ADROID Conference Series: Engineering and Technology
Volume 2, Issue 3, pp.83-92, 2026

Street Eye: A Smart Vision-based System for Road Surface Monitoring Using Neural Networks

B. Srinivasa Rao, Chilakalapudi Malathi, Kodavalluri Dhanush, Mohammed Sohail, Kethepalli Jagadeesh, Shaik Faisal

Department of Information Technology, Lakireddy Bali Reddy College of Engineering, Mylavaram, Andhra Pradesh, India

buragasrinivasarao@gmail.com, chmalathi542@gmail.com, kodavalluridhanush@gmail.com, sohailmohammed778866@gmail.com, jagadeeshk0606@gmail.com, faisalsk743@gmail.com

Abstract. Potholes on the road surface significantly impact road safety, vehicle performance, and passenger comfort. Traditional road inspection systems rely on manual surveying or sensor-based approaches, which are inefficient, expensive, and not suitable for large-scale, real-time road inspection. Recently, computer vision and deep learning technologies have made it possible to automatically analyze road conditions from visual data. In this paper, we introduce Street Eye, a smart, vision-based road-surface monitoring system that uses deep neural networks for automatic pothole detection. The proposed system uses a YOLO-based object detection algorithm to detect potholes from continuous video streams captured by a camera. For each detected pothole, geolocation data such as latitude, longitude, and timestamp are extracted and recorded in a centralized database. A location comparison function is designed to detect road users approaching potholes already detected, enabling real-time alerts. With the integration of real-time detection, geolocation data management, and alerting services, the proposed system is expected to improve road safety and facilitate intelligent transportation infrastructure by efficiently and cost-effectively monitoring road conditions.

Keywords: Pothole Detection; Road Surface Monitoring; Computer Vision; Deep Learning; YOLO; Neural Networks; GPS Geotagging; Geospatial Database; Real-Time Alert System.

1. INTRODUCTION

Road travel has been identified as playing an important role in the development process and in our lives. The condition of the road has been identified as crucial in ensuring the smooth flow of traffic and the level of safety. Potholes have been identified as one of the most important issues in conducting road surveys, leading to traffic jams and accidents, especially when the weather is bad. Currently, road surveys rely on manual checks, which are inefficient because they are slow and do not yield immediate results, nor do they support long-term identification of potholes. In order to improve the shortcomings associated with the process of conducting road surveys using a manual check, an automated process has been developed by researchers. This process uses a machine that surveys the road by causing it to vibrate using an accelerometer and gyroscope, which can identify defects on the road, but the efficiency of the process depends on the speed of the vehicle and the suspension system. The use of image processing methods in conducting road surveys has been widely accepted, whereby the edges, thresholding, and texture of the road can identify potholes, but the process may not yield the desired results due to the presence of shadows.

Image processing has been more effective in road surveys, especially after the introduction of computer vision and the use of deep learning. Convolutional Neural Networks (CNNs) have been widely used in image processing for the analysis of images. The features and objects in the image can be identified through the use of multiple layers in the network. The use of the YOLO object detection technique has been more effective in image processing. The object in the image can be scanned once through the single-stage object detection technique. This technique has been more effective in video processing.

The use of deep learning has been more effective in road surveys. The use of deep learning in identifying potholes in the road has been more effective. The identification of potholes in the road has been more accurate and reliable.

In order to solve the aforementioned problems, the paper has proposed a smart vision-based road surface monitoring system that incorporates the concept of pothole detection along with deep learning and efficient management of geospatial data. Street Eye is the proposed smart road surface monitoring system that incorporates a neural network using YOLO for efficient and effective detection of potholes from the video frames received by the system from the camera attached to a vehicle. The video received by the system is analyzed real-time to identify irregularities that occur on the road surface even under varying environmental conditions. Additionally, metadata regarding the latitude, longitude, and timestamp of the pothole is available with the proposed system. These details are stored in a database to efficiently organize and retrieve the history of road condition [5]. Moreover, the proposed system incorporates efficient facilities that aid in the constant monitoring of the distance between the location of the road user and the location of the pothole. Such facilities aid in increasing road safety, awareness of the road user, and smart technologies used in road transport.

