ORIGINAL RESEARCH

HRNET: AI-on-Edge for Mask Detection and Social Distancing Calculation

Kinshuk Sengupta[1] · Praveen Ranjan Srivastava[2]

Abstract

The purpose of the paper is to provide innovative emerging technology framework for community to combat epidemic situations. The paper proposes a unique outbreak response system framework based on artificial intelligence and edge computing for citizen centric services to help track and trace people eluding safety policies like mask detection and social distancing measure in public or workplace setup. The framework further provides implementation guideline in industrial setup as well for governance and contact tracing tasks. The adoption will thus lead in smart city planning and development focusing on citizen health systems contributing to improved quality of life. The conceptual framework presented is validated through quantitative data analysis via secondary data collection from researcher's public websites, GitHub repositories and renowned journals and further benchmarking were conducted for experimental results in Microsoft Azure cloud environment. The study includes selective AI models for benchmark analysis and were assessed on performance and accuracy in edge computing environment for large-scale societal setup. Overall YOLO model outperforms in object detection task and is faster enough for mask detection and HRNetV2 outperform semantic segmentation problem applied to solve social distancing task in AI-Edge inferencing environmental setup. The paper proposes new Edge-AI algorithm for building technology-oriented solutions for detecting mask in human movement and social distance. The paper enriches the technological advancement in artificial intelligence and edge computing applied to problems in society and healthcare systems. The framework further equips government agency, system providers to design and construct technology-oriented models in community setup to increase the quality of life using emerging technologies into smart urban environments.

Keywords Deep learning · Edge-computing · Social and industrial safety

Introduction

The recent outbreak of coronavirus SARS-CoV-2 infection early detected in December 2019 in Wuhan, China [1]. The magnitude of infectious spread has affected more than 3.2 million peoples causing 239 K deaths, according to the European Centre for Disease Prevention and Control. In the latest report, the total death caused by fever is 64.7% and 52.9% due to cough [1]. The role of flattering the curve via quarantine and preventing community spread using a respiratory surgical mask or N95 mask have found significance in controlling the spread in previously published literature [1]. The literature had shown evidence indicating the use of a surgical mask reduces the transmissibility per individual by preventing the droplets transmission in both laboratory and clinical contexts [2]. In the compliant scenario for industrial workplaces, airports, and places for community gathering possess the highest risk of spread without prevention. Public health authorities have approached to contain the virus spread via isolation, personal protection, and hygiene compliance [3], social distancing, contact tracing, and surveillance application [4].

✉ Kinshuk Sengupta
   efpm01.004@iimrohtak.ac.in

   Praveen Ranjan Srivastava
   praveen.ranjan@iimrohtak.ac.in

1 Microsoft Corporation, New Delhi, India

2 Indian Institute of Management, Management, City Southern Bypass, Sunaria, Rohtak, Haryana 124010, India

Recent surveys and literature have studied how to handle community gatherings to prevent the global spread of COVID-19. Several research gaps that need to address for the response to COVID-19 need to be discussed in the current situation, such as developing an ethical framework to control and contain spread in such states. This also involves devising appropriate ways to prevent and control the infection by identifying optimal personal protective equipment (PPE's), and after that, understand behavior among various vulnerable groups [4]. Many recent research works propose policymakers to make masks as an official guideline to stop spread in the community [2]. Hence, the need to find sustainably designed AI-powered technology solution like robotics, self-explainable digital solutions [5] to tackle the post-pandemic situation in society and industrial setup needs drastic attention to counter the ripple effect of COVID-19 in economic circumstances. Policymakers and industries would need an efficient solution within the industrial and societal structure to trace and track people and prevent droplet spread [6]. An older research review conducted randomized studies-controlled trials using masks and found that a low-cost intervention would be useful to break the transmission of the respiratory viruses [7]. The summary of existing studies is depicted in below Table 1.

The past studies and research work conducted by authors hold few limitations from a conceptual framework point of view. The evidence found literature depicts the need to conduct further research focusing on devising a mature technology model infusing Artificial intelligence to respond to the outbreak in the current situation addressing the gap (Fig. 1). The paper focuses on building a technology framework that can be extended for implementation in public and industrial setup to detect people with or without facemask and, therefore, eases out tracing and further control spread. The current research work on how studying the impact of IoT-based smart disease surveillance systems can help in controlling the spread of the pandemic [8]. Policymakers need to think about enforcing and implementing smart urban and industrial planning solutions across the geographies post-lockdown situation to resume economic activities. In the past, Bibi and Krogstie [9] studies have directed toward a new generation of urban planning tools for improving mobility and accessibility; this can further be applied in combating pandemic situation across the globe.

The Need of Technology Framework

The world had seen unprecedented times before as well due to other epidemic health crises caused by Ebola, H1N1, SARS-CoV1, Zika Virus, and many others that have caused a drastic impact on the economy and social wellbeing. However, now since the exponential spread of COVID-19 globally has created a need for responses from governments due to heavy loss in GDP ranging from 3 to 6%, or even more depending on the country [10] and economy predicted by the model developed by Wang and Yu [1]. The section discusses the impact on the economy due to lockdown and subsequently discuss the need for a responsive framework from industrial, citizen, and government's viewpoint aiding to early lockdown policy formulation and strategy by government and industries.

Industrial Point of View

Globally, government has taken measures to cut the spread by various measures like prohibiting gatherings of the crowd, remote health advice, creating mobile surveillance app for tracking, healthcare system to detect COVID using quantum machine learning methods [11] and tracing people with the virus from studies conducted by Keeling et al. [12]. Though the measures that have been taken till date can help settle the pandemic stronger strategies are envisioned in reducing transmission from community and household with better support for home diagnosis facility, and dealing with the economic consequences of absence from jobs and work for individuals with experiences from past recessions researchers has suggested the impact on the economic backbone can go lower or persistent [13].

Further, the discussion and studies done by Wittkowski [14] and Kissler et al. [15], state that effectiveness on lockdown strategy is not known to impact the spread of the virus. Industries, such as hospitality, the airline, have taken a significant hit, followed by agriculture taking a global drop of more than 20% in demand, and manufacturing had shown a large drop in overall demand [16]. In the present situation with lockdown, corporates and industries have started adopting digital platforms Gaines-Ross [17]. These industries cannot operate with work from the home policy by companies, and due to COVID-19 business has seen disruption in staffing shortage due to self-isolation and lockdown across the globe. Discussions around lifting lockdown in a phased manner and through early planning can help revive economic activities to a great extent, which later would demand effective contact tracing mechanism within industrial setup as well as community, religious places [12]. Hence, the digital solution would play a vital role to support post-lockdown phases in the overall situation, where systems like monitoring, surveillance, detection would need to be developed leveraging IoT, big data systems and AI as the core technology to be in need [18].

