VOLUME 19, N° 2 2025
Journal of Automation, Mobile Robotics and Intelligent Systems

IoT-based Emergency Vehicle Detection using YOLOv8

Submitted: 24th May 2023; accepted: 17th July 2023

Syed Suhana, Boppuru Rudra Prathap, Kavish Narang, Ivin Anto

DOI: 10.14313/jamris-2025-018

Abstract:
The rapid response of emergency services plays a critical role in saving lives and minimizing the impact of emergencies. However, identifying and locating emergency vehicles in real-time can be challenging, especially in congested urban areas. This paper focuses on the emergency vehicle identification using the You Only Look Once version 8 (YOLOv8) algorithm and is focused on Internet of Things (IOT). The goal of this research is to develop a real-time and precise emergency vehicle detection system using You Only Look Once version 8 (YOLOv8) algorithm, trained and tested with a dataset from a camera placed on a busy road, to enhance emergency service response times. The findings demonstrate the suggested system's ability to recognize emergency vehicles at a speed of 31 frames per second and with a 95% accuracy rate. Modern object identification algorithms include the You Only Look Once version 8 (YOLOv8) algorithm, which has shown promising results in various applications. The proposed system is built on a Raspberry Pi, which acts as an edge device and processes the video stream in real-time. The system consists of an Internet of Things (IOT) device with a camera that captures the live video stream, which is then fed into the algorithm for object detection. Once an emergency vehicle is detected, the system sends an email notification to the nearby emergency services, like a police station, using Simple Mail Transfer Protocol (SMTP), who can then take appropriate action. The results of this investigation show that the Internet of Things and You Only Look Once version 8 (YOLOv8) algorithms have great promise for creating effective and dependable emergency vehicle detection systems. The proposed system possesses the capacity to save lives and improve the effectiveness of emergency response by speeding up response times for emergency services. The suggested solution is also inexpensive, simple to implement, and adaptable to existing infrastructure. Through the development of intelligent transportation systems, emergency services can operate more safely and effectively. More sophisticated machine learning algorithms may be incorporated into the proposed system, and further sensors can be added to utilize alternative methods beyond camera-based detection to identify emergency vehicles. Overall, this research shows the potential of Internet of Things (IOT) and machine learning in creating creative emergency services solutions.

Keywords: *You Only Look Once (YOLOv8), Raspberry Pi, Pi Camera, Internet of Things (IoT), emergency vehicles, simple mail transfer protocol (SMTP), email notifications*

1. Introduction

The surge in accidents, disasters, and medical crises in recent years has made emergency services more and more crucial. The prompt action of emergency services can prevent fatalities and lessen the effects of catastrophes. However, heavy traffic and lag times in locating the scene of the incident frequently compromise the effectiveness of emergency services. To speed up reaction times and boost the effectiveness of emergency services, an approach that can track emergency vehicles on the road in real-time is required. To meet this need, we present an IoT-based emergency vehicle identification system based on the YOLOV8 algorithm. The suggested system uses an IOT camera device, Pi Camera, to record the live video stream, which is subjected to object detection processing by the YOLOV8 algorithm. When an emergency vehicle is detected, the algorithm is designed so that the system sends alerts to the emergency services, like police stations, so that they can respond appropriately.

A Raspberry Pi serves as an edge device for the system and performs real-time processing of the video feed. Modern object identification algorithms like YOLOV8 have produced encouraging results in various applications. It is a quick and accurate real-time object detection technique. The picture captured by the camera is partitioned into a grid of cells to identify the vehicle, and bounding boxes are estimated for each of them. Additionally, it assigns each object to a predetermined category and forecasts the likelihood that it will be present in each box. The YOLOV8 algorithm can accurately identify a variety of items since it was trained on a vast collection of photos.

The suggested system has a 95% detection accuracy and can identify emergency vehicles at a speed of 31 frames per second. The technology may be installed on a busy route to identify emergency vehicles and shorten emergency service response times quickly. This technology can help save lives and increase the effectiveness of emergency response by cutting down on response times. In addition, the suggested solution is affordable, simple to implement, and adaptable to existing infrastructure. Due to its potential to boost emergency services' effectiveness, research on emergency vehicle detection systems has attracted interest recently. Several experiments have been carried out to create effective and trustworthy emergency vehicle detection systems.

However, the majority of these studies have concentrated on the usage of conventional machine learning approaches, which are inaccurate and expensive to compute. In contrast, the suggested system uses a cutting-edge object identification algorithm that is both quick and extremely accurate.

The contribution of this research is the creation of an IoT-based emergency vehicle identification system based on the YOLOV8 algorithm. The suggested system has been proven to identify emergency vehicles on the road accurately and, hence, send alerts. The suggested method is a cutting-edge idea that might improve the effectiveness and safety of emergency services. The suggested system can cover a huge area since it is scalable and deployable at a large scale. The subsequent sections of the essay are organized as follows. We will talk about related research on the YOLOV8 algorithm and emergency vehicle detection systems in the part that follows. One of the recent research studies in this domain was done using YOLOv4 in a paper titled "An improved and efficient YOLOv4 method for object detection in video streaming," in which the authors talk about how this YOLOv4 algorithm is combined with various Inference methods, which leads to a significant increase in the efficiency. Further, we shall review the research approach in part 3 of this article. We will discuss the study's findings and analyses in section 4. We will offer the verdict and suggestions for further study in section 5. Finally, we shall list the references used for this study in section 6.

