National College of Ireland

Social Distancing and Face Mask Detection using Deep Learning and Computer Vision

MSc Research Project
MSc in Data Analytics

Isha Shete
Student ID: 18181678

School of Computing
National College of Ireland

Supervisor: Vladimir Milosavljvec

National College of Ireland

MSc Project Submission Sheet

School of Computing

Student Name: Isha Shete

Student ID: 18181678

Programme: MSc in Data Analytics **Year:** 2019-2020

Module: Research Project

Supervisor: Vladimir Milosavljvec

Submission Due Date: 23/09/2020

Project Title: Social Distancing and Face Mask Detection using Deep Learning and Computer Vision

Word Count: 6453 **Page Count** 22

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.
ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature: Isha Shete

Date: 22/09/2020

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST

Attach a completed copy of this sheet to each project (including multiple copies)	□
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	□
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	□

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Social Distancing and Face Mask Detection using Deep Learning and Computer Vision

Isha Shete
18181678

Abstract

A novel virus has caused a world pandemic and huge life losses. Declared by the World Health Organization (WHO), this coronavirus originated from Wuhan, China in late December 2019. Upon thorough research, the virus has been observed as pathogenic and transmissible by air or by coming in close contact with an infected person. To avoid the spread of this virus, many measures have been suggested, such as maintaining a social distance, that is, maintaining a proper physical distance between people and lessening close contact with each other, and wearing a face mask to avoid the droplets from transmitting through the air. Therefore, this research paper focuses and aims its study towards implementing a Social Distancing and Face Mask Detection System. This system will implement object detection and facial recognition in the video footages of pedestrians. Pre-trained models such as the YOLOv3, ResNet Classifier and DSFD are used. People violating the minimum distance were detected as well as faces without face-masks were detected. An overall results board is displayed in the output containing the number of people violating or non-violating the respective measures. After implementing and deploying the models, this research project achieved a confidence score of 100%. Therefore, this research project concludes with proven facts that social distancing and wearing face masks helps reduce the spread of the virus and thus builds a model to help detect these measures.

Keywords **–** COVID-19, Social Distancing, Pedestrian Detection, Face Mask Detection, ResNet Classifier.

Contents

1. Introduction

1.1 An Overview of COVID-19

Ever since being discovered first in Wuhan, China in December 2019, the whole world has been affected by the **CO**rona **VI**rus **D**isease **2019**, or COVID-19, causing the World Health Organization (WHO)[1] to declare a pandemic. The coronavirus is a novel disease and it has spread rapidly over 213 countries and territories resulting in an immense impact upon mankind. As of August 16, there are 21,641,369 active cases which might even increase in just a matter of time. With more than 769,481 deaths being reported to date[2], this virus is said to cause a pandemic. The origin of this virus is most likely said to be from bats or snakes but there has been no intermediate host detected yet. The Chinese researchers and health authorities assume that the virus may have been originated and spread by an animal species that was infected and spread to humans via being trafficked illegally in the seafood market in Wuhan. Researchers are still trying to discover the origin of this virus. The coronavirus causes respiratory and gastrointestinal infections and can be categorized into four different types which are Deltacoronavirus, Gammacoronavirus, Alphacoronavirus and the Betacoronavirus. Birds are infected by the first two types whereas mammals are infected by the latter two (Fong, Dey and Chaki, 2020). While medical researchers are primarily working and preparing to develop a vaccine or any type of drug to help in the prevention of this virus, there are other possible social ways that may help in limiting the rapid spread of the virus within individuals.

According to the researchers, this virus can spread while coming in close contact with the person who is affected. Thus, it being alike all the other infectious respiratory diseases, researchers suggest close contact be withdrawn under any and all circumstances as the possibility of the virus transmitting through airborne ways and droplets is quite likely to occur. Therefore, to overcome the spread of coronavirus, researchers have come up with a new social norm, Social Distancing or Physical Distancing. This measure has been recognised and acknowledged to be a very useful measure in controlling the widespread of the virus. It ensures minimal human interaction within close distances on a national scale. The evaluation of the ways in how this social measure has been effective is investigated by studies by analyzing the confirmed and death cases in Italy and Spain during the national lockdown and following the upliftment of the lockdown with the help of statistics on a time-series (Tobías, 2020; Saez, Tobias, Varga and Barceló, 2020). The outcomes depicted that there were considerably lower reported cases after implementing social distancing in the two countries.

[1] https://www.who.int/

[2] https://www.worldometers.info/coronavirus/

1.2 Background and Motivation

Social distancing has been an effective social measure in terms of normalizing the curve. It helps avoid any direct contact amongst individuals and assists in reducing any sorts of transmission of the droplets containing the virus via the human breath. Some studies are determined towards the exploration of trajectories that transmit the droplets via human breathing ways such as coughing, talking, sneezing and eating. While a few studies considered the droplet size to be correlated with the pathogens of the contagious disease, where they considered large droplets to be the ones that were to carry the microorganisms from the infected person (Christian et al., 2004; Mangili and Gendreau, 2005), other studies assumed small droplets of the particles that were in the nucleus form to be the ones to spread distantly. This was a topic of debate for a long time as questions like "what is a safe distance to maintain social distancing?" and "how further can the droplets spread?" arose. Hence, getting a considerable answer to all these questions was difficult as other factors also mattered, such as transmission through the surrounding air by evaporation. Factors like biological, medical and engineering also needed to be considered to answer the above-mentioned questions. Thus, after considering all these factors, the suggested distance for a safe social distance was recommended to be at least 6 feet, that is, 2 meters. Some believed that this distance was not enough and needed to be increased (Setti et al., 2020). Although, the final distance was considered to be 2 meters and was implemented worldwide making it a successful parameter, now the distance has been reduced to 1 meter[3] in some of the recovering countries. Therefore, considering all these factors, this research project proposes an AI based detection system that helps in detecting any kinds of violations. This research intends on building a model that can be applied in real-time systems and thus help in avoiding the spread of the virus.