Citizen's Point of View

It is important to understand the responsibility of individuals they have to play in controlling the overall spread following

Table 1 Compare and contract review implication and discovering significant scope for GAP analysis

S. No.	Research article	Research objective	Methodology/technique	Limitations	Implications
1.	Defending against the Novel Corona-virus (COVID-19) outbreak: how can the internet of things (IoT) help to save the world? https://doi.org/10.1016/j.hlpt.2020.04.005	Studied how IoT-based smart disease surveillance systems can act as a potential solution to control the current pandemic	Literature review	Limited to the theoretical model	Directing toward conducting more research on automated and effective alert systems for detection and control of the virus
2.	COVID-19: toward controlling of a pandemic https://doi.org/10.1016/S0140-6736(20)30673-5	Article and WHO review for ways and recommendation on controlling COVID-19	Literature review	Need of conceptual framework	A key outcome is to develop a frame-work for outbreak response
3.	The role of masks and respiratory protection against SARS-CoV-2 https://doi.org/10.1017/ice.2020.83	Identifying the role of mask and personal protection again COVID-19 spread	Literature review	Need empirical evidence	N95 and Surgical mask acting are classified as key aspect to fight SARS-CoV-2
4.	Wearing face masks in the community during the COVID-19 pandemic https://doi.org/10.1016/S0140-6736(20)30918-1	Studies whether only wearing mask to control spread of virus or together with social distancing and personal hygiene is also important	Literature review	Need empirical evidence	No empirical evidence of masks in infection control
5.	Face masks against COVID-19: an evidence review https://doi.org/10.20944/preprints202004.0203.v1	Study the factor of lowering community transmission using face masks	Literature review	Need empirical evidence	The review focus on providing evidence from literature to use mask for controlling spread and help frame policy around use of non-medical masks in public
6.	Rational use of face masks in the COVID-19 pandemic https://doi.org/10.1016/S2213-2600(20)30134-X	Study the need of face mask in community settings	Literature review	Limited to the theoretical recommendations from WHO and health agencies	Face masks are recommended by WHO, ICMR and other health agencies to prevent potential asymptomatic or presymptomatic transmission
7.	COVID-19: protecting worker health https://doi.org/10.1093/annweh/wxaa033	Discuss use of PPE, effect of wearing mask and social distancing	Literature review	Need empirical evidence	Debates on urgent need for research on control measures to protect workers and prevent spreading
8.	Scientific and ethical basis for social distancing interventions against COVID-19 https://doi.org/10.1016/S1473-3099(20)30190-0	Discuss impact of social distancing on spread of virus	Literature review	Need empirical evidence	Focuses to create evidence-based intervention for public communication
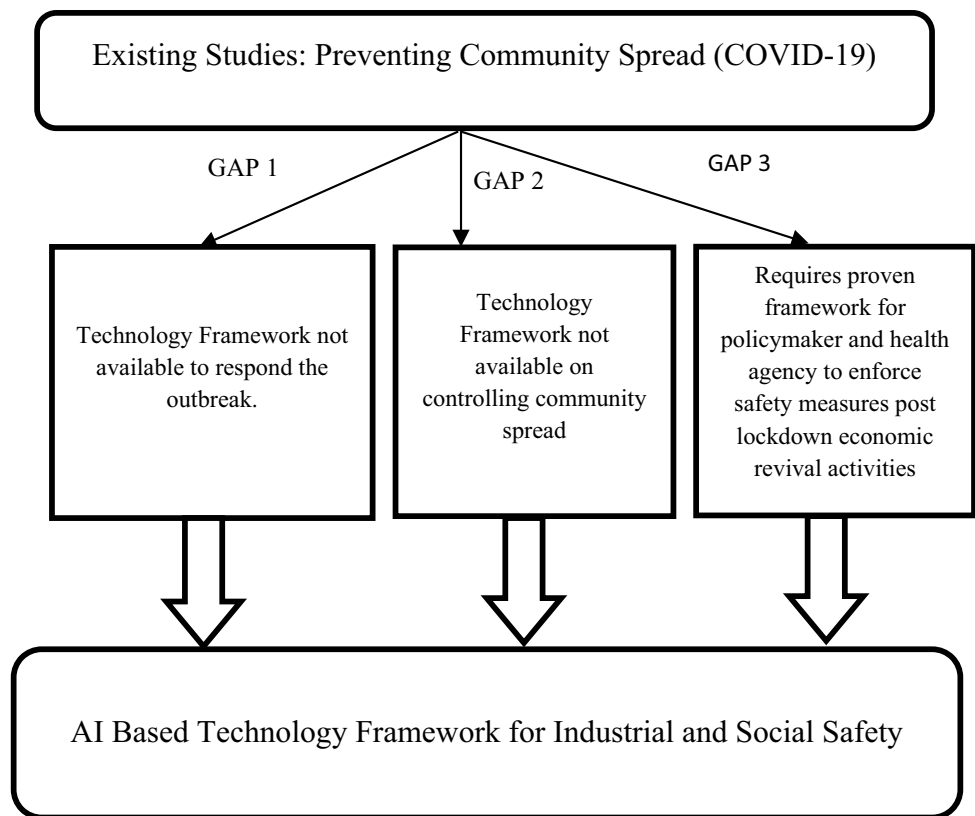
the guideline from local government and WHO recommendation of practicing social distancing, personal hygiene, and wearing the mask in public places [19]. Despite communication from health bodies, individual's awareness around the COVID-19 threat was limited, and people did not adhere to social distance as practice despite mandatory steps outlined as controlling mechanisms by WHO. Social distancing could be an effective way to reduce mortality and spread rate in any setup due to the droplet nature of the virus that can sustain in the environment [20].

The recommendation from health agencies and published government guidelines, people would strictly need to follow social distancing and follow respiratory hygiene [21] when in public. Since vaccine development is a time-consuming process, the spread can only be curb through social reformation and with individual efforts within both societal and industrial setup.

Government's Point of View

Some Asian countries had shown success in controlling the pandemics through testing, contact tracing, and quarantine strategy along with moderate or strong social distancing measures. China leverages intrusive surveillance technology for controlling the virus's spread tracking monitor citizens to establish safety protocol. Similarly, many other countries are now adopting use of technology in many form

to communicate and trace people during this pandemic, like Ministry of Electronics and communication, India has mandated use of mobile tracking application 'Health Bridge' for contact tracking according to official website of MeitY (mygov.in). EU Members States, backed by the commission, have rolled out mobile app 'eHealth' for automated contact tracking which is more efficient compared to manual effort which is time-consuming and expensive.

The government needs to be responsive about to pandemic situation with the citizens, create awareness campaigns, social drills through border forces, and strengthening the disasters-humanitarian coordination. Policymakers need to create effective models for forecasting, helping in making the right decisions in a timely way, even with such uncertainty around COVID-19 containment. Their post-lockdown strategy is a more crucial stage for government to enforce tough decisions around individuals responding to how best to prevent transmission through governed actions by industries and individuals returning to work [19]. Additionally, looking at the economic and social aspects, measures based on isolation are not sustainable in the long run. An extended drawn economic shutdown would create negative health consequences [22]. Several studies depict the role of technology it has played during earlier epidemics, to battle the current situation policymakers and health agencies need to lay a stronger documented lockdown exit strategy keeping post-implication in view and therefore leverage a strong

technology framework for contact tracing and surveillance [17].

Proposed Model and Framework

In this section, the paper proposes the need for building surround outbreak response system (Fig. 2), aiding in tracking and tracing safety-related concern in industrial and societal setup. The overall need for contact tracing is important to control infectious diseases and its spread [23]. The framework would encompass video feeds from a surveillance camera and IoT edge devices placed inside the industrial setup or public places to track people's movement. The architecture proposed here is a hybrid design approach to facilitate feeds from existing cameras as well as IoT devices with edge computing environments on the cloud. Light edge devices, such as Intel Movidius, Nvidia Jetson, or heavy edge devices, like Nvidia Tesla or Intel FPGA and Cloud Environment for training and testing large-scale object detection model.

The framework is focused on leveraging edge computing for detecting face masks in certain environmental setup (workshops, hospitals, industrial premises). Edge computing can deliver swift localized events, near real-time insights, and reduction in overall cost due to efficient local data management and operations [24].

The positive side of deploying deep learning models on edge is to tackle bandwidth-related challenges and providing extensive data security being PII in nature Satyanarayanan et al. [25]. The edge devices can hold relatively lighter deep learning models, process raw information in a smaller size of image frames [24]. The research work extends the concept of multimodal face detection and tracking of people in workshops, public and community places like temples, mosques, government offices, public transports, and offices. The paper stirs work by the researcher [26] using facial-recognition technique to examine the spacio-temporal behavior of individuals. Further, the research work performs advanced experiments in a controlled environment to deduce optimal algorithms for object detection and estimate social distancing criteria from image frames. The sub-sections describe the data collection strategy for the experimental design, feature engineering stages applied, and algorithm implementation benchmarking.