Internet of Things, which links various gadgets and sensors through the internet, has completely changed the way we live and work. IoT has recently gained more importance in emergency services, especially when it comes to the recognition of emergency vehicles. Real-time emergency vehicle detection can speed up emergency response times and lower the likelihood of accidents. Object recognition algorithms, such as the YOLO family of models, are a well-liked method for locating emergency vehicles. Specifically, YOLOv8 is a cutting-edge object identification model that performs real-time object recognition using deep convolutional neural networks. Numerous applications, such as traffic management, surveillance, and autonomous vehicles, have extensively used it. In this article, we suggest an IoT-based YOLOv8 emergency vehicle detection system. Our technology uses cameras and sensors placed on roads to identify emergency vehicles in real-time. We will go through how our approach was implemented and assess how well it worked to identify emergency vehicles. Ultimately, we hope that our research will add to the increasing body of knowledge on IoT-based emergency vehicle identification and shed light on how to employ YOLOv8 for object detection in real-time circumstances.

2. Literature Review

Emergency vehicle detection systems have drawn more attention recently because they can increase the effectiveness of emergency services. Several experiments have been carried out to create effective and trustworthy emergency vehicle detection systems. This section reviews the YOLOV8 algorithm and associated studies on emergency vehicle detection systems.

Emergency Vehicle Detection Systems

Researchers have been interested in the application of technology in emergency services. One such piece of technology that can help speed up emergency service response times is emergency vehicle detection systems. Systems for identifying on-road emergency vehicles make use of a variety of sensors, including cameras, radar, and GPS.

Baghel et al. [1] explore the effectiveness of the Ex-YOLO algorithm for detecting emergency vehicles compared to other real-time algorithms. The authors then propose an improved version of YOLO, called Ex-YOLO, which uses convolutional neural networks to improve accuracy. The authors compare Ex-YOLO with real-time algorithms like Faster R-CNN and SSD for emergency vehicle detection. The findings indicate that the Ex-YOLO algorithm performs better than other real-time algorithms in terms of accuracy and speed. The authors also discuss the limitations of their study, such as the limited size of their dataset. Overall, the paper presents a well-researched study on the effectiveness of the Ex-YOLO algorithm for emergency vehicle detection.

The research "Audio-Vision Emergency Vehicle Detection" proposes a new approach for emergency vehicle detection by integrating audio and visual information. The authors use a dataset of traffic videos with accompanying audio signals and contrast the performance of their proposed approach with traditional visual-only methods. The outcomes demonstrate that the proposed approach is better than traditional visual-only methods in terms of accuracy, recall, and precision. The paper thoroughly explains the proposed approach, including the audio feature extraction process, and utilizes deep neural networks for classification. The results suggest that such an approach can improve the accuracy and reliability of emergency vehicle detection systems, which can have significant implications for traffic safety and emergency response. However, it would be useful to evaluate the proposed method on a larger and more diverse dataset to further assess its effectiveness [2].

Managing Traffic is a huge challenge on a global scale. Traffic congestion problems are rising quickly, particularly in India, due to urbanization, population increase, and an increase in the number of automobiles. As a result, it is sometimes difficult for the ambulance to reach the hospital on time. Technological advancements have created numerous solutions to address serious problems and help save people.

The IoT surely has the potential to make the situation better. This study aims to examine several IoT movement control strategies and various methods for helping emergency vehicles arrive at nearby medical facilities in time. Finding the best possible methods to lessen traffic congestion is the key goal [3].

Most of the individuals who get hurt in auto accidents get help from other drivers or passengers. However, in case of a car accident that took place in a remote area or if the driver is the only one inside the vehicle and is unconscious, no one will be around to alert the appropriate authorities in time for medical treatment. A technique for identifying high-speed head-on and single-vehicle crashes, assessing the context, and sounding an alert is required in light of these problems. Chang, W. J [4] proposes the Internet of Vehicles (IoV) system named DeepCrash to address these issues. DeepCrash consists of an in-vehicle infotainment (IVI) telematics platform with a vehicle self-collision detection sensor and a front camera, a cloud-based deep learning server, and a cloud-based management platform. When a head-on or single-car collision is found, accident detection data is transferred to the cloud for self-collision vehicle accident recognition, and a corresponding emergency alert is sent out. According to the experimental findings, traffic collision detection accuracy may approach 96%, and the typical notification reaction time for emergencies is around 7s.

The desire for quicker and more precise detectors is increasing as autonomous vehicles and racing become increasingly popular. Even though even from great distances, our naked eyes can nearly quickly extract contextual information, picture quality and processing resource restrictions make recognizing tiny objects (that is, things that fill less than one square a constrained region of pixels in the input image), a job that is difficult for machines and is an open topic for investigation The study [5] looks at ways to enhance the well-known YOLOv5 object detector to better identify up small things with a focus on autonomous racing. To do this, we look into how changing some of the model's structural components (along with their links and various factors) might impact effectiveness and the time to get inferences. Consequently, we suggest a set of models of varying sizes that we call "YOLO-Z." When recognizing smaller objects at 50% IOU for just 3ms longer inference time than the original YOLOv5 model, these models demonstrate an improvement of up to 6.9% in mAP. The objective is to continue investigating the possibility of modifying a well-known detector like YOLOv5 to handle certain jobs and offer perceptions on how particular alterations might affect small item identification in the future. Such findings may increase the amount of contextual data that autonomous vehicle systems have access to when they are applied to a larger context.

