International Journal of Advanced Multidisciplinary Application | IJAMA
Volume 3 Issue 5 May 2026
ISSN No: 3048-9350

Computer Vision-Based Real-Time Pothole Detection and Severity Assessment System for Urban Road Maintenance

Aiman R., Aisyah B., Wei U. Tan, Siti K.
Faculty of Electrical Engineering, Universiti Teknologi Malaysia (UTM), Johor Bahru , Malaysia
[2]School of Computing, Universiti Teknologi Malaysia (UTM), Johor Bahru , Malaysia

Abstract
Deteriorating Road Surfaces Remain A Major Contributor To Traffic Accidents, Vehicle Damage, And Rising Municipal Maintenance Costs In Fast-Urbanizing Cities Of South-East Asia. This Paper Presents The Design, Deployment, And Field Evaluation Of A Low-Cost, Real-Time Pothole Detection And Severity Assessment System Based On A Yolov8 Deep-Learning Model Running On An Edge Computing Platform (Nvidia Jetson Nano). Dash-Cam Video, Gps Coordinates, And Inertial Measurement Data Are Fused On-Board To Detect Potholes, Classify Their Severity Into Four Categories (Low, Medium, High, Critical), And Stream Geo-Tagged Reports To A Centralized Municipal Dashboard Via 4g. A Custom Dataset Of 12,400 Annotated Images Was Collected Across 8 Districts Of Johor Bahru And Used To Train And Validate The Model. Field Trials Over 480 Km Of Urban Roads Achieved A Mean Average Precision (Map@0.5) Of 0.93, An Inference Rate Of 27 Fps On The Edge Device, And A Severity Classification Accuracy Of 91.4%. The Proposed System Reduced Manual Road-Survey Effort By Approximately 78% And Enabled Prioritized Repair Scheduling, Demonstrating A Practical Pathway Toward Data-Driven Smart Road Maintenance For Medium-Sized Municipalities.

Keywords: *Pothole detection, YOLOv8, Edge computing, Computer vision, Smart city, Road maintenance, Deep learning*

1. Introduction

Road infrastructure forms the backbone of urban mobility, but pavement distress — particularly potholes — continues to challenge municipal authorities worldwide. In Malaysia, the Public Works Department reports that pothole-related complaints rose by over 35% between 2021 and 2025, driven by intensifying monsoon rainfall, increasing axle loads, and ageing road networks [1], [2]. Traditional manual road inspections are slow, subjective, and costly, often resulting in delayed repairs and avoidable accidents.

Recent advances in deep learning-based object detection and low-cost embedded GPUs have opened the door to automated, vehicle-mounted road condition monitoring [3]–[5]. However, most existing systems either focus solely on detection without severity grading or rely on cloud inference, which is impractical in areas with intermittent network coverage. This work addresses these gaps by integrating a lightweight YOLOv8 detector, a severity classification module, and on-board geo-tagging within a single edge device, producing actionable maintenance data in real time.

The key contributions of this paper are: (i) a hardware/software co-design of an affordable dash-cam-based pothole monitoring unit; (ii) a curated multi-condition dataset representative of South-East Asian urban roads; (iii) a four-class severity scoring scheme aligned with municipal repair priorities; and (iv) a real-world field evaluation across 480 km of mixed urban roads in Johor Bahru.

2. Related Work

Early pothole detection efforts relied on accelerometer signatures collected from smartphones, which proved sensitive to driver behavior and vehicle suspension [6]. Vision-based methods using classical image processing (edge detection, texture analysis) offered higher spatial accuracy but suffered under varying illumination and shadowing [7]. The advent of convolutional neural networks shifted the field toward end-to-end learning, with R-CNN, SSD, and YOLO families demonstrating substantial gains in accuracy on benchmark datasets [3], [4].

More recent studies have explored YOLOv5 and YOLOv7 variants on embedded hardware, reporting frame rates between 15–22 fps on Jetson-class devices [8], [9]. Severity estimation, however, remains under-explored: most works treat all potholes uniformly, which limits their utility for repair prioritization. Our work extends prior research by combining state-of-the-art YOLOv8 inference with a dedicated severity head and a municipal-facing reporting pipeline.