System Design: Edge-Cloud Computing

The section describes the foundation of edge computing architecture for large-scale image analytics and inferencing problems. The system architecture for edge computing environment is a three-tier architecture [27] described in Figs. 3 and 14, respectively. The key advantages of leveraging edge computing are:
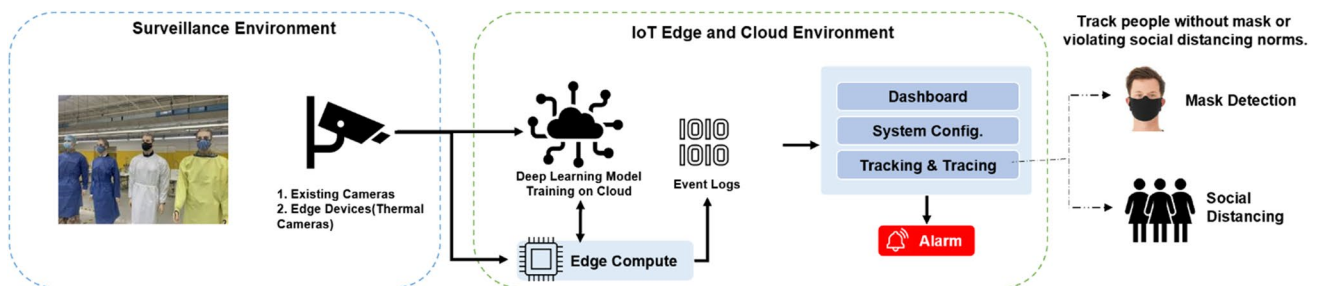

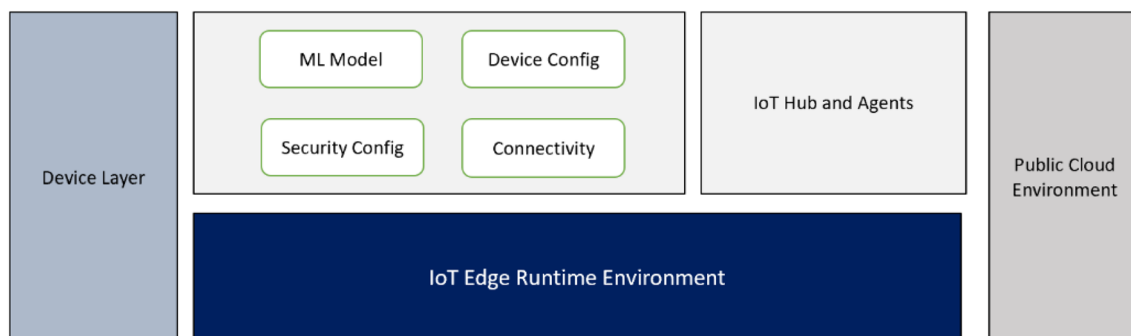**Fig. 2** The high-level design of outbreak response system for tracking and tracing

Fig. 3 Three-tier edge computing architecture

- Low-latency access due to localization of compute environment, storage, and networking locally.
- Lessened bandwidth utilization due to intelligent aggregation and filtering of data to be transited for the purpose of training complex models.
- Localization of models needs intermittent access to the internet.
- Localization of Machine learning inferencing for models trained on public cloud.

The architecture explains the design of edge computing environment, the device layer connects edge devices supporting local compute to run deep learning models as compact models, provides support for connected device configuration to connect multiple mash of edge cameras, security configuration to extend the data privacy together mapped to surveillance cameras IoT edge runtime. The machine learning model training to happen in the public cloud environment and tiny trained model snapshot via docker containers would be deployed using IoT Hub and Agents within the device environment [28]. While edge computing delivering resilience to overall system design from computation standpoint, the challenge of training large deep learning models and deploying for inferencing is more trivial task. The next section analyzes the relevant deep learning models for fast training and deployment with relevant benchmarking of performance and accuracy conducted on different datasets.

AI Models

A combinatory system using IoT, Edge, and Artificial Intelligence would provide the enterprise with a wide range of new services and business opportunities for industries and helps companies create new value [29]. The working description of the model is represented as a flow diagram in Fig. 4. The model needs edge cameras deployed within the manufacturing floor shops or inside office areas continuously

infer leveraging the machine learning models built in. The device capturing the area motion feeds converts video to image frames and keep in small compact patches.

The role of edge cameras is to collect, and process feeds locally using the deployed DL models. The sequence of inferencing happens at the edge environment executing people and mask detection ML module, image segmentation ML module, and social distancing classifier from the segmented image. The processed feeds are stored in a native cloud environment for further alerting and tracking mechanisms through MIS dashboards. The next subsequent sections discuss the overall system design from IT implementation perspective and algorithms evaluated for the completeness of the framework.

The application of deep convolutional neural networks has attained state-of-the-art results in solving computer vision problems like object detection, semantic segmentation, human pose estimation (Fig. 5). The mask detection is a subset domain of object detection technique, and the social distancing is of semantic segmentation. The object is the problem domain to determine where objects are located within an image, i.e., called object localization and which class each object belongs to called object classification Zhao et al. [30]. Semantic segmentation deals with the problem of assigning a class label to each pixel within an image [28]. The conceptual descriptions of models selected for conducting the benchmarking are elucidated in subsequent sections below.

Region-Based Convolutional Neural Network (R-CNN)

The model uses selective search technique Uijlings et al. [31] in contrast of comprehensive search method in an image to detect region proposals. The initialization happens over a small region in an image merging all the regions with a hierarchical grouping. The final group are the boxes containing the entire image. The discovered
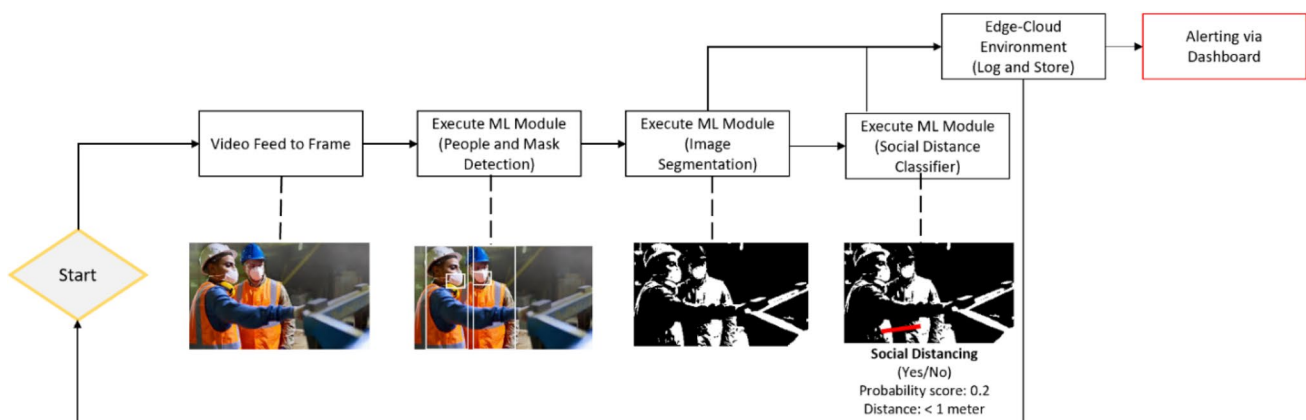

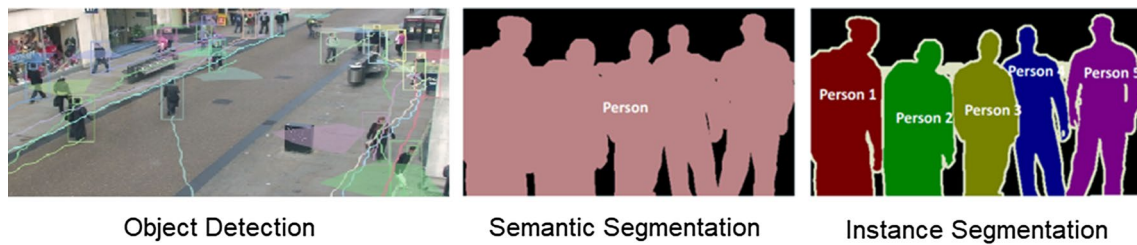**Fig. 4** Flow diagram of contact tracing model (mask detection and social distancing)

Fig. 5 Object detection and segmentation

regions are further combined based on color spaces and similarity metrics. The output are rare numbers of region proposals which could comprise of an object by melding small regions. How the algorithm would function is described in Fig. 6 below.

