A Survey on Facemask Detection to Prevent COVID-19 Disease using Computer Vision and Deep Learning: Algorithms, Frameworks, Research and Implementation Challenges

Akhil Chawla, Praagna Prasad M, Mohana
Electronics and Telecommunication Engineering, RV College of Engineering®
Bengaluru, Karnataka, India.

***Abstract*- The on-going global Covid-19 pandemic has impacted everyone's life and brought economies to a stand still. World Health Organization (WHO) and Governments all over the world have found that social distancing and donning a mask in public places has been instrumental in reducing the rate of COVID-19 transmission. Stepping out of her homes in a face mask is a social obligation and a law mandate that is often violated by people and hence a face mask detection model that is accessible and efficient will aid in curbing the spread of disease. Detecting and identifying a face mask on an individual in real time can be a daunting and challenging task but using deep learning and computer vision, establish tech based solutions that can help combat COVID-19 pandemic. In this paper, we have established various deep learning architecture, designs, models and parameters that are needed to establish a real time working model to identify a person who is donning a face mask and a face mask violator. The main outcome of this survey is to bring light to the various deep transfer learning algorithms and parameters that helps to build a face detection model in real time for various public environment or places.**

Keywords— Faskmask Detection, COVID-19 Disease, Deep Learning, WHO, ICMR, Computer Vision, Artificial Intelligence.

I. INTRODUCTION

The COVID-19 pandemic has instilled fear and anxiety among the general population as nature of the disease is random and unpredictable. This virus has killed more than a million people around the globe, and it is expected to rise and continue in the same way leading to the death of millions of people. India is on the verge of a second wave despite a large-scale vaccination drive. Vaccinating the whole population will take a lot of time, perhaps years and there is no evidence on reinfection or on the long-term protection against the virus , So the most extensively practiced method to curb the transmission of Covid-19 virus is to adopt safety measures recommended by WHO and ICMR. Like maintaining social distance, washing hands regularly, and wearing a proper mask that covers mouth and nose at all times. A mask of any type gives 98% protection against virus droplets spreading through their mouth. However, we observe that a lot of people in public places don't wear their mask properly and even if they are donning a face mask, it is not worn in a way that covers both nose and mouth. Thus, endangering others around them. A lot of carelessness has led to the subsequent second wave in India today and people still refuse to wear a mask properly while stepping out. While our healthcare system is overwhelmed and the governments are imposing lockdowns, stepping out of our homes can be a dangerous act even if it's an essential need. To contribute to assist government and the officials in making sure that people abide by the basic rule i.e., to ensure people wear a face mask properly in public at all times, for which come up with a design for a Face Mask Detection using Deep Learning Technique and computer vision. Aim is to identify an individual and recognize if he is wearing a mask properly covering his mouth and nose by leveraging deep learning algorithms. As stated by WHO, respiratory droplets and any type of physical contact are the two ways of coronavirus spread, and to prevent virus from the spread, medical masks are best and only solution.

Mask is important because of two reasons: The virus spreads directly from infected person's sneeze or cough and if he/she is masked, they will prevent themselves from transmitting to others. Secondly, the mask will stop transmission of the virus in case you touch a virus-contaminated surface and then your mouth or nose. The nature of spread of the disease is rampant, no one anticipated that we would be engulfed by a worldwide pandemic but with the power of technology, we can innovate to prevent the spread of the disease and hopefully see an end to it. With the on-going crisis, many companies and governments have used the power of technology, particularly Artificial intelligence (AI) to help combat the disease. We have seen robots delivering food, medicines etc, to infected individuals in hospitals. AI has been used extensively to help predict origin of the disease and also in helping find the cure and there has been ample research regarding technology based solutions to detect Covid-19 through X-rays and CT scans.

Machine learning and AI concepts are utilized in developing tracking software including wearable tech gadgets to make sure people aren't violating their mandatory quarantine rule. Certain countries like Taiwan integrated the database of their national medical insurance with database of their immigration and customs, hence predicting possible infected passengers who might have contracted it while travelling from a Covid-19 hotspot. AI and technology have had a significant impact in combating the disease. AI and ML had been adapted into a few healthcare applications although the trend seems to be rather

slow and isolated. As the pandemic comes to an end and our memories fade, AI will eventually be on the forefront of automating healthcare industries and an integral part in healthcare systems and governments around the world.

