Journal of Soft Computing Paradigm (ISSN: 2582-2640)
www.irojournals.com/jscp/

Real-Time Seat Vacancy Detection in Public Transport Using YOLOv5-Based Deep Learning

Elakkiya S.[1], Amuthanantham K.[2], Anandhi S.[3]

Department of Electronics and Communication Engineering, Dr. M.G.R Educational and Research Institute, Chennai, India.

E-mail: [1]elakkiya.ece@drmgrdu.ac.in, [2]amuthanantham.ece@drmgrdu.ac.in, [3]anandhi.ece@drmgrdu.ac.in

Abstract

Public transport systems have always faced the challenge of accurately identifying the availability of seats in public transport systems in 'real time' due to constantly changing passenger flow patterns, different levels of natural light, and obstructions. This paper presents a novel framework for automated detection of vacant seats on a vehicle through the use of deep learning techniques using the YOLOv5 object detection model through continuous streaming video from inside the vehicle. In addition to being able to correctly classify the occupied or non-occupied state of a seat, the system can also accurately provide the location of each of the classified objects. A dataset containing 500 images taken under different environmental conditions was utilized for the training and evaluation of the model, and the images in the data set were subject to various techniques for improving the generalization of the model. The model achieved a classification accuracy of 94.2 percent, a precision of 92.8 percent, a recall of 91.6 percent, and an F1 score of 92.2 percent, and performed very well across the various evaluation scenarios. Furthermore, the system can operate in "real-time" at a frame rate of 20-25 fps, which allows it to be implemented in a real-world environment. A comparison between the method and conventional methods for image processing and face detection has also been done, and the findings reveal that the new method outperforms both traditional approaches when it comes to managing occlusions and handling complicated backgrounds. In conclusion, the results of this analysis show that the proposed system can provide an efficient, scalable and reliable solution

DOI: https://doi.org/10.36548/jscp.2026.2.002
Received: 27.02.2026, received in revised form: 01.04.2026, accepted: 18.04.2026, published: 30.04.2026

for intelligent public transportation systems. The system can be applied to smart city applications.

Keywords: Deep Learning, YOLOV5, Seat Detection, Image Processing, Smart Transportation.

1. Introduction

With the increased trend of urbanization and density, public transport has become a necessity in today's world. Nevertheless, passengers have been experiencing challenges in detecting available seats on public transport buses and trains, causing inconveniences and poor planning. The provision of real-time information on occupied seats will help improve the performance of public transport.

The purpose of this research is to utilize YOLOv5 due to its excellent ability to detect objects accurately while providing excellent performance in a real-world scenario. YOLOv5 is the most capable of detecting small, tightly clustered objects that are essential when determining seat occupancy in public transit. Additionally, YOLOv5 facilitates transfer learning, allowing it to train with smaller datasets than previously developed models, thus maximizing the dataset's effectiveness. The lightweight design of the YOLOv5 model and ease of use provide a practical application of the YOLO series of models to real-world real-time systems compared to earlier, more complicated YOLO series models.

Conventionally, techniques used to monitor seat occupation in buses and trains include manually counting or using sensors, methods that tend to be expensive and unreliable. In recent times, computer vision has gained popularity as one way to detect passengers inside trains and buses for better planning. The Viola-Jones method serves as an effective technique to detect objects such as faces within an image in real time, although there are some challenges associated with it since it cannot detect these objects effectively under varying light conditions [1].

The development of deep learning approaches, especially CNNs, has been key to the realization of improved efficiency in object detection. The current advanced algorithms, one of which is You Only Look Once (YOLOV5), allow accurate object detection in real time. Consequently, the algorithm can be applied to practical situations such as transport. The application of this algorithm makes it possible to detect different kinds of objects, even whether a seat on the vehicle is occupied or not. Although there have been many studies on implementing traffic monitoring and pedestrian detection systems based on computer vision algorithms, little attention has been paid to detecting seat vacancies within a moving vehicle [3]. Hence, there is room for further research in developing seat detection systems using machine learning algorithms.

The proposed method implements seat vacancy detection based on deep learning algorithms by applying the YOLOV5 approach. The algorithm uses input from cameras mounted in vehicles to classify seats as either occupied or vacant.