Fast R-CNN

Like R-CNN, the paper also benchmarks Fast R-CNN. The advantage of Fast R-CNN (Fig. 7) is to reduce overall model computational expense that was happening in R-CNN version. Here, instead of multiple ConvNet being applied for each region, a single ConvNet takes the entire patch of image. A layer of regions is detected with a search method and fed into a fully connected layer creating feature vector and passed through SoftMax classifier to predict object.

Faster R-CNN

Fast R-CNN has some limitations, the search algorithm used in region detection is slow and the model replaces it with fast neural network with a novel concept of regional proposal networks. The model considers each location in previous feature-map and consider r different boxes. Then from each box, the outputs containing an object or not are selected and fed into Fast R-CNN.

YOLO (You Only Look Once)

YOLO was developed by Facebook AI research group [. The architecture of YOLO is much faster that can process in real time at 45 fps. The proposed model considers object detection as a regression problem to spatially separated bounding boxes and associated class probabilities. Not like region proposal or sliding window methods, YOLO observes an entire image during the training
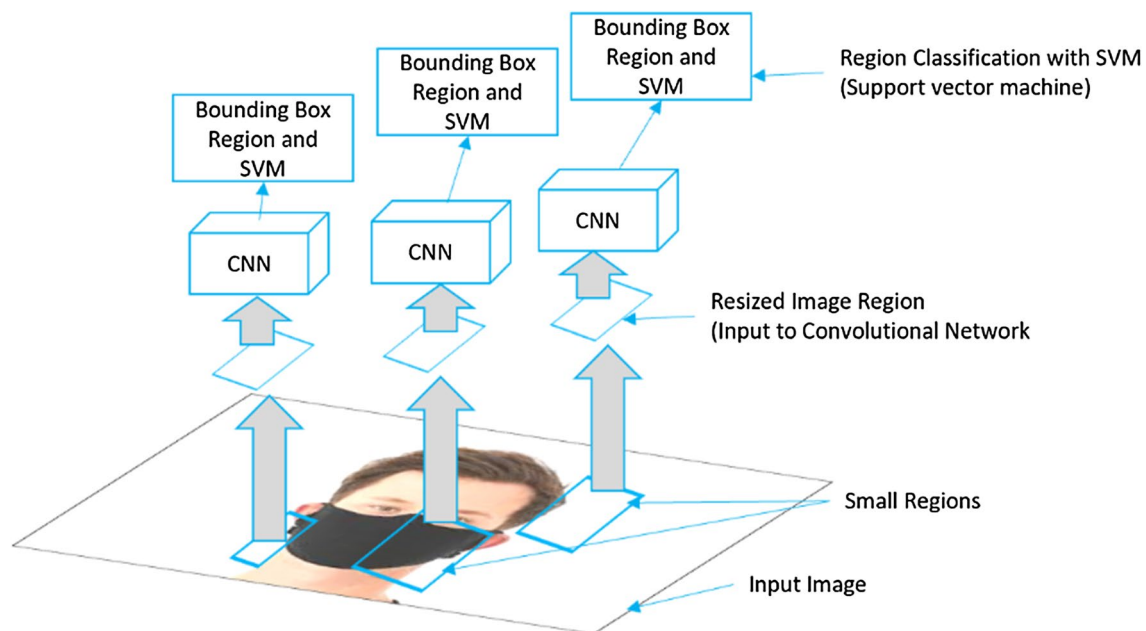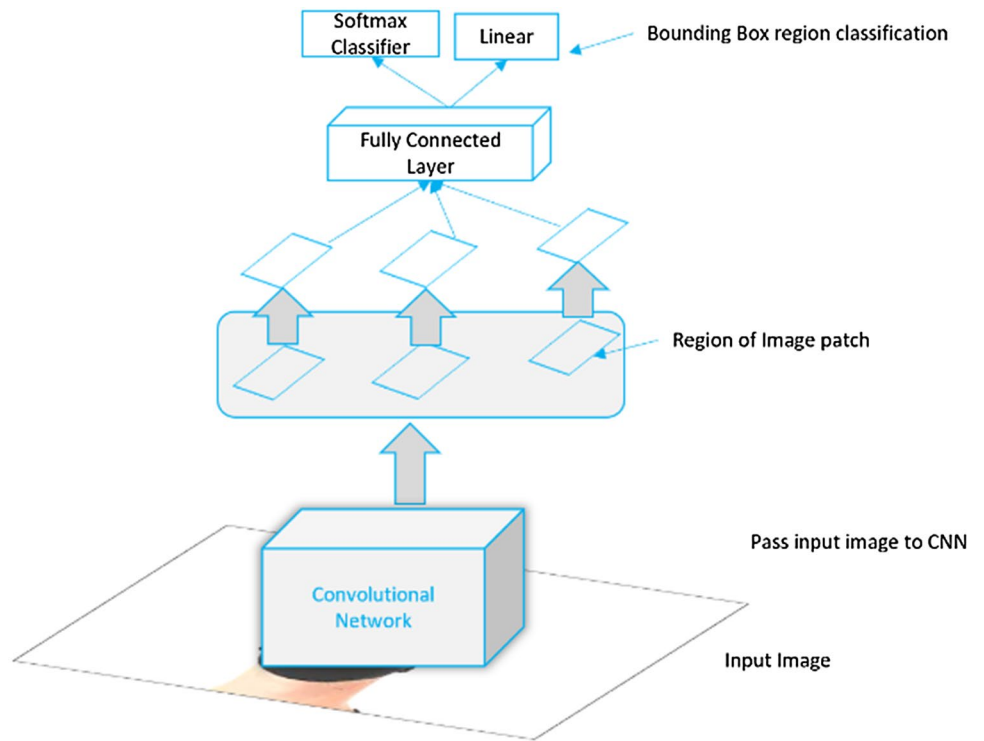

Fig. 6 R-CNN model

Fig. 7 Fast R-CNN model

process and therefore implicitly encodes contextual information about classes as well as their appearance. The architecture of YOLO is described in Fig. 8 below. The method does pre-training of convolutional layers on ImageNet classification task with 224 × 224 input image resolution and thereafter fold twice the resolution for detection.