1.3 Research Objective

This research paper proposes a novel method of an Artificial Intelligence-based real-time social distancing and face mask detection system. In this research paper, pre-trained deep neural network models such as ResNet Classifier, DSFD, and YOLOv3 are used in the detection of individuals and face masks with the help of bounding boxes. The system also measures the number of people present in the camera footage.

Thus, this paper contributes to the following sections:

- Developing a novel and unique implementation of a real-time based social distancing detection as well as face mask detection system.

- Applying various deep learning models to get a more fast and accurate result.

[3] https://www.who.int/westernpacific/emergencies/covid-19/information/physical-distancing

- Analyses carried out on a less crowded region such as the Oxford Town Centre data set.

The following research paper document is structured as follows: (1) Section 2 contains a brief study regarding the work related to this this domain. (2) Section 3 consists of the inspiration, motivation, and the process of this research, that is, the methodology of this research. (3) Section 4 contains the implementation of this research project, that is, applying and understanding the deep neural network models. (4) Section 5 is a segment where the evaluations have been conducted to examine the performance of the models applied. (5) Section 6 consists of the results and conclusion that has been led by this research.

2. Related Work

2.1 Social Distancing for COVID-19

In order to control the current situation regarding the widespread of COVID-19, social distancing is emphasized in limited spaces. The main issue related to this strategy was figuring out what the safest distance was and the answer to that was very crucial in reopening schools and businesses. A study was introduced that utilized a mixed Wells-Riley model to predict the infection probability related to the airborne virus, being the main reason for implementing social distancing (Sun and Zhai, 2020). Their study depicted the distance 1.6m to 3m to be the safest possible social distance while considering the air transmission of the exhaled droplets from coughing, talking, sneezing and so on. An 8.2m distance was suggested if a calm air environment was to be considered. By validating their model, they came to the conclusion that maintaining social distancing can reduce the widespread of the virus by 20% to 40% in the first 30 mins.

While, the other study focused on detecting the mode of its transmission so that appropriate control measures could be applied to slower the spread (Richard et al., 2020). The study, with the help of ferrets, showed that SARS-CoV-2 could be transmitted easily via any kinds of direct contact and through the air as well. The ferrets could be seen getting infected within 1 to 3 days and between 3 to 7 days after being exposed. The study provided experimental proof of the virus' transmission to be airborne, therefore, supporting the implementation of social distancing on a community level that is being currently practiced in several countries.

Another research was conducted to study the effect of social distancing measures that were carried out in 10 countries (Thu et al., 2020). The authors focused on the statistics related to the confirmed cases and the death cases due to COVID-19 in mainly 10 deeply affected countries such as Spain, Italy, Iran, and so on. They focused on how these countries responded to the pandemic between January 11 and May 2, 2020. Social distancing and its relation to COVID-19 cases were used to analyze how effective this measure was. According

to their study, it took 1 to 4 weeks ever since the declaration of social distancing measures was announced to show the signs of decreasing cases. The results varied from country to country due to their different approaches.

Similarly, another research was conducted to study the role of the cultural dimension that was present in the practice of social distancing in the world (Huynh, 2020). The author drew data from Google community and the cultural factors of 58 countries from Hofstede. Thus, the author came up with a conclusion that proposed effective measures to contain the virus by focusing on the role of uncertainties.

On the other hand, a research paper had a different approach to test the significance of social distancing. They formulated a model to rapidly test the individuals that were considered to be asymptomatic (Aldila et al., 2020). By applying their model on the data that they had acquired, they found out that the individuals who had self-isolated or maintained social distancing reduced or delayed the time of an outbreak. Therefore, they concluded by stating that if a strict social distancing measure was carried out in all the countries, it would help in avoiding the outbreak of the virus.

Some researchers in their study were determined to investigate whether the social distancing measures were effective and whether relaxing some measures was possible or not to avoid a second wave (Wu et al., 2020). The research that they had conducted confirmed that social distancing measures that were carried out helped reduce the transmission of the virus and brought down the infection contact rate. In their study, they depicted that improving the case detection rate was necessary to reduce the reproduction of the virus and also there are many other protections measures available that should be carried out in order to improve the personal protection to recompense the strict social distancing measures.

2.2 Pedestrian Detection

Modern times and digital media evolution have led to object detection becoming popular in the computer vision area. Specifically, pedestrian detection has increased, drawing attention towards our everyday lifestyle. With the help of pedestrian detection and it being successful it has led to intelligent driving systems, intelligent monitoring systems, and more. Although there are numerous algorithms available that are deep learning-based it is still a big challenge. Due to the blockage caused by the background, or the changing postures of pedestrians every second impacts significantly on the detection results. Therefore, a study proposed a method based on Mask R-CNN that helped reduce the external factors influence while detecting (Yu et al., 2020). The method experimented on the pedestrian's dataset showed that the proposed algorithm achieved a better detection accuracy than the other convolutional methods.

There was a similar study to the above-mentioned research where they introduced a CNN based pedestrian detection system designed from scratch without utilizing any standard module (Pranav and Manikandan, 2020). The recognition accuracy that they achieved was between 96.73% to 100%. Another aspect of their real-time pedestrian detection system was to help the non-autonomous vehicles as a driver assistant.