This paper describes the use of deep learning concepts to identify people wearing a face mask properly covering nose and mouth from those who are violating this norm which will help to prevent COVID-19 disease. To assist governments in identifying covid-19 protocol violators. Using Deep Transfer Learning Approach to tackle the problem, DL is a subset of ML which are based on ANN and divided into supervised, semi-supervised and unsupervised. While transfer learning is an approach in DL where knowledge can be transferred from pre trained highly efficient models trained by researchers leading to high efficiency and reduced computational complexity. Use of Python along with cross platform computer vision library, OpenCV which help to develop real-time computer vision applications. TensorFlow and Keras that provided a Python interface for artificial neural networks.

II. LITERATURES SURVEY

S. Sen [1] implemented a mask detection system. Model uses PyTorch library and OPENCV of python for mask detection along with MobileNet deep learning framework. The accuracy of 79.24 % was achieved. I. Shete [2] The system deployed makes use of pre-trained models like YOLOv3, ResNet Classifier and DSFD for object detection and facial recognition by determining the minimum distance between people and faces without face mask, achieved a confidence score of 100%. P. Nagrath [3] presented a model for detection of face masks, which uses the concepts of deep transfer learning, and tools such as TensorFlow, Keras, and OpenCV. System gave an accuracy of 0.9264 and F1 score 0.93.V. Vinitha [4] presented DNN model for face mask detection. A high accuracy was achieved from the model by tuning the hyper parameters. I. Venkateswarlu [5] presented a pre-trained MobileNet model for face mask detection. Model gave an accuracy of 99% and 99.56% on two publicly available datasets respectively. Vinay Sharma [6] proposed a model that uses the YOLOv5 and TensorFlow technologies for processing the images and real-time videos. The model has been tested using YOLOv5s as well as YOLOv5x on Google Colabs. Y. Said [7] deployed a face mask detection system based on CNN and YOLO object detection framework to implement it on an embedded low power device. The model was compared against state-of-the-art works such as Retina Face Mask, SSD gave a higher Precision and Recall value of 94.6% and 95.8% respectively.

Siegfried [8] proposed a comparative study for the usage of different deep learning methods used for facemask detection like MobileNetV2, ResNet50V2, and Xception and combined them with a facial image recognition model called Caffe Model. The dataset used for this study was gathered from Bing Search API, Kaggle Dataset, and Real World Masked Face Dataset. Xception turned out to be the most stable and efficient model with an accuracy of 98%. G. Jignesh Chowdary [9] proposed an automation model based on a deep transfer learning to automate identification of unmasked people. To increase data for better training and testing, Image augmentation technique is adopted. Performance metrics used were Accuracy, Precision, Matthews Correlation Coefficient, Classification Loss. The model outperformed on all metrics and achieved an accuracy of 99.9% on training set and 100% on test set. T. Quang [10] described a real time face mask detection model which can alarm when it detects an unmasked person. The proposed model is also enhanced to detect an incorrectly worn facemask or something else used as a face mask. The Haar cascade classifier plays a vital role along with the YOLOV3 algorithm to detect the faces and the masks respectively. The model is trained and tested on the MAFA dataset. The system was deployed in real time through a webcam and used metrics such as Precision and Recall to determine its performance, which can out to be 89% and 90.1% respectively.

Haoxiang Li [11] explained t a model cascade architecture which is built on Convolutional neural networks with very powerful discriminating capabilities along with high performance. The model also has a CNN based calibration stage to improve localization effectiveness. D. Chen [12] proposed CNN, developing a supervised transformer network to handle the pose variation issues during the detection. The model has a multi-task RPN, for the prediction of candidate face area amplifying the facial marks and a Region Based Convolutional Neural Networks (RCNN) that verifies valid faces. The model is trained to best scale itself, to determine face patterns. F. Schroff [13] explained about a system called the FaceNet that learns to detect face on a Euclidean space where the distance corresponds to the similarity between faces. This model can be implemented as a feature vector and can assist majorly in face recognition, verification and clustering. Susanto [14] has employed YOLO V4 algorithm, which detects if a person has worn the mask correctly and if he has not worn it, it can be detected even when said person is moving. The model is expected to be useful in real-time applications to detect all the types of face masks available in the commercial market. YOLOV4 capable of identifying an individual who is wearing a face mask and who is not. YOLOV4 algorithm consists of a two stage detector that is able to run twice as fast as any other neural network. Sharun Akter [15] Model deals with YOLOV3 algorithm. YOLO darknet framework is one of the first frameworks that employs face mask detection using image visualization method. YOLOv3 has access to CNN because of hidden layers which fetch the algorithm and it possesses the ability to identify and conceptualize any given image. Model was found to have an accuracy rate of 96% and the average loss had declined to 0.073 and the mean average precision score was found to be 0.96.