Semantic Segmentation Using HRNetV2

The HRNetV2 (Fig. 9) recently has advanced in semantic segmentation problem, helping segment objects in an image. The method uses high-resolution representation by combining up sampled representations from all parallel convolutions in contrast to choosing representation from the high-resolution convolution in deep high-resolution representation learning model making it powerful for segmentation
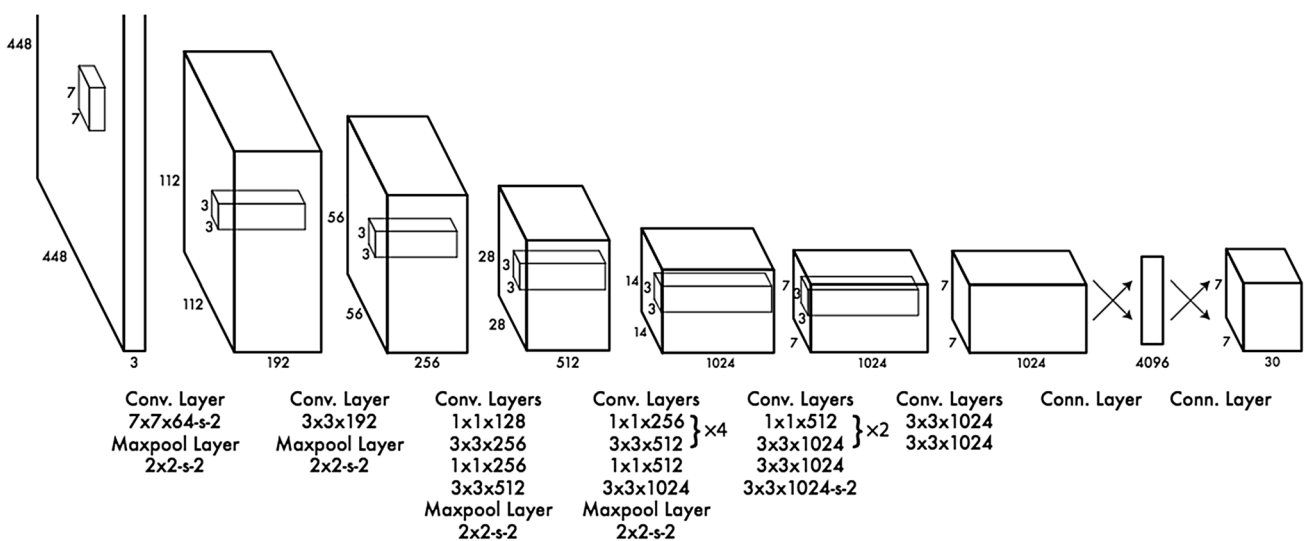

Fig. 8 Yolo network has 24 convolutional layers followed by two fully connected layers
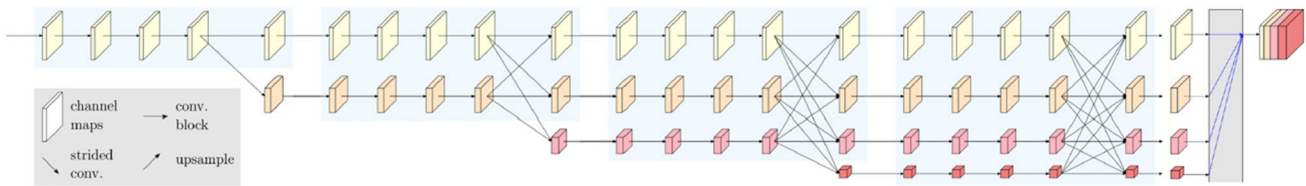
Fig. 9 High-resolution network architecture Sun et al. [32]

problems. The scope of the model can further be extended to calculate distance between two segmented using distance measure functions, such as Euclidian distance, hamming distance, partition distance measure functions [33].

Methods

The section discusses the approach pursued for model development. Specifics on the overall data collection process followed by key algorithmic implementation for the study are outlined in upcoming sections.

Data, Pre-processing and De-biasing

This section illustrates the data assimilation, pre-processing steps examined for the design of the experiments. The data collection was performed using open-source datasets repositories like msropendata.com, data.world and datasetsearch.research.google.com, pjreddie.com, town center dataset and image segmentation model on PASCAL VOC dataset for overall experimentations and modeling. Two sets of datasets were collected for the problem, (a) images of individual personals with mask, without masks in group and unaccompanied. (b) Images representing group of people in a cluster with some distance and others without distance observed. A de-noising was performed on images to scale the image quality and resizing was performed for fitting the model for end experiments and create fairer model.

A known problem with image datasets is determining existence of visual representation bias of women over men. This is widespread in the history of journalism and advertising [34–36]. Hence, as key step during pre-processing stage, an appropriate step was taken for producing fair samples toward specific gender shade. An automated labeling tool was used to provide label to the images tagged as 'with-mask' or 'without-mask' and correspondingly label as 'following-distance' or 'not-following-distance'. The selected images from the large, collected sample corpora were first labeled using the outline tool. Post labeling a bootstrapping, resampling technique was applied on the entire image dataset to set free from any representational bias [37]. For gender-specific biases, a de-biasing method was employed

to identify the male–female representation by a separate model used for identifying male and female in each image and later providing statistical estimation of frequency of the entire dataset. The process was an important step recently researchers have demonstrated how many algorithms have seen demonstrating biases [38]. Crawford and Paglen [39] deliberately argues how ImageNet dataset that is extensively used for training image-labeling algorithms exhibit biases in more than 14+ million images collected through web scrapping from the internet.

Experiments

The section illustrates details around the actual experimental setup and specifics related to model implementation, algorithmic implementation of social distance calculation process, and their evaluation criteria have been discusses. The experiment design phase involved selecting a base HRNET model for initial trials and rapidly performed A/B testing across other models using the similar set of datasets collected.

Mask Detection

The experiment adopted Yolo, Microsoft Computer Vision, Fast-RCNN, RCNN to bench-marked across samples of data chosen for identifying objects in an image. The machine learning models were trained on the collected dataset, Pascal VOC dataset and COCO dataset [40]. In previous work, authors have proposed transfer learning approach [41] fine-tuned on the MobileNetV2 model Yadav [42]; however, the architecture is not tested on large-scale data problems on edge systems. The paper further evaluates models for runtime as well as fastest inferencing on commercial edge devices. Here as part of experiments, Microsoft Vision AI DevKit (Fig. 10) was used to simulate the environment. The simulation environment architecture is depicted in Fig. 11 below.

The simulation tasks leveraged Vision AI dev kit to deploy custom YOLO model trained on collected datasets consisting of images of individuals with and without mask using Azure ML services. The model was further tested on real subjects in a lab environment setup for evaluation.
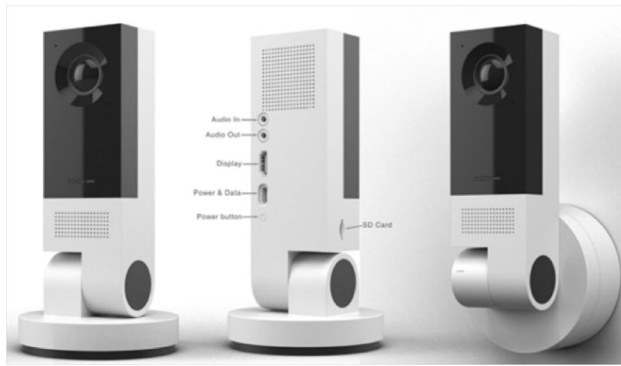
Fig. 10 Azure IoT Vision AI Development KIT (Vision AI Development Kit-Qualcomm Developer Network)

Social Distance Calculation

The paper introduces a distance calculation algorithm to calculate the social distance score of a segmented image using HR-Net segmentation model. The core segmentation model is developed using HRNET and object contextual representation transformer architecture [43]. The transformer pipeline is illustrated in Fig. 12. The calculation algorithm design is shown below in Table 2.

The designed calculation algorithms were further deployed as custom model following mask detection method in the AI dev kit environment for simulations. The scope of the simulation was limited to limited objects and human identifications in a certain environmental setup chosen for this study. The models were evaluated as cross-reference benchmarking activities discussed in upcoming results sections.

Results and Discussion

The paper discusses the use of deep learning algorithms to accomplish two problems, first is detecting people with or without a mask and then predicts if social distancing is observed as image segmentation problem. The experimental results validate the performance of HRNetV2 for generating predictions for measuring social distancing score. The paper compares the computational performance of various algorithms suitable for edge inferencing and the overall accuracy of algorithms for the problem domain. The methodology adopted for benchmarking in the paper is based on secondary data collection and experimental run using existing benchmarking tools. There has been significant research work done in the space of object detection, and semantic segmentation within deep learning space and models like HRNET, Yolo, R-CNN, F-CNN, Mask-RCNN (refer to Appendix 1 for nomenclature) have proven to have yielded significant breakthrough results in image and video analytics space [44].