To help improve the performance of autonomous vehicles and their driving safety, improving pedestrian detection system is necessary. However certain factors like blurry or decrypted images in the hazy weather makes it difficult to detect the pedestrians. Thus, a paper came up with a solution by presenting three novel approaches using deep learning (Li et al., 2020). They used the linear bottleneck skills and the depthwise separable convolutions to help reduce the number of parameters and the computational costs. They collected images of pedestrians in hazy weather and were further augmented to improve the database. Their results showed that the models they used could detect pedestrians present in hazy weather.

Saeidi and Ahmadi, (2020) proposed a novel model based on the changes found in the camera viewing angles. Their proposed method was integrated with a Single-Shot Detector (SSD) model and a set of parallel Fast R-CNNs to detect pedestrians accurately. Their approach also introduced a novel training method which was based on the camera angles that sought for the best region of interest (ROI). The method was then evaluated and they found it to be better than the other conventional methods.

Pedestrian detection and their future trajectories detection are demanding tasks for various applications, like the autonomous driving systems. There are methods that perform these tasks by dividing these two tasks into two categories, that is, first the detection and then the prediction. A novel research paper presented an end-to-end two-staged network called the Spatio-Temporal-Interactive Network or the STINet (Zhang et al., 2020). In their work, they modelled the pedestrians which was in a 3D format and achieved the temporal information of every individual. Their model also interacted with objects using an interaction graph, that helped them gather information about the neighbouring objects.

2.3 Face Detection

Due to the coronavirus, the WHO has suggested some effective measures and one of them is wearing a face mask in public. A new research paper proposed a hybrid model that included both deep as well as machine learning techniques (Loey et al., 2021). The proposed model included two components, first, with the help of Resnet50, the component was designed to extract the features. And the second component performed classification of the face masks using ensemble methods, SVM and decision trees. Their results showed that the accuracy achieved for SVM was 99.64%, 99.49% and 100% respectively.

Factors such as expression, change in posture, illumination and various other factors influence the face recognition systems due to the lower capability of the single CNN. A study proposed a new face recognition method that was based on the local binary pattern (LBP) and parallel ensemble learning of CNN (Tang et al., 2020). They used the LBP operator to extract features of the facial textures followed by implementing 10 CNNs with 5 distinct network structures for the further extraction of the features in order to train, improve and get the classification result. Later with the help of parallel ensemble learning, the final facial recognition results were generated. They achieved a 100% and a 97.51% of recognition rate by applying this method on their respective datasets. The experiments helped in the improvement of the accuracy as well as the tolerance level of the occlusions present in an image.

The need for automatic face-mask detection systems in today's time is a necessity. Therefore, a proposal was proposed by a research that suggested a face attribute recognition as well as a face detection algorithm (Wang et al., 2020). The proposed method contained mechanisms such as spatial and channel attention and fused feature pyramid and also adopted segmentation for weakly supervised learning. the accuracy they achieved was 99.50% and this they have opened the algorithms for the public to help increase its social value and its application value.

Most of the detection methods detect faces by applying a bounding box around the face which sometimes cannot segment the face from the background image. Hence to overcome this, (Lin et al., 2020) proposed a model based on an improved Mask R-CNN, which included face segmentation and detection in one framework that aimed to obtain a more fine-grained face. They called this model the G-Mask. In their paper, they used the Resnet-101 for feature extraction, and then used the RPN for the generation of the ROIs through an FCN. Later they used the Generalized Intersection over Union (GIoU) as the bounding box loss function which helped in improving the detection accuracy. When compared with the other conventional methods such as Faster R-CNN, and other algorithms, they found out their model to be more promising.

Mangmang (2020) aimed in his study to detect faces and whether they were wearing a mask or not. He also aimed his study towards detecting whether individuals were wearing a mask properly or not, that is, was the nose visible, were the nostrils not covered, and all. The model contained three inception networks, two fully connected layers, three max-pooling layers, and also a dropout layer. In his study, the model was successful and gave an accuracy of classification of 98.5%.

In a research paper, they proposed a technique that was able to detect masked faces by applying four steps (Deore et al., 2016). The four steps were the calculation and estimation of the distance from the camera, facial detection, eye detection and eye line detection. The paper used commonly available algorithms for face and people detection. Their proposed method turned out to be way simpler than the conventional methods thus making real-time detection feasible.

3. Methodology

3.1 Data Collection and Preparation

This research paper uses a two pedestrian videos, the first is obtained online from YouTube[4] which is provided by BriefCam[5] and the other video footage of the Oxford Town Centre[6]. The first dataset contains a CCTV footage of pedestrians in an area walking on the pavement and the other dataset contains a video of people walking in a busy downtown centre in Oxford, England. The Oxford Town Centre dataset has been utilized multiple times in multiple research projects. These are open datasets and can be used for various developments and research projects in the area of object detection, facial recognition, and many such other projects.

This dataset is a very unique dataset, in the sense that it uses video footage straight out of a public CCTV camera that on the contrary was assigned for public safety reasons. In this video, it shows that the pedestrians are walking or acting in a normal and unrehearsed manner. These pedestrians are just normal people walking on the road or pathway minding their own business.

Although these people do not know about the research projects, they were aware of being under supervision cameras and it is with their consent that this dataset has been created hence not violating any ethical issues.

3.2 Data Pre-processing

This research first downloads the video footages that are available online. The video footage contains a fixed camera that detects individuals in a region of interest (ROI) and measures their distances in real-time without recording any sort of data. Moreover, this research proposes a novel approach towards detecting people and whether or not they are violating any social distancing regulations.

While detecting the interpersonal distances between the individuals present in the video, with the help of facial detection, their faces are detected to verify whether the individuals are wearing a mask or not.