2. LITERATURE REVIEW

The state of the roads is a crucial aspect that has a direct influence on our lives, both in terms of safety and the economy. The World Health Organization states that the state of the road is the main cause of accidents and damages to vehicles. Potholes in the road are a dangerous aspect, as they affect everyone driving on the road in a negative way, especially when there is an irregular pattern in the maintenance of the road. Thus, to avoid accidents and improve road safety, it is important to monitor the state of the road and detect damage before it occurs. However, the report is only able to provide data, trends, and recommendations, but not an automated system to monitor the state of the road [6].

Huang et al. proposed a vision-based pothole detection system that used traditional image processing techniques to detect potholes in the road. It analyzed the change in texture, edges, and irregularities in the road to detect potholes. Although the proposed system was successful in detecting potholes, it was highly sensitive to light, shadows, and road texture. Furthermore, the system was not a real-time system and performed poorly in an urban environment [7].

Redmon et al. proposed a new model for object detection, which is referred to as YOLO. YOLO stands for You Only Look Once. YOLO is a simplified object detection model. Unlike other object detection models, which follow a step-by-step approach for object detection, this model follows an object detection approach that is referred to as a regression problem. This model treats the image as a whole entity for object detection. This model is an object detection model that runs in real-time with low latency compared to other object detection models. This model runs once through the image to perform object detection. The major problem with this model is that it is not able to detect smaller-sized objects such as potholes.

To avoid the limitations of this model and other YOLO models, Bochkovskiy et al. proposed a new object detection model referred to as YOLOv4. This model is able to detect objects accurately without compromising its performance. This model is tested for real-time object detection. This model performs well, proving its reliability for object detection. This model still lacks the capacity to deal with domain-related issues. This model still does not have the capacity to deal with issues related to road surface and geolocation.

A real-time pothole detection technique has been proposed by Shaghouri et al. using deep learning models that combine convolutional neural networks. It compares road images received from cameras fixed to the vehicles. It is useful for increasing accuracy compared to other conventional vision-based techniques. Moreover, it shows that it is possible to learn complex features of potholes using deep learning models. However, it is not useful for the actual application of the technique for the benefit of users because it does not use geospatial data[10].

POT-YOLO has proposed an improved pothole boundary detection technique by including an edge segmentation feature with YOLOv8 models. It is useful for improving the pothole detection technique by reducing false alarms from road markings and shadows. However, it is not useful for the actual application of the technique for the benefit of users[11].

The instance segmentation techniques are used in the detection of potholes by using the YOLOv8 networks. These instance segmentation techniques are useful in the accurate localization and segmentation of the detected potholes. These accurate estimations are useful in determining the size of the detected potholes. These accurate estimations are useful in determining the severity of the detected potholes. However, these instance segmentation techniques are less optimized and require more computational power[12].

From the review of literature, it is evident that there are developments in the vision-based and deep learning methods that are used for the detection of potholes. The traditional methods that are used for the detection of potholes are highly prone to environmental factors. However, the latest methods that are used for the detection of potholes are efficient. The YOLO networks are useful for the real-time processing of video feeds. These networks are useful for the real-time processing of video feeds. Thus, it is evident that the YOLO networks are useful for real-time detection of potholes. However, it is evident that there is less research done in the detection of potholes with respect to geolocation tagging, data storage, and real-time alerts to users. Thus, it is evident that there is an urgent need for developing a comprehensive framework that is the motivation behind conducting the proposed research.

3. DATA COLLECTION & PREPROCESSING

The process of collecting the data involves an embedded vision system that includes a Raspberry Pi 5 as a main processing unit. A Raspberry Pi camera module connects to the board and captures images and video frames of the road surface in real-time as the system moves. To enable accurate spatial referencing of the images captured, a GPS module connects to the embedded vision system. The GPS module captures the latitude and longitude coordinates of each image.