Edge devices work on local inferencing or performing real-time inference on the camera itself. The technology has no transmission delay, and errors can be debug faster than the previous method, it becomes essential to evaluate DL models that are computationally inexpensive in training and deployment as well as provide better efficacy to the problem domain. The paper discusses the benchmarks for MM detection toolbox presented in Table 3. The MM detection toolbox is developed by Multimedia lab from The Chinese University of Hong Kong. It compares Mask RCNN and RetinaNet to help evaluate the best deep learning method to adopt for faster computing on GPU based cloud environment and edge-based local inferencing.

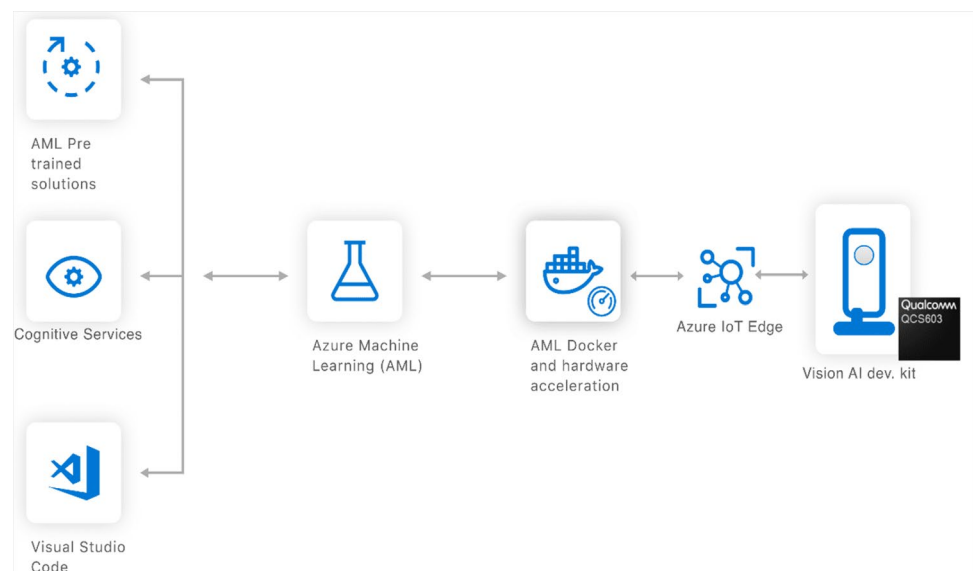**Fig. 11** Edge simulation architecture

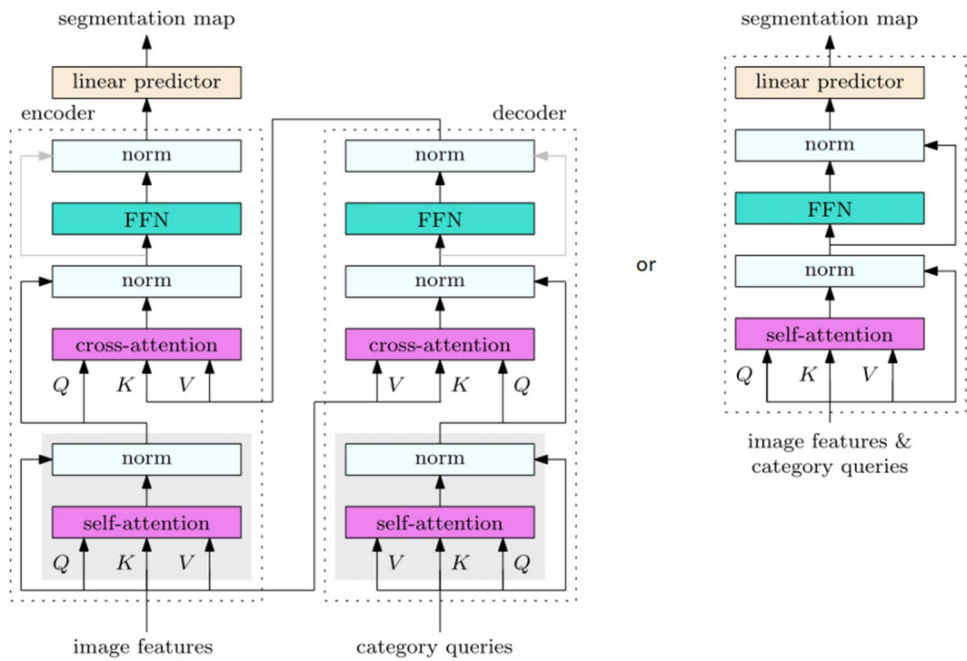**Fig. 12** HRNET + object contextual representation transformer pipeline

Table 2 Social distance calculation algorithm

Step 1: The model at first is set by the weights pre-trained on ImageNet dataset

Step 2: The semantic segmentation of an image frame is obtained from the above step

Step 3: The segmented images (Fig. 13) is further taken as input for edge detection using Canny, dilation and erosion for removing any gap between object edge

Step 4: Detect contours for shapes of the objects in the edge-map using findContours method in OpenCV

Step 5: Loop over contours individually, then rotated bounding box is calculate of the contour using minAreaRect and BoxPoints method in OpenCV

Step 6: Re-ordering the contours to organize in defined top-left, top-right, bottom-right and bottom-left order to draw the rotated bounding box and then calculate the center of the bounding box

Step 7: To calculate distance between each object, the algorithm starts considering each contour starting with left-most as initial reference, then keeps on calculating the mid-point between top-left and top-right points followed by top-right and bottom-right points

Step 8: In final stage, Euclidean distance is calculated between mid-points for final handling of reference object reconstruction
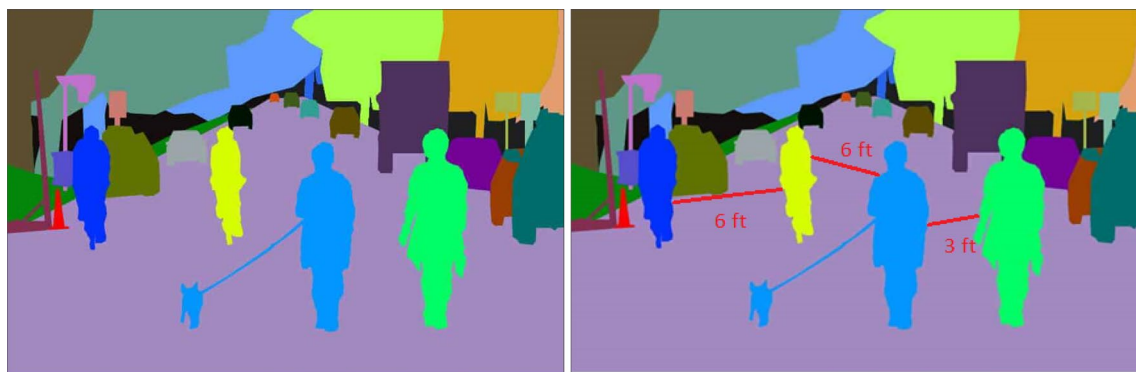

Fig. 13 Segmented image (left) + distance calculation between each object (right)

Table 3 MM detection analysis [45]

Model	Train (iter/s)	Inf (fps)	Mem (GB)	AP_{box}	AP_{mask}
Mask RCNN	0.43	10.8	3.8	37.4	34.3
Mask RCNN	0.436	12.1	3.3	37.8	34.2
Mask RCNN	0.744	8.1	8.8	37.8	34.1
Mask RCNN	0.646	8.8	6.7	37.1	33.7
RetinaNet	0.285	13.1	3.4	35.8	–
RetinaNet	0.275	11.1	2.7	36	–
RetinaNet	0.552	8.3	6.9	35.4	–
RetinaNet	0.565	11.6	5.1	35.6	–

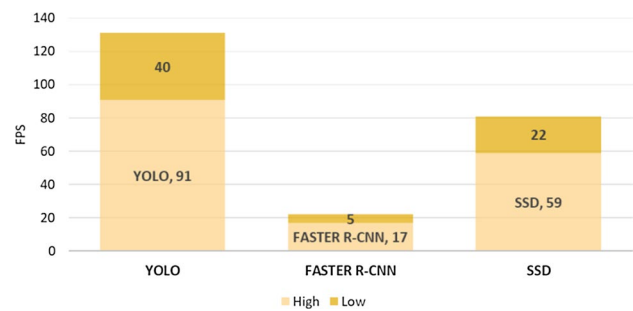**Fig. 14** FPS capability of various object detectors

The paper further contrasts overall performance of processing frames per second (highest and lowest) of various object detection models (Fig. 14) on Town Centre Dataset and image segmentation model on PASCAL VOC dataset. The visual object classes' dataset is standardized image data used for object recognition problem. The benchmarks of real-time systems on PASCAL VOC 2007 and 2012 dataset are represented below in Table 4. The analysis provides higher degree of comparison on algorithms to be selected and studied further.

