© INTERNATIONAL JOURNAL FOR RESEARCH PUBLICATION & SEMINAR
ISSN: 2278-6848 | Volume: 14 Issue: 03 | April - June 2023
Paper is available at http://www.jrps.in | Email : info@jrps.in
Refereed & Peer Reviewed

Special Edition
NCASIT 2023, 29th April 2023
Department of Computer Engineering,
St. Vincent Pallotti College of Engineering & Technology, Nagpur,

ANALYSING THE SAFETY OF THE ENVIRONMENT BY DETECTING AND COUNTING PEOPLE

Pendyala Himaja[1]*,
penhimaja@gmail.com,
Kalyani Rajput[1]*,
kalyanirajput092@gmail.com

Kanishk Kalkar[1]*,
kanishkkalkar@gmail.com
Karan Parate[1]*
karan21parate@gmail.com

Prof. Prajakta Kharwandikar[2]
pkharwandikar@stvincentngp.edu.in

1. Undergraduate student, Department of Computer Engineering, St. Vincent Pallotti College of Engineering and Technology.*

2. Assistant Professor, Department of Computer Engineering, St. Vincent Pallotti College of Engineering and Technology.

***ABSTRACT*:**
Person detection and analysing the safety of the environment is a crucial topic in computer vision and image processing. It involves detecting people in images, videos or real-time, counting them, and assessing the safety of the environment based on the count. This technique has diverse applications in security, transportation, and healthcare. Recent advances in deep learning-based object detection algorithms, like YOLOv3 and CNN, have made person detection and counting more accurate and efficient. Safety analysis can be done by comparing the number of people in specific areas with a threshold value and taking necessary actions to enhance security. This paper reviews state-of-the-art person detection and counting techniques and discusses their applications in safety analysis. Future directions include real-time processing and integration with other sensing technologies.
Keywords: Person Detection, Person Count, Analysing the safety, YOLOv3, CNN, COCO dataset, OpenCV, pyqt5, imutils.

I. INTRODUCTION

Person detection and analysis of the safety of the environment have been important research areas in computer vision and artificial intelligence for many years. The safety of the environment is a critical concern for individuals, organizations, and governments worldwide. One of the key aspects of ensuring safety is the ability to detect and analyse the presence of people in various environments. The ability to detect people and analyse their behaviour via image, video, or real-time is critical for many applications such as surveillance, security, and autonomous navigation.

In recent years, there has been a significant increase in interest in developing automated systems that can accurately detect and count the number of people in complex environments. This is due to the growing need for advanced security systems in public places, such as airports, train stations, stadiums, etc.

Person detection refers to the process of identifying the presence of people in a given environment. Moreover, analysing the safety of the environment, which includes identifying areas and predicting potential risks based on the count, is crucial for ensuring public safety in crowded areas. This requires advanced machine learning algorithms to process large amounts of data and identify patterns that can help predict risks. Convolutional neural networks have been used earlier for object detection. As there are various object detection techniques available, the YOLOv3 and the SSD, have attracted more researchers for object detection. Both methods can work for the image as well as for the video detection of the objects [1]. These techniques enable the detection of people in various settings, such as indoor and outdoor environments, and in different lighting conditions.

In this paper, we present a comprehensive review of recent research on person detection and the analysis of the safety of the environment. We will discuss various techniques for detecting a person, including traditional computer vision methods and deep learning-based approaches like object detection

© INTERNATIONAL JOURNAL FOR RESEARCH PUBLICATION & SEMINAR
ISSN: 2278-6848 | Volume: 14 Issue: 03 | April - June 2023
Paper is available at http://www.jrps.in | Email : info@jrps.in
Refereed & Peer Reviewed

Special Edition
NCASIT 2023, 29th April 2023
Department of Computer Engineering,
St. Vincent Pallotti College of Engineering & Technology, Nagpur,

techniques i.e. YOLOv3 and CNN, and various datasets like the COCO dataset.

II. OBJECTIVE

The goal of our project is to assess environmental safety. In this project, first a threshold is to be set, then detection techniques are used to detect a person, and finally, the total number of people can be determined. The environment is analysed based on the set threshold to determine whether the area is safe or unsafe. The goal of this project is to display a warning if the area becomes unsafe due to overcrowding.

III. LITERATURE REVIEW

According to H. Zhao et al., accurately detecting faces is an essential task. Given that they employ two stages of detection, the quicker RCNN algorithms perform well. The shortcomings of YOLOv3 were eliminated by employing the suggested strategy, which is YOLOv3 face-based on YOLO v3. The method that was demonstrated yielded better reliable results for face detection [2].

