Crowd Detection Using Artificial Intelligence

Shital Dnyaneshwar Cheke
B. Kolte College
of Engineering, Malkapur,

Manjiri U Karande Padm. Dr. V.
Padm. Dr. V. B. Kolte College
of Engineering, MalkapurIndia.
India.

Abstract: - The rapid expansion of the global population has led to an upsurge in public gatherings, giving rise to concerns about overcrowding and safety. Among these safety issues, crowd-smashing accidents have emerged as unexpected and swiftly escalating situations that pose significant risks to the general public. Manually predicting and managing such chaotic scenarios presents numerous challenges. As witnessed during the COVID-19 pandemic, maintaining physical distance is crucial in curbing the transmission of viruses from person to person. The World Health Organization (WHO) mandates limiting the number of people in a given space. To address these challenges, AI-based monitoring systems now incorporate cutting-edge object detection algorithms, with YOLOv8 being a prominent example. This advanced approach provides real-time crowd density assessments and instant alerts to authorities, enabling swift action and accident prevention. Human detection and crowd counting are fundamental tasks in computer vision, serving practical purposes such as surveillance, security, crowd management, and traffic analysis. Deep learning models, particularly the You Only Look Once (YOLO) approach, have achieved remarkable success in these domains. In our research paper, we delve into the current state of the art in human detection and crowd counting using YOLOv8, discussing both its advantages and limitations. Notably, our proposed model extends beyond crowd counting by detecting abnormal activities within the crowd, including weapons, fires, falls, and smoke. By identifying potentially hazardous crowd densities and promptly detecting abnormal incidents, our system not only prevents disasters like crowd-smashing, detecting physical distances but also strengthens overall security measures.

Keywords: - Yolo, Overcrowding, Convolutional Neural Network (CNN), Real- time video processing, ROI, LOI, Object Detection, Ultralytics

I. INTRODUCTION

According to a report published by the World Health Organization (WHO), the COVID-19 pandemic has affected millions of people worldwide, resulting in significant fatalities. For an extended period, countries suspended routine activities thatformed the basis of their livelihoods. Gradually, countries have resumed daily life while adhering to health protocols prescribed by the WHO [1]. One crucial health measure is crowd prevention, achieved by determining the maximum number of people an area can accommodate based on its size. When the number of individuals per square meter exceeds a certain threshold, appropriate actions can be taken. Manual monitoring of crowd density is challenging in various locations, including banks, train stations, shopping malls, and schools. To address this, computer- based systems utilize cameras to provide real-time control. In our study, we select a region within a video recording and estimate its size. A threshold value is set, and the system counts the people in that area. If the crowd exceeds the area's capacity, it is marked in red; otherwise, it is marked in green. Deep learning is a prominent field within machine learning, with algorithms that have found successful applications across various domains. These applications include image processing, machine translation, and natural language processing. However, deep learning extends beyond these areas [1]. For instance, in recent years, deep learning algorithms have been used for sound classification, language recognition, and even cancer diagnosis. When it comes to person detection, there are two primary methods: Region of Interest (ROI)and Line of Interest (LOI). In the ROI-based approach, the goal is to estimate the number of people within a limited region. Conversely, in the LOI-based method, we estimate the number of people crossing a predetermined line. Deep learning models play a crucial role in the ROI-based system. They are frequently employed in object detection and classification tasks. Researchers have increasingly turned to deep learning for people counting challenges as well. In our study, we utilized deep learning methods with an ROI-based approach. Specifically, we perform object detection to locate objects within an image and classify each object. The detected objects are enclosed in bounding boxes, and their corresponding classes are estimated. The You Only Look Once (YOLO) models, which leverage deep learning techniques, are commonly used for object detection. Detecting and counting people in various spaces have become critical tasks for video surveillance systems. As a result, there has been a surge in research studies in recent years. YOLO models are frequently employed to detect people, count them, and even measure the distance between individuals [1].

