Based On Segmented Video Image Processing, Person Detection for Social Distancing and Safety Violation Alert

Deepak P.[1], Dr. Durgam Kumaraswamy[2]

[1]Research Scholar, OPJS University, Churu Rajasthan
[2]Research Supervisor, OPJS University, Churu Rajasthan

Abstract

Both humans and animals are harmed by the broad family of coronavirus viruses. The coronavirus identified as Covid-19 initially made its way to Wuhan, China, in December 2019. We can alter the pixels to either improve the image's quality or get more relevant data out of them. Other than measuring social distance, the device can also identify persons in restricted zones and utilise that information to provide warnings. The program's efficacy for both goals has been tested through some analysis. According to the results, the distance tracking system tested indoors in a controlled setting and achieved 100% accuracy, whereas testing outside required tougher input footage and the system only managed 56.5% to 68% accuracy. Contrarily, it was discovered that the safety violation alert feature based on segmented ROI had obtained greater accuracy, ranging from 95.8% to 100% for all of the examined input films.

Keywords: *Covid-19, Machine Learning, YOLOv3, video capture module, speak module*.

1. INTRODUCTION

Both humans and animals are harmed by the broad family of coronavirus viruses. The coronavirus identified as Covid-19 initially made its way to Wuhan, China, in December 2019. As the outbreak spread quickly throughout the world, the World Health Organization proclaimed it to be a pandemic (WHO). While Covid-19 is still a problem in many nations, incidents in Malaysia have begun to level off. The Recovery Movement Control Order (RMCO) was made public by the Malaysian government on June 10, 2020, in an effort to flatten the curve of Covid-19 cases in Malaysia. The majority of the economic sectors have been reopened, and Malaysian citizens are now able to travel and carry out their everyday activities within the conditions of the new normal. In order to prevent the spread of viruses, people must abide by the MOH and WHO guidelines when they are outside. Implementing social distance between persons with at least one meter apart is one of the best-known methods for halting the spread of Covid-19.

By using pixels, an image is a binary representation of visual data. It is packed with data that may be used to carry out several practical tasks in a variety of applications, including the medical industry, object detection, video surveillance systems, computer vision, pattern recognition, remote sensing, etc. We can alter pixels using computer algorithms to either improve the image quality or extract valuable information from it. Segmentation is a crucial and difficult step in the operation of digital image processing, which comprises several stages. Image segmentation works to break down large pictures into manageable parts that share common features and behaviors. Segmentation's purpose is to take a complex visual representation and simplify it so that it can be properly categorized and analyzed. Using image segmentation, we may pinpoint items in an image and determine their boundaries. Next, label every single pixel in the picture. Similar labels on adjacent pixels suggest that they have a common characteristic.

In this analysis, we take a look at the different segmentation methods and the algorithms that power them. We evaluate the segmentation done using pixel attributes to evaluate the local and global properties in order to identify how to speed up the segmentation process by upgrading the existing ACM model along with local and global characteristics. More iterations are required for the expectation of maximization method to segment the image. We can reduce the number of calculations by basing modifications on features of likelihood. Successful segmentation is communicated by means of an EM-based localization, 3D U-Net, and an EM-based attention mechanism. Color pictures may be divided using the RGB Histogram and the Firefly algorithm. By integrating the SCT-I, SCT-V, and SAMFO-TH algorithms, we can effectively address the multilayer threshold issue in an

RGB colour picture. Color object detection can be recognized using color hymnographies. The image can be divided into sections using the example blocks as a basis. This approach makes use of the benefits of the Creminis algorithm as well as the Watershed image segmentation technique. The process of extracting text information from an image is impacted by numerous variables. When an image is recorded from a live video camera, there is greater noise. For the purpose of extracting text information, the difficulties of segmentation are dealt with by the Encoder-Decoder architecture. Faster text retrieval is possible using the Instance Segmentation Network method. If you want to give them a complete idea of how image segmentation works, you need to explain it, this study compares and examines the various methods.