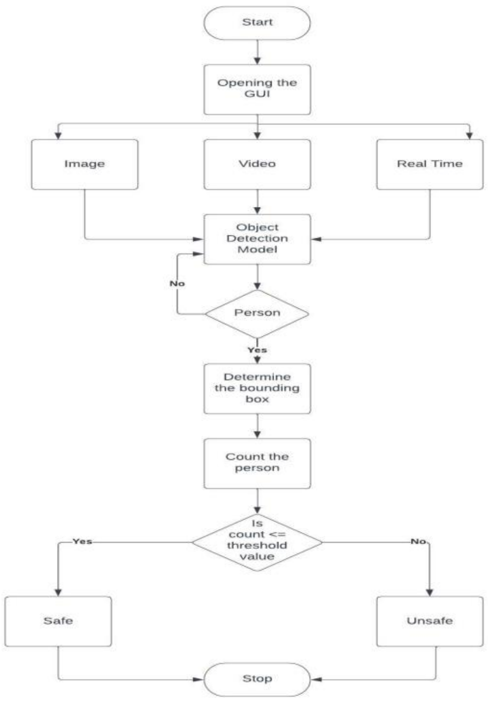
A. Rastogi et al. present a method for detecting real-time behaviour in egg breeders. They used the YOLO v3 for object detection in this case. The manual training of the model was completed. The entire dataset was gathered by hand. The outcome was satisfactory [3].

Sindagi et al. presented an end-to-end cascaded CNN that learns the crowd density map as well as a high-level global before aid in the prediction of density maps from images with large variations in scale and appearance. The high-level prior is a crowd count classification, in which crowds are classified into several groups based on the number of people present. Unfortunately, in this case, aerial images captured by drones were not considered in these works. Similarly, they proposed complex methods that may be too costly for real-time applications [4].

In [5], to increase the localization accuracy of the deep network model, Hsieh et al. recommended injecting spatial layout information. The authors tackled the issue of automobile counting in photos of parking lots to gauge the efficacy of their approach. Yet, in non-uniform and informal human-packed scenarios, having precise spatial layout information is an assumption that is not met.

By utilising multi-scale prediction and enhancing the fundamental classification network, Redmon et al. proposed YOLOv2 [6] and YOLOv3 [7] with quick detection speeds. In [8], authors offer YOLOv4, which makes use of a cross-stage partial connection and an additional bounding box regressor based on the Intersection over Union (IoU) in their backbone design, as well as mosaic and cut-mix data augmentation. All of the aforementioned techniques, nevertheless, are solely dependent on the detection choice made for each frame, therefore a fast-processing speed approach is still needed to increase accuracy.

Fig. 1 Flowchart of Project



IV. PERSON DETECTION

Object detection and person detection are two of the most recent and popular features of artificial intelligence and computer vision technologies. Object detection is a two-step procedure. First to train the model to be used and second, the original data set is used to perform object detection [10]. Person detection is based on object detection

Special Edition
NCASIT 2023, 29th April 2023
Department of Computer Engineering,
St. Vincent Pallotti College of Engineering & Technology, Nagpur,

systems that can detect human classification, which means they have the data and training to classify the detected object as human. Person detection is the process of identifying people in an image, video or real-time via webcam or CCTV. There are various techniques for person detection, including traditional computer vision methods and deep learning-based approaches.

Computer vision (CV) is a branch of computer science concerned with enabling computers to comprehend images. It is a topic of study that aims to create techniques that allow computers to "see" and interpret the content of digital pictures like photographs and movies [9]. Deep learning-based approaches, on the other hand, have shown significant improvements in person detection accuracy in recent years. These methods use convolutional neural networks (CNNs) to learn features directly from the data and can handle complex and diverse environments. The most popular deep learning-based person detection methods are based on the Single Shot Multibox Detector (SSD) or the You Only Look Once (YOLO) architecture.

V. PERSON COUNTING

Person counting refers to the task of counting the number of people in the scene [11]. The number of persons in the image, video stream or real-time can be counted once they have been identified. Counting the humans present in the videos has various applications in intelligent systems [12]. In this study, we used deep learning methods that make use of a convolutional neural network.

A person detection algorithm must be applied to the pre-processed picture or video stream. Once the people are detected, we can count the number of people in the image or video stream. This can be achieved by drawing a bounding box around each detected person and counting the number of boxes.