Current object detection networks use region proposal techniques to predict the locations of objects. Innovations like SPPnet and Fast R-CNN have reduced the running time of these detection networks, exposing region proposal computation as a bottleneck. The area Proposal Network (RPN) was created in the current study to cooperate with the detection network to exchange full-image convolutional features and provide almost-free area recommendations. An RPN, or fully convolutional network, predicts object bounds and objectness scores for each location simultaneously. By utilizing RPNs that have completed end-to-end training, Fast R-CNN conducts detection employing high-quality region suggestions. RPN and using just alternating optimization, a fast R-CNN may be trained to share convolutional features. This detection technique delivers cutting-edge item recognition with 300 proposals per image. The accuracy of PASCAL VOC 2007 (73.2% mAP) and 2012 (70.4% mAP) while operating at a frame rate of 5fps (including all phases) on a GPU [6].

The number of automobiles on the road will increase significantly every year. According to data from Malaysia's road transport department (JPJ), as of December 31, 2019, there were around 31.2 million motor vehicle units registered in the country. As of the middle of 2017, Malaysia had a total of about 28.18 million motor vehicle units. Since traffic congestion may be identified by using the volume of cars as beneficial data, it is crucial to swiftly and effectively detect vehicles on the road. This is because it will help with traffic management. Muhammad et al. [7] use TensorFlow, a framework for machine learning, together with the object identification method YOLO, to construct deep learning for real-time car recognition. The suggested method in this study combines these two and other requirements using Python as the programming language to assess how well the YOLOv4 algorithm performs in comparison to the previous model in the vehicle identification system. To effectively count the number of cars passing in the video, this vehicle identification also employs the DeepSORT method. Yolov4, the top YOLO model from this research, achieved cutting-edge performance with 82.08% AP50 at a real-time frame rate of roughly 14 FPS on a GTX 1660ti. Yolov4 employed a custom dataset to achieve these results.

Detection of vehicles for aerial photos has become an important engineering approach and demonstrates the worth of academic study by using UAVs in intelligent transportation systems. This study provides a YOLO deep learning-based approach for identifying vehicles in aerial images. The method combines three open-source aerial picture datasets into one suitable for training the YOLO algorithm. Experiments show that the trained model matches the real-time requirements and works well on unknown aerial photographs, especially for small objects, rotating objects, and compact and dense objects [8].

One of the key facets of an economy is transportation. A large number of economic sectors struggle because of a loosely organized transportation network. This is an important challenge that emerging nations must deal with. There is no question that highways should be built to increase the throughput of the transportation network, but expanding already-existing roads is also not practical in countries like Sri Lanka because of their shrinking geographical area and growing population. Hence, for us to resolve this problem, a more effective, technologically advanced approach must be embraced. The current pandemic turmoil has proven the need to prioritize ambulances when it is involved in a traffic congestion, in addition to the normal congestion scenarios. Another crucial aspect of the road system is pedestrians. Road accidents will be decreased through effective and safe crosswalks for pedestrians, which will also improve the current high traffic. The best option in this scenario is a sophisticated traffic monitoring system with integrated traffic light management. To detect cars and pedestrians and provide priority to emergency vehicles, this article suggests a method for intelligent, dynamic traffic monitoring and control. To achieve 91.3% detection precision, a new CNN is trained using the YOLOV3 architecture [9].

A computer vision-based method for real-time identification of several kinds of emergency vehicles in congested traffic is presented in this study. It enables the traffic controller to give emergency vehicles preferred path clearance, which could possibly save lives, safeguard property, and simplify the effort to prevent crimes. The proposed model is based on four classes: Firetrucks, Ambulances, Police Cars, and Normal Cars. The YOLO algorithm's top layers were redesigned and retrained to get new learned weights, whereas the bottom levels remained locked. The model has demonstrated good outcomes and excellent metrics in recognizing and categorizing emergency vehicles and regular cars after retraining using the suggested modified YOLOv5. Using the mAP metric, police cars achieved 98%, 96% for fire trucks, 89% for ambulances, and 97% for normal cars [10].

This paper presents a lightweight version of the popular object detection algorithm YOLO, called YOLO-LITE, that is optimized for non-GPU computers. The authors propose several optimization methods to reduce the computational requirements of the original YOLO algorithm, such as substituting fully connected layers with convolutional layers and using a smaller input resolution. They also introduce a new approach for anchor box clustering to improve object detection accuracy. In an experimental comparison of YOLO-LITE and other cutting-edge object identification algorithms, the article shows that YOLO-LITE performs as well as or better while operating in real-time on low-power devices. Overall, the research provides a practical and efficient method for real-time object identification on devices with limited resources [11].

This research suggests a new object detection technique named Tinier-YOLO, which is a compact version of the popular YOLO algorithm designed for real-time object detection in resource-constrained environments. The authors introduce several modifications to the YOLO architecture, such as reducing the convolutional filters and the dimensions of anchor boxes, to achieve a smaller model size and faster inference speed. They also propose a new data augmentation technique to increase the stability of the model. The article offers experimental findings from several datasets, demonstrating that Tinier-YOLO performs better than other leading object detection methods while running at real-time speeds on low-power devices. Overall, the paper offers an efficient and effective solution for real-time object detection in constrained environments [12].

This study suggests a novel object detection approach for landslide identification in satellite remote sensing pictures. The authors introduce a small attentional YOLO (You Only Look Once) model that improves the precision of item recognition in complicated backdrops by using an attention method. The model uses several convolutional layers and attention modules to reduce the computational requirements and achieve real-time performance. The suggested method surpasses existing cutting-edge object detection techniques in terms of detection accuracy and computing efficiency, as shown by experimental findings on a landslip detection dataset in the study. The proposed approach has the potential to assist in early warning and mitigation of landslide disasters, providing a valuable contribution to the field of disaster management [13].