Mohammad Marufur Rahman [16] proposed a model which is built on the concepts of deep learning that was trained with a dataset of images of individuals wearing a face mask and unmasked Individuals gathered from multiple sources. The proposed model design was found to have a 98.7% rate of accuracy on Identifying people who have donned a face mask against those who haven't worn a facial mask. CCTV cameras mounted across the city were used to capture data in real-time from different crowded and public places in the city. Using the video footage captured from the CCTV cameras, the images were analyzed, identified and captured and these extracted images were used to detect masked people from unmasked people. CNN algorithm used for feature extraction process and then these features extracted from images are discovered through a number of hidden layers. The model recognizes Individuals violating Covid-19 safety norms by not wearing a face mask properly, this data is notified to the local civic body and they are encouraged to act on it.

S. Sanjaya. [17] Proposed a model in which the face mask detection used, created using a popular machine learning algorithm which employs image classification methods such as MobileNetv2. MobileNetV2 is a type of feature extractor based on CNN that was created by Google to improve Performance.

A. Lodh [18] had proposed a model focuses on increasing the accuracy of prediction by focusing on straightening its rate of accuracy along with detection probability and was achieved through MobileNet_V2. This model also has the features to detect mask violators and send them an email notification. This model has accuracy and lesser time complexity because of implementing the model through MobileNet Algorithm. S. Islam1 [19] model is based on face mask detection developed using multiple CNN classifiers. CNN is chosen as it gives high accuracy and it is faster than others. This model proposes a face mask detector that will be able to identify a face mask on a person regardless of the arrangement and motion and is used to train it in an appropriate neural network to get accuracy results. Taking an RGB input image, output is obtained from any of the orientation. The crux of the algorithm used lies in feature extraction and prediction of classes applied to the images. A. das [20] system built for this model employs a model containing two 2D convolution. Proposed model a total of two datasets were used. From the first dataset, 1376 images out of 690 images were images of people wearing a face mask on them and the rest of 686 images were of people who were unmasked. Second dataset was taken from Kaggle which had a total of 853 images, some of these images have their head turned, tilted and slanted in a crowded space with people in the frame wearing different types of commercial face masks that have their own design, pattern and shape. The proposed model is then trained using specific parameters and then tested using these two datasets.

M. Yan [21] model presents a SSD model, the model revolves three important ideas: 1) It proposes a lightweight network architecture leveraging SSD for feature extraction, and separable convolution network for improving the detection speeds and for easier implementation in real-time. 2) It also proposes a module which strengthens the deep learning features from CNN model for an accurate depiction of small objects 3) Also aims at obtaining a large dataset by real time collection of images from super markets. The experiment results yield high detection precision and real time performance. G. Cao [22] developed a model to detect tiny objects using a SSD. The system implements multi-level complex features in SSD to achieve high accuracy. The design features two fusion models which are the combinational module and element-sum module. Implementation of the model gives high mAP on PASCAL VOC2007 datasets achieving speeds of 43 and 40 FPS respectively. T. Kong [23] model present's a deep neural network, based on the HyperNet, for finding proposal regions for the dataset along with OD. Hyper Feature design development is focused upon, which aims to develop on the high resolution features of the image and also provides an efficient framework for complex CNN features. The architecture gives high recalls and produces a small number of object region proposals. HyperNet is efficiently developed to achieve very high object detection accuracy. R. Girshick [24] proposed a method uses high-capacity convolutional neural networks (CNNs) to localize and segment objects. The model observes a significant performance boost by implementing domain-specific fine-tuning. Per-category segmentation accuracy on VOC 2011 dataset was observed to be best among other available methods across each of the 20 PASCAL classes. D. Erhan [25] deployed a deep learning network model for OD, which constructs anchor boxes, defining the probability of the presence of the required object in that box. High accuracy and efficiency was achieved by the model on two major publicly available datasets. Precision vs Recall graphs of various state-of-the-art models along with the one proposed were given for both the ILSVRC-2012 and VOC 2007 to support the results of the model's efficient performance. S.Abbas [26] implemented a model which implements deep learning model and classifies the data into two categories masked and unmasked faces. A YOLO face detector is also implemented in another approach to carry out facemask detection. The faces are categorized into masked and unmasked using a newly designed fast CNN algorithm achieving an accuracy of 99.5%. P. Khamlae [27] developed a CNN which was trained on images which had a very high resolution. The trained model gave an accuracy rate of 81 %. The model was tested and implementation in real time. G. Yang [28] proposed a model for object detection, which implements deep transfer learning along with the YOLOV5 algorithm which is considered the most powerful objection detection algorithm at present. A. Das [29] described a facemask detection model which is trained on images. The model leverages the Sequential Convolutional Neural Network to obtain optimized values of parameters to detect facemask at a higher accuracy while preventing overfitting. The accuracy obtained by the model was 95.77%.