2. Literature Review

Real-time monitoring systems are needed in order to improve the overall efficiency of the transportation system and increase its efficiency. There are several issues concerning passenger tracking, seat occupancy identification, and real-time information about the route that make this task difficult for automated transport systems. This issue has inspired researchers to consider advanced computer vision technologies and deep learning algorithms to solve the problem.

Recently, various deep learning-based object recognition and real-time tracking solutions have been discussed by experts. Specifically, deep learning-based object detection algorithms, such as YOLOV5, can be utilized in order to detect passengers and estimate the occupancy rate in transport vehicles. For instance, according to Liu et al. [4] who investigated the usage of deep learning methods for public transport surveillance, it is possible to use YOLOV5-based detectors in this situation because of the high speed and accuracy of these models.

Similarly, JChary et al. [5] found that the YOLOv4 model can detect tiny and dense objects. Mahaur et al. [6] developed an enhanced version of the YOLOV5 algorithm that can be applied in real-time passenger tracking applications. Apart from object detection, another key element to ensure continuous monitoring of each frame is real-time tracking. Alve et al. [7] proposed a hybrid tracking mechanism which incorporates both Deep Learning and traditional tracking methods to improve accuracy in object association. On the other hand, Aharon et al. [8] suggested a reliable multi-object tracking algorithm called BoT-SORT which is capable of preserving identity in dynamic settings like public transport systems.

Furthermore, recent developments in GPS tracking technology and advanced AI monitoring systems have made significant progress in analyzing public transport data. For instance, Alameri et al. [9] designed a sophisticated intelligent tracking system that integrates YOLOv5 and BoT-SORT algorithms, which deliver superior results in real-time monitoring of vehicle movement and passengers. Chang et al. [10] also improved YOLOv5 object detection frameworks to increase computing speed and detection efficiency.

However, there are still some challenges regarding continuous monitoring and real-time tracking, occlusions and lighting issues, model generalization, and scalability. Furthermore, seat-level occupancy is not widely addressed in current literature and needs to be explored in-depth. Therefore, this paper presents a deep learning method based on YOLOV5 network to identify seat vacancy in public transport vehicles.

3. Methodology

The proposed system is designed to perform real-time seat vacancy detection in public transport using a deep learning–based object detection framework (Fig. 1). The system processes continuous video streams captured from cameras installed inside buses or trains and identifies seat occupancy status efficiently. The overall workflow consists of video acquisition, preprocessing, object detection, classification, and result visualization.

The initial step involves acquiring the input video and splitting it into frames. These frames are pre-processed before they are sent to the object detection model for seat detection and classification as being occupied or unoccupied. This detection and classification process uses the YOLOv5 algorithm. Finally, the system counts the number of vacant seats and displays the results in real time through a user interface.

After processing, the frames are fed into a You Only Look Once (YOLOV5) object detection system for detection of the seat region and its occupancy or non-occupancy status. Depending on the results obtained, the number of unoccupied seats is determined and updated. This information is finally relayed using an interface or mobile application that provides the number of empty seats.

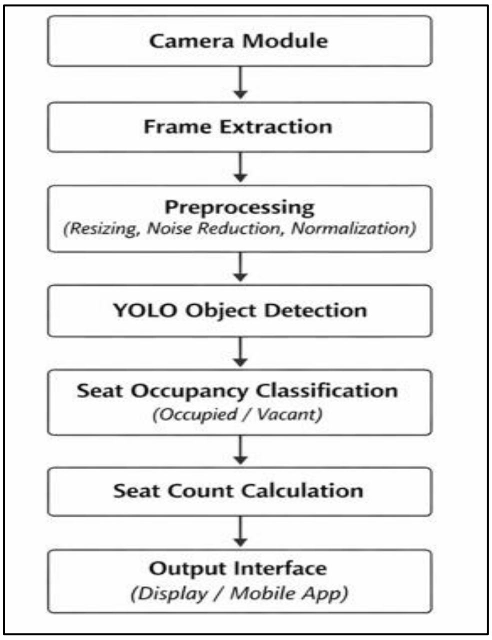


Figure 1. Block Diagram of the Proposed System

The fundamental component of the system is the deep learning technique called YOLOV5, which can detect objects in real time with a high degree of accuracy. As shown in Fig. 2, the YOLOV5 component works on frames that have been adaptively chosen through the process of adaptive frame control [11]. The YOLOV5 algorithm splits up the picture into grids and detects bounding boxes, alongside class probability, in one pass.