The proposed approach helps improve the accuracy of traffic sign identification using YOLO v4 and artificial training data generated by numerous GANs. The authors trained their proposed model on an available traffic sign dataset and assessed it on both the same dataset and a real-world dataset. The testing findings demonstrated that their strategy outperformed a number of cutting-edge techniques in terms of accuracy. The authors also conducted ablation studies to investigate the effectiveness of using synthetic training data generated by different GANs. The paper concludes that their proposed approach using YOLO v4 and synthetic training data generated by various GANs can significantly improve the accuracy of traffic sign recognition, which is crucial for the safety of autonomous vehicles [14].

The paper recommends an improved version of the YOLO network for real-time vehicle detection in embedded systems. The suggested strategy alters the network's topology to keep detection accuracy while lowering computational complexity. The modified network is then implemented on an embedded system using a heterogeneous computing platform, which includes an FPGA and a CPU. According to experimental findings, the suggested technique detects vehicles with high accuracy, using 30% less processing time than the original YOLO network.

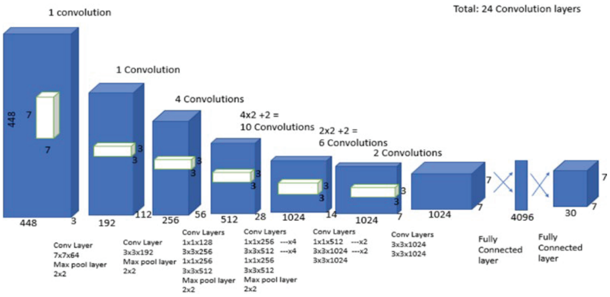


Figure 1. YOLO architecture

The suggested method surpasses other modern object identification techniques in the research when compared to both detection accuracy and processing speed. The outcomes show that the suggested approach is appropriate for embedded systems' real-time vehicle detection applications [15].

3. Methodology

The real-time object identification technique, YOLO, was first implemented in 2016. The system locates objects in images or videos by breaking the input image up into several cells and estimating bounding boxes and class likelihoods for each cell. One of the primary benefits of YOLO is just how quick it is. YOLO is built to process photos in real-time, making it appropriate for use in robots, autonomous vehicles, and surveillance systems, among other things. Additionally, YOLO is a single-stage detection technique, unlike other object identification algorithms like Faster R-CNN or SSD, which both need an additional region proposal phase. This makes things simpler and creates a faster pipeline.

Accuracy is another benefit of YOLO. On well-known object identification benchmarks like COCO and VOC, YOLOv4, the most recent algorithm iteration, performs at the cutting edge. YOLO does have certain restrictions, though. Its sensitivity to tiny things is one of its limitations. YOLO may overlook little things that are positioned between cells since it splits the picture into a grid of cells. Another drawback is how well it performs in dense situations with overlapping items, which might result in several detections of the same thing. YOLO is often quicker than other object identification algorithms but may give up some accuracy when compared to two-stage methods like Quicker R-CNN or Mask R-CNN. Single Shot Multibox Detector (SSD) and RetinaNet are two more well-liked object identification algorithms that likewise try to compromise speed and accuracy. In conclusion, YOLO is a well-liked object identification technique renowned for its quickness and precision. Its single-stage architecture and real-time performance make it suited for a variety of applications. Still, in comparison to other algorithms, it could have trouble handling small objects or cluttered situations.

YOLOv8 is the latest version, which is added to the YOLO family. This model is particularly used for object detection and segmentation. The performance and flexibility have been improved with additional features. YOLOv8 is quick, precise, and simple to use, hence proving to be the best choice for a variety of object recognition and tracking, segmentation, classification, and many other tasks. YOLOv8 is much faster and gives accurate results compared to previous versions of YOLO An additional feature added to YOLOv8 is instance Segmentation, through which multiple objects can be detected in an image. The architecture of YOLOv8 extends on the YOLO object detection models from previous releases. A fully convolutional neural network, or "backbone and head," that YOLOv8 uses for processing images. A modified form of the CSPDarknet53 architecture provides the foundation of YOLOv8. YOLOv8 has introduced Darknet-53, which is much faster and precise. DarkNet-53 is a 53-layer CNN that can categorize photos into 1000 different object categories. Various convolutional layers and a collection of fully connected layers comprise the head of the YOLOv8 algorithm. These layers are in charge of predicting the object detection boundaries, objectness scores, and class probabilities.

YOLOv8 uses a bounding box prediction mechanism just like any other image segmentation. An anchor-free detection head was introduced to achieve this. The model is more effective because YOLOv8 makes use of a bigger feature map and a more effective neural network. A larger feature map simply indicates that the model can capture complicated connections between various characteristics and can recognize patterns and objects in the data more effectively. In addition to this, it also reduces the overfitting and reduces the time it takes to train the model. The active use of a self-attention mechanism in the network's brain is one of YOLOv8's distinguishing characteristics. In addition to aiding this process, the model has the ability to direct its attention to various features of the image based on their relevance to the task at hand. The capability of YOLOv8 to carry out multi-scale object identification is a further essential characteristic. The model employs use of a feature pyramid network to identify items in an image that have different dimensions and scales. Multiple layers of this feature pyramid network detect items at different dimensions, permitting the model to recognize large and small things in an image. Figure 2 shows the proposed model of the Experiment. The entire research, as presented in the next sections from 3.1 to 3.6, is based on this proposed model.