According to records, the covid-19 virus, which causes harm to humans when it spreads, was first discovered in Wuhan, China, in December 2019. Since then, Since the World Health Organization has announced an imminent epidemic, numerous countries have declared states of emergency in an effort to contain the virus. The illness has claimed several lives and sent many more to hospitals for treatment. The high incidence of pregnancy and the very high number of newly reported instances of COVID positivity each day. Major concepts and recommendations have been published by the WHO and ICMR, and people should maintain social distance from one another to avoid contracting COVID-19, wear masks, and routinely sterilize or wash their hands. Today, human daily activities are returning to normal. According to information from WHO and ICMR, the coronavirus spreads from one person to another through tiny droplets from the mouth and nose while they chat, cough, or sneeze. In other words, social distance is the greatest strategy for minimizing contact with those who have been impacted by maintaining a minimum distance of one meter between individuals.

2. LITERATURE REVIEW

S. PRABU et.al (2021) Object identification, computer vision, video surveillance, and other applications all rely heavily on image processing techniques, which are an integral aspect of today's computer technologies. Image segmentation is a crucial step in the processing of images. Image segmentation is the process of dividing an image into several, relatively small pieces known as segments. In order to interpret photos, image processing makes it possible to express images more simply. For the purpose of segmenting images depending on a certain pixel feature, numerous algorithms have been developed. Different segmentation methods can be studied, analyzed, and finally listed along with their comparisons in this work. The accuracy and performance of segmentation methods across a range of image processing domains can be improved with the help of this comparison research.

K. Jeevitha et.al (2020) The importance of image-processing techniques has increased across a wide range of applications as a result of advancements in computer technology. In image processing, image segmentation is crucial. The term "image segmentation" describes the division of an image into various sections that differ and are comparable in terms of things like color, intensity, and texture. For picture segmentation, various algorithms and methods have been developed. This document examines and compiles a few of the image segmentation technologies.

Afiq Harith Ahamad et.al (2020) If individuals don't take action to halt the spread of the virus, the worldwide Covid-19 pandemic situation will be unmanageable, and the coronavirus case curve will continue to rise. One of the most important things to do during these outbreaks is to make sure that there is enough space between individuals to prevent the spread of disease. In order to prevent physical contact between individuals, this study suggests the detection of persons through social distance monitoring. This research uses the object tracking model Mobile Net Single Shot Multi box Detector (SSD) and the OpenCV library for image processing to locate and identify individuals in certain scenes. The recorded video's pixel values will be compared to the computed distances between the persons in the footage. A distance is calculated between each person based on their position relative to the segmented tracking area's core points and the overlapping border. Alerts or cautions can be sent out if it is determined that the space between persons is unsafe. Other than measuring social distance, the device can also identify persons in restricted zones and utilise that information to provide warnings. The program's efficacy for both goals has been tested through some analysis. According to the results, the distance tracking system tested indoors in a controlled setting and achieved 100% accuracy, whereas testing outside required tougher input footage and the system only managed 56.5% to 68% accuracy. On the other side, using segmented ROI, it was found that the safety violation warning function achieved excellent accuracy, with results ranging from 95.8% to 100% across all tested input films.

M.Vedraj et.,al (2021) Today, the world is experiencing a severe increase in the spread of the COVID-19 virus, and this is frequently due to people not taking the greatest preventive precautions, such as wearing a mask while outside, keeping a safe distance from others, and frequently disinfecting their hands and belongings. The study

makes recommendations for detecting social distance between people utilizing open cv and machine learning in order to measure the distance between them and stop the transmission of the corona virus. the coronavirus This detection tool was created to look for social distance between individuals in crowded places and warn whether a certain distance is maintained. It frequently achieves this by combining video input from a live camera using YOLOv3. Numerous modules on this device, including as the object detection, video recording, distance calculation, and communication modules, were designed to prevent people from maintaining social distance.