The camera in the embedded vision system captures images of the road surface in real-time. The images are given a time stamp and referenced with their GPS coordinates. This enables accurate spatial and time-based mapping of the road surface. The images are then sent to a backend server for processing and analysis. After being received at the backend server, the images are subject to an initial filtering process. This process filters out blurred images, repeating images, and poor-quality images that could interfere with model performance. The images are then labeled to identify the areas of interest for the object detection model. The labeling process involves using tools such as LabelImg and Roboflow.

The dataset has been annotated and is in the YOLO format. Each image has a corresponding text file that contains class IDs and normalized coordinates. The dataset is compatible with the YOLOv8 Nano model used in the detection step.

Before we start with the training process, several preprocessing steps are taken that make the model more robust and versatile. These steps include resizing the image to a fixed size that is compatible with the YOLOv8 Nano model. Normalization of pixels is also performed so that the model converges better during training. Augmentation techniques such as rotation, scaling, flipping, and changing brightness are applied.

The final dataset is divided into training, validation, and test datasets so that unbiased performance evaluation can be carried out. The model is now trained using the following command : yolo task=detect mode=train model=yolov8n.pt data=data.yaml epochs=60 batch=8 imgsz=640 patience=20.

The model is now validated with new images so that its performance can be evaluated. The command used for validation is: yolo task=detect mode=predict model=runs/detect/train/weights/best.pt source=test_images/.

The model is now evaluated using the following command : yolo task = detect mode = val model = runs/detect/train/weights/best.pt data = data.yaml.

The overall system is developed by making use of modern technologies in the field of software to process the data in an efficient manner. In the system, Python is used for the processing of the images and the objects, and the object detection is done by making use of YOLOv8 Nano. Node.js is used in the backend for managing the API requests, data transmission, and communication with the embedded device and the front-end application. MongoDB is used for the storage of the data obtained from the object detection, GPS coordinates, etc., and MongoDB Compass is used for the management of the database. In the system, the front-end interface is developed by making use of React Native to display the mobile view of the detected anomalies on the road, and Postman and Visual Studio Code are used for testing the system to ensure the reliable performance of the system.

4. PROPOSED SYSTEM

The system proposed in this paper has been named. It is called Street Eye. It has been defined as a smart vision-based intelligent system for the monitoring of the road surface. The system proposed in this paper aims at detecting the potholes on the road along with their geospatial information. The system proposed in this paper would comprise a system architecture with different layers. These layers would comprise the Processing Layer, Data and Application Layer, and User Layer.

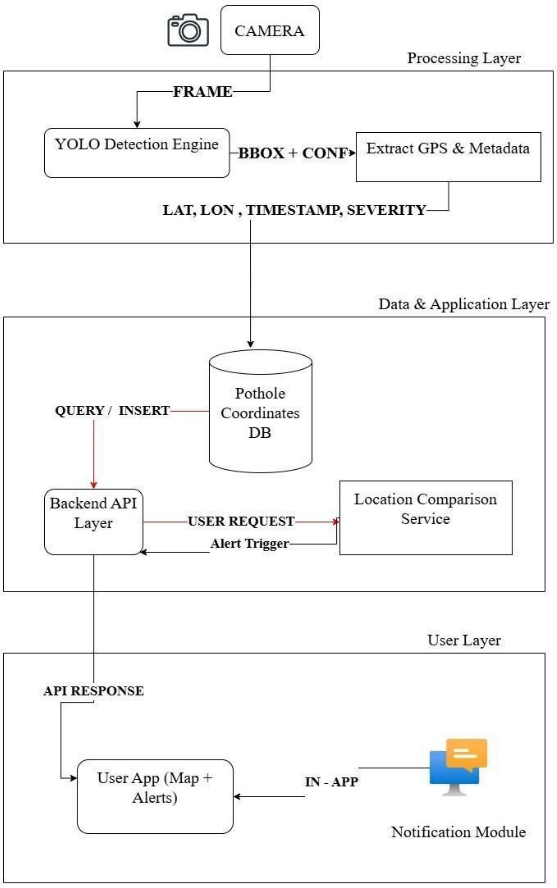

Fig. 1. Overall Architecture Street Eye

A. System Overview

The system starts with the acquisition of video frames depicting road surface conditions using a camera mounted on a moving object. The system further processes these video frames using a deep learning object detection model, which detects potholes. Upon detecting a pothole, it collects pertinent metadata such as geographical coordinates and timestamps. The metadata is stored in a central database. The metadata is used to verify the location of users and send notifications when they are approaching areas that are likely to have potholes.

