Social Distance Monitoring and Mask Detection Using Deep Learning Techniques

M Sneha[1], Prasad A M[2]

[1]Student, M.Tech, Computer Science Engineering, Dayananda Sagar College of Engineering
[2]Professor, Department of Computer Science Engineering, Dayananda Sagar College of Engineering

---

Abstract - *The proposed system uses a combination of image processing and machine learning techniques to analyze video feeds and image sequences from surveillance cameras and other visual sources. Accurately and efficiently identify people in a scene using the YOLOv5 algorithm, a state-of-the-art deep learning object detection model. By analyzing the spatial relationships between detected individuals, the system can infer whether social distancing guidelines are being followed. This methodology includes several critical steps. The system then uses image preprocessing techniques to improve the quality of the input image and extract meaningful features. Combining background subtraction with foreground segmentation can help identify areas of interest populated by people. These regions are input into his YOLOv5 model for object detection. The system uses a geometric analysis approach to determine if social distancing is being followed. By estimating the distance between each pair of detected individuals, the interpersonal distance is calculated and compared to a pre-defined social distance threshold. Violations are reported and displayed in real time so that relevant authorities are notified in a timely manner.*

Key Words: **YoloV5 algorithm, social distancing, image processing, real-time detection**

1. INTRODUCTION

The outbreak of the novel coronavirus disease (COVID-19) in 2019 has had a major impact on societies around the world. To curb the spread of the virus, health officials and governments have taken various measures, including social distancing. The Objective of social distancing is to reduce close physical interactions between individuals, thereby curbing transmission of the virus. Social distancing in public places is essential to ensure public safety and stop the spread of the pandemic. However, manually monitoring compliance with social distancing guidelines can be labor-intensive, time-consuming, and error-prone. To meet this challenge, researchers and engineers have developed automated social distancing detection systems using advanced technologies such as image recognition and deep learning algorithms. This research focuses on utilizing image recognition technology and his YOLOv5 algorithm, a state-of-the-art deep learning model, to enable accurate real-time monitoring of social distancing compliance in various public environments. The development of a real-time social distancing system using image recognition technology and his YOLOv5 algorithm is expected to help curb the spread of infectious diseases like COVID-19. The proposed system provides a cost-effective, accurate and efficient solution for enforcing social distancing guidelines in various public settings by automating the surveillance process. Its applications span multiple areas such as public health, workplaces, events, transportation, retail and education, making it a valuable tool for promoting safety and mitigating the impact of pandemics and infectious diseases. As technology continues to advance, such systems could play an important role in shaping a safer and more secure future for society.

1.1 PROBLEM STATEMENT

This issue explores the need for automated systems to enforce social distancing protocols in public spaces during the COVID-19 pandemic. The proposed system aims to detect adherence to social distancing guidelines in real time using image recognition technology and his YOLOv5 algorithm. By analyzing video feeds and image sequences from surveillance cameras, the system can identify people and assess their spatial relationships to determine if social distancing is being followed. The proposed system aims to detect adherence to social distancing guidelines in real time using image recognition technology and his YOLOv5 algorithm. By analyzing video feeds and image sequences from surveillance cameras, the system can identify people and assess their spatial relationships to determine if social distancing is being followed.

1.2 OBJECTIVES

1) Development of real-time social distance detection system: The main purpose of this research is to design and develop a real-time social distance detection system using image recognition technology and his YOLOv5 algorithm. The system processes video feeds or series of images from surveillance cameras to identify people at crime scenes.
2) Fine Tuning of YOLOv5 Social Distance Detection Algorithm: To achieve accurate and efficient social distancing detection, the YOLOv5 algorithm has been improved using a dedicated dataset created specifically for this purpose. This dataset contains a variety of

scenarios in which individuals exhibit both social distancing compliant and non-compliant behaviors.

3) Image preprocessing and feature extraction: We employ efficient image preprocessing techniques to improve the quality of the input images and extract features relevant to social distance analysis. Background subtraction and foreground segmentation techniques are used to identify regions of interest populated by an individual to ensure accurate inputs to his YOLOv5 model for object detection.
4) Geometric analysis to assess social distancing: We apply a geometric analysis approach to estimate the distance between each pair of detected individuals. By calculating interpersonal distances, the system compares them to pre-defined social distancing thresholds to determine compliance. Violations are immediately reported, displayed in real time, and relevant authorities are notified.
5) Performance evaluation and comparative analysis: The developed social distancing detection system undergoes rigorous performance evaluation using various datasets including both controlled environments and real-world scenarios.