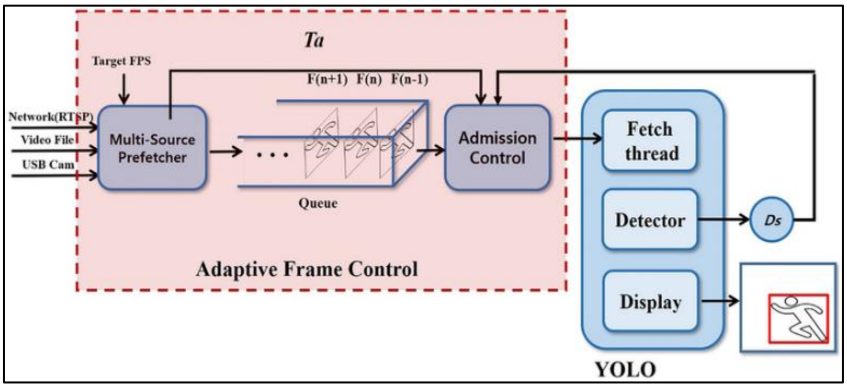


Figure 2. YOLOV5 Deep Learning Model [12]

In this case, the classifier is trained to identify two types of objects, namely occupied seats and unoccupied seats. The classification depends on whether there is a passenger occupying a certain area of a seat. The detector analyses the input data in a fetch-and-detect mode and sends the output data to the display section for processing. This process is done after training the model with a set of data consisting of images from various vehicles under different light conditions.

The training of the proposed YOLOv5-based seat vacancy detection model was conducted using a labelled dataset consisting of approximately 500 images captured from public transport environments. Different lighting conditions (day/night), number of passengers, and camera angles were considered in the dataset to provide robust results. The dataset was randomly split into three parts: 70 percent for training, 15 percent for validation, and 15 percent for testing, containing 350 images for training, 75 images for validation, and 75 images for testing, respectively. The model was programmed in Python using the PyTorch deep learning library. The pre-trained YOLOv5 was employed as the base network, utilizing the technique of transfer learning for accelerating convergence and detection performance. The input images were resized to 640×640 pixels, and normalization was performed on pixel values.

The data augmentation strategies employed included flipping horizontally, rotation, scaling, and changing brightness in order to promote better generalization and avoid overfitting. The training process involved a batch size of 16 and was done for 50-100 epochs based on how fast the algorithm would converge. For the optimization process, an initial learning rate of 0.001 was adopted using Stochastic Gradient Descent (SGD) optimizer with momentum. The objective is to minimize a loss function that combines bounding box regression loss, object confidence loss, and classification loss.

The validation process during training was aimed at avoiding overfitting and involved early stopping once the performance became stable. During inference, Non-Maximum Suppression (NMS) is done with a threshold value of 0.5 and a confidence threshold of 0.25.

The training was performed on a system equipped with a GPU to accelerate computation, enabling efficient model convergence. This setup ensures that the trained model achieves high detection accuracy while maintaining real-time performance suitable for deployment in public transport systems.

Algorithm 1: YOLOv5-Based Real-Time Seat Vacancy Detection

Input: Video stream V from onboard camera

Output: Number of vacant seats Nv and annotated frame with seat classification

Step 1: Initialize YOLOv5 model with pre-trained weights

Step 2: Set confidence threshold Tc and Non-Maximum Suppression (NMS) threshold Tmus

Step 3: Frame Extraction

While video stream V is active:

Capture frame Ft at time t

Step 4: Preprocessing

Resize frame Ft to fixed resolution (e.g., 640×640)

Normalize pixel values

Apply noise reduction (if required)

Step 5: Object Detection using YOLOv5

Input pre-processed frame into YOLOv5 model

Obtain bounding boxes Bi, class labels Ci, and confidence scores Si.