Further to the analysis, below are training comparisons conducted on Cityscape's dataset. The model training and testing were conducted with an image with input size of 512×1024 and 1024×2048. The models were set with the weights pre-trained on the ImageNet on Small (Table 6) and large model (Table 6). The analysis outcome is represented using mean Intersection over Union (mIoU). Intersection over Union evaluation metric is used for object detector on a database to measure the accuracy represented as area of overlap by area of union.

The model comparison conducted encompasses small and large models. The small model is trained on smaller set of parameters ranging from 15 to 150 parameters on different models taken for benchmarking evaluated on mIoU shown in Table 5, whereas large model in Table 6, illustrates the algorithmic comparison on large parameters set 65.8 million. The analysis however limits to AI models on a limited dataset and, future empirical research work needs to be conducted to validate large-scale deployment and adoption across various societal and industrial setup.

Societal Implication

The situation of pandemic and COVID-19 situation across the globe has forced the government agencies, industrial ecosystem, and educational setup to comply and reinforce standard safety protocols for individuals to break the spread of the virus. There has been sheer need to build technology solutions for different setup to reduce the impact of spread

[46]. The study provides relevant technical solutions for policymakers and technology innovators to build systems around the problem to reduce adverse impact in post-pandemic. The study provides future directions to adopt new technology systems and help design frameworks to keep safety as key aspect while planning for return-to-work or return-to-school stages when currently both are prohibited due to surge in cases worldwide.

Conclusion

To summarize, the paper discusses an applicable framework using artificial intelligence on edge computing for adopting an outbreak response system for contact tracing, mask detection, and detecting social distancing measures from video feeds in the surveillance systems. The benchmarks conducted on selected models to gage the performance and accuracy yield a deeper investigational analysis to further help in the overall implementation of such a solution in industrial practice. Overall YOLO outperforms in object detection task and is faster enough for complex edge inferencing and HRNetV2 outperform semantic segmentation problem applied to solve social distancing prediction.

The study provides directional support for new technology adoption for policymakers from identified gaps in the current situation of COVID-19 spread. It discusses various

Table 4 Performance and speed comparison of models on PASCAL VOC dataset on edge compute

Model	Speed (real-time)	mAP (%)	FPS
YOLO	Yes	63.4	45
Fast- YOLO	Yes	52.7	155
YOLO-VGG-16	No	66.4	21
Fast R-CNN	No	70.1	0.5
Faster R-CNN VGG-16	No	73.2	7
Faster R-CNN ZF	No	62.1	18

Table 5 Small model

Selected model(s)	Number of parameters	Multi-scale	Flip	Distillation	mIoU
ICNet	–	No	No	No	70.6
ResNet18 (1.0)	15.2	No	No	No	69.1
ResNet18 (1.0)	15.2	No	No	Yes	72.7
MD (enhanced)	14.4	No	No	No	67.3
MD (enhanced)	14.4	No	No	Yes	71.9
SQ	–	No	No	No	59.8
CRF-RNN	–	No	No	No	62.5
Dilation10	140.8	No	No	No	67.1
MobileNetV2Plus	8.3	No	No	No	70.1
MobileNetV2Plus	8.3	No	No	Yes	74.5
HRNetV2-W18-Small-v1	1.5 M	No	No	No	70.3
HRNetV2-W18-Small-v2	3.9 M	No	No	No	76.2

Table 6 Large model

Selected model(s)	Number of parameters	Multi-scale	Flip	mIoU
HRNetV2-W48	65.8 M	No	No	80.9
HRNetV2-W48	65.8 M	No	No	81.2
HRNetV2-W48	65.8 M	Yes	Yes	80.5
HRNetV2-W48	65.8 M	Yes	Yes	81.1
HRNetV2-W48	65.8 M	Yes	Yes	81.5
HRNetV2-W48	65.8 M	Yes	Yes	81.9

GPU: graphical processing unit	TP/TN: true positive/true negative
SGD: stochastic gradient descent	DL: deep learning
IoT: internet of things	PASCAL: pattern analysis statistical modeling and computational learning
FPS: frames per second	VOC: visual object classes

aspects of new technology development from government point of view that could be leveraged for contact tracking and tracing in community setup once the lockdown is lifted to resume the economy. Not only the government, but the framework would also assist and encourage industry practices to adopt and build stronger surveillance systems within the workplace environment for deeper adherence to social distancing, physical hygiene, and further thermal screening using thermal cameras to help in combating spread to a large extent.

Declarations

Appendix 1: Terminology Classification

CNN: convolutional neural network	FPGA: field programmable gate array
R-CNN: regions with CNN features	YOLO: you only look once; an object detection system trained on COCO dataset
HRNet: high-resolution networks	mPA: mean average precision
COCO: common objects in context	FPS: frames per second

References

1. Wang Q, Yu C. The role of masks and respiratory protection against SARS-CoV-2. Infect Control Hosp Epidemiol. 2020. https://doi.org/10.1017/ice.2020.83.
2. Howard J, Huang A, Li Z, Tufekci Z, Zdimal V, van der Westhuizen HM, von Delft A, Price A, Fridman L, Tang LH, Tang V, Watson GL, Bax CE, Shaikh R, Questier F, Hernandez D, Chu LF, Ramirez CM, Rimoin AW. An evidence review of face masks against COVID-19. In: Proceedings of the National Academy of Sciences of the United States of America. 2021;118(4):e2014564118. https://doi.org/10.1073/pnas.2014564118.
3. Semple S, Cherrie JW. Covid-19: protecting worker health. Ann Work Expo Health. 2020. https://doi.org/10.1093/annweh/wxaa033.
4. Bedford J, Enria D, Giesecke J, Heymann DL, Ihekweazu C, Kobinger G, et al. COVID-19: towards controlling of a pandemic. Lancet. 2020;395(10229):1015–8.