VI. ANALYSIS OF THE SAFETY OF THE ENVIRONMENT

First, set the threshold value for detection. Detect people and find the count based on appropriate boundary selection [1]. Once the number of people present in the area is known, crowd density analysis can be performed to identify potential hazards and risks. High crowd densities can lead to potential risks such as trampling, stampedes, or the spread of infectious diseases. Based on the crowd density analysis, the level of risk in the environment can be assessed.

VII. REQUIREMENTS

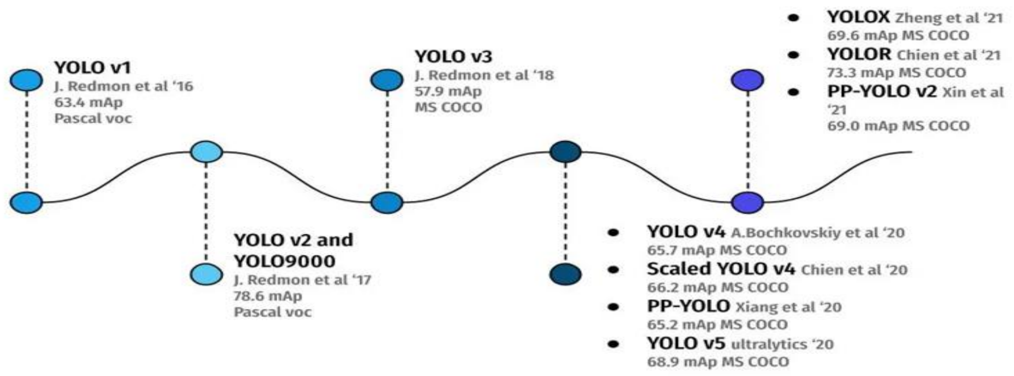
1. YOLO

To detect objects utilising the You Only Look Once (YOLO) technique. This methodology has a lot of benefits over other object-detecting strategies. While other algorithms, like Convolutional Neural Networks and Fast Convolutional Neural Networks, only partially examine the image, YOLO does so by using convolutional networks to predict the bounding boxes and the class probabilities for these boxes. As a result, YOLO detects the image more quickly than other algorithms. Over the years, several versions of YOLO have been developed to improve the accuracy and speed of the algorithm.

Fig. 2 The different versions of YOLO[15]

© INTERNATIONAL JOURNAL FOR RESEARCH PUBLICATION & SEMINAR
ISSN: 2278-6848 | Volume: 14 Issue: 03 | April - June 2023
Paper is available at http://www.jrps.in | Email : info@jrps.in
Refereed & Peer Reviewed

Special Edition
NCASIT 2023, 29th April 2023
Department of Computer Engineering,
St. Vincent Pallotti College of Engineering & Technology, Nagpur,



In this project, we have used YOLOv3. YOLOv3 was released in 2018 and further improved the accuracy and speed of the algorithm. It introduced several new features such as multi-scale prediction, feature concatenation, and spatial attention. YOLOv3 achieved an average precision of 87.4% on the COCO dataset. This technique focuses on object recognition and speed of detection without achieving absolute object localization.

Object detection in YOLOv3:

- The CNN's pre-training for image classification.
- Break the image up into small cells. If the centre of the object is within the cell, then this cell is in charge of detecting that object.
- Calculating the probability of classification of the object in the bounding box using the box's location and the confidence score.
- The dimensions of the bounding box are (x, y, width, height), and the range of the width and height is (0,1).
- The degree to which the cell matches the object determines the confidence score of the cell.
- For the purpose of producing the prediction tensor of size, the final layer of CNN (Pre-trained) is adjusted [1].

Object counting in YOLOv3:

- Set the detection threshold value as 6.
- Use the YOLO v3 Model to detect persons by retraining the Region-based Convolutional Network (R-CNN).
- Use the proper boundary selection to find people and count them.
- Process the information related to the previous step's count.
- The results of the people count in the form of people inside and outside the boundary [1].

2. CNN

Machine learning includes convolutional neural networks, also known as CNNs. It is a subset of several artificial neural network models that are employed for diverse purposes and data sets. [13] A CNN is a particular type of network design for deep learning algorithms that are utilised for tasks like image recognition and pixel data processing. Although there are different kinds of neural networks in deep learning, CNNs are the preferred network architecture for identifying and recognising objects. They are therefore ideally suited for computer vision (CV) activities and for applications where accurate object recognition is crucial, such as facial and self-driving automobile systems. Feature extraction and object recognition algorithms were developed, as well as computer vision tools like OpenCV [9].