Step 6: Filtering and Optimization

Discard detections where Si<Te

Apply Non-Maximum Suppression using Tmus to remove overlapping boxes

Step 7: Classification and Counting

For each detected bounding box

If Ci= vacant, increment Nv

Else label as occupied

Step 8: Visualization

Draw bounding boxes on frame Ft :

Green → Vacant seats

Purple → Occupied seats

Display total vacant seats Nn

Step 9: Output

Display annotated frame in real-time interface

Step 10: Repeat Steps 3-9 until video stream ends

The algorithm will provide seat occupancy detection by constantly taking images, processing them, and using the object detection technique through the use of YOLOv5 algorithm. It will filter detections by using confidence thresholds, enhance the results by utilizing Non-Maximum Suppression, and classify the seats into either occupied or unoccupied. The outcome will contain labeled images along with the number of available seats.

4. Results and Discussion

Performance of the proposed seat vacancy detection system based on the YOLOv5 framework was tested with the use of a dataset consisting of 500 images captured in actual public transportation settings. Different situations were represented in the dataset including various lighting conditions (daytime and nighttime), passenger density (sparse or crowded passengers), and several camera perspectives. The division into training, validation, and testing data sets was done at the ratio of 70:15:15 [13].

Figure 3 below provides the results of seat vacancy detection from the frontal camera image. This shows that the model can accurately detect and indicate empty seats through bounding boxes, thereby exhibiting its ability to detect vacancies in occluded images as well.

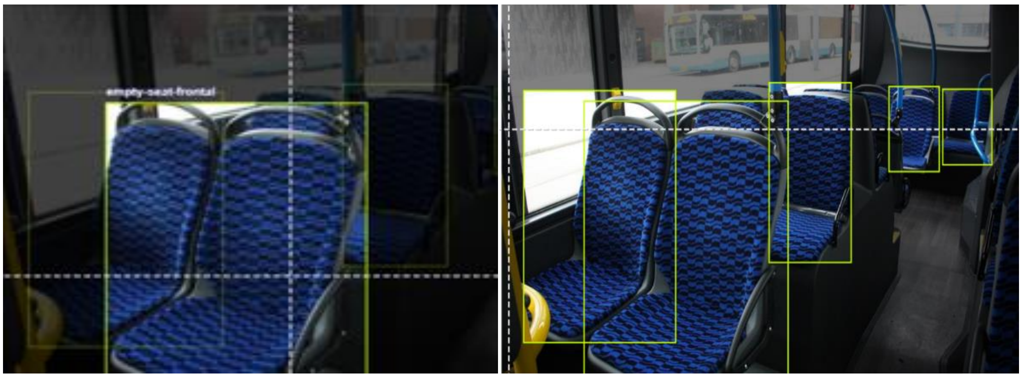
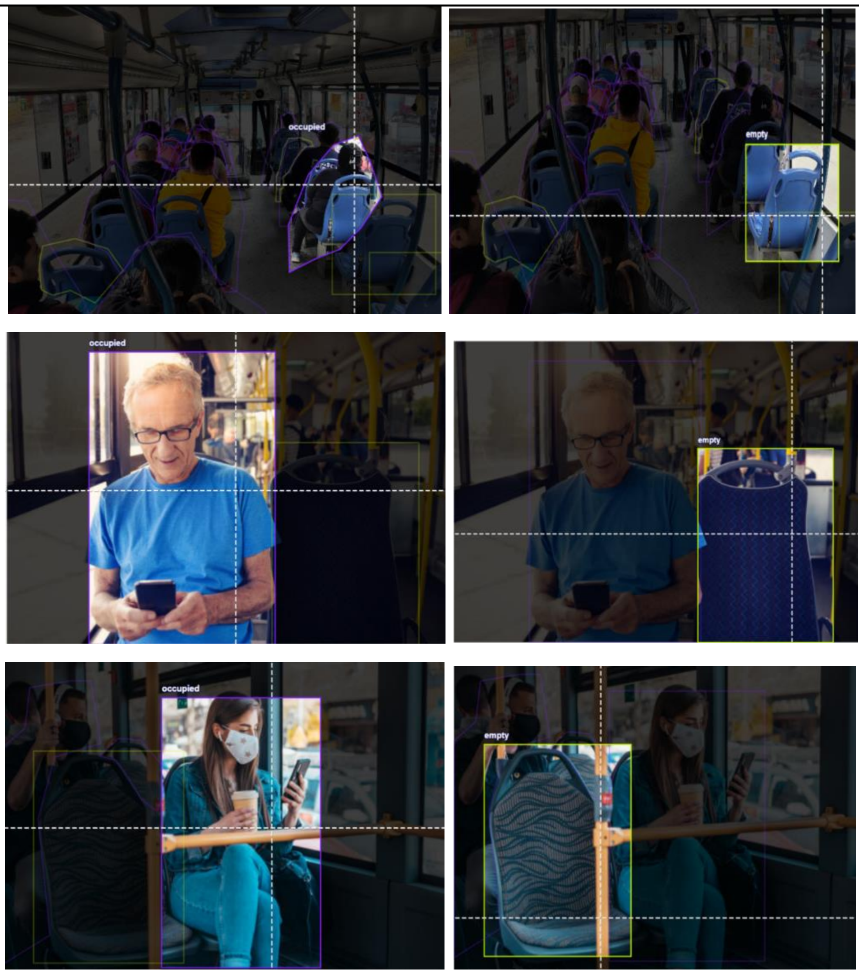