2. RELATED WORK

Existing systems often use state-of-the-art object detection algorithms such as YOLOV3 (You Only Look Once), Faster R-CNN (Region-based Convolutional Neural Networks), or SSD (Single Shot Multibox Detector). These algorithms are trained to recognize people and other related objects in images or video frames. Many systems use pre-trained deep learning models to recognize objects in images. These models are often trained on large datasets containing annotated images of people and objects, so they can transition well to new scenarios. To adapt the object detection algorithm to social distancing detection, fine-tuning is performed on a dedicated dataset containing examples of compliant and non-compliant social distancing scenarios. Fine-tuning helps optimize the model for accurately detecting social distancing violations. Image preprocessing techniques are applied to improve the quality of input images or video images. Techniques such as background subtraction, image resizing, and noise reduction are commonly used to improve the accuracy of the recognition process. This is important for monitoring social distancing adherence during the pandemic. This research report presents a social distancing monitoring system that combines YOLOv3 with image segmentation technology. The authors focus on recognizing people in public spaces and analyzing their spatial relationships to assess social distancing compliance. Detected violations are often visualized by drawing bounding boxes around people who are too close. Alerts can be generated to notify security personnel or relevant authorities of social distancing violations. Existing systems are evaluated against various performance indicators such as accuracy, precision, recall, and F1 score. These systems are typically designed for use in various public spaces such as airports, shopping malls, public transport stations, and workplaces. It can be integrated into your existing surveillance system or operated as a standalone solution. Privacy considerations are important as social distancing systems incorporate video surveillance.

3. PROPOSED WORK

YOLOv5 uses deep neural networks as its backbone to extract high-level features from input images. The network architecture is based on CSPDarknet53, a lightweight and efficient variant of the Darknet backbone used in his previous YOLO version. CSPDarknet53 uses the concept of inter-stage sub-connections to improve information flow and reduce computational complexity. A backbone network processes an input image through multiple convolution layers to extract hierarchical features at different scales. These features capture contextual and semantic information in images, enabling models to recognize objects of varying size and complexity. The following points epitomize the main factors of this approach. Deep literacy attracts attention in mortal recognition Purpose of discovery. Develop a social distance discovery tool that can descry the distance between people who should be defended. Evaluation of bracket results by assaying real- time videotape streams camera. This proposed system makes it delicate for people to get together and fraternize. Numerous people gather in sacred places bad bad. Every country in the world is now in a state of Lockdowns forcing people to stay home. But ultimately people will start doing it Visit further public places, religious places, sightseer spots and further.

YOLOv5 introduces a new neck architecture called PANet (Path Aggregation Network). PANet is a feature pyramid fusion technique that combines features from different layers of backbone networks to improve object detection performance at different scales. This allows the model to effectively handle objects of varying sizes. This research focuses on utilizing image recognition technology and his YOLOv5 algorithm, a state-of-the-art deep learning model to enable accurate real-time monitoring of social distancing compliance in various public environments.

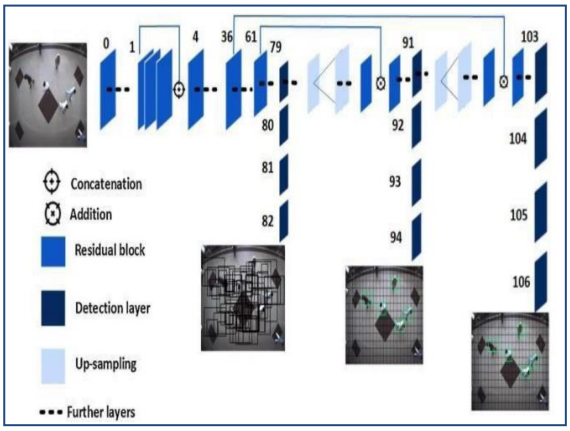

Figure 3.1 YOLO V5 Network Architecture

International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 10 Issue: 07 | July 2023 www.irjet.net p-ISSN: 2395-0072

4. CENTROID ALGORITHM

In this work, YOLOv5, which improves prediction accuracy, is used for human detection. Especially for small format objects. The main advantage is that the network is customized. Structures for multi-scale object recognition. In addition, it is also used to classify objects various independent logistics instead of softmax. This function is in the model training is done using convolution layers, also called residual blocks. A block consists of many layers of convolution and skipping connections. A unique feature is that detection is performed on her three different scales. Practice convolution layers with specific steps to down sample feature maps and transfer properties of immutable size. 3 feature maps are used for the object recognition. Trained using an overhead data set. Transfer learning for this purpose.