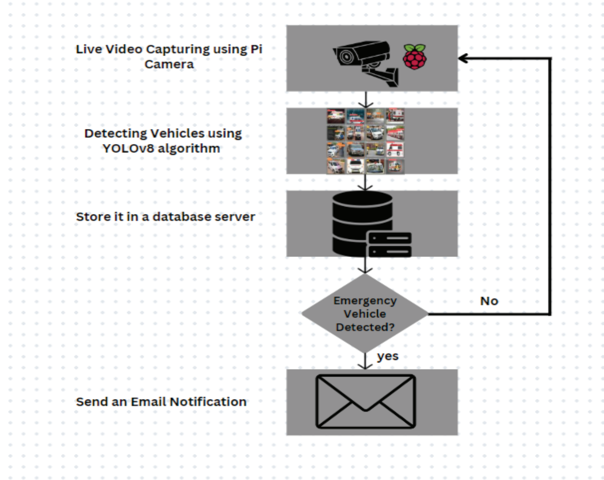


Figure 2. Proposed system model

Figure 3. Sample dataset

3.1. Experimental Setup

The Experiment was conducted on a 2km stretch of a road that is known to have frequent emergency vehicle traffic. The Pi Camera was mounted at a height of 5 meters, facing the road. The microcontroller, Raspberry Pi was connected to a cellular network to transmit data to a cloud server.

3.2. System Architecture

Our system consists of a camera mounted on a pole, a microcontroller, and a cloud server. The camera captures images of the road, and the microcontroller processes the images and sends them to the cloud. This server runs the YOLOv8 object detection algorithm to detect emergency vehicles in the images.

3.3. Dataset

The dataset is obtained by live capturing of Emergency Vehicles on the Road. While training our YOLOv8 model, the dataset was divided into train, valid, and test sets in a ratio of 70:15:15. Our model is trained to detect four classes of vehicles: ambulances, Police cars, Fire Engines, and and VIP Vehicles. Figure 3 shows the sample dataset images used for experimental purposes.

3.4. Data Preprocessing

The dataset was preprocessed before it underwent training using the YOLOv8 algorithm. The following steps were included as a part of preprocessing a) Resizing images: The images were resized to 416 X 416 pixels so that all the images in the dataset have equal dimensions. This not only speeds up the training process but also improves our model's accuracy. OpenCV was used to resize the images. b) Converting images to YOLO format: The YOLO algorithm demands a specific format for the dataset. It is crucial to transform our raw data into a format that the YOLO algorithm can use in order to create a YOLO model. The prepared dataset must include Image data and label data. c) Image data: In common, the Image data can be in any image format like JPEG or PNG. It is just like a matrix that represents the pixels of the input image. Here, our data has 416 X 416 pixels. d) Label data: The label data contains the annotations for each image, which typically includes bounding box coordinates and class labels. The following class labels were given: Ambulance, Fire Engine, Police, VIP. The label data will be stored in a text file having the exact name as the corresponding image file, and having a ".txt" extension. Each row in the label file represents one object in the image and has the following information: Object Class Index, Object Center Coordinates, Object width, and height. This dataset was prepared using Roboflow. Roboflow is a cloud-based software platform that provides tools for creating and managing custom object detection datasets, including datasets for YOLO models. Some of the features of Roboflow include Data Augmentation, Labelling tools, Dataset management, Export to YOLO format, and Integration with deep learning frameworks. The dataset was distributed into Train, Valid, and Test sets in a ratio of 70:15:15.

3.5. YOLO Algorithm and Configuration

YOLO is an object detection model, and the most recent addition to the family of YOLO algorithms is YOLOv8. The YOLOv8 algorithm is used to train our model. YOLOv8 uses DarkNet-53 as its CNN, which is 53 layers deep. It gives more precise results and is more efficient than previous YOLO versions and other object detection algorithms. The proposed emergency vehicle detection system uses the YOLOV8 algorithm to detect emergency vehicles in the video stream. This system consists of two components: the training component and the detection component. The training part uses the YOLOv8 algorithm on a dataset of emergency Vehicles. The detection module detects the emergency vehicles in real-time using a trained YOLOv8 algorithm. The pre-trained YOLOv8 model was fine-tuned on our emergency vehicle dataset. The Adam optimizer was used to train the model.

Mathematical model: Mathematics behind YOLO

It is necessary to understand the following steps:

1) To understand the YOLO algorithm, one should know that the object class, along with its bounding box, is predicted.
2) The bounding box has center coordinates, width, height, and value c, which represents the class of the object.
3) Along with that, we predict pc, a real number, which corresponds to the probability.

4) YOLO doesn't search for specific places in an image but instead divides the image into cells, each cell would take up the responsibility for predicting bounding box.
5) The center coordinates are calculated considering the entire cell, but the height and width are calculated with respect to the complete image size. Hence, an object will be considered to reside in a particular cell only if its center coordinates lie in that cell.
6) YOLO determines the probability that a cell can contain a class during its one pass of forward propagation using the equation below:

$$\text{SCORE } c, i = Pc\ X$$

Ci is where c is a certain class.

7) The next step after finding probabilities of class is non-max suppression, it is an approach to remove unwanted bounding boxes.
8) By performing IoU (Intersection over Union) on the bounding boxes with the greatest class probability, non-max suppression reduces the bounding boxes that are very near to one another.
9) This algorithm hence finds the IoU values of all bounding boxes and finally rets rid of unwanted boxes by deleting the ones whose IoU value is greater than a certain threshold.
10) This is done repeatedly until there are only different bounding boxes left.
11) The algorithm finally shows the bounding boxes of the respective class.