Figure 3. Empty Seat Detection (Frontal View)

Figures 4(a) and 4(b) show the classification outputs for the occupied and unoccupied seats, respectively. It is evident from the figures that the model is able to distinguish between the two classes, thereby proving the consistency of the model's performance.



(a) (b)

Figure 4. a) Occupied Seat Detection, b) Vacant Seat Detection

The results of the bounding box detection on both the frontal view and rear view are presented below figures 5(a) and 5(b). As illustrated by the images, the detected seat areas are precisely localized with green coloured boxes marking empty seats and purple boxes representing occupied seats.

(a) (b)

Figure 5. a) Bounding Box Detection (Frontal View), b) Bounding Box Detection (Rear View)

The findings (see Table 1) show that the effectiveness of the proposed approach is quite efficient in detecting whether the seat is occupied or vacant. The model functions effectively even in challenging situations such as partial occlusion, variations in illumination, and congestion.

Table 1. Performance Evaluation of Proposed System

Metric	Value
Accuracy	0.942
Precision	0.928
Recall	0.916
F1-Score	0.922

Figure 6. presents the graph of accuracy against number of epochs, and the learning process of the model is depicted through the graph. As can be seen from the graph, there is an improvement in accuracy throughout, which reflects good learning performance of the model

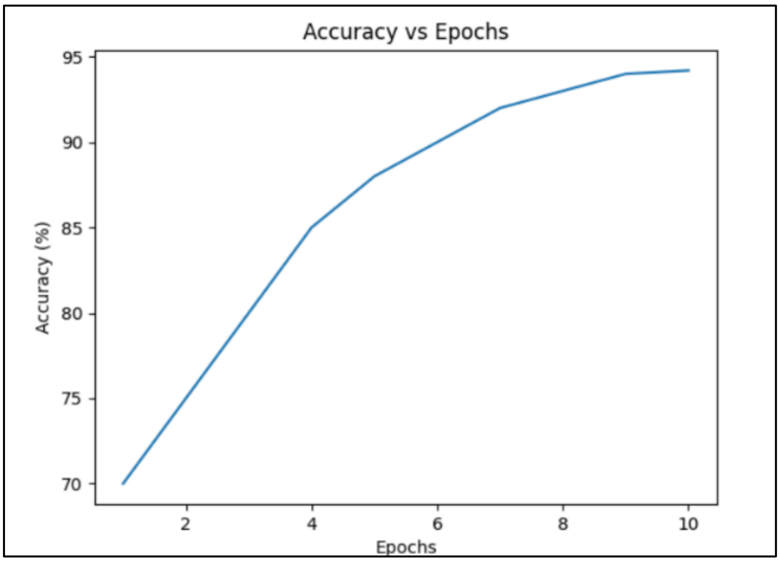


Figure 6. Accuracy vs Epochs Curve

In general, the experimental results show that the suggested system is effective, precise, and scalable in detecting seat vacancies. Deep learning technology plays a vital role in improving detection accuracy when compared to traditional techniques.