[4]https://www.youtube.com/watch?v=aUdKzb4LGJI
[5] https://www.briefcam.com/
[6] https://megapixels.cc/oxford_town_centre/

4. Implementation

4.1 Object Detection and Tracking

Although there has been a confusion between the two terms Image Classification and Object Detection, often meaning them to be the same, they are completely different. Image Classification performs identification of objects in images while Object Detection performs identification of the objects including its location in the images. Both of these terms are widely popular in Computer Vision tasks (Merkulova, Shavetov, Borisov and Gromov, 2019). They can be used in every field possible such as healthcare, defence, sports, and various other industries.

The next question that arises is whether Object Detection and Tracking are the same terms or not. Yes, Object Detection and Tracking are two very similar terms in the way they are functioned. They are basically designed to perform the same functionality but with a little difference. Object Detection is used to detect objects present in an image or in multiple images where an object is stationary while Object Tracking performs detection on videos, that is, it keeps a track of the following object detected while it is moving (Porikli and Yilmaz, 2012). A video is a combination of fast-moving frames and thus identification of the objects and their location from every frame is performed by Object Tracking.

Object detection can be stated as a fundamental problem in the computer vision and the images domain. It intends to detect objects in the video that belong to specific classes such as humans, vehicles, and more. The deep neural network models like CNN have dominated the benchmarks of object detection. Pre-trained models like the MS COCO[7], has more than 896K objects and over 123K images in the training and validation set and almost 80K images with more than 80 categories in the testing set. With the help of supervised learning techniques like data augmentation, these models are trained.

The best approach to building a model to perform object detection is with the help of a Sliding Window technique (Glumov, Kolomiyetz and Sergeyev, 1995). The sliding window approach is where an image is divided into particular sizes and regions and then according to the region, the respective classes are classified as shown in Figure 1. Here, in this research, the sliding window will detect all the people present in the video footage and form a bounding box around it. A CNN classifier will state a confidence value where it will represent on the certainty that the window contains an individual or not. Then for each and every region, a CNN will be passed which will extract the features and then further pass it on to the classifier and the regressor.

[7] https://cocodataset.org/#home

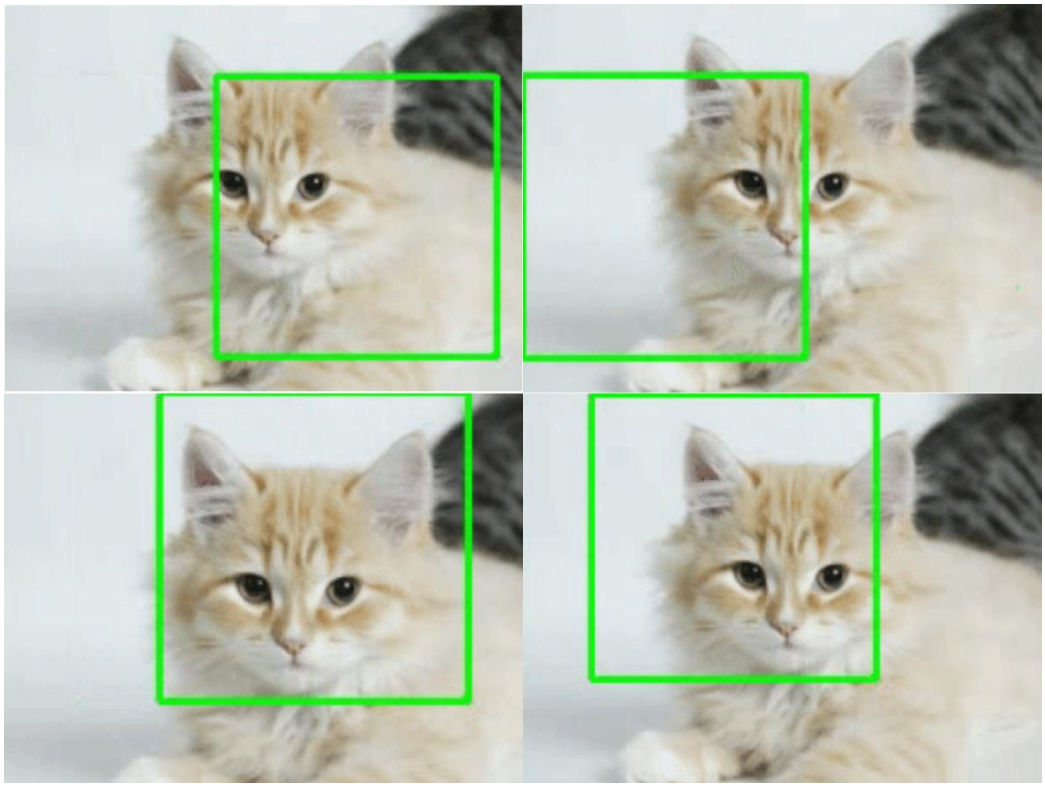


Figure 1: Sliding window technique

The state-of-the-art approach of this paper is capable of deploying pre-trained models. Pre-trained models can achieve good results to detect 2D object even with the different image qualities, angles, and camera models. Hence, this paper uses state-of-the-art pre-trained deep learning models that are used in detecting pedestrians and their face masks. These models can be used in the monitoring of social distancing as explained in the following sections.

A Social Distancing Detection and a Face Mask Detection model is trained and deployed to identify whether people are violating the social distancing measures and wearing a mask or not. YOLOv3 along with DBSCAN clustering is used for recognizing the potential intruders. The face mask classification model uses the ResNet50 to train and deploy the model in order to identify the people whether they are with or without a mask. The project description is given below.