5. Srivastava PR, Sengupta K, Kumar A, Biswas B, Ishizaka A. Post-epidemic factors influencing customer's booking intent for a hotel or leisure spot: an empirical study. J Enterp Inf Manag. 2021. https://doi.org/10.1108/JEIM-03-2021-0137.
6. Greenhalgh T, Schmid MB, Czypionka T, Bassler D, Gruer L. Face masks for the public during the COVID-19 crisis. BMJ. 2020;369:m1435. https://doi.org/10.1136/bmj.m1435.
7. Jefferson T, et al. Physical interventions to interrupt or reduce the spread of respiratory viruses. The Cochrane Database Syst Rev. 2011.
8. Siddikur Rahman Md, Peeri NC, Shrestha N, Zaki R, Haque U, Hamid SHA. Defending against the Novel Coronavirus (COVID-19) out-break: how can the internet of things (IoT) help to save the world? Health Policy Technol. 2020. https://doi.org/10.1016/j.hlpt.2020.04.005.
9. Bibi SE, Krogstie J. Smart, sustainable cities of the future: an extensive interdisciplinary literature review. Sustain Cities Soc. 2017;2017(31):183–212.
10. Fernandes N. Economic effects of coronavirus outbreak (COVID-19) on the world economy (March 22, 2020). 2020. https://doi.org/10.2139/ssrn.3557504.
11. Sengupta K, Srivastava PR. Quantum algorithm for quicker clinical prognostic analysis: an application and experimental study using CT scan images of COVID-19 patients. BMC Med Inform Decis Mak. 2021;21:227. https://doi.org/10.1186/s12911-021-01588-6.
12. Keeling MJ, Hollingsworth TD, Read JM. Efficacy of contact tracing for the containment of the 2019 novel coronavirus (COVID-19). J Epidemiol Community Health 2020;74:861–866.
13. Correia S, Luck S, Verner E. Pandemics depress the economy; public health interventions do not: evidence from the 1918 flu. Mimeo. 2020.
14. Wittkowski K. The first three months of the COVID-19 epidemic: Epidemiological evidence for two separate strains of SARS-CoV-2 viruses spreading and implications for prevention strategies. 2020. https://doi.org/10.1101/2020.03.28.20036715.
15. Kissler SM, Tedijanto C, Goldstein E, Grad YH, Lipsitch M. Projecting the transmission dynamics of SARSCoV-2 through the postpandemic period. Science (New York, N.Y.), 2020;368(6493):860–868. https://doi.org/10.1126/science.abb5793.
16. Nicola M, Alsafi Z, Sohrabi C, Kerwan A, Al-Jabir A, Iosifidis C. et al. The socio-economic implications of the coronavirus pandemic (COVID-19): A review. Int J Surg 2021;78:185–193. https://doi.org/10.1016/j.ijsu.2020.04.018.
17. Ting DSW, Carin L, Dzau V, et al. Digital technology and COVID-19. Nat Med. 2020;26:459–61. https://doi.org/10.1038/s41591-020-0824-5.
18. Statista (2020) COVID-19/coronavirus. Accessed 31 Mar 2020.
19. Gaines-Ross L. "Reputation Warfare". Harvard Bus Rev. 2010;88(12):70–76.
20. Anderson R, Heesterbeek H, Klinkenberg D, Hollingsworth T. How will country-based mitigation measures influence the course of the COVID-19 epidemic?. Lancet. 2020;395(10228):931–934. https://doi.org/10.1016/s0140-6736(20)30567-5.
21. Leung NHL, et al. Respiratory virus shedding in exhaled breath and efficacy of face masks. Nat Med. 2020. https://doi.org/10.1038/s41591-020-0843-2.
22. Gilbert M, Dewatripont M, Muraille E. et al. Preparing for a responsible lockdown exit strategy. Nat Med. 2020;26:643–644. https://doi.org/10.1038/s41591-020-0871-y.
23. Chen H, Yang B, Pei H, Liu J. Next generation technology for epidemic prevention and control: data-driven contact tracking. IEEE Access. 2019;7:2633–42. https://doi.org/10.1109/ACCESS.2018.2882915.
24. Ananthanarayanan G, Bahl V, Cox L, Crown A, Nogbahi S, Shu Y. Demo: video analytics - killer app for edge computing. In: The 17th annual international conference on mobile systems, applications, and services (MobiSys'19), 17–21 Jun 2019, Seoul, Republic of Korea. ACM, New York, NY, USA. https://doi.org/10.1145/3307334.3328589.
25. Satyanarayanan M, Simoens P, Xiao Y, Pillai P, Chen Z, Ha K, Hu W, Amos B. Edge analytics in the internet of things. IEEE Pervas Comp. 2015;14:24–31.
26. Shi W, Cao J, Zhang Q, Li Y, Xu L. Edge computing: vision and challenges. IEEE Internet Things J. 2016;3:637–46.
27. Ganansia F, Carincotte C, Descamps A, Chaudy C. A promising approach to people flows assessment in railway stations using standard CCTV networks. In: Proceedings of the transport research arena (TRA) 5th conference: transport solutions from research to deployment, Paris, France, 14–17 Apr 2014.
28. Ray P. A survey on Internet of Things architectures. J King Saud Univ Comput Inf Sci 30(3):291-319. https://doi.org/10.1016/j.jksuci.2016.10.003
29. Hui G. How the internet of things changes business models. Harvard Bus Rev. 2014;95:1–5.
30. Zhao ZQ, Zheng P, Xu T, Wu X. "Object Detection With Deep Learning: A Review," in IEEE Transactions on Neural Networks and Learning Systems. 2019. vol. 30, no. 11, pp. 3212–3232. https://doi.org/10.1109/TNNLS.2018.2876865.
31. Uijlings JRR, van de Sande KEA, Gevers T. et al. Selective Search for Object Recognition. Int J Comput Vis. 2013;104:154–171. https://doi.org/10.1007/s11263-013-0620-5.
32. Sun K, Xiao B, Liu D, Wang J. Deep High-Resolution Representation Learning for Human Pose Estimation. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2019. pp. 5686-5696. https://doi.org/10.1109/CVPR.2019.00584.
33. Chen K, Wang J, Pang J, et al. MMDetection: open MMLab detection toolbox and benchmark. https://arxiv.org/abs/1906.07155. 2019.
34. Porumbel D, Hao J, Kuntz P. An efficient algorithm for computing the distance between close partitions. Discrete Appl Math 2011;159(1):53-59. https://doi.org/10.1016/j.dam.2010.09.002
35. Goffman E. Gender advertisements. New York: Macmillan International Higher Education; 1976.
36. Aiello AE, Perez V, Coulborn RM, Davis BM, Uddin M, Monto AS. Facemasks, Hand Hygiene, and Influenza among Young Adults: A Randomized Intervention Trial. PLoS ONE. 2012;7(1):e29744. https://doi.org/10.1371/journal.pone.0029744.
37. Xi N, Ma D, Liou M, Steinert-Threlkeld ZC, Anastasopoulos J, Joo J. Understanding the political ideology of legislators from social media images. 2019. https://arxiv.org/abs/1907.09594. Accessed 10 Oct 2020.
38. Duguid J. The size and the duration of air-carriage of respiratory droplets and droplet-nuclei. Epidemiol Infect. 1946;44:471–9.
39. Becker HS. Photography and sociology. Stud Vis Commun. 1974;1(1):3–26.
40. Ferree MM, Hall EJ. Visual images of American Society: gender and race in introductory sociology textbooks. Gend Soc. 1990;4(4):500–33.
41. Crawford K, Paglen T. Excavating AI: the politics of images in machine learning training sets. Ai & Society. 2021;1–12.
42. Alghamdi A. Impact of the COVID-19 pandemic on the social and educational aspects of Saudi university students' lives. PLoS One. 2021;16(4): e0250026. https://doi.org/10.1371/journal.pone.0250026.
43. Keniya R, Mehendale N. Real-time social distancing detector using social distancing net-19 deep learning network. 2020. https://doi.org/10.2139/ssrn.3669311. SSRN: https://ssrn.com/abstract=3669311. Accessed 7 Aug 2020.

44. Yadav S. Deep learning based safe social distancing and face mask detection in public areas for COVID-19 safety guidelinesadherence. Int J Res. Appl Sci Eng Technol. 2020;8(7):1368–1375.

45. Anfinrud P, Bax C, Stadnytskyi V, Bax A. Could SARS-CoV-2 be transmitted via speech droplets? (2020). https://doi.org/10.1101/2020.04.02.20051177.

46. Redmon J, Divvala S, Girshick R, Farhadi A. "You Only Look Once: Unified, Real-Time Object Detection," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), (2016). pp. 779–788. https://doi.org/10.1109/CVPR.2016.91.