3. OPEN CV

A free software library for computer vision and machine learning is called OpenCV (Open-Source Computer Vision Library). Computer vision is essential to artificial intelligence or serves as its main foundation. The OpenCV branch highlighted the problems that had to be fixed in order to take a picture and detect whether the movement had taken place. Real-time operations, which are crucial in

© INTERNATIONAL JOURNAL FOR RESEARCH PUBLICATION & SEMINAR
ISSN: 2278-6848 | Volume: 14 Issue: 03 | April - June 2023
Paper is available at http://www.jrps.in | Email : info@jrps.in
Refereed & Peer Reviewed

Special Edition
NCASIT 2023, 29th April 2023
Department of Computer Engineering,
St. Vincent Pallotti College of Engineering & Technology, Nagpur,

today's systems, are made possible by the big open-source package OpenCV for computer vision, machine learning, and image processing. It can be used to identify objects, people, and even human handwriting in still images and moving pictures. When used in conjunction with other libraries like NumPy, such as Python, the OpenCV array structure can be processed for analysis [9]. The library, which contains a thorough blend of both traditional and cutting-edge computer vision and machine learning approaches, contains more than 2500 optimised algorithms. For the purpose of processing images, OpenCV was developed. Each function and data structure were created with the image processing programmer in mind.

4. COCO DATASET

COCO (Common Objects in Context) is a large-scale image recognition, segmentation, and captioning dataset. In this project, the COCO dataset is used. The COCO object detection model has been used in comparing the images data set [14]. A sizable object identification, image segmentation, and captioning dataset called MS COCO were released by Microsoft. The COCO dataset is frequently used by machine learning and computer vision experts for a variety of computer vision projects. The COCO dataset consists of over 330,000 images, each annotated with object labels, object segmentations, and object captions. The object labels are from 80 categories, including people, animals, vehicles, and household items. The object segmentations are provided as polygonal masks that outline the object boundaries. Understanding visual situations is one of computer vision's main objectives; it entails identifying the items that are present, localising them in 2D and 3D, figuring out their characteristics, and defining the relationships between them. As a result, the dataset can be used to train algorithms for object detection and object categorization.

5. PYQT5

PyQt5 is a Python binding for the popular Qt library. Qt is a cross-platform application framework for creating desktop and mobile applications. PyQt5 provides an interface to the Qt library for Python programmers, allowing them to develop graphical user interfaces (GUIs) using Python. PyQt5 provides a wide range of features and modules for developing GUI applications, including support for buttons, labels, menus, text boxes, and many other GUI widgets. It also includes support for multimedia, database integration, and web browsing. In this project, PyQt5 is used to develop desktop applications with graphical user interfaces. With PyQt5, we created a window, dialogues, and menus, and customize the appearance and behaviour of the application.

6. IMUTILS

Imutils is a Python library that provides a set of utility functions for image processing and computer vision tasks. It is designed to simplify the use of OpenCV, a popular computer vision library, by providing a more Pythonic interface for common tasks. Imutils include a range of functions that are commonly used in image processing, such as resizing, rotating, cropping, and thresholding images. It also includes functions for working with contours, detecting edges, and performing perspective transformations. Imutils is designed to work seamlessly with OpenCV, and many of its functions are optimized for use with OpenCV. It provides a set of convenience functions that make it easier to perform common tasks with OpenCV, such as reading and displaying images and converting between colour spaces.

VIII. TEST CASES

CASE 1:

In case study (Fig.3), we took an image containing two people, and we used it to test our application. With the help of object detection technique i. e. YOLOv3, person detection was done. A bounding box detected the people in the picture. The output indicated that the region was safe and the person count was shown as 2. Since we've established a threshold value of 6. Therefore, if the count exceeds the threshold value, the output will be displayed as unsafe; conversely, if the value is less than 6, the output will be displayed as safe.

Special Edition
NCASIT 2023, 29th April 2023
Department of Computer Engineering,
St. Vincent Pallotti College of Engineering & Technology, Nagpur,

Fig.3. Case 1: Image detecting two persons

CASE 2:
In case study (Fig.4), we took a picture of a group of 8 individuals to test our application. Person detection was achieved using the YOLOv3 algorithm. The people in the image were detected using a bounding box. Although there were 8 people in the image, a bounding box only spotted 7 of them, hence the count was shown as 7. This resulted from the overlapping of two person. YOLOv3 may fail to detect individuals when two people are too close together, especially if they share a similar appearance or colour. This happens because the algorithm may perceive the two people as a single object, leading to inaccurate detection or no detection at all. While YOLOv3 is a powerful object detection algorithm, it has its limitations, and detecting objects that are too close together or partially occluded can be challenging. The threshold value was set as 6 and there were 7 individuals detected in the bounding box so the results showed that the area was unsafe.