4.2 Pedestrian Detection

You Only Look Once Version 3[8] is a real-time, state-of-the-art object detection system which is pre-trained on the COCO[9] dataset. It has a resolution of 416*416 and is used in this research for the purpose of obtaining bounding boxes of individuals in the videos.

[8] https://pjreddie.com/darknet/yolo/

[9] https://cocodataset.org/#download

There are many ways for gaining the position of an individual such as obtaining the center of the bounding box, that is, the midpoint of the bounding box. Other options include OpenCV's bird's eye perspective. This project focuses on the first method to obtain the distances between every individual. First, the model detects individuals in the frames and their faces. It then puts a red or green bounding box around the individual and his/her face to determine if they are safe or not and wearing a mask or not. In order to reduce the complexity of the project, the regions that do not contain the object can be discarded. The process of extraction of such regions is called Region Proposals. These algorithms are proposed so that the Regions of Interest (ROIs) can be selected. Therefore, one of the best approaches to ROI can be found by deploying the Region-based Convolutional Neural Network (R-CNN) (Liao et al., 2019).

4.3 Face Mask Detection

Dual Shot Face Detector is utilized in this research to detect the individuals' faces. It is a method which derives from SSD and offers a Feature Enhance Module (FEM) to transfer original features to expand the single shot to the dual shot detector (Li et al., 2019). Conventional face detectors such as the MTCNN (Xiang and Zhu, 2017) or the Haar-Cascades (Tej Chinimilli et al., 2020) are not useful for this research as it lacks in detecting faces that are in low-resolutions or faces that are covered. DSFD is a bit complex and heavy on the pipeline but it delivers accurate results. It is widely used where detections are in a more large-scale ranged-orientations. As this research is working on video frames, the probability of encountering blurred faces is high and thus with the help of DSFD none of them will be missed. The blurriness could occur due to various reasons such as the face being out of focus or any random rapid movements or noise during the capture of the video.

A somewhat modified ResNet50 model whose layers have been pre-trained on ImageNet is used to classify individuals based on whether they are with a mask or without a mask. Figure 2 depicts the basic pipeline behind the working of the Social Distancing and the Face Mask Classification model. The basic methodology behind this algorithm is to first, divide the video into frames and detect people and their faces in every frame, individually. Later on, the frames are combined which then again forms a video. It works as follows:

1. Capture the video.
2. Read the video by dividing it into a number of frames.
3. Else, detect persons in each frame and get the bounding boxes around them with the help of YOLOv3.
    a. If it reaches to EOF, stop.
4. Further, get the positions of the people with the help of DBSCAN to detect where the clusters are forming.
5. While detecting persons, detect their faces with the help of the DSFD model to detect whether they have masks on or not.
6. With the help of bounding boxes on the person and their faces, measure the distance between them and detect masks on them.

7. Create a results board on top of the video and display the results.
8. Create an output stream and then show the results.
9. Do this for every frame till it reaches to end of file.

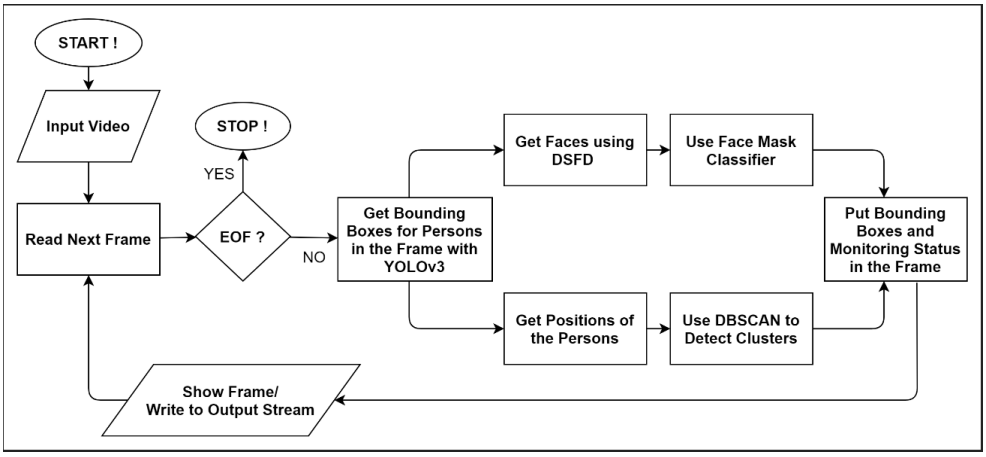
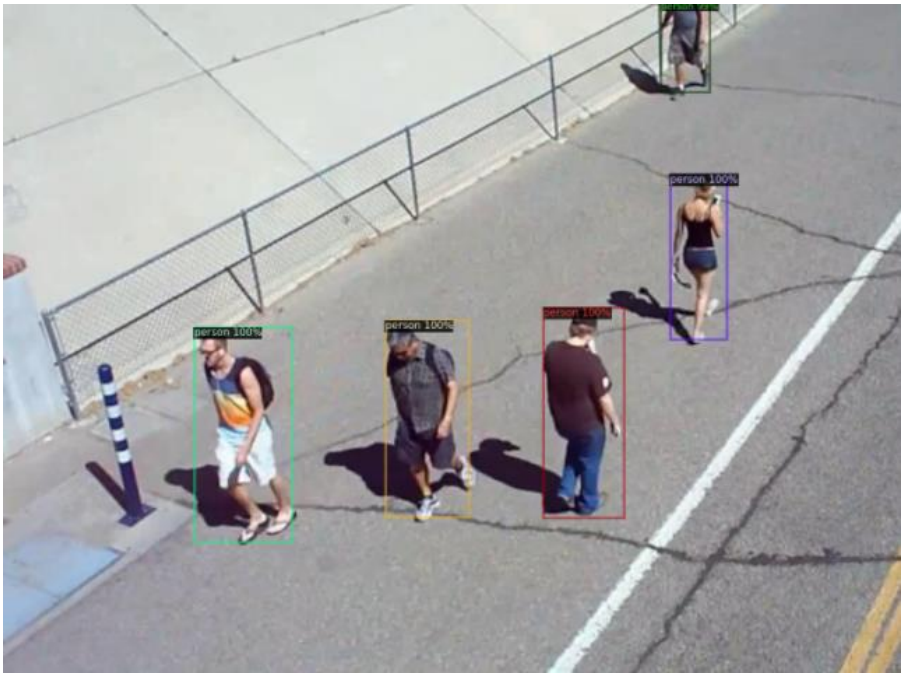