B. Processing Layer

The Processing Layer processes individual frames of the video in real time to retrieve information concerning potholes. Images of the road surface from the video are captured by the camera. Individual frames of the video are then processed by the system. Individual frames of the video are processed by the system using the YOLO engine. It uses its algorithm to identify the potholes by drawing bounding boxes and confidence levels. The severity of the pothole is identified based on the visual attributes of the pothole, which may be the size and shape of the pothole, indicated by the bounding box of the pothole. At the same time, the GPS and metadata extraction module processes information such as latitude, longitude, and timestamps of individual frames of the video. A data object concerning potholes is constructed by using the information derived from individual frames of the video and other data.

C. Data and Application Layer

The Data and Application Layer manages data storage, retrieval, and application logic. The detected pothole data and its metadata are saved in a centralized Pothole Coordinates Database. This database allows for quick querying and historical analysis of road surface conditions.

A Backend API Layer enables communication between the processing layer, database, and application layer. It handles data insertion, query requests, and response generation. A Location Comparison Service compares the user’s location with stored pothole coordinates continuously. If the user’s location matches or approaches a pothole-prone area within a set distance, an alert is generated and sent to the backend system.

D. User Layer
This layer enables the communication between the system and the end user through a mobile/web application. This application will be used by the user to view the location of the pothole on the map. It also sends a notification to the user in real time while he approaches the critical area of the road. These notifications are sent through the Notification Module.
The API response system sends information to the user regarding the condition of the road. This user layer sends information to the user. In this way, the user becomes aware of the situation.

E. Key Advantages of the Proposed System
There are a number of benefits associated with the suggested system. First, the system will include features such as the detection of potholes, geolocation tagging, data management, and alerting. Second, the proposed system will be modular, hence easier to integrate with the smart city infrastructure. Finally, the use of deep learning ensures the reliability of the system despite the environmental conditions.

5. RESULT & DISCUSSION

The pothole detection system was assessed using object detection metrics obtained from the pre-trained backbone network. To assess the pothole detection system's effectiveness in real-world scenarios, we assessed the model on a new dataset which the network was not previously exposed to.

A. Detection Performance Analysis
The Precision-Recall curve indicates a good trade-off between precision and recall at varying confidence levels. The model has a mean Average Precision (mAP) at IoU 0.5 of 0.781. This indicates excellent pothole detection capability at high confidence levels.

