

Terravision-Pothole: A Comprehensive Edge-AI Framework For Real-Time Pothole Detection, Severity Modeling And Geo-Tagged Smart Infrastructure Monitoring

Asst.Professor Anil Londhe

Dept. of Artificial Intelligence & Data Science
University of Mumbai, Mumbai, India

Mr. Ayush Singh

Dept. of Artificial Intelligence & Data Science
University of Mumbai, Mumbai, India

Mr. Sagar Singh

Dept. of Artificial Intelligence & Data Science
University of Mumbai, Mumbai, India

Mr. Sujal Suthar

Dept. of Artificial Intelligence & Data Science
University of Mumbai, Mumbai, India

Mr. Raj Upadhyay

Dept. of Artificial Intelligence & Data Science
University of Mumbai, Mumbai, India

Abstract: Road accidents caused by potholes, debris, stalled vehicles and unmarked road irregularities remain a major challenge for transportation safety worldwide. Traditional navigation systems rely heavily on manual reporting and therefore cannot provide automated detection of hazardous road conditions in real time. This paper presents TerraVision-Pothole, an intelligent Edge-AI based road hazard detection framework designed to automatically identify potholes and road anomalies using computer vision techniques. The system processes video streams captured from vehicle-mounted dashcams and analyzes them using deep learning models such as YOLOv8 for object detection. Image preprocessing operations are performed using OpenCV to enhance detection accuracy. The framework integrates GPS-based geo-tagging and cloud-based analytics to visualize detected hazards on an interactive monitoring dashboard. To ensure user privacy, the system incorporates automatic face and license plate blurring mechanisms before transmitting visual data. Experimental evaluation demonstrates that the system achieves approximately 85% detection accuracy while maintaining reliable real-time performance. The proposed framework contributes toward the development of intelligent transportation infrastructure by enabling automated road monitoring, proactive hazard detection, and data-driven maintenance planning.

Index Terms—Road Hazard Detection, Computer Vision, YOLOv8, Edge Computing, Intelligent Transportation Systems, Road Safety.

I. INTRODUCTION

Road infrastructure plays a critical role in modern transportation systems by enabling efficient mobility and economic development. However, deteriorating road conditions such as potholes, cracks, debris accumulation and poorly marked speed breakers often create hazardous driving environments. These road anomalies are responsible for a large number of accidents, vehicle damage and traffic disruptions.

In many regions, road maintenance authorities rely on manual inspection techniques to identify damaged road segments. These methods involve periodic road surveys conducted by maintenance personnel. Although effective in certain cases, manual inspection methods are labor-intensive, time-consuming and unable to provide continuous monitoring of road conditions.

Modern navigation applications such as Google Maps and Waze allow users to report hazards manually. However, such systems depend heavily on user participation and may not detect hazards immediately. Consequently, many dangerous road conditions remain unreported until accidents occur.

Recent advances in artificial intelligence and computer vision provide an opportunity to develop automated road monitoring systems. Deep learning algorithms are capable of analyzing visual data and identifying patterns associated with road defects. Convolutional Neural Networks (CNNs) have demonstrated exceptional performance in tasks such as object detection, image classification and scene understanding.

Among the various deep learning approaches, the YOLO (You Only Look Once) family of object detection models has gained popularity due to its ability to perform real-time detection with high accuracy. The latest YOLOv8 architecture improves detection performance by optimizing feature extraction and bounding box prediction.

The TerraVision-Pothole system leverages these technologies to build an intelligent road monitoring framework. The system analyzes video streams captured from vehicle-mounted cameras and automatically detects road hazards. The detected hazards are geo-tagged using GPS data and transmitted to a cloud-based monitoring dashboard where authorities and drivers can visualize hazard alerts.

The key contributions of this research include:

- Development of an automated pothole detection framework using deep learning.

- Integration of edge computing to enable real-time hazard detection.
- Implementation of severity classification based on bounding box analysis.
- Deployment of a geo-tagged hazard monitoring dashboard.

I. LITERATURE SURVEY

Automated road hazard detection has gained significant attention with the advancement of intelligent transportation systems and computer vision technologies. Early research relied on traditional image processing techniques such as edge detection, thresholding and morphological filtering to detect cracks and potholes in road surfaces. Although computationally efficient, these approaches were highly sensitive to illumination changes and variations in road texture.

The development of deep learning significantly improved visual recognition capabilities. Krizhevsky et al. [5] introduced deep convolutional neural networks through the AlexNet architecture, demonstrating the ability of deep models to learn hierarchical visual features from large-scale datasets. Later architectures such as ResNet [9] improved feature extraction and classification performance in complex visual environments.

Several studies have applied deep learning models for pavement defect detection. Zhang et al. [6] proposed a CNN-based framework for road crack detection using high-resolution road images. Similarly, Maeda et al. [7] created a road damage detection dataset using smartphone images and trained deep neural networks to classify different road defects including potholes and cracks.

Object detection models significantly improved real-time detection performance. Redmon et al. [1] introduced the YOLO (You Only Look Once) framework which performs object detection in a single neural network pass. Region-based models such as Faster R-CNN [8] also demonstrated strong detection accuracy but required higher computational resources.