II. MATERIALS AND METHODS

Certainly! In this study, our focus lies on computing the area of a region with predefined boundaries within an area. Additionally, we aim to determine the maximum number of people that can occupy this specific region. To achieve this, we utilize cameras that observe the area. Specifically, we employ the YOLOv8 model for detecting and counting people within the designated area [1].

III. BACKGROUND STUDY AND RELATED WORK

Social distancing has emerged as a reliable technique to curb the spread of infectious diseases. Against this backdrop, in December 2019, when COVID-19 surfaced in Wuhan, China, authorities implemented social distancing as an unprecedented measure on January 23, 2020. Within a month, the outbreak in China reached its peak during the first week of February, with 2,000 to 4,000 new confirmed cases daily. Subsequently, for the first time since the outbreak, there was a glimmer of hope: no new confirmed cases were reported for five consecutive days until March 23, 2020. This underscores the effectiveness of social distancing measures initially enacted in China, which were later adoptedworldwide to combat COVID-19.

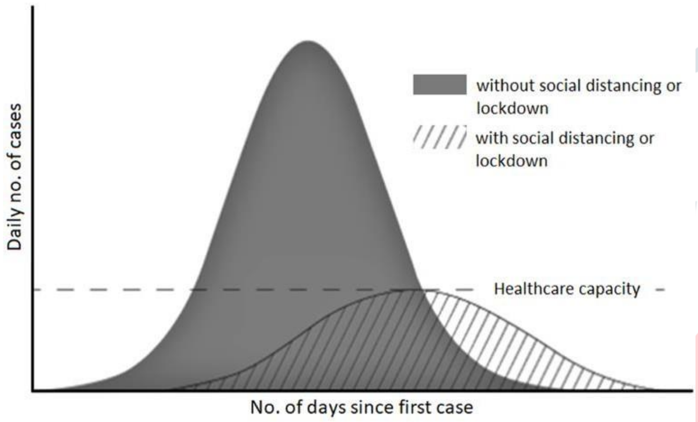

Fig No. 1 Social distancing vs healthcare capacity. [12]

A. Investigating the Impact of Social Distancing Measures on COVID-19 Spread

In an effort to understand the effects of social distancing measures on the spread of the COVID-19 epidemic, researchers employed synthetic location-specific contact patterns. These patterns were used to simulate the ongoing trajectory of the outbreak using Susceptible-Exposed-Infected-Removed (SEIR) models. Notably, the study highlighted that prematurely and abruptly lifting social distancing restrictions could lead to an earlier secondary peak in infections. However, this secondary peak could be mitigated by gradually easing interventions. While social distancing is essential for flattening the infection curve, it also comes with economic challenges. The United States faced a unique situation where lack of consensus among policymakers hindered early adoption of social distancing measures. Consequently, ongoing harm to public health ensued. Despite the impact on economic productivity, researchers continue to explore ways to mitigate losses caused by social distancing.

B. Exploring the Relationship Between Social Distancing and Economic Status

In light of the current context, we can investigate the link between the stringency of social distancing measures and the economic conditions of a region. Research suggests that allowing intermediate levels of activity while avoiding a massive outbreak is a prudent approach. Since the onset of the novel coronavirus pandemic, countries worldwide have turned to technology-based solutions to combat the outbreak. Notably, developed nations like India and South Korea have harnessed GPS technology to track the movements of suspected or infected individuals. This monitoring helps assess the risk of exposure among healthy people. For instance, in India, the government employs the Arogya Setu App, which utilizes GPS and Bluetooth to locate the presence of COVID-19 patients in the vicinity. The app also encourages others to maintain a safe distance from infected individuals. Conversely, law enforcement agencies have adopted alternative methods, such as deploying drones and other surveillance cameras. These tools help detect mass gatherings of people, enabling regulatory actions to disperse crowds. While manual intervention in critical situationscan aid in flattening the curve, it also introduces unique threats to public safety and poses challenges for the workforce [10].