Figure 2: Social Distancing and Face Mask Detection Pipeline

5. Evaluation

After implementing and deploying the models, this research project achieved a confidence score of 100%. The **confidence score**[10] is the probability of a bounding box containing an object. Figure 3 shows the confidence score achieved.



[10] https://blog.zenggyu.com/en/post/2018-12-16/an-introduction-to-evaluation-metrics-for-object-detection/

Figure 3: Confidence Score

Intersection over Union (IoU)[9] can be defined as the area of the intersection divided by the area of the union of a predicted bounding box (Bp) and a ground-truth box (Bgt), as shown in Equation 1:

$$IoU = \frac{area(B_p \cap B_{gt})}{area(B_p \cup B_{gt})} \quad (1)$$

Equation 1: IoU

A detection is said to be a True Positive (TP) when it satisfies these three conditions:

- When the confidence score s higher than the threshold value
- When the predicted and the ground truth classes match.
- When the IoU of the predicted bounding box is greater than the ground truth threshold.

If either of the latter two conditions is violated, it produces False Positives (FP). If an IoU is implied as 1, it means that the ground-truth and the predicted bounding boxes overlap perfectly. Setting a threshold value beforehand can determine whether the object detection is legitimate or not. Confidence score and IoU are evaluation standards that are used to determine whether the object detected is a TP or a FP.

Non-maximum suppression (NMS) is an essential post-processing method that is used in computer vision. It is used in transforming multiple object windows that are imprecise and unnecessary into a smooth single anchor box for every object that is detected (Rothe, Guillaumin and Van Gool, 2015). This research paper makes use of the NMS function. Figure 4 shows the difference between MS and NMS.

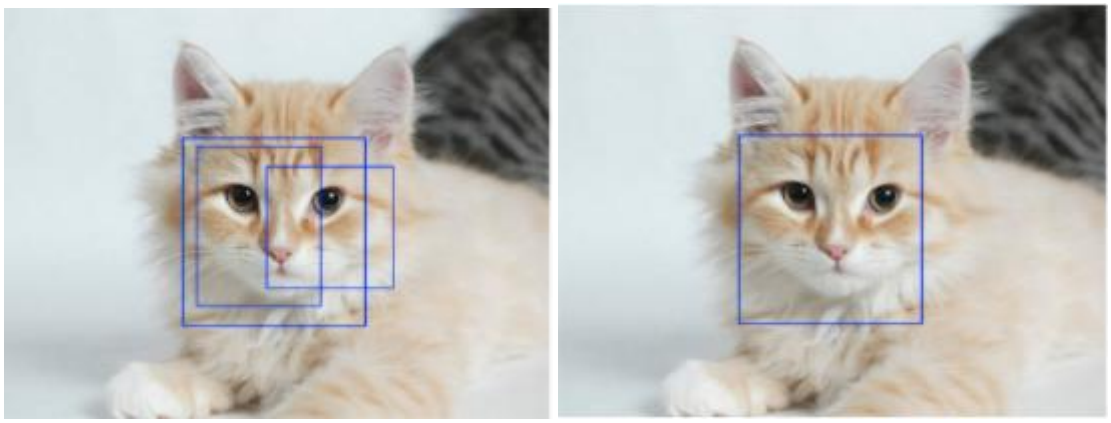


Figure 4: Left-side: Maximum suppression; Right-side: Non-Maximum Suppression

Figure 5(a) and Figure 5(b) show the accurate results obtained after implementing the models on the two datasets. On the top of the frame, there is a box that contains all the information of the objects present in the video frame. It shows the total number of people, number of people safe and unsafe, that is, whether they are implementing social distancing or not, and the number of people wearing a mask, not wearing a mask, and unknown in case the person's face is not facing towards the camera. Here we can see that the people have been identified perfectly with bounding boxes around them. The green box means they are safe, or within the social distancing limit, that is 1m. The red box tells that the people are unsafe. The small box on the faces of the people detect whether they are wearing a mask or not. Red means they are not wearing a mask and green means that they are wearing a mask.