3. METHODOLOGY

This study uses Python 3, OpenCV, and the Caffe framework to build the suggested notion. The image processing techniques that will be covered in more detail in this part are implemented using the OpenCV library.

The major purpose of this system is to analyse recorded video footage in order to identify persons, and then evaluate those detections for potential security breaches or social animosity. Therefore, it all starts with scanning individual frames from a live video source. This is depicted in Fig. 1, which provides a flowchart of the entire series of events. The object detection framework is the study's most significant component. This is as a result of the portion of the study that concentrates on locating a person from the input frame. Therefore, it's crucial to use the best object detection model to prevent any issues with human detection.

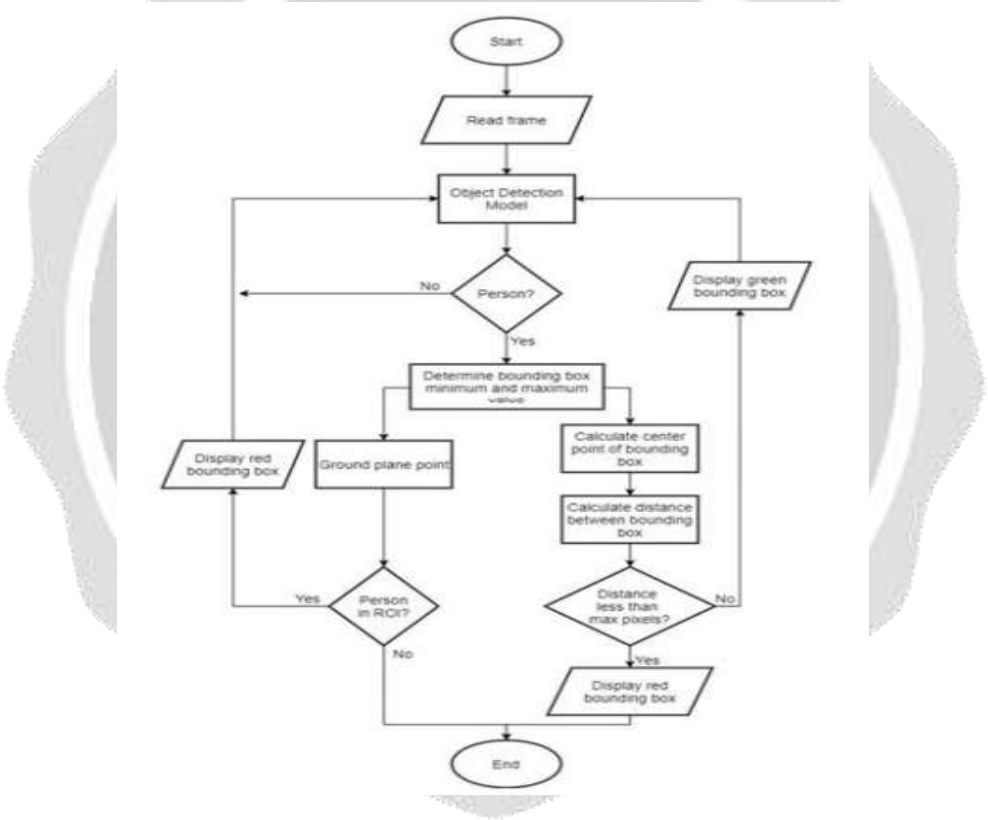

Fig. 1 Person detection for social distancing and safety violation alert based on segmented ROI flowchart.

A. Object Detection Model

Caffe, a deep learning model framework, was used to power the object identification model in this investigation. Due to the speed of execution, Mobile Net SSD was selected as the model.

B. Threading Parallelism

Because of Python 3's threading capabilities, many parts of a programme may run at once. Reduce the time needed to conduct object detection for each frame in this analysis by using threading. The frame will be processed simultaneously with object detection using a multi-threading method.