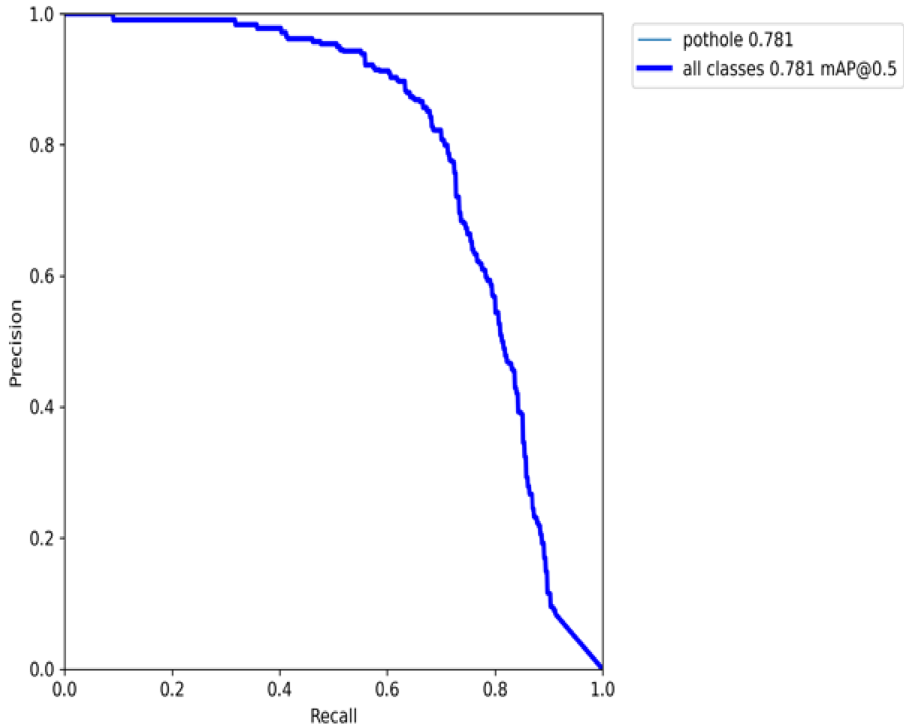

Fig. 2. Precision–Recall (PR) curve

The precision-confidence curve increases steadily with the increase in confidence level, and the curve reaches a perfect
1.00 precision at a 0.967 confidence level. In short, the more the confidence level, the more accurate the prediction.

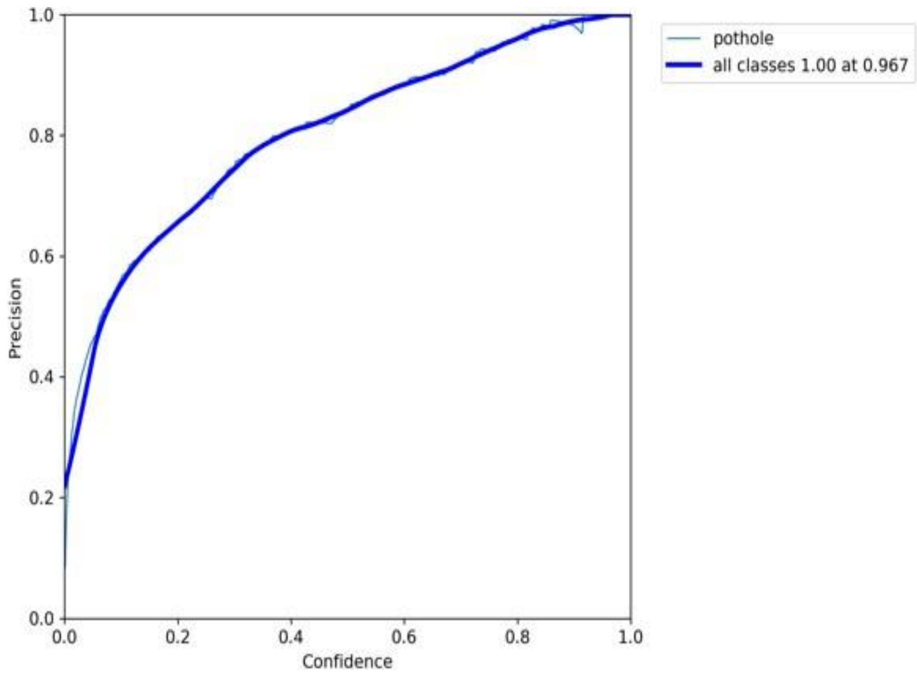

Fig. 3. Precision–Recall (PR) curve

The F1-Confidence curve illustrates the balance in the model. When the confidence level is close to 0.414, the F1-score increases to over 0.75. Here, precision and recall are in harmony. This is the confidence level that I would recommend using to deploy the model.