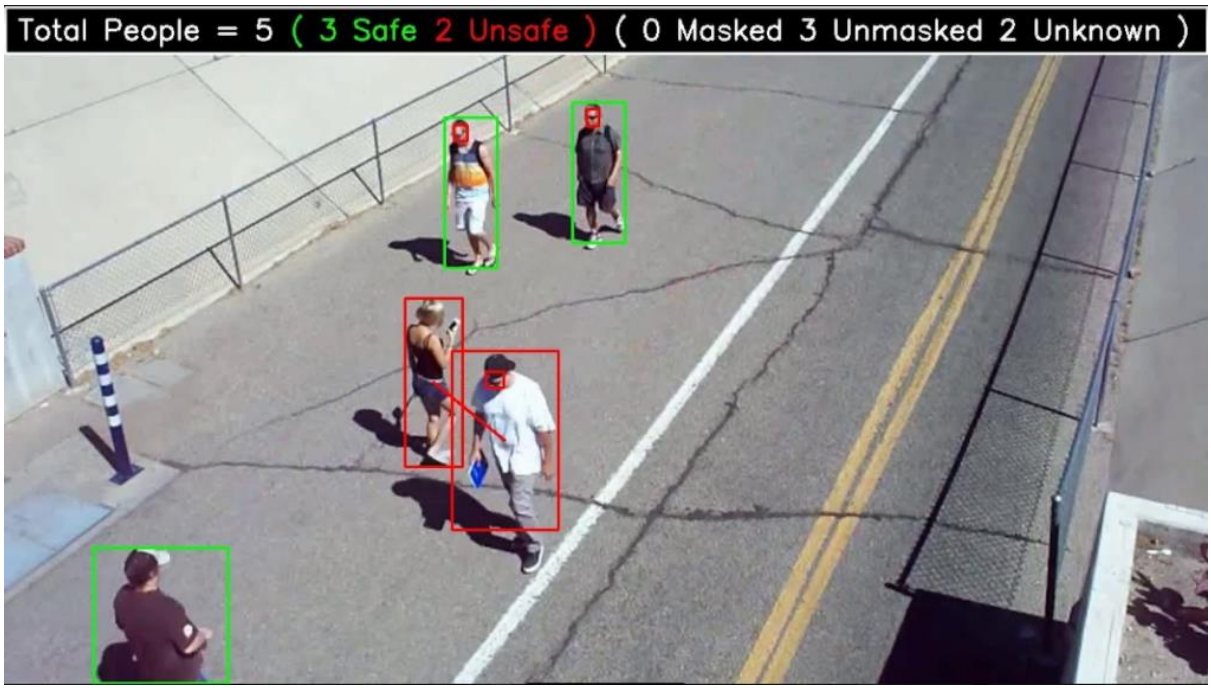


Figure 5 (a): Outcome of the proposed model

Figure 5 (b): Outcome of the proposed model

6. Conclusion

Thus, by understanding and studying about the virus, this paper can conclude two things: social distancing can reduce the spread of the coronavirus; face coverings help prevent the infectious disease to transmit via the air. Therefore, to support this study, this research proposed an AI-based real-time approach towards the detection of social distancing and face mask. Crowd density was examined by gaining the ROI of the video frames and the count of people violating and non-violating the measures was also shown. The results obtained were accurate and real-time based. The pandemic is continuing to increase and is still going on while studying this research. Many countries are still in the quarantine phase whereas some countries have terminated the quarantine phase and are stuck to their social distancing measures. Social distancing may be implemented forever in some countries and people may have to live with the new normal. Russia has declared to have found a vaccine to cure the virus and has started the production of it[11]. If it is successful, the world might finally be able to go back to how it was.

7. Future Work

In this research, the people who have close relationships or who know each other and are walking together are detected as they are violating the social distancing measure. Some argue

[11] https://www.abc.net.au/news/2020-08-16/coronavirus-update-russia-starts-manufacturing-vaccine/12562298

that they should practice social distancing in public areas while some argue that they need not. Hence, this is one of the limitations in this research and can be considered for future work. This research does not make use of the bird's eye view function and hence can be considered for future work as well.

Some places where the virus transmission is high, people are required to wear double face masks, and thus to detect whether a person is wearing a double face mask can be achieved in the future. If in the future there is no need for social distancing, this project can be reutilized and repurposed for other applications. Museums have a 2-3 feet distance policy between the artefact and the individual and thus it can be used to detect this distance.

8. Acknowledgment

I would like to thank my Research Supervisor Vladimir Milosavljvec for his relentless support and his constant motivation to encourage me to work on the research project. For a period of 13 weeks, he assisted and supervised me clearing all my doubts and helping me achieve my goal. This research work has been inspired by the works of Andrew Ng and his Landing AI team. I would like to thank my family and friends for their love, appreciation and support.

9. References

Aldila, D., Khoshnaw, S., Safitri, E., Anwar, Y., Bakry, A., Samiadji, B., Anugerah, D., GH, M., Ayulani, I. and Salim, S., 2020. A mathematical study on the spread of COVID-19 considering social distancing and rapid assessment: The case of Jakarta, Indonesia. *Chaos, Solitons & Fractals*, 139, p.110042.

Christian, M., Loutfy, M., McDonald, L., Martinez, K., Ofner, M., Wong, T., Wallington, T., Gold, W., Mederski, B., Green, K. and Low, D., 2004. Possible SARS Coronavirus Transmission during Cardiopulmonary Resuscitation. *Emerging Infectious Diseases*, 10(2), pp.287-293.

Deore, G., Bodhula, R., Udpikar, V. and More, V., 2016. Study of masked face detection approach in video analytics. *2016 Conference on Advances in Signal Processing (CASP),*.

Fong, S., Dey, N. and Chaki, J., 2020. An Introduction to COVID-19. *Artificial Intelligence for Coronavirus Outbreak*, pp.1-22.

Glumov, N., Kolomiyetz, E. and Sergeyev, V., 1995. Detection of objects on the image using a sliding window mode. *Optics & Laser Technology*, 27(4), pp.241-249.

Huynh, T., 2020. Does culture matter social distancing under the COVID-19 pandemic?. *Safety Science*, 130, p.104872.

Li, G., Yang, Y. and Qu, X., 2020. Deep Learning Approaches on Pedestrian Detection in Hazy Weather. *IEEE Transactions on Industrial Electronics*, 67(10), pp.8889-8899.