C. Masking frame for ROI area estimation

A small picture component is used in the image processing technique known as masking to change a bigger image. Masking entails changing a portion of an image's pixel values to zero and a different backdrop value, as in (1).

$$f_{mask} = f - \text{mask}(f) \qquad (1)$$

It will do this by masking the ROI area and isolating the image, f. A video, to provide just one example, is a set of pictures that run in order for a certain amount of time. In this experiment, we will use the OpenCV masking approach to generate regions of interest (ROIs) in each input frame.

D. Determine person location

An individual's bounding box and associated segment are located by comparing each ground plane point to the ROI range.

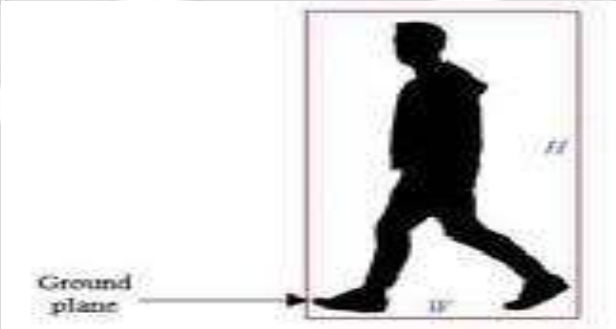

Fig. 2. The ground plane of a human walking model.

In order to monitor a specific region, such as a high-risk area or areas of interest for an organization, surveillance cameras are typically installed in high positions called overhead cameras. For this purpose, comparing the ground plane for the detection box rather than its centre point, is more appropriate.

E. Calculate the center point of a bounding box

The midpoint equation, as in, is used to locate the centre point, C (X, Y), of the bounding box for the identified individual (2).

$$C(x,y) = \left(\frac{x_{min}+x_{max}}{2}, \frac{y_{min}+y_{max}}{2}\right) \qquad (2)$$

The bounding box's centre will be determined by taking the minimum and maximum values for its width (MIN AND X MAX) and height (YMIN AND YMAX).

F. Calculate distance between bounding box

To measure the distance, ($_{MIN}$Y,$_{MIN}$X) $_1$C, AND C_2 (X $_{MAX}$,Y_{MIN}), the distance equation as in 3.6 is applied to each identified individual in the scene.

$$d(C_1, C_2) = \sqrt{(x_{max} - x_{min})^2 + (y_{max} - y_{min})^2} \qquad (3)$$

In this research, researchers compare two sets of bounding boxes by focusing on their respective centres. After getting the value for the centres, the algorithm will decide whether the distance is less than or more than 300 pixels.

4. DATA ANALYSIS

All of the code for this project's system has been written in Python 3 and uses OpenCV for image processing and the Caffe object identification model framework. Analyses have shown that this artificial system works, and findings have been made. The primary method for person detection in this investigation was the Mobile Net trigger SSD Caffe model. The primary clip is taken from the larger scenario, which was shot in the family room from above, in order to make adjustments to the programmed.

A. Masked ROI

The masked ROI is made up of two sections as a first step in identifying areas of interest. First, as can be seen in Fig. 3, the frame's foreground has been masked. In the front, we see the region's return on investment (ROI) where a person's detection will result in the required ground plane point value being returned. The mask used in the backdrop is made by inverting the foreground mask. Figure 4 shows how the foreground and background frames were combined to get the desired return on investment.

Fig. 3. OpenCV masking step to get foreground frame.

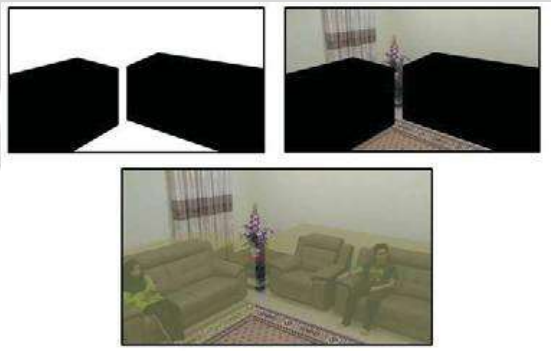

Fig. 4. Masked background frame with combination of foreground frame.