III. FACEMASK DETECTION ALGORITHMS

A. Region-based Convolutional Neural Networks (R-CNN)

It is a type of CNN algorithm that finds its use primarily in object classification and objection detection, where with the help of bounding boxes, multiple objects present in an image are identified and only a single object that is accepted will be of dominance in that region and this is carried out using selective search. It has multiple successors to it which have improved on the previous version, Fast R-CNN, Faster R-CNN and Mask R-CNN.

B. Region-based Fully Convolutional Network (R-FCN)

R-FCNs, are a type of region-based object detector, which understands and analyzes in feature maps and applies convolution to recreate position-sensitive score maps with depth $k\times k \times (C+1)$.

C. Spatial Pyramid Pooling (SPP-net)

SPP-Net is a convolutional neural network, this algorithm implements a spatial pyramid pooling methodology which aims to eliminate the fixed-size constraint of the network. It allows variable size input images to CNN and can be used for Classification and Object Detection.

D. Single Shot Detector (SSD)

SSD has no delegated region proposal network and predicts boundary boxes and classes directly from feature maps.

E. You Only Look Once (YOLO)

It's one of the most exciting and latest algorithm, it is fast real-time multi-object detection algorithm which deals with regression rather than choosing prominent parts of an image. In addition, it also predicts classes and boundary boxes for the whole image in just one run of the algorithm. YOLO works in

a method where images are split into 19x19 grid cells. Each of these 19x19 cells is then used to predict K bounding boxes and establish the probability of a cell containing a certain class. The grid cell is assigned a class with the maximum probability, based on the k bounding boxes and this process is carried out for all the grid cells that are present in the image.

F. Blitznet

The prominent feature of this algorithm is it performs object detection and semantic segmentation side by side in one forward pass which enables real-time computations. Object detection and semantic segmentation leads to higher accuracy as well as the benefit of having a single network perform several tasks.

IV. Implementation of Facemask Detection Algorithms

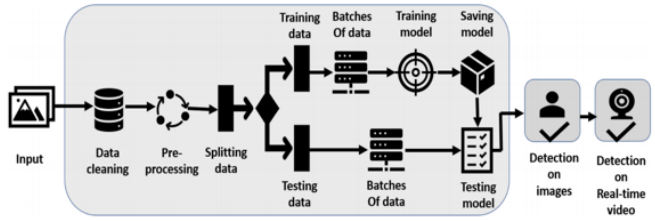

Fig.1. Block diagram of facemask detection

Fig.2. Flowchart for implementation of facemask detection using image dataset.

Figure 1 shows the general block diagram of facemask detection. Detailed flow of implementation of facemask detection using image dataset as shown in figure 2. The model is first fed with the dataset containing images of both masked and unmasked people in varied environments and postures. Then train the model on the given dataset. Detection code then turns on the video stream, captures frames continuously from video stream to be fed to the model for real time detection. Bounding boxes of a certain height and width drawn using object detection process. This collected data is passed through MobileNetV2 model layers after processing, which classifies image as a person wearing a face mask or a person is unmasked. If the person is masked, a green anchor box is displayed and red if not wearing a mask with the accuracy of detection for the same tagged on the anchor box.

V. Facemask Detection Models and Frameworks

TABLE I. FACEMASK DETECTION MODELS AND FRAMEWORKS USED IN COMPUTER VISION

Haar-cascade	Random Forest	Vanilla CNN
ResNetl8	MobileNet v2	DenseNet161
Inception v3	VGG - 19	VGG - 16
ResNet - 152	ResNet - 101	Xxceptioon
Inception v4	ResNet - 50	DenseNet - 121
DenseNet - 201	DenseNet - 169	ResNet - 34
MobileNet v2	MobileNet v1	ResNet - 18
GoogLeNet	ENet	Fd - MobileNet
BN-NIN	ShuffleNet	SqueezeNet
BN-AlexNet	AlexNet	

Table I shows the various facemask detection models and frameworks can be used to implement facemask detection algorithm real time at public places using computer vision and OpenCV.