Li, J., Wang, Y., Wang, C., Tai, Y., Qian, J., Yang, J., Wang, C., Li, J. and Huang, F., 2019. DSFD: Dual Shot Face Detector. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR),*.

Liao, G., Yang, G., Tong, W., Gao, W., Lv, F. and Gao, D., 2019. Study on Power Line Insulator Defect Detection via Improved Faster Region-Based Convolutional Neural Network. *2019 IEEE 7th International Conference on Computer Science and Network Technology (ICCSNT),*.

Lin, K., Zhao, H., Lv, J., Li, C., Liu, X., Chen, R. and Zhao, R., 2020. Face Detection and Segmentation Based on Improved Mask R-CNN. *Discrete Dynamics in Nature and Society*, 2020, pp.1-11.

Loey, M., Manogaran, G., Taha, M. and Khalifa, N., 2020. A hybrid deep transfer learning model with machine learning methods for face mask detection in the era of the COVID-19 pandemic. *Measurement*, 167, p.108288.

Mangili, A. and Gendreau, M., 2005. Transmission of infectious diseases during commercial air travel. *The Lancet*, 365(9463), pp.989-996.

Mangmang, G., 2020. Face Mask Usage Detection Using Inception Network. *Journal of Advanced Research in Dynamical and Control Systems*, 12(SP7), pp.1660-1667.

Merkulova, I., Shavetov, S., Borisov, O. and Gromov, V., 2019. Object detection and tracking basics: Student education. *IFAC-PapersOnLine*, 52(9), pp.79-84.

Porikli, F. and Yilmaz, A., 2012. Object Detection and Tracking. *Studies in Computational Intelligence*, pp.3-41.

Pranav, K. and Manikandan, J., 2020. Design and Evaluation of a Real-time Pedestrian Detection System for Autonomous Vehicles. *2020 Zooming Innovation in Consumer Technologies Conference (ZINC)*.

Richard, M., Kok, A., de Meulder, D., Bestebroer, T., Lamers, M., Okba, N., Fentener van Vlissingen, M., Rockx, B., Haagmans, B., Koopmans, M., Fouchier, R. and Herfst, S., 2020. SARS-CoV-2 is transmitted via contact and via the air between ferrets. *Nature Communications*, 11(1).

Rothe, R., Guillaumin, M. and Van Gool, L., 2015. Non-maximum Suppression for Object Detection by Passing Messages Between Windows. *Computer Vision -- ACCV 2014*, pp.290-306.

Saeidi, M. and Ahmadi, A., 2020. A novel approach for deep pedestrian detection based on changes in camera viewing angle. *Signal, Image and Video Processing*, 14(6), pp.1273-1281.

Saez, M., Tobias, A., Varga, D. and Barceló, M., 2020. Effectiveness of the measures to flatten the epidemic curve of COVID-19. The case of Spain. *Science of The Total Environment*, 727, p.138761.

Setti, L., Passarini, F., De Gennaro, G., Barbieri, P., Perrone, M., Borelli, M., Palmisani, J., Di Gilio, A., Torboli, V., Fontana, F., Clemente, L., Pallavicini, A., Ruscio, M., Piscitelli, P. and Miani, A., 2020. SARS-Cov-2RNA found on particulate matter of Bergamo in Northern Italy: First evidence. *Environmental Research*, 188, p.109754.

Sun, C. and Zhai, Z., 2020. The efficacy of social distance and ventilation effectiveness in preventing COVID-19 transmission. *Sustainable Cities and Society*, 62, p.102390.

Tang, J., Su, Q., Su, B., Fong, S., Cao, W. and Gong, X., 2020. Parallel ensemble learning of convolutional neural networks and local binary patterns for face recognition. *Computer Methods and Programs in Biomedicine*, 197, p.105622.

Tej Chinimilli, B., T, A., Kotturi, A., Reddy Kaipu, V. and Varma Mandapati, J., 2020. Face Recognition based Attendance System using Haar Cascade and Local Binary Pattern Histogram Algorithm. *2020 4th International Conference on Trends in Electronics and Informatics (ICOEI)(48184)*,.

Thu, T., Ngoc, P., Hai, N. and Tuan, L., 2020. Effect of the social distancing measures on the spread of COVID-19 in 10 highly infected countries. *Science of The Total Environment*, 742, p.140430.

Tobías, A., 2020. Evaluation of the lockdowns for the SARS-CoV-2 epidemic in Italy and Spain after one month follow up. *Science of The Total Environment*, 725, p.138539.

WANG, Y., ZHANG, X., YE, J., SHEN, H., LIN, Z. and TIAN, W., 2020. Mask-wearing recognition in the wild. *SCIENTIA SINICA Informationis*, 50(7), pp.1110-1120.

Wu, J., Tang, B., Bragazzi, N., Nah, K. and McCarthy, Z., 2020. Quantifying the role of social distancing, personal protection and case detection in mitigating COVID-19 outbreak in Ontario, Canada. *Journal of Mathematics in Industry*, 10(1).

Xiang, J. and Zhu, G., 2017. Joint Face Detection and Facial Expression Recognition with MTCNN. *2017 4th International Conference on Information Science and Control Engineering (ICISCE)*,.

Yu, W., Kim, S., Chen, F. and Choi, J., 2020. Pedestrian Detection Based on Improved Mask R-CNN Algorithm. *Advances in Intelligent Systems and Computing*, pp.1515-1522.

Zhang, Z., Gao, J., Mao, J., Liu, Y., Anguelov, D. and Li, C., 2020. STINet: Spatio-Temporal-Interactive Network for Pedestrian Detection and Trajectory Prediction. *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*.