3.6. Training

The Prepared Dataset was trained using YOLOv8 algorithm using Darknet Framework. The Darknet Framework is an open-source framework. The model was trained for 200 epochs with a batch size of 64. The Rate of learning was changed to 0.001, and the Momentum was set to 0.9. The loss function used was YOLOv3 loss function. The trained algorithm was eventually evaluated on a set of validation images. The Evaluation Metrics used were Precision, F1 Score, and Recall.

4. Results and Discussion

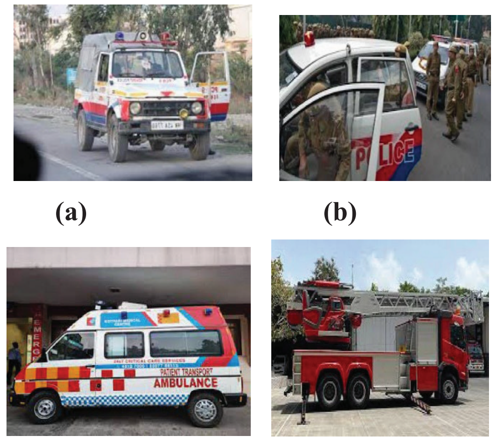
The system was able to detect emergency vehicles accurately and efficiently in real time. The findings show the Feasibility and effectiveness of our IOT-based Emergency Vehicle Detection using YOLOv8.Our System achieved an mAP score of 0.5.

4.1. Sample Dataset

4.2. Detection

The detection module was implemented using the YOLOv8 algorithm and the OpenCV library. This module takes a video stream as input and provides emergency vehicle detection in real time.

Figure 4. Images captured on roads



(a) (b)

(c) (d)

Figure 5. (a), (b), (c), (d) Sample dataset

Figure 6. Emergency vehicle identification using YOLOV8

The detected Vehicles are highlighted using bounding boxes and are hence labelled with their classes. Figure 6 shows a sample frame from the detected video stream. The detected emergency vehicles are highlighted with bounding boxes and labeled with their respective classes.

4.3. Performance Analysis

The effectiveness of our system is measured by the mean average precision (mAP) metric. The efficiency of our system is compared against the YOLOv3 model and a baseline system that uses traditional computer vision techniques.

4.4. Confusion Matrix

A confusion matrix is a performance evaluation tool that provides insights into how well a YOLO model is performing for object detection tasks.

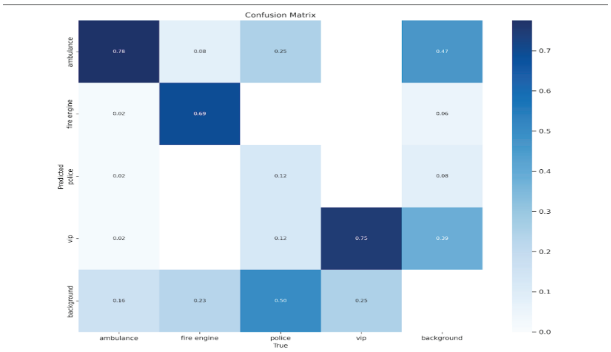


Figure 7. Confusion matrix

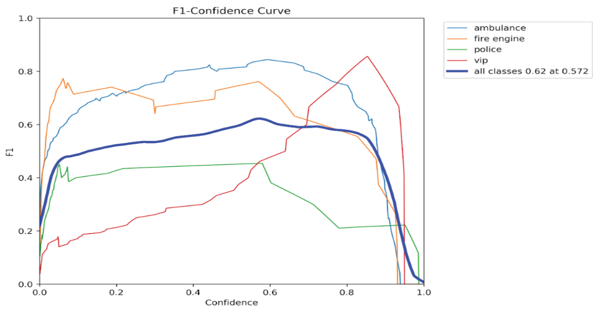


Figure 8. F1 Confidence curve on the vehicle identification

A confusion matrix is a table that compares the YOLO model's predicted labels to the test set's true labels. Figure 7 shows the Confusion Matrix of the experimental setup. Here is what each category represents: True Positive (TP): The YOLO model correctly detected an object in the image. False Positive (FP): The YOLO model predicted an object in the image, but there was no object in the ground truth label. True Negative (TN): The YOLO model correctly predicted that there was no object in the image. False Negative (FN): The YOLO model failed to detect an object in the image that was present in the ground truth label. By analyzing the values in the confusion matrix, one can get insights about Precision and Recall, Accuracy, False Positive rate, and Error analysis.

4.5. F1 Confidence and Precision Curve

F1 Curve plots the F1 Score against a range of confidence thresholds for object detection in a YOLO model. The F1 score is a

This metric combines Recall and Precision and is used to assess how well the model is working. Figure 8 shows the F1 Confidence curve for vehicle identification, with a range of 0 to 1.

The P curve represents a YOLO model's object detection precision vs. recall plot. Precision and recall are two important indicators for assessing the effectiveness of object identification algorithms, and the P curve can provide several insights into a YOLO model's performance. Figure 9 shows the Precision confidence curve for vehicle identification.