VI. 2D and 3D Facemask Detection

2D V/S 3D Recognition

- 3D images are represented by point clouds, meshes, volumes while 2D images have pixel grid representations.
- The computation and memory resource requirement is higher for 3D as an extra dimension is added.
- It is hard to achieve one unified framework for 3D object detection as there is a difference in distribution of the objects between indoor and outdoor scene areas.
- 3D data being sparse for outdoor scenes makes detection tasks more challenging, compared with the dense 2D images.
- Unavailability of large size labelled datasets, which are extremely important for supervised based algorithms, are still under construction for 3D images compared to well-built and extensively available 2D datasets.

3D Facemask detection Dataset

3D object recognition in computer vision, is a technique of recognizing and determining 3D information of an object, by analyzing the pose, volume, or shape, of the objects from a photograph or range scan. The object to be identified is provided as an input to the computer vision system and then detection of the same object is done in a video stream, in a controlled environment. This can be done both offline and in real time. 3D Facemask detection can be carried out using remote photoplethysmography, Contrastive Context-Aware Learning, etc. Table II. Shows the list of 3D facemask detection.

TABLE II. 3D FACEMASK DETECTION DATASET

3DMAD	3DFS-DB	BRSU
MARsV2	SMAD	MLFP
ERPA	WMCA	HiFiMask
CASIA-SURF 3DMask		

Popular dataset used for facemask detection using computer vision:

- UFMD
- RMFD
- RWMFD
- Face mask
- MaskedFace-Net
- EfficientNet-b0 to b3
- Joseph Nelson Roboflow
- Prasoonkottarathil Kaggle
- Ashishjangra27 Kaggle
- Andrewmvd Kaggle
- MAFA_data
- Wider Face dataset
- Face Mask Label Dataset (FMLD)

Performance parameter used for facemask detection:

$$\text{ACCURACY} = \frac{(TP+TN)}{(TP+TN+FP+FN)}$$

$$\text{PRECISION} = \frac{(TP)}{(TP+FP)}$$

$$\text{SENSITIVITY} = \frac{TP}{TP+FN}$$

$$\text{SPECIFICITY} = \frac{TN}{TN+FP}$$

$$\text{IoU} = \frac{TP}{TP+FP+FN}$$

$$\text{MCC} = \frac{(TP\,x\,TN)-(FP\,x\,FN)}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}}$$

VII. IMPLEMENTATION METHODS AND SOLUTIONS TO PREVENT COVID-19 DISEASE

Wear a mask when you step out of home, adapt the double-masking method if possible. Use a surgical mask followed by a cotton mask. Studies show that the transmission of virus between 2 individuals who have double masked and are interacting with 6 ft distance between them is almost close to 0% which heightens the importance of wearing a face mask. Isolating for symptomatic will help prevent the spread of virus to immediate family and friends. The US and UK have set an example by combatting the second wave by aggressive vaccination drives, getting vaccinated at the earliest can help combat the spread of disease. Technology has always served as a medium in solving global issues and with the on-going pandemic, technology-based solutions are the need of the hour. Using and implementing large scale AI based models that replicate human labour and monitoring will be able to reduce the pressure on essential workers and keep them safe. A face detection system can be helpful for the police and task force and a safer alternative for them to make sure citizens are abiding by the rules. Figure 3 shows the different Methods to prevent covid-19 disease. Fgure 4. Shows the architecural design using computer vision to prevent covid-19 disease atpublic placess.

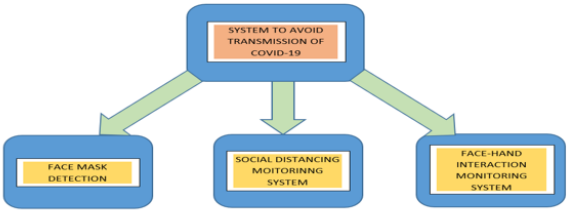

Fig.3. Methods to prevent covid-19 disease.

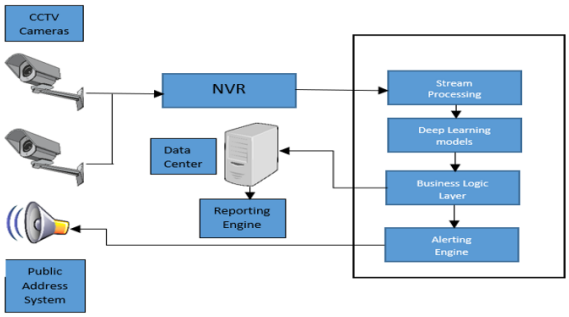

Fig. 4. Block diagram of safety measure and to prevent covid-19 disease .