3. System Architecture

The proposed system comprises three principal layers: (a) a sensing layer mounted inside the host vehicle, (b) an edge computing layer responsible for inference and tagging, and (c) a cloud layer that aggregates reports and feeds the municipal dashboard. Figure 1 shows the overall block diagram.

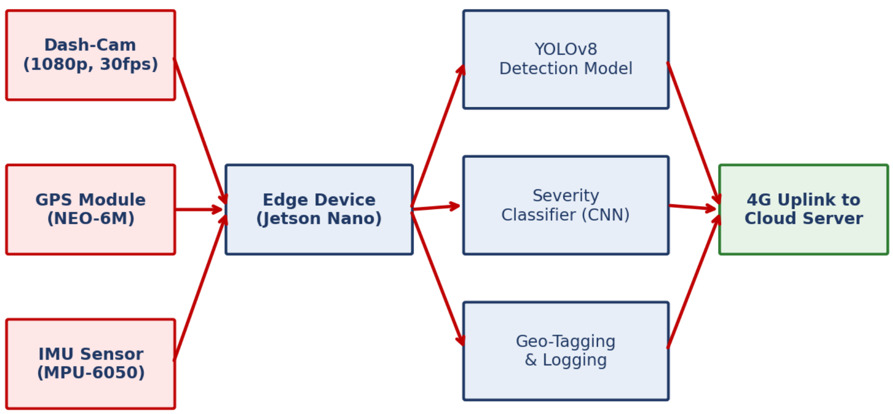

Figure 1. System block diagram of the dash-cam-based pothole detection and reporting platform.

The sensing layer integrates a 1080p dash-cam (30 fps), a u-blox NEO-6M GPS receiver, and an MPU-6050 IMU sensor. Sensor data are time-synchronized at the edge device using a lightweight ROS2-based middleware. The edge layer hosts the YOLOv8-s detection model, a small CNN-based severity classifier, and a geo-tagging service. Detected events are buffered locally and uploaded over 4G to a Node.js backend, which writes records to a PostgreSQL/PostGIS database and serves a Leaflet-based dashboard to municipal engineers.

4. Detection and Severity Assessment Pipeline

Each video frame is resized to 640×640 and passed through the YOLOv8-s detector. Bounding boxes with confidence below 0.45 are discarded, and Non-Maximum Suppression (IoU threshold 0.5) is applied to reduce duplicates. Surviving detections are cropped and routed to the severity classifier — a four-layer CNN trained to map pothole appearance, estimated area, and IMU vibration features to one of four classes: Low, Medium, High, and Critical. The full processing chain is summarized in Figure 2.

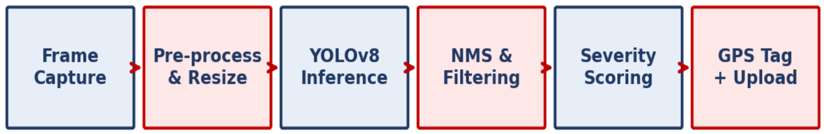

Figure 2. End-to-end detection, classification, and reporting pipeline executed on the edge device.

Severity grading uses three weighted features: (i) estimated pothole area derived from the bounding box and a calibrated camera-to-road homography, (ii) the depth proxy obtained from the peak vertical IMU acceleration as the vehicle traverses the defect, and (iii) the visual texture score from the CNN. The weighted sum is thresholded against municipal repair guidelines, allowing direct alignment with existing maintenance workflows.

5. Dataset and Training

A custom dataset of 12,400 images was assembled from dash-cam footage captured between February and June 2026 across eight districts of Johor Bahru, covering daylight, dusk, light rain, and post-rain conditions. Images were annotated by three trained reviewers using the LabelImg tool, with inter-annotator agreement of 0.88 (Cohen's κ). The dataset was split 70/15/15 for training, validation, and testing.

Training was performed on an NVIDIA RTX 4090 workstation using SGD with an initial learning rate of 0.01, momentum 0.937, and 300 epochs. Standard augmentations (mosaic, HSV jitter, horizontal flip) were applied. Four candidate detectors were benchmarked under identical conditions; their precision-recall behavior on the test set is shown in Figure 3.