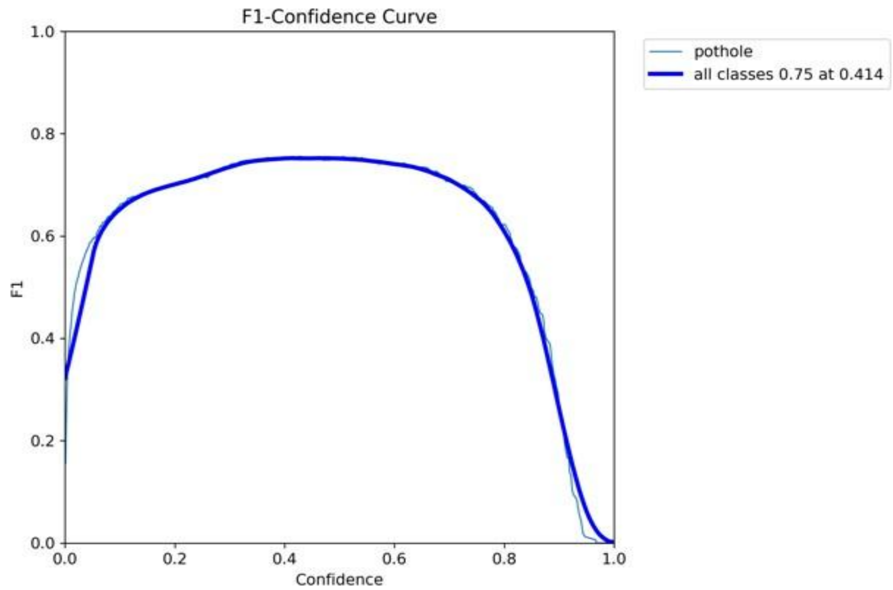

Fig. 4. F1–Confidence curve

As we can see from the Recall-Confidence curve above, the model has a maximum recall rate of 0.88 when the confidence threshold is close to zero. That is to say, when the threshold is low, the model is able to detect most pothole instances. As we expected, the recall rate decreases when the confidence threshold increases.

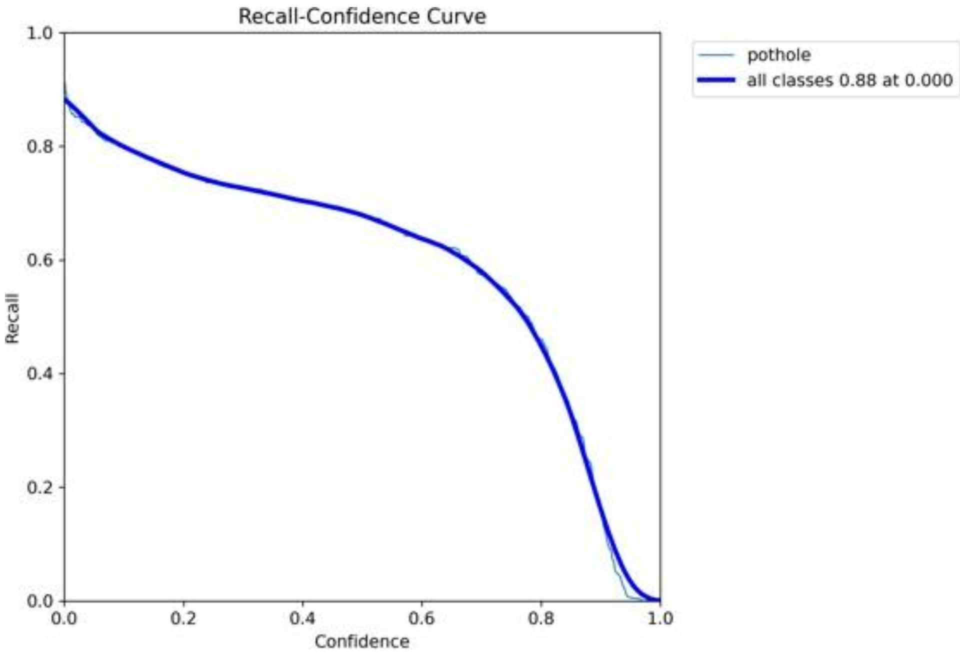

Fig. 5. Recall–Confidence curve

B. Qualitative Detection Results

Apart from the numerical results, we also test the pothole detector on actual images of the road. After training, the YOLOv8 Nano model works well in detecting and outlining the potholes. This model works well for holes of varying sizes and depths, whether they contain water or have rough and smooth asphalt. The bounding box closely fits the pothole. The confidence levels range from
0.46 to 0.90. This indicates the efficiency of the model under varying conditions of light, shadows, and road textures.