Further improvements to object detection models were introduced through YOLOv4 [2] and the recent YOLOv8 architecture [3], which provide improved detection accuracy and faster inference speeds. These characteristics make YOLOv8 suitable for real-time edge-based applications.

Edge computing has also become an important component in intelligent transportation systems. Shi et al. [10] highlighted the role of edge computing in enabling scalable Internet-of-Things (IoT) applications by processing data closer to the source and reducing latency.

Despite these advancements, many existing approaches focus on individual components of road monitoring. There remains a need for an integrated system that combines real-time hazard detection, edge processing and centralized monitoring. The proposed TerraVision-Pothole system addresses this gap by integrating deep learning-based detection with edge computing and cloud-based analytics for automated road infrastructure monitoring.

II. SYSTEM ARCHITECTURE

The TerraVision-Pothole framework follows a multi-layer Edge-to-Cloud architecture consisting of three primary components.

A. Edge Processing Layer

The edge layer is responsible for capturing road video streams and performing real-time analysis. Dashcams mounted on vehicles continuously capture road footage. Video frames are extracted and preprocessed using OpenCV operations such as noise reduction, image normalization and contrast enhancement.

The preprocessed frames are passed to the YOLOv8 detection model which identifies potential hazards including potholes, debris and stalled vehicles.

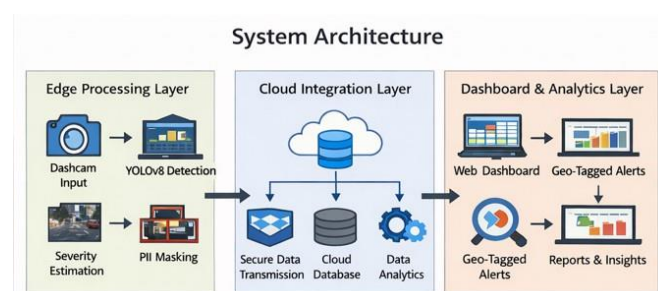


Fig. 1. Comprehensive Edge-Cloud Architecture of TerraVision-Pothole

B. Cloud Integration Layer

After hazards are detected at the edge layer, relevant information is transmitted to a cloud server through secure communication protocols. The transmitted data includes hazard type, GPS coordinates, timestamp and severity level.

The cloud layer stores this information in a centralized database where it can be used for analytics and monitoring purposes.

C. Dashboard and Analytics Layer

The dashboard layer provides a web-based interface for visualizing detected hazards. An interactive map displays hazard locations using color-coded markers representing severity levels.

This dashboard enables transportation authorities to monitor road conditions, analyze hazard distribution patterns and prioritize maintenance operations.

III. METHODOLOGY

The TerraVision-Pothole system follows a multi-stage computer vision pipeline designed for real-time hazard detection from vehicle-mounted cameras. The methodology consists of five major stages: video frame acquisition, image preprocessing, deep learning-based hazard detection, severity estimation, and geo-tagged alert generation.

A. Video Frame Acquisition

Road footage is captured using vehicle-mounted dash cameras that continuously record the road environment. The captured video stream is divided into individual frames at regular intervals for analysis.

Let the input video stream be represented as:

$$V = \{F_1, F_2, F_3, \dots, F_n\} \quad (1)$$

where:

- V represents the input video stream
- F_i represents the i^{th} extracted frame
- n represents the total number of frames captured

Each frame F_i is passed to the preprocessing module before being analyzed by the detection model.

B. Image Preprocessing

Image preprocessing improves the visual quality of frames

and reduces noise that may negatively affect detection accuracy. The preprocessing stage includes noise filtering, normalization, and contrast enhancement.

Let the raw image frame be represented as:

$$I_r = f(x, y) \quad (2)$$

where:

- I_r represents the raw input image
- x and y represent pixel coordinates in the image

After preprocessing operations, the enhanced image is obtained as:

$$I_p = P(I_r) \quad (3)$$

where:

- I_p represents the processed image
- $P(\cdot)$ denotes the preprocessing function including noise reduction and normalization

These preprocessing steps improve the robustness of the detection model under varying lighting conditions.

C. Hazard Detection using YOLOv8

The processed image frame is analyzed using the YOLOv8 deep learning model. YOLOv8 is a single-stage object detection network capable of predicting object locations and classes in real time.

For each detected hazard, the model predicts a bounding box defined as:

$$B = (x, y, w, h) \quad (4)$$

where:

- B represents the predicted bounding box
- x represents the horizontal coordinate of the bounding box center
- y represents the vertical coordinate of the bounding box center
- w represents the width of the bounding box

- h represents the height of the bounding box

In addition to bounding box coordinates, the model predicts a confidence score:

$$C = P(\text{Object}) \times \text{IoU} \quad (5)$$

where:

- C represents the confidence score
- $P(\text{Object})$ represents the probability that an object exists in the bounding box
- IoU represents the Intersection over Union between predicted and ground truth boxes

The Intersection over Union metric is calculated as:

$$\text{IoU} = \frac{\text{Area}(B_{\text{pred}} \cap B_{\text{gt}})}{\text{Area}(B_{\text{pred}} \cup B_{\text{gt}})} \quad (6)$$

where:

- B_{pred} represents the predicted bounding box
- B_{gt} represents the ground truth bounding box

A higher IoU value indicates better overlap between predicted and actual object regions.