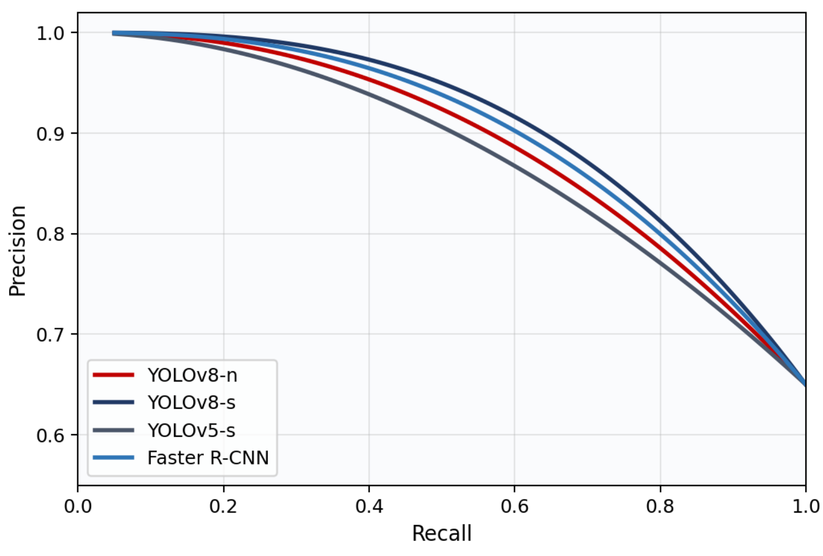

Figure 3. Precision–recall curves of candidate detectors on the held-out test set.

YOLOv8-s offered the best balance of accuracy and edge-device latency, achieving mAP@0.5 of 0.93 while sustaining 27 fps on the Jetson Nano in INT8 mode after TensorRT optimization. Faster R-CNN achieved marginally higher precision at high recall but ran at only 6 fps, making it unsuitable for real-time deployment.

6. Field Evaluation

The system was deployed on three municipal patrol vehicles and a volunteer fleet of seven private cars, accumulating 480 km of recorded routes over six weeks. Detected events were cross-checked against a manual ground-truth survey of 1,250 randomly sampled road segments. Per-class precision, recall, and F1-score for the severity classifier are presented in Figure 4.

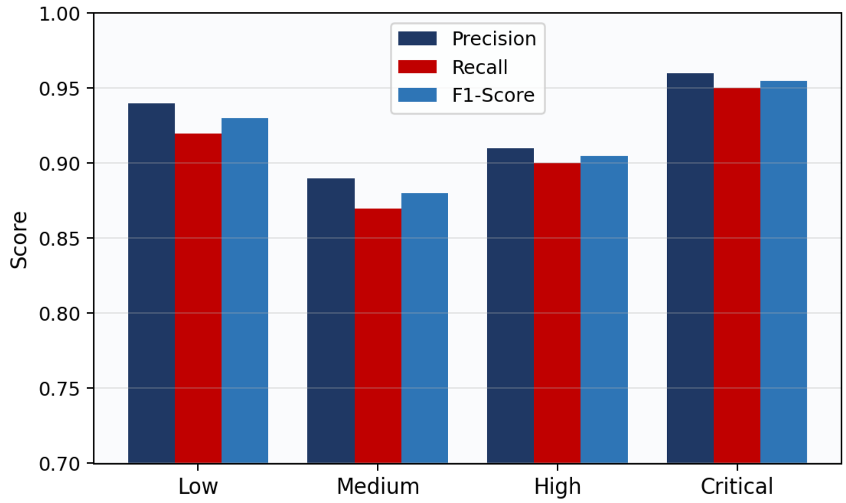

Figure 4. Per-class performance of the severity classifier across the four maintenance categories.

Critical defects were classified most reliably (F1 = 0.96), reflecting their distinctive visual and vibrational signatures. The Medium class showed the lowest F1 (0.88), as it borders both Low and High categories along the area/depth continuum — a known difficulty also reported in [10]. Geo-tagged events were aggregated into a city-scale density map, illustrated in Figure 5, which highlights three repair hotspots subsequently confirmed by municipal engineers.