VIII. UPCOMING TECHNOLOGY FOR FACEMASK DETECTION RESEARCH

A. *Facemask detection using Generative Adversarial Networks (GANs)*

Face mask detection can be carried out from speech analysis and processing, by understanding the variation in sound parameters while talking with a mask on. A method based on data augmentation while training a Generative Adversarial Networks (GANs) with cycle- consistency loss can be effective for classification of speech into masked and unmasked. The process involves converting the original and converted speech data into spectrograms which are provided to a deep learning neural network (Ex. ResNet), classification can be carried out after further processing. This face mask detection method also has multiple applications in forensic investigations, person recognition via speech while wearing a mask, etc.

B. *Facemask detection using Transfer learning approach*

To implement image classification algorithms, deep neural networks are preferred choice as they have higher rate of performance accuracy as compared to the rest of the algorithms but consequently trying to train a deep neural network is complex requiring higher computational power more resources which makes it not only expensive but also time consuming and hence deep learning was evolved to transfer learning. Transfer learning is a methodology in which you can transfer the learnings and patterns of a previous network with respect to parameters and adapt it to another model that has a different task. The main advantage of this is that, in transfer learning the model has increased performance even if it was trained on a smaller dataset. Transfer learning is a method in which a model which was trained for one AI related task is modified and then implemented for another AI related task. It aids as a great solution for when there is insufficient data to carry on a machine learning task and it also improves performance when we are modelling the target task. The main advantage of using transfer learning is that we don't have to start the process from scratch. Instead you start from patterns that have been learned while trying to solve a different task and you can leverage the previous training of the model and adapt it to fit for another

task. Transfer learning results in decreasing the overall training time for a neural network model and it also can generate a result in lower generalization error.

C. Facemask detection using Edge computing

Serverless edge-computing is one of the latest evolutions in infrastructure of edge-computing, in this process end-users are able to directly access computing resources. Most of the available face mask detection models in the market are integrated with complex hardware components or software's and therefore is not as accessible to the public. A need for a more appealing and light-weight device arised to enable fast and convenient face mask detection and a solution was found in edge-computing based in-browser applied to face mask detection .This can be accessed on any device like phone, computer and tablets that have a stable net connection and a web browser. For face mask detection, the option of cloud computing based solution is limited by the availability of stable internet for sufficient bandwidth to access real time video for monitoring and it is also very expensive to avail cloud services By contrast, Wear Mask server less edge-computing design is significantly cheaper as it cuts down on hardware costs and it is more accessible to the public. The edge-computing design also has a massive advantage as it is significantly low in privacy risk.

D. Facemask detection using mobile phones

Facemask detection using mobile phones is often the foremost accessible and successful prototype because it is accessible to everyone. This method has high efficiency and is backed by a large amount of research.

IX. Key Players in Facemask detection research and development

With the on-going pandemic, Tech companies and industry innovators have been making some advancements in AI and ML models to suppress the spread of the COVID-19 disease. Some of the examples are

NVIDIA CLARA GUARDIAN - is an application developed by NVIDIA, which aims at simplifying the process of development and deployment of smart sensors used in AI for healthcare services. Clara guardian is loaded with a collection of healthcare-specific pre-trained models that are backed by GPU application frameworks, toolkits and SDKs.

Leeway Hertz - developed a facial recognition system powered by AI to detect whether a person is masked or not. The developed application can connect to any existing or newly developed IP face mask detector cameras. App users also have an option to add their faces and phone numbers, to send them an alert in case of a mask violation.

Trident's VQI - a technology for Face Mask Detection is an image analytics solution, powered by AI and Computer Vision, aimed at providing a solution to identify the Covid-19 norms violators. It leverages the power of AI and its ability to detect violations such as the use of improper face masks and ensures the use of face masks at all times. It also accounts for instances of violation of the social distancing norms. The proposed system can be very effective in crowded areas like Hospitals, vaccination centers etc.

MOBISOFT technologies - deployed a model which leverages AI technology to detect if a person is masked or not. It's used by connecting the app to any surveillance camera at your premise. The system is developed in a way where it sends an alert to the administrator if someone is unmasked on the premises.

Zicom - developed a system which checks if a person entering the premises is wearing a face mask or not and a thorough temperature check is also carried out. They have integrated a medical grade IR temperature sensor to detect body temperature and it also acts as an admission tool by restricting the entry of people with abnormal temperature.

Softweb Solutions - deployed a number of covid-19 specific models that aims to prevent the spread of the disease. For example: Wearable bands for employees and doctors that detects body temperature and notifies employers and managers when an employee shows abnormal temperature, A social distancing detection solution model which uses occupancy sensors to measure the distance between individuals in queues, public places etc.