C. Human Detection in Visual Surveillance Systems: - Challenges and Approaches.

Human detection using visual surveillance systems is a well-established area of research. However, it heavily relies on manual methods for identifying unusual activities, which inherently limits its capabilities. Recent advancements underscore the need for intelligent systems that can automatically detect and capture human activities. Despite being an ambitious goal, human detection faces several constraints. These include:

1. Low-resolution video: Often, surveillance footage lacks high resolution, making accurate detection challenging.

2. Varying articulated pose: Human poses can vary significantly, affecting detection accuracy.

3. Clothing variations: Different clothing styles and colors further complicate the task.

4. Lighting and background complexities: Changing lighting conditions and cluttered backgrounds introduce noise.

5. Limited machine vision capabilities: The inherent limitations of machine vision systems impact performance.

To enhance detection accuracy, prior knowledge of these challenges is crucial. Detecting moving objects involves two stages: - object detection and object classification.

D. Object Detection Techniques: -

Background Subtraction: This method computes the difference between the current frame and a background frame (usually the first frame). It operates at the pixel or block level. Popular approaches include:

1. Adaptive Gaussian mixture models
2. Temporal differencing
3. Hierarchical background models
4. Warping background
5. Non-parametric background modeling

Optical Flow: In this technique, flow vectors associated with an object's motion are characterized over a time span. It identifies regions in motion within a sequence of images. While these methods contribute to human detection, ongoing research aims to overcome limitations and improve performance. As surveillance systems evolve, addressing these challenges becomes critical for effective human activity recognition [10].

E. Advancements in Object Detection Techniques: - YOLO vs. CNN

Researchers have highlighted certain limitations of optical flow-based techniques. These methods tend to have computational overheads and are sensitive to motion-related outliers, including noise, color variations, and lighting changes. However, recent developments have efficiently addressed object detection challenges. Over the past decade, several advanced techniques have emerged, including:

1. Convolutional Neural Networks (CNN)

2. Region-based CNN

3. Faster Region-based CNN

These approaches utilize region proposal techniques to generate an objectness score before classifying objects. Subsequently, bounding boxes are generated around the objects of interest for visualization and statistical analysis. While these methods are efficient, they suffer from longer training times. In contrast, the You Only Look Once (YOLO) approach takes a different route. Instead of relying solely on classification, YOLO employs a regression-based method to separate bounding boxes dimensionally and interpret their class probabilities. This unique approach offers a fresh perspective on object detection, emphasizing efficiency and accuracy [10].

F. Advancements in Object Detection: Speed and Efficiency Trade-offs

In this method, the designed framework efficiently divides an image into several portions, each represented by bounding boxes along with class probability scores. This approach significantly improves speed while maintaining efficiency. The detectormodule demonstrates powerful generalization capabilities,representing the entire image [10]. Recent research has explored various applications, with crowd counting emerging as a promising area. For instance:

1. Researchers focused on crowd detection and person count by proposing multiple height homographies for head-top detection. They successfully addressed occlusion challenges associated with video surveillance applications.

2. Another study developed an electronic advertising application based on the concept of crowd counting.

Similarly, a vision-based people counting model was proposed, utilizing inputs from stationary cameras. Background subtraction techniques were employed to train the model on crowd appearance and foreground shape in videos. These efforts contribute to enhancing surveillance systems and addressing real- world challenges related to crowd management [10].

IV. PROPOSED ARCHITECTURE

"Object detection enables machines to understand and interact with the visual world."

YOLOv8 the Cutting-Edge in Real-Time Object Detection. The latest addition to the YOLO series, YOLOv8, marks a significant advancement in real-time object detection capabilities. Researchers and developers now have access to state-of-the-art accuracy and speed, making YOLOv8 the preferred choice for applications in robotics, autonomous driving, and video surveillance.

YOLO	YEAR	MAIN ADVANCEMENTS
YOLOv1	2015	Introduction of real-time object detection using a grid-based approach
YOLOv2	2016	Incorporation of anchor boxes, feature pyramid networks, and multi-scale prediction
YOLOv3	2018	Improvements in accuracy and speed with the introduction of Darknet-53 and multiple detection scales
YOLOv8	2021	State-of-the-art advancements in real-time object detection with improved accuracy and speed

Table No. 1 THE EVOLUTION OF YOLO [11]

With each iteration, YOLO (You Only Look Once) has relentlessly pushed the boundaries of object detection in computer vision. Driven by continuous research and innovation, the evolution from YOLOv1 to YOLOv8 stands as a testament to the collective efforts of researchers and practitioners. These advancements enable real-time object detection systems to operate with unparalleled efficiency and accuracy.