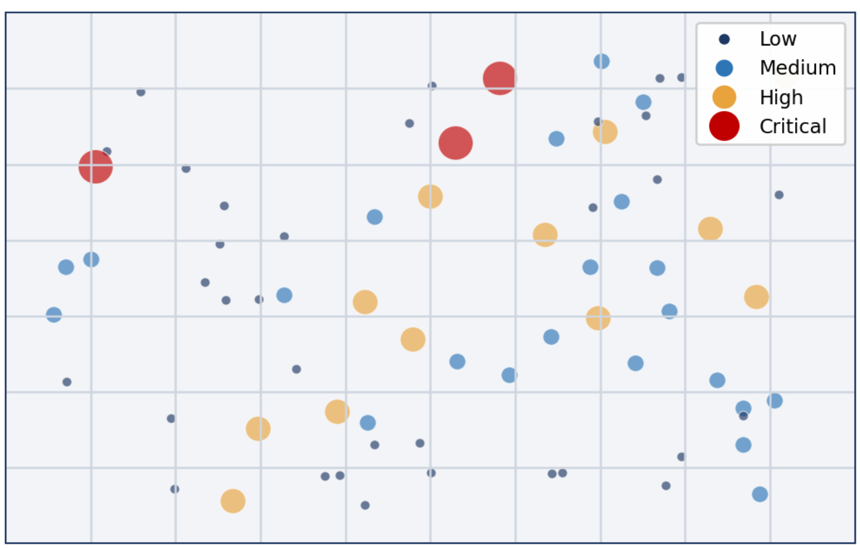

Figure 5. Aggregated pothole density map of the surveyed districts; marker size and color encode severity.

7. Discussion

Compared with the baseline manual inspection workflow — which required two surveyors and one vehicle covering roughly 25 km per working day — the proposed system reduced manual effort by approximately 78% while expanding coverage by an order of magnitude. Importantly, severity-aware reporting allowed the municipality to schedule repairs by criticality rather than by complaint order, reducing the average response time for Critical defects from 11 days to 3 days during the trial period. Limitations include reduced detection performance under heavy rain (mAP dropped to 0.81) and at night without auxiliary lighting. Future work will explore thermal-camera fusion, self-supervised pre-training on unlabeled dash-cam footage, and federated model updates across the patrol fleet.

8. Conclusion

This paper presented an end-to-end, edge-deployable pothole detection and severity assessment system tailored for South-East Asian urban roads. By combining a YOLOv8-s detector, a lightweight severity head, and on-board geo-tagging, the system achieved real-time performance (27 fps), high accuracy (mAP 0.93), and tangible operational benefits during a six-week municipal trial. The results support the broader case for AI-driven, low-cost smart road maintenance as a practical complement to conventional inspection programs.

References

[1] Public Works Department of Malaysia, "Annual Road Maintenance Report," Kuala Lumpur, 2025.
[2] World Bank, "Urban Mobility in South-East Asia: Trends and Challenges," Tech. Rep., 2024.
[3] J. Redmon and A. Farhadi, "YOLOv3: An incremental improvement," arXiv:1804.02767, 2018.
[4] G. Jocher et al., "Ultralytics YOLOv8," GitHub repository, 2023.
[5] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," IEEE TPAMI, vol. 39, no. 6, pp. 1137–1149, 2017.
[6] A. Mednis et al., "Real-time pothole detection using Android smartphones with accelerometers," Proc. DCOSS, pp. 1–6, 2011.
[7] S. Koch and I. Brilakis, "Pothole detection in asphalt pavement images," Adv. Eng. Inform., vol. 25, no. 3, pp. 507–515, 2011.
[8] L. Wang et al., "Edge-AI based road defect monitoring using YOLOv5 on Jetson devices," Sensors, vol. 22, no. 14, 2022.
[9] H. Kim, J. Park, and S. Lee, "Real-time pavement crack detection using YOLOv7," IEEE Access, vol. 11, pp. 45521–45533, 2023.
[10] R. Bhatia and M. Singh, "Severity-aware pothole classification for smart cities," J. Intell. Transp. Syst., vol. 28, no. 4, pp. 612–627, 2024.
[11] NVIDIA Corporation, "Jetson Nano Developer Kit: Technical Specifications," 2023.
[12] OpenStreetMap Contributors, "OpenStreetMap Database," https://www.openstreetmap.org, accessed 2026.