An approach is chosen that increases the efficiency of the model. With transfer learning the model is additionally trained without losing any valuable information of the model existing model. Additionally, an additional trained overhead dataset is added existing architecture. In this way the model was pre-trained, the information is retrained to provide better and faster recognition results for both. In the architecture shown in Figure 3.1, we used a single-level network over the input image. During extraction, the architecture takes full advantage of convolution layers and class prediction. Connected layers are used. For human identification, the input frame is split as follows: The area of S × S, also called a grid cell. These cells refer to bounding boxes. Estimation and class probabilities. Predict the probability of hitting the center of the person's bounding box is within a grid cell:

$$Conf(p)=Pr(p)\times IOU(pred,actual)$$

Pr(p) indicates whether there is an occupant within the detected bounding box. The value of Pr(p) is 1 for yes and 0 for no. Determining IoU (default, actual. The intersection point on the union of the actual bounding box and the predicted bounding box. then, manually labeled ground truth boxes (actual) in the training data set represented by BoxT, the predicted bounding box is displayed as a BoxP region representing the clipping region. Predict and set the tolerance for each recognized person in the input frame. A confidence value is applied after prediction to achieve the best bounding box. H, W, X, Y are estimated for each predicted bounding box. The bounding box looks like this: Coordinates are defined by x,y, width and height are determined by w,h. model produces the predicted bounding box values threshold is defined process high confidence values and discard low confidence values. A non-maximal suppression is used to derive the final position parameters of the detected objects. Finally, the loss function for the detected bounding boxes is calculated. Given the loss function is the sum of three functions: regression, classification, and trust. Classification is lost if an object is detected in each grid cell. Computed as the squared error of the conditional probabilities.

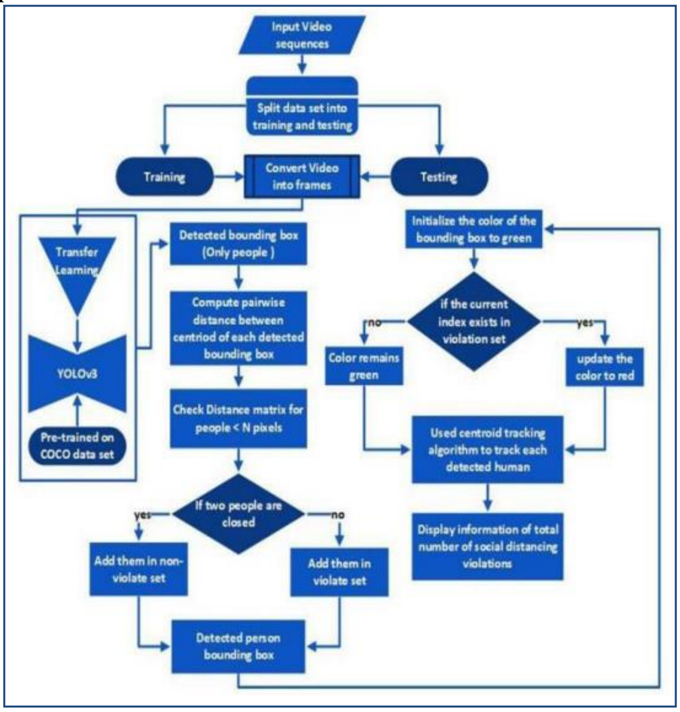

Figure 4.1 Flowchart for of overhead view of social distancing framework

Object detection with YOLO (You Only Look Once) involves several key steps to accurately identify and locate objects in images or videos. Here is the step-by-step process of object detection using YOLO. Collect a data set of images or video frames and annotate them with labeled bounding boxes around objects of interest. Each bounding box must be associated with an appropriate class label. Based on your needs, choose the appropriate YOLO variant (such as YOLO V5, YOLOv4, or YOLOv5) and download pre-trained weights or configuration files from the official YOLO repository. Resize the input image or video image to a fixed size compatible with the input requirements of the YOLO model. Commonly used sizes include 416 x 416 pixels or 608 x 608 pixels. Load the YOLO model using the configuration file and pretrained weights obtained in step 2. Depending on your programming language, you can use libraries such as Darknet (for YOLO V5 and YOLOv4) or PyTorch/TensorFlow (for YOLOv5).