A. Main Features of YOLOv8 for Object Detection: -

YOLOv8 boasts a multitude of powerful features that position it asan exceptional choice for object detection tasks. Whether you require pre-trained models or seek to create custom models tailored to specific object types, YOLOv8 provides a comprehensive suite of capabilities to meet your needs.

1. Pre-trained Models: - YOLOv8 provides the advantage of leveraging pre-trained models. These models have already been trained on extensive datasets, such as COCO (Common Objects in Context). As a result, they have acquired the ability to identify and classify a diverse array of objects. This versatility makes YOLOv8 well-suited for a wide range of object detection applications.

2. Custom Models: - YOLOv8 not only provides pre-trained models but also empowers users to create custom models tailored to their specific object detection requirements. This involves the crucial process of data preparation, where you meticulously select and label the desired object types within your training dataset. By training a custom model, you can achieve higher accuracy and precision for object detection tasks that are unique to your specific application domain.

3. Data Preparation: - Data preparation plays a pivotal role in training custom models using YOLOv8. This process entails meticulous curation and precise labeling of the training dataset, ensuring that the model receives accurate examples of the desired object types. Thoughtful data preparation significantly impacts the

effectiveness and performance of the object detection model.

4. Web Application Support: - YOLOv8 takes a significant leap by supporting the development of web applications for real-time object detection. Its seamless integration into web browsers empowers users to create robust and intuitive interfaces without requiring additional software installations.

V. RESULT

Let's examine the functioning of YOLO by observing theoutcomes of image classification shown in figures 2 and 3. Ourgoal is to identify the objects within the image. Image classificationallows us to discern that there is a person in the image. However, employing object localization provides us with additional details regarding the type and position of the object within the image, indicated by a bounding rectangle.

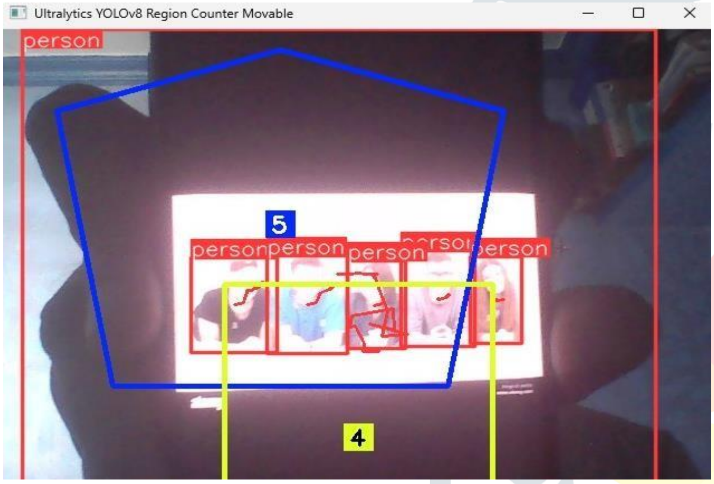

Fig 2 Image Classification and Localization

Now, in case of yolo algorithm we have a vector like this

$$\begin{bmatrix} Pc & 1 \\ Bx & 50 \\ By & 30 \\ Bw & 60 \\ Bh & 80 \\ C1 & 0 \\ C2 & 1 \end{bmatrix}$$

Where, Pc = Existence of a dog or a person
Bx = Co-ordinate of x axis to the centre
By = Co-ordinate of y axis to the centre
Bw = Width of red box
Bh = Height of red box
C1 = Class of dog
C2 = Class of person

To simultaneously identify the object and its bounding box, we can utilize a neural network that has been trained for this purpose. As part of a supervised learning task, it's necessary to label each image in our dataset with bounding boxes. Since neural networks require numerical data, these bounding boxes must be transformed into vector form, as illustrated.

