susanto S, Putra FA, Analia R, Suciningtyas IKLN (2020) The face mask detection for preventing the spread of COVID-19 at Politeknik Negeri Batam. In: 2020 3rd International Conference on Applied Engineering (ICAE), pp 1–5. <https://doi.org/10.1109/ICAE50557.2020.9350556>
96. Tang H, Peng A, Zhang D, Liu T, Ouyang J (2019) SSD real-time illegal parking detection based on contextual information transmission. Computers, Materials & Continua 61(3):293–307.
97. Teboulbi S, Messaoud S, Hajjaji MA, Mtibaa A (2021) Real-time implementation of AI-based face mask detection and social distancing measuring system for COVID-19 prevention. Sci Program. <https://doi.org/10.1155/2021/8340779>

98. Tian Y, Yang G, Wang Z, Wang H, Li E, Liang Z (2019) Apple detection during different growth stages in orchards using the improved YOLO-V3 model. *Comput Electron Agric* 157:417–426. <https://doi.org/10.1016/j.compag.2019.01.012>
99. Tomè D, Monti F, Baroffio L, Bondi L, Tagliasacchi M, Tubaro S (2016) Deep convolutional neural networks for pedestrian detection. *Signal Process Image Commun* 47:482–489. <https://doi.org/10.1016/j.image.2016.05.007>
100. Trabelsi A, Elouedi Z, Lefevre E (2019) Decision tree classifiers for evidential attribute values and class labels. *Fuzzy Sets Syst* 366:46–62. <https://doi.org/10.1016/j.fss.2018.11.006>
101. Vijitkunsawat W, Chantngarm P (2020) Study of the performance of machine learning algorithms for face mask detection. In: 2020 - 5th international conference on information technology (InCIT). <https://doi.org/10.1109/InCIT50588.2020.9310963>
102. Voulodimos A, Doulamis N, Doulamis A, Protopapadakis E (2018) Deep learning for computer vision: a brief review. *Computational Intelligence and Neuroscience* 2018:1–13. <https://doi.org/10.1155/2018/7068349>
103. Wan S, Goudos S (2019) Faster R-CNN for multi-class fruit detection using a robotic vision system. *Comput Netw*. <https://doi.org/10.1016/j.comnet.2019.107036>
104. Wang R, He X (2019) Face detection based on template matching and neural network. In: 2019 international conference on communications, information system and computer engineering (CISCE). <https://doi.org/10.1109/CISCE.2019.00127>
105. Wang J, Jiang S, Song W, Yang Y (2019) A comparative study of small object detection algorithms. In: 2019 Chinese control conference (CCC). <https://doi.org/10.23919/ChiCC.2019.8865157>
106. Wang J, Luo J, Liu B, Feng R, Lu L, Zou H (2020) Automated diabetic retinopathy grading and lesion detection based on the modified R-FCN object detection algorithm. *IET Comput Vis*. <https://doi.org/10.1049/iet-cvi.2018.5508>
107. Wang Z, Wang G, Huang B, Xiong Z, Hong Q, Wu H, Yi P, Jiang K, Wang N, Pei Y, Chen H, Yu M, Huang Z, Liang J (2020, 2003) Masked face recognition dataset and application. *arXiv:09093v2*
108. Wu Di WT, Liu Q, Yang Z (2020) The SARS-CoV-2 outbreak: what we know. *Int J Infect Dis* 94:44–48. <https://doi.org/10.1016/j.ijid.2020.03.004>
109. Wu X, Sahoo D, Hoi SCH (2019) Recent advances in deep learning for object detection. *Neurocomputing* 396:39–64. <https://doi.org/10.1016/j.neucom.2020.01.085>
110. Xiang J, Zhu G (2017) Joint face detection and facial expression recognition with MTCNN. In: 2017 4th international conference on information science and control engineering (ICISCE), pp 424–427. <https://doi.org/10.1109/ICISCE.2017.95>
111. Xu M, Chen D, Zhou G (2020) Real-time face recognition based on Dlib. In: Yang CT, Pei Y, Chang JW (eds) *Innovative computing*. Lecture notes in electrical engineering, vol 675. Springer. https://doi.org/10.1007/978-981-15-5959-4_177
112. Xue B, Hu J, Zhang P (2020) Intelligent detection and recognition system for mask wearing based on improved RetinaFace algorithm. In: 2020 2nd international conference on machine learning, big data and business intelligence (MLBDBI). <https://doi.org/10.1109/MLBDBI51377.2020.00100>
113. Yadav S (2020) Deep learning-based safe social distancing and face mask detection in public areas for COVID-19 safety guidelines adherence. *International Journal for Research in Applied Science and Engineering Technology*. <https://doi.org/10.22214/ijras.2020.30560>
114. Yadav N, Binay U (2017) Comparative study of object detection algorithms. *International Research Journal of Engineering and Technology (IRJET)* 4(11)
115. Yin S, Li H, Teng L (2020) Airport detection based on improved faster RCNN in large scale remote sensing images. *Sensing and Imaging*. <https://doi.org/10.1007/s11220-020-00314-2>
116. Yu J, Zhang W (2021) Face mask wearing detection algorithm based on improved YOLO-v4. *Sensors* 21(9):3263. <https://doi.org/10.3390/s21093263>
117. Yu P, Zhu J, Zhang Z, Han Y (2020) A familial cluster of infection associated with the 2019 novel coronavirus indicating possible person-to-person transmission during the incubation period. *J Infect Dis* 221(11):1757–1761. <https://doi.org/10.1093/infdis/jiaa077>
118. Zhang K, Zhang Z, Li Z, Qiao Y (2019) Joint face detection and alignment using multitask cascaded convolutional networks. *IEEE Signal Process Lett* 23(10):499–503. <https://doi.org/10.1109/LSP.2016.2603342>
119. Zhang R, Li Y, Zhang AL, Wang Y, Molina MJ (2020) Identifying airborne transmission as the dominant route for the spread of COVID-19. *Proc Natl Acad Sci* 17(26):14857–14863. <https://doi.org/10.1073/pnas.2009637117>
120. Zhang N, Luo J, Gao W (2020) Research on face detection technology based on MTCNN. In: 2020 international conference on computer network, electronic and automation (ICCNEA). <https://doi.org/10.1109/ICCNEA50255.2020.00040>

121. Zhiqiang W, Jun L (2017) A review of object detection based on convolutional neural network. In: 2017 36th Chinese Control Conference (CCC). <https://doi.org/10.23919/ChiCC.2017.8029130>
122. Zhong Z, Wang M, Shi Y, Gao W (2018) A convolutional neural network-based flame detection method in video sequence. *SIViP* 12:1619–1627. <https://doi.org/10.1007/s11760-018-1319-4>
123. Zidi S, Moulahi T, Alaya B (2018) Fault detection in wireless sensor networks through SVM classifier. *IEEE Sensors J* 18(1):340–347. <https://doi.org/10.1109/JSEN.2017.2771226>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.