Fig. 6. Qualitative Pothole Detection Results

When you see a section of the road filled with potholes, that section is referred to as a "rough road." For your information, there is a four-level warning system. This warning level varies depending on the density of potholes, the confidence of the model regarding their presence, and their distance from you.
The main parameters on which the alert score is dependent are as follows:
(i) the number of potholes found in the section
(ii) the confidence of the model regarding their presence
(iii) the distance of the potholes from you.

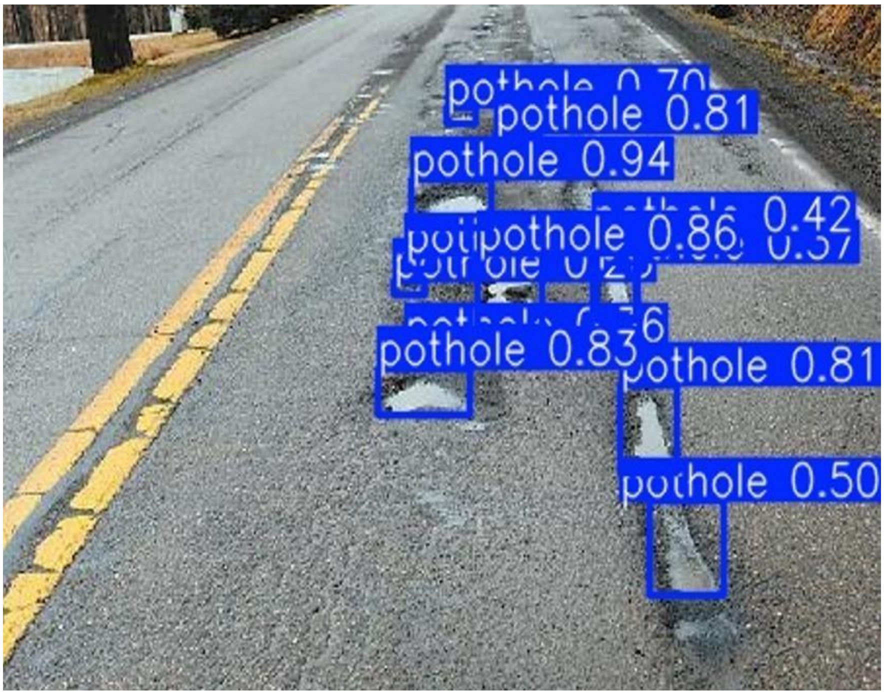

Fig. 7. Multiple Pothole Detection in a Frame

If a few potholes are encountered and the distance to the user is greater than 100 meters, then a Yellow Alert is declared. This is a low-risk situation. An Orange Alert is declared if a moderate number of potholes are encountered and the distance to the user is less than 100 meters. This is a situation that might cause a hazard.
If the user is very close to the potholes, then a Red Alert is declared. This is because this is a situation of immediate risk. This alert is declared to let the user know that he/she is about to encounter a dangerous situation. Lastly, if a few potholes are encountered in a small area, then a Dark Red Alert is declared. This is because this is a situation of a severely damaged and rough road. This alert is declared to let the user know that this is the highest risk.

C. GPS Delay Time Analysis

Real-time pothole detection systems need to handle GPS delays carefully to provide timely and reliable alerts. The proposed system estimates the distance between the user and detected potholes using GPS data. Therefore, the alert delay must be adjusted based on vehicle speed. A GPS delay time study was conducted for speeds from 10 km/h to 100 km/h, looking at the time needed to cover short distances of 5 m and 10 m.

TABLE I
GPS DELAY TIME ANALYSIS.

Speed (Km/h)	Speed (m/s)	Time for 5m	Time for 10m	Suggested Delay
10	2.7	1.8s	3.7s	2 – 3 sec
20	5.5	0.9s	1.8s	1 – 1.5 sec
30	8.3	0.6s	1.2s	0.8 – 1 sec
40	11.1	0.45s	0.9s	0.6 – 0.8 sec
50	13.8	0.36s	0.72s	0.5 – 0.7 sec
60	16.6	0.3s	0.6s	0.4 – 0.6 sec
70	19.4	0.26s	0.52s	0.35 – 0.5 sec

80	22.2	0.23s	0.45s	0.3 – 0.45 sec
90	25.0	0.20s	0.40s	0.25 – 0.4 sec
100	27.7	0.18s	0.36s	0.22 – 0.35 sec

However, when the speed is low, more time is taken to cover shorter distances. This will ensure the utilization of a high value of delays. For instance, when the speed of the vehicle varies from 10-20 km/h, the range of delays from 1-3 seconds is utilized in providing accurate GPS signals and timely warnings. At high speeds of 30-50 km/h, the time taken for the vehicle to cover a certain distance is reduced. In this case, the value of the delays is reduced. The range of the delays is set at 0.5-1 second.