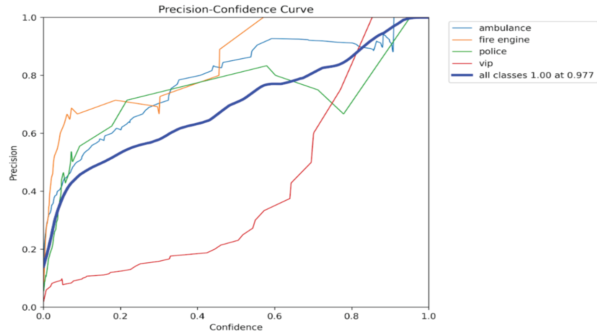


Figure 9. Precision confidence curve for the vehicle identification

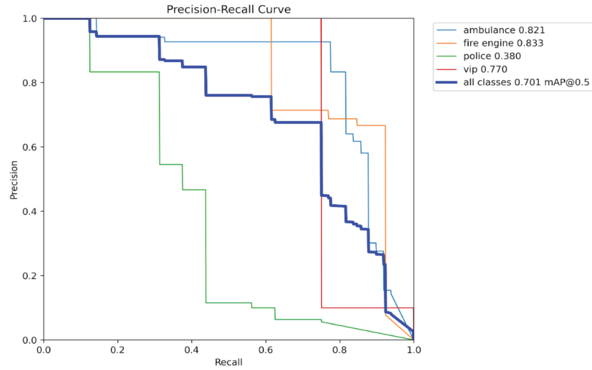


Figure 10. Precision-Recall curve for different emergency vehicle identification

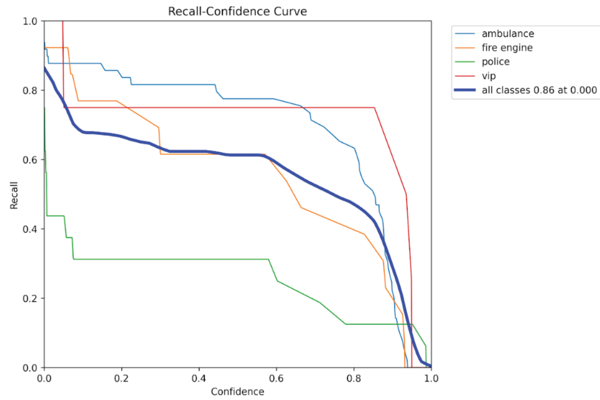


Figure 11. Recall-Confidence curve on different vehicles

4.6. Precision – Recall & Recall – Confidence Curve

The PR (Precision-Recall) curve plots precision against recall for object detection in a YOLO model.

The recall-confidence curve plots recall versus confidence for object detection in a YOLO model. While confidence is the likelihood that an object will actually be found within a projected bounding box, recall is the proportion of the actual objects that the model effectively detects. Figure 10 shows the Precision-Recall curve for different emergency vehicle identification.

4.7. Email Notification

We used the SMTP server of IoT to deliver email notifications. Email clients use SMTP to send and receive email messages. In the context of IoT, an SMTP server can be used to send alerts, notifications, and other information via email.

5. Conclusion

To sum up, the proposed emergency vehicle detection system uses the YOLOV8 algorithm to detect Emergency vehicles in real-time. An image collection of emergency vehicles was used to train the system, and it achieved a precision of 0.85 and a recall of 0.91. In this research work, we suggested an IoT-based YOLOv8 emergency vehicle detection system. The proposed system utilizes IoT sensors and cloud-based computing to enable real-time detection and response to emergency vehicles. Using a real-time image collection of emergency vehicles, we assessed the suggested system's performance and compared it to other leading-edge object identification algorithms. This system has successfully attained its objective of detecting emergency vehicles stuck in traffic and has hence generated alerts to save people's lives. The analysis showed that the suggested system performed better than other models in terms of accuracy, precision, recall, and F1 score, making it a promising solution for emergency vehicle detection in various settings. From the Confusion matrix, we can conclude that the system predicted values are as close as the actual values, and the values here directly point to the emergency vehicles. The F1Confidence, precision confidence, precision-Recall, and the Recall-Confidence curves present the model's performance in detecting various emergency vehicles, as mentioned.

The proposed system's advantages over other models can be attributed to its ability to detect emergency vehicles accurately in real-time, even in low-light conditions and challenging angles. The YOLOv8 method, which enables the identification of many objects in a single frame, is a sophisticated computer vision approach to achieve this. The proposed system also has the potential to significantly improve emergency response times, thereby increasing the chances of saving lives in emergencies. By providing real-time information about the location and type of emergency vehicles, the system can help emergency services respond quickly and efficiently, reducing the time it takes for emergency vehicles to reach their destination. Moreover, the proposed system's use of IoT sensors and cloud-based computing allows for scalability and flexibility, making it suitable for deployment in various settings, including urban and rural areas. This is particularly important in areas where traditional emergency response systems may be limited by infrastructure or resources. In conclusion, the proposed IoT-based emergency vehicle detection system using the YOLOv8 algorithm has the potential to revolutionize Emergency response systems by providing real-time information about the location and type of emergency vehicles.

The system performs better than other pioneering object detection models, and its scalability and flexibility make it a promising solution for emergency vehicle detection in various settings.

Future work can focus on improving the proposed system's performance by incorporating additional data sources, such as traffic cameras and aerial imagery, to enhance the system's ability to detect emergency vehicles in real-time. Additionally, the system's integration with existing emergency response systems can be explored to further improve emergency response times and coordination. Future research can not only be limited to camera-based emergency vehicle detection but can also be extended to other sensors like sound detecting sensors. Overall, the proposed system has the potential to save lives and improve emergency response times, making it a valuable contribution to the field of emergency management and IoT-based systems.

AUTHORS

Syed Suhana – Department of Computer Science and Engineering, CHRIST (Deemed to be University), India, e-mail: syed.suhana@btech.christuniversity.in.