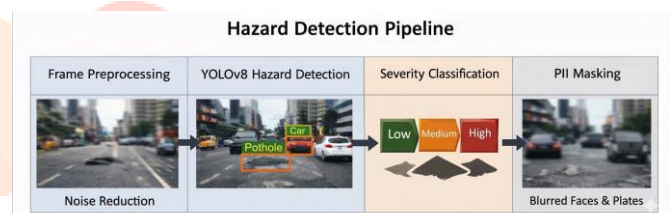


Fig. 2. Edge-to-Cloud Architecture of TerraVision-Pothole

D. Severity Classification

Once a pothole is detected, the severity level is estimated based on the spatial dimensions of the detected region. The size of the pothole can be approximated using the area of the bounding box.

$$A_b = w \times h \quad (7)$$

where:

- A_b represents the area of the bounding box
- w represents the bounding box width
- h represents the bounding box height

Two threshold values T_1 and T_2 are used to categorize potholes into severity levels:

- Low Severity: $A_b < T_1$
- Medium Severity: $T_1 \leq A_b < T_2$
- High Severity: $A_b \geq T_2$

where:

- T_1 represents the threshold for low severity potholes
- T_2 represents the threshold for high severity potholes

This classification enables transportation authorities to prioritize road repairs based on the severity of detected hazards.

V. CONCLUSION

This research presented TerraVision-Pothole, an intelligent road hazard detection framework based on computer vision and edge computing technologies. The system automatically detects road hazards using YOLOv8 and visualizes them through a geo-tagged dashboard.

Experimental results demonstrate reliable detection performance with approximately 85% accuracy. The integration of edge processing and cloud analytics enables real-time hazard monitoring and improved road safety.

Future work may focus on improving detection accuracy under adverse weather conditions and integrating predictive analytics for infrastructure maintenance.

REFERENCES

- [1] Redmon, J., Divvala, S., Girshick, R., and Farhadi, A. (2016). You Only Look Once: Unified Real-Time Object Detection. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).
- [2] Bochkovskiy, A., Wang, C. Y., and Liao, H. Y. M. (2020). YOLOv4: Optimal Speed and Accuracy of Object Detection. arXiv:2004.10934.
- [3] Jocher, G., Chaurasia, A., and Qiu, J. (2023). YOLOv8: Ultralytics Object Detection Model. Ultralytics Documentation.
- [4] Bradski, G. (2000). The OpenCV Library. Dr. Dobb's Journal of Software Tools.
- [5] Krizhevsky, A., Sutskever, I., and Hinton, G. (2012). ImageNet Classification with Deep Convolutional Neural Networks. NIPS.
- [6] Zhang, L., Yang, F., Zhang, Y. D., and Zhu, Y. J. (2019). Road Crack Detection Using Deep Convolutional Neural Network. IEEE Access.
- [7] Maeda, H., Sekimoto, Y., Seto, T., Kashiyama, T., and Omata, H. (2018). Road Damage Detection and Classification Using Deep Neural Networks with Smartphone Images.
- [8] Ren, S., He, K., Girshick, R., and Sun, J. (2015). Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks.
- [9] He, K., Zhang, X., Ren, S., and Sun, J. (2016). Deep Residual Learning for Image Recognition.
- [10] Shi, W., Cao, J., Zhang, Q., Li, Y., and Xu, L. (2016). Edge Computing: Vision and Challenges. IEEE Internet of Things Journal.

Severity Classification Levels



Fig. 3. Severity Classification Model

E. Geo-Tagged Alert Generation

After detection and severity estimation, the system generates geo-tagged hazard alerts. Each detected hazard is associated with location coordinates obtained from the GPS module integrated with the edge device.

Each hazard event is represented as:

$$H = (L_{lat}, L_{lon}, t, S) \quad (8)$$

where:

- H represents the hazard record
- L_{lat} represents the latitude coordinate
- L_{lon} represents the longitude coordinate
- t represents the detection timestamp
- S represents the severity level

These geo-tagged alerts are transmitted to the cloud server, where they are stored and visualized through an interactive dashboard for monitoring road conditions.



Fig. 4. Geo-Tagged Hazard Dashboard

IV. RESULTS AND DISCUSSION

TABLE I
SYSTEM PERFORMANCE EVALUATION

Model	Accuracy	Recall	Precision	mAP@50
Face Blur	85%	84%	87%	85%
Hazard Detection	85%	80%	89%	89.7%
Vehicle Detection	83%	85%	81%	92.9%
Number Plate Detection	83%	80%	86%	85.8%

The hazard detection model achieved an accuracy of 85% and precision of 89%. The mAP@50 score of 89.7% indicates strong localization performance of the YOLOv8 model.

Privacy protection modules also performed reliably with 85% accuracy for face blurring and 83% for number plate detection.

These results demonstrate that the proposed system can detect hazards effectively while preserving user privacy.