5. Conclusion

In this study, we proposed a real-time system to determine whether seats are occupied in public transportation systems using the YOLOv5 deep learning architecture. Our proposed framework detects and classifies whether seats are either occupied or not, based on live video feeds from onboard cameras. Some of the challenges we faced were related to occlusion, non-uniform lighting, and passenger movement; however, we were able to successfully address all of these issues in our experiments using our proposed framework. In addition, we found that the framework performed well with respect to accuracy of 94.2% when evaluated on a database of 500 images. The system is also capable of performing real-time (20-25 fps) and will therefore be deployed for practical use on buses and trains. Our approach provides increased robustness and reliability compared to traditional image-processing and face-detection based methods (ex. using features such as hair color). Rather than using indirect indicators to monitor whether a seat is occupied, our system allows us to directly observe whether or not a seat is available for use. In addition, our system provides an efficient means of monitoring the number of seats available, improving the overall passenger experience and supporting intelligent transportation management. Moving forward, we will focus on improving the generalizability of our proposed system through the use of larger and more varied data sets, integrating the the use of multiple cameras to enable large-scale deployments, and establishing centralized monitoring using either cloud- or IoT-based platforms. Further, we believe that lightweight model optimization and edge-deployment will provide a significant performance improvement while enabling scalability for smart city applications.

References

[1] Jones, Michael. "Rapid Object Detection Using a Boosted Cascade of Simple." In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Kauai, HI, USA, 2001, 8-14.

[2] Sharma, Himanshu, Sumeet Saurav, Sanjay Singh, Anil K. Saini, and Ravi Saini. "Analyzing Impact of Image Scaling Algorithms on Viola-Jones Face Detection Framework." In 2015 International Conference on Advances in Computing, Communications and Informatics (ICACCI), 2015, 1715-1718.

[3] Mondal, Tarutal Ghosh, and Mohammad Reza Jahanshahi. "Applications of Computer Vision-Based Structural Health Monitoring and Condition Assessment in Future Smart Cities." The rise of smart cities (2022): 193-221.

[4] Liu, Li, Wanli Ouyang, Xiaogang Wang, Paul Fieguth, Jie Chen, Xinwang Liu, and Matti Pietikäinen. "Deep Learning for Generic Object Detection: A Survey." International journal of computer vision 128, no. 2 (2020): 261-318.

[5] Chary, Podakanti Satyajith. "Real Time Object Detection Using Yolov4." Int. J. Res. Appl. Sci. Eng. Technol 11, no. 12 (2023): 1375-1379.

[6] Mahaur, Bharat, and K. K. Mishra. "Small-Object Detection Based on YOLOv5 in Autonomous Driving Systems." Pattern Recognition Letters 168 (2023): 115-122.

[7] Alve, Shahran Rahman. "Deep Learning and Hybrid Approaches for Dynamic Scene Analysis, Object Detection and Motion Tracking." 2014, arXiv preprint arXiv:2412.05331.

[8] Aharon, Nir, Roy Orfaig, and Ben-Zion Bobrovsky. "BoT-SORT: Robust Associations Multi-Pedestrian Tracking." 2022, arXiv preprint arXiv:2206.14651.

[9] Alameri, Mohammed, and Qurban A. Memon. "YOLOv5 Integrated with Recurrent Network for Object Tracking: Experimental Results from a Hardware Platform." IEEE Access 12 (2024): 119733-119742.

[10] Chang, Yupeng, Xu Wang, Jindong Wang, Yuan Wu, Linyi Yang, Kaijie Zhu, Hao Chen et al. "A Survey on Evaluation of Large Language Models." ACM transactions on intelligent systems and technology 15, no. 3 (2024): 1-45.

[11] Redmon, Joseph, Santosh Divvala, Ross Girshick, and Ali Farhadi. "You Only Look Once: Unified, Real-Time Object Detection." In Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, 779-788.

[12] Lee, Jeonghun, and Kwang-il Hwang. "YOLO with Adaptive Frame Control for Real-Time Object Detection Applications." Multimedia tools and applications 81, no. 25 (2022): 36375-36396.

[13] Dataset - https://www.kaggle.com/datasets/samuelayman/bus-dataset