When the speed is increased and is greater than 60 km/h, the vehicle is able to reach the zones in a shorter time. In this case, a low value of delays is required in providing timely warnings. In this case, the range of the delays is set at 0.22-0.5 seconds. It is clear that the GPS delay technique based on the speed of the vehicle is highly effective in providing accurate GPS signals and timely warnings of the proposed system.

6. CONCLUSION

Although it is evident that the proposed Street Eye system is effective in detecting potholes, it is still possible to make some improvements that will make it even more effective and efficient. One of the possible improvements is that it should be possible for it to detect other types of road issues aside from potholes. It will then be possible to analyze road issues completely. It is also possible to make use of deep learning models that will make it possible to detect road issues accurately. It is possible to make use of data from multiple sensors that will make it possible to accurately detect road issues even during bad weather. It is possible to make some improvements to the proposed system that will make it possible for it to be more efficient when it is used for real-time purposes, which will make it possible to avoid using cloud computing. It is possible to make use of data from other vehicles that will make it possible to create road condition maps for the entire city.
From an application perspective, it is possible to make use of the proposed system to make it possible to optimize routes by avoiding roads with potholes. It is possible to make use of it for predictive maintenance of roads by road authorities. It is possible to make use of advanced analytics that will make it possible to predict when potholes will be formed so that road maintenance will be prioritized. It is possible to make it possible to create an even more comprehensive intelligent transportation system that will make it possible to make road travel safer, smarter, and more sustainable.

Conflict of Interest

The authors declare no conflicts of interest regarding the current research.

References

[1] World Health Organization, Global Status Report on Road Safety, WHO Press, 2018.
[2] Huang, X., et al., “Automatic Pothole Detection Using Vision-Based Techniques,” Journal of Advanced Transportation, 2018.
[3] Redmon, J., Divvala, S., Girshick, R., Farhadi, A., “You Only Look Once: Unified, Real-Time Object Detection,” CVPR, 2016.
[4] Bochkovskiy, A., Wang, C. Y., Liao, H. Y. M., “YOLOv4: Optimal Speed and Accuracy of Object Detection,” arXiv, 2020.
[5] Shaghouri, A. A., et al., “Real-Time Pothole Detection Using Deep Learning,” arXiv, 2021.
[6] Peden, M., et al., World Report on Road Traffic Injury Prevention, World Health Organization, Geneva, 2004.
[7] Maeda, H., Sekimoto, Y., Seto, T., Kashiyama, T., Omata, H., “Road Damage Detection Using Deep Neural Networks with Images Captured Through a Smartphone,” Computer-Aided Civil and Infrastructure Engineering, 2018.
[8] Ren, S., He, K., Girshick, R., Sun, J., “Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks,” IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017.
[9] Kodabagi, B., Bhavana, V., “POT-YOLO: Real-Time Road Potholes Detection Using Edge Segmentation-Based YOLOv8 Network,” IEEE Sensors Journal, 2024.
[10] Van, D. D., “Application of Advanced Deep Convolutional Neural Networks for the Recognition of Road Surface Anomalies,” Engineering, Technology & Applied Science Research, 2023.

[11] Shravan, A. P., Sowmya, V., "YOLOv8-Based Visual Detection of Road Hazards: Potholes, Sewer Covers, and Manholes," 5th International Conference on Computing and Communications Applications and Technologies (I3CAT), 2023.
[12] Neha, Murari, B. K., "Enhanced Roadway Pothole Detection Using YOLOv8 Instance Segmentation," International Journal of Engineering Research & Technology (IJERT), 2024.