X. Research and real time implementation challenges

A. Illumination

Object appearance can be impacted significantly by slight changes in pose and illumination. Specular light causes false detections that come under the Illumination challenges and problems which can be introduced from weather changes, along with many other visual challenges.

B. Pose

Human pose estimation is one of the most common challenges in computer vision. It is an important area of study as this technology can be useful for a number of applications such as Human activity and movement interpretation, Augmented Reality, Robotics, Gaming, Animation, etc. For example, human pose estimation allows for higher-level reasoning in the context of human-computer interaction and activity recognition

C. Occlusion

Occlusion basically describes a situation where an important feature of an object is not visible or hidden due to an obstruction placed in the field of view of the sensor. Occlusions are very common in the real world, causing difficulties in building robust object detection models. Texture-rich objects have distinctive local features which makes them easy to detect under severe occlusions, whereas large uniform regions without texture can be found in many man-made objects, making it difficult to detect them. These objects are characterized by their shapes, curves and profile which are ambiguous and need assessments from multiple views. Detection of textureless objects increases the complexity as they require multiple viewpoints which can have a high percentage of occlusion. Solutions to deal with this problem can vary from one problem to another.

D. Low Resolution

Low Resolution poses a challenge as extremely small objects may contain only a few pixels within the bounding box and the resolution is very low and the features won't be detected accurately, to overcome this it is advised to increase the resolution of your images which in turn leads to an increase in the richness of features that your detector can catch from that small box.

E. Model Complexity

In real-time application to meet the demands for video processing, an object detection algorithm has to be extremely fast in addition to accurately classifying important objects. The more number of layers lead to higher accuracy but decreased

speed and vice versa. It's a tradeoff that should be made depending on the factors and type of model. Poor quality of data can also impact the final detection model negatively, we need to be mindful of the features and to eliminate redundant and irrelevant features are necessary to overcome model complexity challenges.

F. Imbalance in dataset

Foreground-to-Background imbalance, is the most commonly known imbalance problem in object detection, which is due to the imbalance in the number of positive to negative examples in the dataset. As in an image, there might be only a few positive examples but multiple negative examples can be extracted. Imbalance problem is one of the major problems faced by many models, if not dealt with properly, it can cause a great impact on the performance and accuracy of the model.

G. Classifier struggle with classifying some face mask types

This could pose a challenge as the model focuses on face detection and the classifier should be able to identify the different types of masks and to be able to differentiate between masks.

XI. CONCLUSION

COVID-19 pandemic has impacted the way people interact with each other and definitely has changed the social patterns that existed before the pandemic. According to data published by WHO and various researchers, COVID-19 transmission is significantly higher in crowded places and between a groups that is not wearing face masks. In this paper, we have highlighted various multiple object detection models including their architecture design, pattern recognition and the parameters that are used in implementation of these algorithms. Various performance metrics used in deep learning and computer vision were also highlighted and drawn inferences from. Documented the integration of edge computing for face mask detection models, in real time. Discussed 3D dataset and how it compares with 2D dataset, multiple datasets for 3D are also listed which is a new zone of research and it has a lot of scope in the near future for further implementation. Documented face mask detection using mobile phones in this paper which has the potential to be the most accessible and popular mode of face mask detection as it doesn't need any additional hardware components to access the platform. This paper also contains multiple challenges and setbacks that occur during implementation with their possible solutions to overcome these challenges. This paper also addresses how GAN algorithm can be used for face mask detection through speech analysis and processing, by understanding the variation in sound parameters while talking with a mask on. Key players and industry leaders are creating and developing new technology to combat the spread of covid-19. Corporates and companies can implement these smart solutions to protect their employees and ensure a safe working environment.

REFERENCES

[1] Sneha Sen et all., "Face Mask Detection System for COVID_19 Pandemic Precautions using Deep Learning Method", *International Journal of Emerging Technologies and Innovative Research,2020.*

[2] https://www.researchgate.net/publication/347439579_Social_Distance_Monitoring_and_Face_Mask_Detection_Using_Deep_Neural_Network

[3] P. Nagarath et all., "A real time DNN-based face mask detection system using single shot multibox detector and MobileNetV2", *Sustainable Cities and Society, 2021.*

[4] V.Vinitha et all., "COVID-19 FACEMASK DETECTION WITH DEEP LEARNING AND COMPUTER VISION" *International Research Journal of Engineering and Technology (IRJET),*2020.