Fig.4. Case 2: Image detecting seven instead of eight persons.

Sr. No	*PARAMETERS*	*CASE 1*	*CASE 2*
1	Input	Image containing two people	Image containing eight people
2	Threshold	6	6
3	Number of bounding boxes	2	7
4	Person Count	2	7
5	Analysis of the environment	Safe (Person Count < Threshold)	Unsafe (Person Count > Threshold)
6	Analysis	The count value is less than 6, the output will be displayed as safe.	This was caused by the overlap of two bounding boxes. When two people are too close together, YOLOv3 may fail to detect them. As a result, the threshold value was set to 6, and there were 7 people spotted in the bounding box, indicating that the region was unsafe.

Table No.1. Comparison Table of Case 1 and Case 2.

IX. PROJECT ACCURACY

	Precision	Recall	Accuracy
Image	80%	80%	75%
Video	100%	75%	66%
Real-Time	75%	75%	71%

© INTERNATIONAL JOURNAL FOR RESEARCH PUBLICATION & SEMINAR
ISSN: 2278-6848 | Volume: 14 Issue: 03 | April - June 2023
Paper is available at http://www.jrps.in | Email : info@jrps.in
Refereed & Peer Reviewed

Special Edition
NCASIT 2023, 29th April 2023
Department of Computer Engineering,
St. Vincent Pallotti College of Engineering & Technology, Nagpur,

Table No. 2. Result of Tested Dataset

Precision = (True Positive/ (True Positive + False Positive))
Recall = (True Positive/ (True Positive + False Negative))
Accuracy = ((True Positive + True Negative)/ (True Positive + True Negative + True Positive + False Positive))

Fig. 7. Person detection and analysing the safety of the environment through Webcam.

X. APPLICATION

Application in day-to-day life:
Elevators CCTV Surveillance - Crowd detection can also help optimize elevator dispatching to reduce wait times and congestion. The system can use real-time data on the number of people waiting for elevators and the occupancy status of each elevator to determine the most efficient way to dispatch elevators.
Metro Compartments CCTV-Crowded metro trains can be uncomfortable and even unsafe for passengers. Additionally, the technology can be used to detect overcrowding or people blocking doors, which can lead to safety hazards.
Bridge Supervision - Bridges are critical transportation infrastructures that often experience high volumes of traffic. Crowd detection technology

Fig. 5. Person detection and analysing the safety of the environment through Image.

Fig. 6. Person detection and analysing the safety of the environment through Video.

can be used to monitor traffic flow in real time and detect anomalies, such as accidents or congestion.
Carnivals - Attendees of carnivals and other large-scale events often want to know real-time information about the event, such as the location of attractions, wait times, and traffic conditions. Crowd detection technology can provide this information through digital displays or mobile apps, helping attendees make informed decisions about how to spend their time at the event.
Malls - Malls can become crowded and congested, particularly during peak shopping times. Crowd detection technology can be used to monitor the number of people entering and exiting the mall and detect potential bottlenecks or areas of congestion. This information can be used to adjust staffing levels or to redirect shoppers to less crowded areas of the mall to help manage the crowd.
Sports Stadiums - Planning and logistics are critical components of stadium operations. Crowd detection technology can provide valuable data and insights to stadium operators, such as crowd movement patterns, popular areas, and peak attendance times. This information can be used to optimize the layout of the stadium and improve the overall experience for fans
Classroom - Student safety is important while studying in schools and colleges when a virus outburst like a corona. Crowd detection technology can be used to monitor the number of students in the

Special Edition
NCASIT 2023, 29th April 2023
Department of Computer Engineering,
St. Vincent Pallotti College of Engineering & Technology, Nagpur,

classroom and the crowd can be controlled. This information can be used to maintain the safety of the students on school or college premises.