Person Detection Localization and Social Distancing

The masked ROI is made up of two sections as a first step in identifying areas of interest. The foreground of the frame is the first area that has been masked, as seen in Fig. 4. In the front, we see the region's return on investment (ROI) where a person's detection will result in the required ground plane point value being returned. The mask used in the backdrop is made by inverting the foreground mask. ROI is achieved by combining the foreground and background frames into one image, as seen in Figure 5.

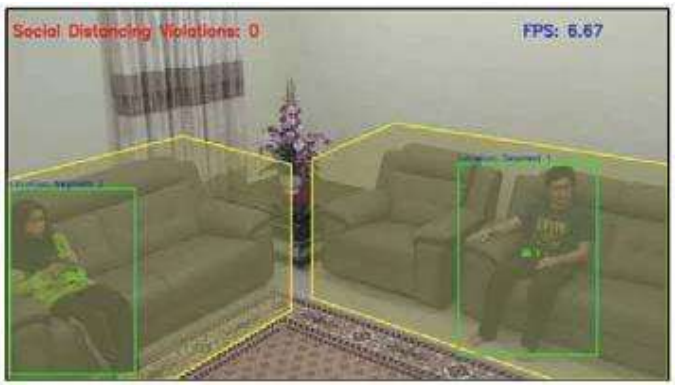

Fig. 5. Green bounding boxes when there is no social distance violation.

As can be seen in Figure 6, the bounding boxes of each potentially identified individual are indicated in green when their distance exceeds the default setting of the social distance range. The standard pixel size for experimental purposes in Figure 6 is 300. Each video input has a unique camera angle, which in turn affects the assumed minimum social distance. When the safe social distance is violated, on the other hand, the bounding boxes of the affected persons become red to signal that they are closer than is considered to be safe. This allows for a clear delineation between those who breach the designated space and those who do not. Figure 7 shows the results.

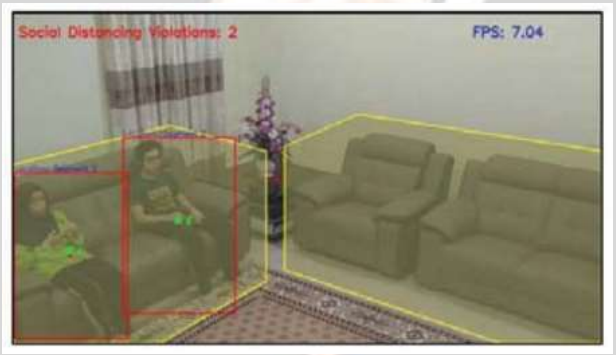

Fig. 6. Red bounding boxes indicate an alert or warning during social distance violations.

To measure how well this system works, its precision is calculated. For social distance monitoring, accuracy is calculated by adding the values of true positive (TP), true negative (TN), false positive (FP), and false negative (FN). The equation used to determine the accuracy is displayed in (4).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \qquad (4)$$

Four separate movies are used to assess the social distance monitoring system's accuracy: a dataset from PETS2009, Town Center, VIRAT S, and a self-shot video. Self-shot video is captured in a controlled, indoor setting. As a result, the system's maximum level of accuracy is predicated on a self-shot video. The system's calculated accuracy is displayed in Table 1

Table 1. Accuracy of Social Distance Monitoring

No	Video	TP	TN	FP	FN	Accuracy
1	Self-taken	2	2	0	0	100%
2	TownCentre	11	19	14	4	62.5%
3	PETS2009	14	38	19	5	68%
4	VIRAT_S	9	4	0	10	56.5%

According to the results, this study was unable to sustain the social distance system's highest level of accuracy. The object detection model, which cannot recognize people in some of the film, is to blame for this issue. In difficult settings, the object detection model has trouble identifying human beings and pinpointing their position. This was found to be the case in many shots throughout the video. As can be seen in Figures 7 and 8, the object detection model has a better chance of picking up on the presence of a human if the camera used to capture the video is placed in close proximity to the object being studied or in a stable interior environment. Therefore, it is necessary to enhance the social distance system for outdoor environments, particularly for videos that feature far-off locations.