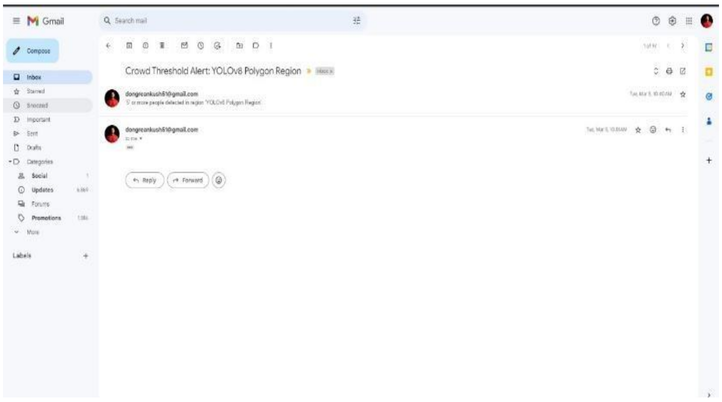

Fig 3 Image Classification and Localization

VI. CONCLUSION

YOLOv8 pioneering Real-Time Object Detection. In the realm of computer vision, YOLOv8 stands as a remarkable leap forward in real-time object detection. Its deep learning model, enriched with an enhanced architecture and cutting-edge features, delivers highly accurate results across diverse domains. With YOLOv8, real-time object detection becomes a tangible reality in applications such as robotics, autonomous driving, and video monitoring. Its ability to swiftly and precisely identify objects within a scene opens up new horizons for industry and research. Building upon the successes of its predecessors, YOLOv8 solidifies YOLO's position as a leading deep learning model for object detection. Its state-of-the-art algorithm and advanced computer vision techniques make it the go-to choice for professionals in the field. As the field of computer vision continues to evolve, ongoing research and improvements will propel real-time object detection systems to unprecedented heights. The future holds promising directions for YOLOv8 and its counterparts as we strive to enhance performance, accuracy, and usability in these essential technologies.

VII. REFRENCES

[1] Mehmet Şirin Gündüz & Gültekin Işık, "A new YOLO-based method for real-time crowd detection from video and performance analysis of YOLO models", J Real-Time Image Proc 20, 5 (2023).

[2] Mubin Modi, Zaid Marouf, Anvay Wankhede and Dr. Shabina Sayed, "A Real-time Crowd Counting application with Live analytics and selective detection using YOLOv4", JETIR May 2021, Volume 8, Issue 5.

[3] Seung-Hyun Lee, Jaekyoung Moon & Minho Lee, "A Region of Interest Based Image Segmentation Method using a Biologically Motivated Selective Attention Model", The 2006 IEEE International Joint Conference on Neural Network Proceedings.

[4] www.encord.com, "YOLO models for Object Detection Explained [YOLOv8]", Nikolaj Buhl, February 8, 2023.

[5] www.alamy.com, Alamy Ltd

[6] www. medium.com

[7] www.labellerr.com

[8] M. Monika1, Udutha Rajender, A. Tamizhselvi and Aniruddha S Rumale, "Real-time Object Detection in videos using Deep Learning models". ICTACT journal on image and video processing, November 2023, volume: 14, issue: 02.

[9] Anuj Kumar Suhane, Aryan Vani Harsh, Parihar, Udit Raghuwanshi, Arjun Nimbark, Lucky Saxena, "Human detection and Crowd Counting using YOLO," www.researchgate.net.

[10] Narinder Singh Punn, Sanjay Kumar Sonbhadra, Sonali Agarwal and Gaurav Rai, "Monitoring COVID-19 social distancing with person detection and tracking via fine-tuned YOLO v3 and Deepsort techniques", 27th April 2021.

[11] www. keylabs.ai.com

[12] Punn NS , Sonbhadra SK , Agarwal S , Rai G, "Monitoring COVID-19 social distancing with person detection and tracking via fine-tuned YOLO v3 and Deepsort techniques", 24 Apr 2021.