Pass the preprocessed image or frame through to the YOLO model. The model outputs predictions in the form of bounding boxes, confidence scores, and class probabilities for each detected object. Non-Maximum Suppression (NMS): Apply non-maximal suppression (NMS) to filter out duplicate redundant bounding boxes. NMS removes duplicate detections by keeping only the most confident bounding boxes among bounding boxes with high overlap. Running NMS gives us a final set of bounding boxes containing the associated class labels and confidence values. Additionally, you can filter out detections below a certain confidence threshold to reduce false positives. Optionally draw bounding boxes and class labels over the original input image or box to visualize the detected objects. This step helps you understand the model's performance and validate the recognition results. To perform real-time object detection on the video stream, repeat steps 3-8 for each image. Depending on the model and hardware used, optimizations such as batching and hardware acceleration (such as GPU) can be used to achieve real-time performance. Evaluate the performance of the YOLO model by comparing the detected objects to the ground truth annotation of the dataset. Common evaluation metrics include mean average accuracy (mAP) and intersection over union (IoU). Fine tuning (optional) If the pre-trained YOLO model does not perform well on a particular dataset or task, consider fine-tuning the model on a custom dataset to improve detection accuracy.

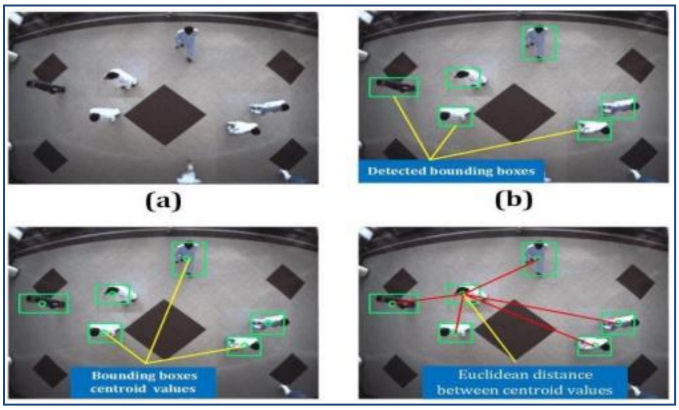

Figure 4.2 Bounding Box and Centroid Detection

5. OBSERVATION & RESULTS

Images can be used to visualize the model recognizing people in different scene locations. As shown in the image, it effectively identifies people with different characteristics and also calculates the social distance between people.

No social distancing violations were detected in the frames in figures (a)–(c) because all subjects were marked with green rectangular boxes in the automated frames. A violation is detected in the frame in figure (e). However, since the number of people present at the crime scene is low compared to figure (b) where everyone is social distancing, no violations are observed for her. In Figures (d)-(f), violations due to close interactions between people are recorded by an automated system. The same behavior can be seen in Figures (g)-(i), but both (g) and (h) have approximately 12 people, and the degree of injury is three times greater in (h) than in (g). In Figures (d)-(f), several people are walking and their entry into the scene is detected and monitored. The framework effectively recognized and demarcated violations of social distancing between people. If people are too close, the box will appear as a red rectangle.