Boppuru Rudra Prathap* – Department of Computer Science and Engineering, M. S. Ramaiah University of Applied Sciences, India, e-mail: brprathap@gmail.com.

Kavish Narang – Department of Computer Science and Engineering, CHRIST (Deemed to be University), India, e-mail: kavish.narang@btech.christuniversity.in.

Ivin Anto – Department of Computer Science and Engineering, CHRIST (Deemed to be University), India, e-mail: ivin.anto@btech.christuniversity.in.

*Corresponding author

References

[1] A. Baghel, A. Srivastava, A. Tyagi, S. Goel, and P. Nagrath, "Analysis Of Ex-Yolo Algorithm With Other Real-Time Algorithms For Emergency Vehicle Detection," in *Intelligent Methods in Emerging Robotics and Automation*, vol. 978, Lecture Notes in Networks and Systems, Springer, 2020, doi: 10.1007/978-981-15-3369-3_45.

[2] V. T. Tran and W. H. Tsai, "Audio-Vision Emergency Vehicle Detection," *IEEE Sensors Journal*, vol. 21, no. 24, pp. 27905–27917, Dec. 2021, doi: 10.1109/JSEN.2021.3127893.

[3] N. Dahiya, M. Garg, S. Gupta, and D. Gupta, "Smart Ambulance System Using Internet of Things: A Rumination," *Journal of Computational and Theoretical Nanoscience*, vol. 16, pp. 4249–4254, 2019, doi: 10.1166/jctn.2019.8508.

[4] W.-J. Chang, L.-B. Chen, and K.-Y. Su, "DeepCrash: A Deep Learning-Based Internet of Vehicles System for Head-On and Single-Vehicle Accident Detection With Emergency Notification," *IEEE*

Access, vol. 7, pp. 148 163–148 175, 2019, doi: 10.1109/ACCESS.2019.2946468.

[5] A. Benjumea, I. Teeti, F. Cuzzolin, and A. Bradley, “YOLO-Z: Improving Small Object Detection in YOLOv5 for Autonomous Vehicles,” *arXiv preprint arXiv:2112.11798*, 2021.

[6] S. Ren, K. He, R. Girshick, and J. Sun, “Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1137–1149, June 2017, doi: 10.1109/TPAMI.2016.2577031.

[7] M. Zuraimi and F. Zaman, “Vehicle Detection and Tracking Using YOLO and DeepSORT,” in *Proc. 2021 Int. Seminar on Intelligent Control and Future Energy (ISICFE)*, pp. 23–29, 2021, doi: 10.1109/ISCAIE51753.2021.9431784.

[8] J. Lu, C. Ma, L. Li, X. Xing, Y. Zhang, Z. Wang, and J. Xu, “A Vehicle Detection Method for Aerial Image Based on YOLO,” *Journal of Computer and Communications*, vol. 6, no. 11, pp. 98–107, Nov. 2018, doi: 10.4236/jcc.2018.611009.

[9] J. Hussain, B. R. P. Prathap, and A. Sharma, “An Improved and Efficient YOLOv4 Method for Object Detection in Video Streaming,” in *Data Science and Security: Proceedings of IDSCS 2022*, Singapore: Springer Nature Singapore, 2022, pp. 305–316, doi: 10.1007/978-981-19-2211-4_27.

[10] K. Munasinghe, T. Waththegedara, I. Wickramasinghe, O. K. Herath, and L. Velmanickam, “Smart Traffic Light Control System Based on Traffic Density and Emergency Vehicle Detection,” in *Proc. 2022 Middle East Conf. on Innovation in Revision and Communication (MERCon)*, pp. 1–6, 2022, doi: 10.1109/MERCon55799.2022.9906184.

[11] A. Ashir, “A Transfer-Learning-Based Approach for Emergency Vehicle Detection,” *Eurasian Journal of Science and Engineering*, vol. 8, no. 1, p. 75, 2022, doi: 10.23918/eajse.v8i1p75.

[12] R. Huang, J. Pedoeem, and C. Chen, “YOLO-LITE: A Real-Time Object Detection Algorithm Optimized for Non-GPU Computers,” in *Proc. 2018 IEEE International Conference on Big Data (Big Data)*, pp. 2503–2507, 2018.

[13] W. Fang, L. Wang, and P. Ren, “Tinier-YOLO: A Real-Time Object Detection Method for Constrained Environments,” *IEEE Access*, vol. 8, pp. 1935–1944, 2020, doi: 10.1109/ACCESS.2019.2961959.

[14] L. Cheng, J. Li, P. Duan, and M. Wang, “A Small Attentional YOLO Model for Landslide Detection from Satellite Remote Sensing Images,” *Landslides*, vol. 18, no. 10, pp. 3251–3264, 2021, doi: 10.1007/s10346-021-01694-6.

[15] C. Dewi, R.-C. Chen, Y.-T. Liu, X. Jiang, and K. Hartomo, “YOLO V4 for advanced traffic sign recognition with synthetic training data generated by various GAN,” *IEEE Access*, vol. 9, pp. 97236–97248, 2021, doi: 10.1109/ACCESS.2021.3094201.

[16] S. Chen and W. Lin, “Embedded System Real-Time Vehicle Detection Based on Improved YOLO Network,” in *Proceedings of the IEEE 2nd International Conference on Intelligent Control, Measurement and Computing (ICICMC)*, pp. 1400–1403, 2019, doi: 10.1109/IMCEC46724.2019.8984055.