The object detection framework's inability to accurately identify the person in the picture is seen in Figures 9 and 10. This makes it hard and inaccurate to determine the distance between each bounding box, which is required by the social distance measurement since it relies on the centre point values of the detecting boxes.

Internal visual input was used to evaluate this feature, and the camera is situated close to the monitored ROI. As a result, it can pinpoint the person's location with accuracy.

Fig. 7. No red bounding boxes on detected persons outside the ROI.

Table 2. Rate of Return-Based Restricted Area Alert Accuracy

No	Video	TP	TN	FP	FN	Accuracy
1	Self-taken	2	2	0	0	100%
2	CamNeT	55	58	0	5	95.8%

5. CONCLUSION

Implementing social distance between persons with at least one meter apart is one of the best-known methods for halting the spread of Covid-19. The numerous picture segmentation strategies discussed in this study are demonstrated and associated, as are the various image segmentation approaches. The accuracy of both aspects has been examined. According to the overall findings, this study appears to have accomplished all of its goals. The results obtained do have certain limits, though. According to the findings of the system's tests, it is difficult for the object detection model used to identify humans to do so accurately in outdoor settings and challenging scenarios with distant views. An improved object detection model can be used in the future to make improvements.

6. REFERENCES

1. K. Jeevitha et.al (2020) a review on various segmentation techniques in image processsing issn 2515-8260
2. Afiq harith ahamad et.al (2020) person detection for social distancing and safety violation alert based on segmented roi
3. D. Ganiger, k. A. Patil, p. Patil, and m. Anandhalli, "automatic of power supply in classroom using image processing," in proceedings -2017 international conference on recent advances in electronics and communication technology, icraect 2017, 2017, pp. 230–234, doi: 10.1109/icraect.2017.16.
4. S. Khedkar and g. M. Malwatkar, "using raspberry pi and gsm survey onhome automation," in international conference on electrical, electronics, and optimization technique, iceeot 2016, 2016, pp. 758–761.
5. S. Suresh, h. N. S. Anusha, t. Rajath, p. Soundarya, and s. V. P. Vudatha, "automatic lighting and control system forclassroom," in proceedings of 2016 international conference on ict in business, industry, and government, ictbig 2016, 2017, doi: 10.1109/ictbig.2016.7892666.
6. V. S. Rao and s. S. Kowshik, "pir sensor characterization and a novel localization techniqueusing pirs," in proceedings of the 14th international conference on information processing in sensor networks, 2015.
7. M. Habaebi, r. Rosli, and m. R. Islam, "rssi-based human presence detection system for energy saving automation," indones. J. Electr.eng. Informatics, vol. 5, no. 4, pp. 339–350, 2017, doi: 10.11591/ijeei. v5i4.356.
8. H. Fan, h. Zhu, and d. Yuan, "people counting in elevator car based on computer vision," in iop conference series: earth and environmental science, 2019, vol. 252, p. 032131, doi: 10.1088/1755-1315/252/3/032131.
9. J. Hariyono and k. H. Jo, "detection of pedestrian crossing road," proc. -int. Conf. Image process. Icip, vol. 2015-decem, no. January, pp. 4585–4588, 2015, doi: 10.1109/icip.2015.7351675.
10. World health organization, "coronavirus disease 2019," coronavirus disease (covid-19) pandemic, 2020. Https://www.who.int/emergencies/diseases/novel-coronavirus-2019 (accessed jun. 19, 2020).
11. M. Ministry of health, "covid-19 malaysia," http://covid19.moh.gov.my/, 2020. Http://covid-19.moh.gov.my/.
12. H. H. Nazi, "social distancing enforced on all rapidkl trains," nst, 2020. Https://www.nst.com.my/news/nation/2020/03/577603/socialdistancing-enforced all-rapid-trains (accessed jun. 30, 2020).