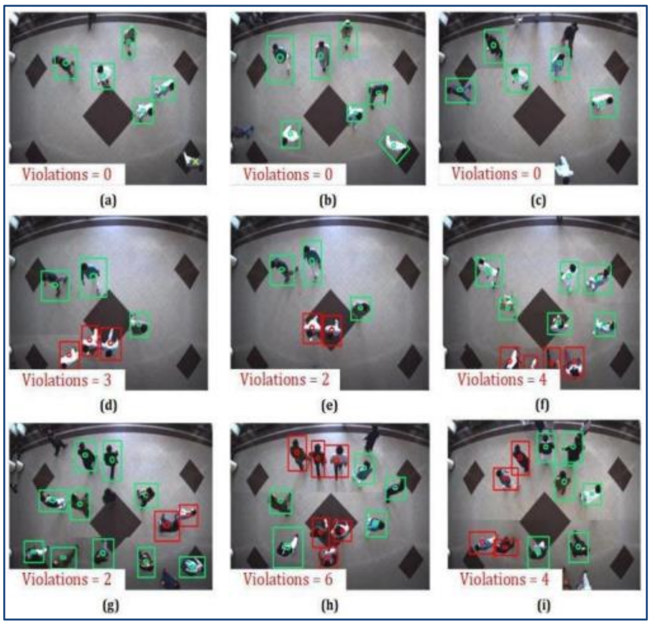

Figure 5.1 Result of social distancing in closed area

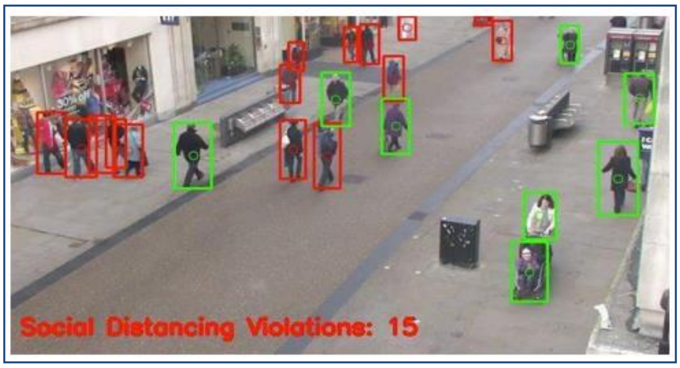

Figure 5.2 Result of Social Distancing in street

Figure 5.3 Result of Face Mask Detection

To validate social distancing violations between people, physical distance approximations to pixels are used and a threshold is defined. The violation threshold is used to check if the distance value violates the social minimum. Whether a distance is set, also, a centroid shadowing algorithm is used to track people in the scene. trials show that the frame can efficiently descry people who violate social distancing and get too close. also, transfer literacy ways ameliorate the overall effectiveness and delicacy of recognition models. For pretrained models without transfer literacy, models with transfer literacy achieve 92 and 95 recognition delicacy. The model has a shadowing delicacy of 95. In the future, it may work better for different inner and out-of-door surroundings. Colorful discovery and shadowing algorithms can be used to track individualities violating or exceeding social distancing thresholds.

6. COCNLUSION

In summary, applying image recognition technology and the YOLOv5 algorithm to monitor social distancing has proven to be a powerful solution to address the challenges posed by the COVID-19 pandemic and similar public health crises. The need for effective measures to enforce social distancing guidelines in public places has led researchers and developers to harness the power of deep learning and computer vision. YOLOv5 has proven to be a good candidate for this task due to its real-time object detection capabilities and high accuracy. By combining YOLOv5 model fine-tuning and input image preprocessing, social distance detection systems can efficiently identify people in crowded scenes.

By analyzing the spatial relationships between detected individuals, the system can determine whether social distancing guidelines are being followed or violated. Geometric analysis allows interpersonal distances to be calculated and compared to pre-defined thresholds, allowing immediate detection of social distancing violations. Existing systems show promising results with real-time monitoring capabilities and accurate detection of social distancing violations. These systems have been deployed in various public spaces such as transportation hubs, workplaces, retail spaces and public events to ensure adherence to social distancing guidelines and curb the spread of infectious diseases.

However, there is room for future improvement and expansion in this area. Ongoing research efforts aim to improve the accuracy and performance of social distance detection algorithms. By fine-tuning the model for large and diverse datasets and considering multi-person tracking techniques, a more robust and comprehensive detection system can be achieved. Additionally, integrating social distancing detection with other modalities such as thermal imaging and depth sensing can further improve accuracy in challenging situations.

7. FUTURE SCOPE

The future scope of 'Social distance detection using image detection techniques and YOLOv5 algorithms' is promising and beyond the current research landscape. As technology continues to evolve and we see new advances in computer vision and deep learning, there are several possible future developments and applications for social distancing detection using YOLOv5.

In summary, the future scope of "social distancing detection using image recognition technology and his YOLOv5 algorithm" is vast and includes various advances in accuracy, real-time processing, analytics, and integration with other technologies. As societies continue to face health challenges and embrace technological advances, such systems will play an important role in ensuring public safety and adherence to social distancing guidelines for years to come.

REFERENCES

1. Redmon, J., & Farhadi, A. (2018). YOLOv3: An Incremental Improvement. arXiv preprint arXiv:1804.02767.
2. Wang, C., Wu, X., Bao, Y., Liu, Q., & Xu, Y. (2020). A Deep Learning-Based Approach for Social Distancing Monitoring during COVID-19 Pandemic. IEEE Access, 8, 124810-124819.
3. Mehta, R., Srivastava, S., & Kapoor, R. (2020). A Deep Learning Approach for Social Distancing Monitoring during COVID-19. International Journal of Advanced Research in Computer Science, 11(5), 68-72.
4. Tadakamadla, S., Jain, G., & Tandon, A. (2020). Social Distancing Detection Using Deep Learning Techniques. In Proceedings of the International Conference on Artificial Intelligence and Signal Processing (AISP).
5. Rahman, S., Mahmud, M., & Maula, S. (2020). A Deep Learning-Based Approach for Social Distancing Monitoring during COVID-19 Pandemic. In Proceedings of the International Conference on Emerging Trends in Engineering, Science, and Technology (ICETEST).
6. Pandey, A., & Singh, R. (2020). Deep Learning-based Social Distancing Monitoring in Public Places using Surveillance Cameras. In Proceedings of the International Conference on Advances in Computing, Communication and Control (ICAC3).
7. Kittur, H. M., Khopkar, P., & Yenpure, R. (2021). Real-Time Social Distancing Detection System using Deep Learning. International Journal of Advanced Studies in Computer Science and Engineering, 10(2), 77-85.