[5] I. Venkateswarlu et all.,"Face mask detection using MobileNet and Global Pooling Block" *Conference on Information & Communication Technology (CICT), 2020.*

[6] https://scholarworks.calstate.edu/downloads/wp988p69r

[7] Y.Said et all.,"Pynq-YOLO-Net: An Embedded Quantized Convolutional Neural Network for Face Mask Detection in COVID-19 Pandemic Era*" International Journal of Advanced Computer Science and Applications (IJACSA), Sustainable Cities and Society*,2021.

[8] https://www.researchgate.net/publication/344239546_Comparative_Study_of_Deep_Learning_Methods_in_Detection_Face_Mask_Utilization

[9] G. Chowdary et all.,"Face Mask Detection using Transfer Learning of InceptionV3,"*arXiv:2009.08369*, 2020.

[10] T. Quang et all.,"Real-Time Face Mask Detector Using YOLOv3 Algorithm and Haar Cascade Classifier",*International Conference on Advanced Computing and Applications (ACOMP),*2020.

[11] H.Li et all.,"A Convolutional Neural Network Cascade for Face Detection" *Conference on Computer Vision and Pattern Recognition (CVPR),* 2015.

[12] D. Chen et all.,"Supervised Transformer Network for Efficient Face Detection" *arXiv:1607.05477*,2016.

[13] F. Schroff et all.,"FaceNet: A Unified Embedding for Face Recognition and Clustering", *arXiv:1503.03832*,2015.

[14] Susanto et all.,"The Face Mask Detection For Preventing the Spread of COVID-19 at Politeknik Negeri Batam"*International Conference on Applied Engineering (ICAE),*2020.

[15] M.Rafiuzzaman Bhuiyan, et all.,"A Deep Learning Based Assistive System to Classify COVID-19 Face Mask for Human Safety with YOLOv3",*International Conference on Computing, Communication and Networking Technologies (ICCCNT),*2020.

[16] Mohammad Marufur Rahman, et all.,"An Automated System to Limit COVID-19 Using Facial Mask Detection in Smart City Network", *International IOT, Electronics and Mechatronics Conference (IEMTRONICS),*2020.

[17] S.Sanjaya et all.,"Face Mask Detection Using MobileNetV2 in The Era of COVID-19 Pandemic", *International Conference on Data Analytics for Business and Industry: Way Towards a Sustainable Economy (ICDABI),*2020.

[18] Anirudh Lodh, et all.,"Prototype for Integration of Face Mask Detection and Person Identification Model – COVID-19" *International Conference on Electronics, Communication and Aerospace Technology (ICECA),* 2020.

[19] Md. Shahriar Islam et all., "A Novel Approach to Detect Face Mask using CNN", Proceedings of the Third International Conference on Intelligent Sustainable Systems (ICISS) 2020.

[20] https://www.researchgate.net/publication/344725412_Covid19_Face_Mask_Detection_Using_TensorFlow_Keras_and_OpenCV.

[21] W. Han et all., "A Mask Detection Method for Shoppers Under the Threat of COVID-19 Coronavirus", International Conference on Computer Vision, Image and Deep Learning (CVIDL),2020.

[22] G. Cao et all., "Feature-Fused SSD: Fast Detection for Small Objects" *International Conference on Graphic and Image Processing (ICGIP),*2017.

[23] T. Kong, et all., "HyperNet: Towards Accurate Region Proposal Generation and Joint Object Detection", Computer Vision and Pattern Recognition (CVPR),2016.

[24] R.Girshick et all., "Rich feature hierarchies for accurate object detection and semantic segmentation Tech report". *arXiv:1311.2524.*

[25] D. Erhan et all., "*Scalable Object Detection Using Deep Neural Networks*", Computer Vision and Pattern Recognition,2020.

[26] S. Abbasi, et all., "A Face-Mask Detection Approach based on YOLO Applied for a New Collected Dataset," *International Computer Conference, Computer Society of Iran (CSICC),* 2021.

[27] P. Khamlae et all., "An Application of Deep-Learning Techniques to Face Mask Detection During the COVID-19 Pandemic," *Global Conference on Life Sciences and Technologies (LifeTech),* 2021.

[28] G. Yang et al., "Face Mask Recognition System with YOLOV5 Based on Image Recognition," 2020 IEEE 6th International Conference on Computer and *Communications (ICCC),* 2020.

[29] https://www.pyimagesearch.com/2020/05/04/covid-19-face-mask-detector-with-opencv-keras-tensorflow-and-deep-learning/.