XI. CONCLUSION

In conclusion, person detection and analysing the safety of the environment based on the count are critical techniques in various applications, such as surveillance, security, and crowd management. With the advancement of computer vision and deep learning techniques, it is now possible to automate the detection and counting of people in real time, making it easier to monitor and improve the safety of different environments. The process involves collecting data from cameras or sensors, pre-processing the data, detecting people using the YOLOv3 algorithm, counting the number of people in specific areas, analysing the safety of the environment based on the count, and implementing necessary actions to improve safety. Moreover, these techniques can be combined with other technologies, such as IoT sensors and artificial intelligence, to provide a more comprehensive safety solution. In conclusion, person detection, counting, and analysing the safety of the environment based on the count can play a crucial role in enhancing safety, security, and crowd management. The future of these technologies looks promising, and it is expected that they will continue to evolve and become more advanced in the years to come, leading to safer and more secure environments for people.

XII. FUTURE SCOPE

Person detection and analysis of the safety of the environment techniques have significant potential for future applications in various domains. Some of the potential future scopes of person detection and counting and analysing the safety of the environment based on the count are:

1. *Smart Surveillance System*: Person detection and analysis of the safety of the environment techniques can be used to develop smart surveillance systems for public spaces, such as airports, railway stations, and shopping malls, to monitor crowd flow and determine the safety of the specific areas.
2. *Safety in Industrial Setting*: Person detection and analysis of the safety of the environment techniques can be used in industrial settings to monitor worker safety, such as detecting workers in hazardous areas.
3. *Crowd Management at Event*: Person detection and analysis of the safety of the environment techniques can be used in event management to monitor the crowd and prevent overcrowding, which can lead to safety hazards.
4. *Social Analytics*: Person detection and analysis of the safety of the environment can be used to analyse customer behaviour in retail environments, such as counting the number of people entering and leaving the store and if it is a safe area during a virus outburst.
5. *Healthcare*: Person detection and analysis of the safety of the environment can be used in healthcare settings to monitor patient movement and ensure that the patients are following safety protocols.

REFERENCES

1. Pooja Gupta, Varsha Sharma, Sunita Varma, "People detection and counting using YOLOv3 and SSD models".
2. H. Zhao, Z. Li, L. Fang, T. Zhang, A Balanced "Feature Fusion SSD for Object Detection, Neural Process". Lett. 51 (3) (2020) 2789–2806, https://doi.org/10.1007/s11063-020-10228-5.
3. A. Rastogi, B.S. Ryuh, "Teat detection algorithm: YOLO vs Haar-cascade", J. Mech.Sci. Technol. 33 (4) (2019) 1869–1874, https://doi.org/10.1007/s12206-019-0339-5.
4. V. A. Sindagi and V. M. Patel, "CNN-based cascaded multi-task learning of high-level prior and density estimation for crowd counting," in 2017 14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS). IEEE, 2017, pp. 1–6.

5. M.-R. Hsieh, Y.-L. Lin, and W. H. Hsu, "Drone-based object counting by spatially regularized regional proposal network," in Proceedings of. The IEEE International Conference on Computer Vision, 2017, pp. 4145–4153.
6. Redmon J, Farhadi A (2017) "Yolo9000: better, faster, stronger". In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 7263–7271.
7. Redmon J, Farhadi A (2018) "Yolov3: An incremental improvement". arXiv:1804.02767.
8. Bochkovskiy A, Wang C-Y, Liao H-YM (2020) "Yolov4: Optimal speed and accuracy of object detection".
9. Dr Yusuf Perwej, Nikhat Akhtar, Shivam Chaturvedi, (2022) "An Intelligent Motion Detection Using OpenCV".
10. W. Fang, L. Wang, P. Ren, "Tinier-YOLO: A Real-Time Object Detection Method for Constrained Environments", IEEE Access 8 (2020) 1935–1944, https://doi.org/10.1109/Access.628763910.1109/ACCESS.2019.2961959
11. Giovanna Castellano, Corrado Mencar, Ciro Castiello, Gennaro Vessio, "Crowd Counting from Unmanned Aerial Vehicles with Fully-Convolutional Neural Networks",2020
12. Y. Wang, Y. Zou, W. Wang, "Manifold-Based Visual Object Counting", IEEE Trans. Image Process. 27 (7) (2018) 3248–3263, https://doi.org/10.1109/TIP.2018.2799328.
13. https://www.techtarget.com/searchenterpriseai/definition/convolutional-neural-network#:~:text=How%20convolutional%20neural%20networks%20work,more%20detailed%20after%20each%20layer.
14. Afiq Harith Ahamad, Norliza Zaini, Mohd Fuad Abdul Latip, "Person Detection for Social Distancing and Safety Violation Alert based on Segmented ROI".
15. https://pyimagesearch.com/2022/04/04/introduction-to-the-yolo-family/