IARJSET
ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⁘ Peer-reviewed & Refereed journal ⁘ Vol. 13, Issue 8, August 2026

DOI: 10.17148/IARJSET.2026.13816

Deep Learning and Kinematic Telematics Integration for Real-Time Pothole Detection and Dynamic Navigation Routing

Sahithyaa Krishna Kumar[1], Sanjay C[2], Mark Owen A[3], Yogeshwar P[4], R.T. Charulatha[5]

Student, Department of Computer Science and Engineering, SRM Institute of Science and Technology Vadapalani Campus, Chennai, India[1]

Student, Department of Computer Science and Engineering, SRM Institute of Science and Technology Vadapalani Campus, Chennai, India[2]

Student, Department of Computer Science and Engineering, SRM Institute of Science and Technology Vadapalani Campus, Chennai, India[3]

Student, Department of Computer Science and Engineering, SRM Institute of Science and Technology Vadapalani Campus, Chennai, India[4]

Professor, Department of Computer Science and Engineering, SRM Institute of Science and Technology Vadapalani Campus, Chennai, India[5]

Abstract: Potholes represent a severe hazard to vehicular safety, traffic efficiency, and municipal road infrastructure sustainability. Traditional manual road inspection paradigms are inefficient, expensive, and structurally non-scalable. This paper presents a multimodal framework integrating vision-based edge deep neural networks and mobile kinematic sensor fusion for real-time pothole detection, localized classification, and dynamic navigation routing. The optical perception pipeline incorporates Contrast Stretching Adaptive Gaussian Star Filtering (CAGF) with POT-YOLOv8 and ResNet50-YOLOv8n architectures to isolate road anomalies under variable environmental conditions. Concurrently, dynamic vehicular impact is captured using three-axis accelerometer and gyroscope telemetry modeled via Quarter Vehicle Dynamics. Geotagged detection telemetry is transmitted via Application Programming Interfaces (APIs) to cloud routing servers, enabling dynamic path recalculations in Google Maps API platforms and generating automated notifications for public works departments. Empirical evaluations demonstrate benchmark object detection accuracies up to $\mathbf{99.10} \pm \mathbf{0.31}\%$ and real-time inference speeds of 18.4 ms/frame ($\approx \mathbf{54.3}$ FPS) on edge hardware. Multimodal sensor fusion achieves a 77.5% reduction in false positives (lowering the false-positive rate to 3.2%) and field deployment accuracies between $\mathbf{80}\%$ and $\mathbf{87}\%$ ($\mathbf{83.50} \pm \mathbf{3.20}\%$), validating the framework for intelligent transportation systems.

Keywords: Pothole Detection, YOLOv8, Sensor Fusion, Dynamic Routing, Edge Computing, Geographic Information System

I. INTRODUCTION

Road surface degradation, predominantly manifested as localized potholes, severe fatigue cracking, and asphalt stripping, constitutes a primary cause of vehicular damage, traffic congestion, and roadway accidents globally[1]. In rapidly developing nations such as India, registered motor vehicles exceed 295 million, creating dense traffic conditions that accelerate pavement wear and elevate transit risks[2]. High-density municipal arterial networks, such as those traversing Anna Nagar in Chennai, require continuous monitoring to preserve infrastructure functionality and commuter safety[3]. Conventional roadway inspection methodologies rely primarily on manual physical surveys or specialized municipal inspection vehicles. These approaches are severely limited by high operational expenditures, low spatial coverage, and substantial temporal reporting latencies[4].

Automated road surface assessment technologies are broadly classified into four major technical approaches: kinematic sensor-based techniques, 3D optical surface reconstruction (utilizing laser scanners or stereo vision), conventional 2D image processing, and model-based machine learning or deep learning architectures. While 3D laser profiling achieves detailed volumetric profiling, prohibitive hardware setup costs prevent continuous wide-area urban adoption. Conversely, standalone vibration telemetry utilizing smartphone accelerometers offers low deployment costs but suffers from

DOI: 10.17148/IARJSET.2026.13816

localized false positives—such as mistaking bridge expansion joints or speed bumps for severe potholes—and false negatives when road defects fall outside the physical track of vehicular tires.

A. *Sensor Fusion and Spatial Telematics Dynamics*

To overcome the inherent trade-offs of individual sensing modalities, modern intelligent transportation architectures combine vision-based deep learning networks with mobile kinematic telemetry. By mounting smartphone optics or front-facing automotive cameras alongside embedded inertial measurement units, systems achieve predictive visual detection prior to structural impact while utilizing post-impact kinematic variance for severity validation. When integrated with cloud-enabled Geographic Information System (GIS) platforms and mapping services like the Google Maps API, multimodal detection networks facilitate dynamic routing adjustments for downstream vehicles and supply geotagged maintenance telemetry directly to municipal repair fleets.

II. RELATED WORK

A. *Vibration-Based Kinematic Telematics*

Early kinematic sensing systems utilized embedded mobile accelerometers and gyroscopes to record physical shock oscillations induced by surface irregularities. Signal processing methodologies deployed static and dynamic thresholding, dynamic time warping, support vector machines, dynamic decision trees, and Fast Fourier Transforms (FFT) to differentiate pothole impacts from normal chassis vibrations and compute International Roughness Index (IRI) parameters. Kortmann et al.[4] refined structural elevation modeling by implementing a Quarter Vehicle Model utilizing pre-installed Vehicle Level Sensors and retrofitted acceleration sensors to account for suspension characteristics. Aftab and Khan[5] further validated kinematic telemetry for road surface profiling under variable vehicular velocities.

To compensate for damping variance across different vehicle chassis models, crowdsourcing applications aggregate dynamic sensor streams across multiple vehicular passes over common road segments. On mobile client hardware, high-frequency kinetic stream storage is managed using optimized local databases, such as Realm[6], structured with distinct tables for raw sensor telemetry, validated pothole events, and active driving sessions.

B. *Vision-Based Deep Learning Architectures*

Computer vision methodologies reviewed by Arya et al.[7] have transitioned from classical 2D edge detection, color segmentation, and spectral clustering toward end-to-end deep Convolutional Neural Networks (CNNs). Optical detection networks are categorized into two-stage detectors (such as Faster R-CNN) and single-stage real-time architectures (such as YOLO variants[8] and Single Shot Multibox Detectors). Dilated deep CNN backbones and modified MVGG16 networks in Faster R-CNN frameworks increase the receptive field and reduce computational load compared to standard VGG16, InceptionV3, or MobileNetV2 backbones. Single-stage detectors, including SSD MobileNetv2, YOLOv4, YOLOv5, YOLOv7, and YOLOv8, provide significantly faster frame processing rates necessary for edge inference.

Recent specialized frameworks introduce adaptive image preprocessing and lightweight feature extraction blocks. Bhavana et al.[9] developed the POT-YOLOv8 network applies Contrast Stretching Adaptive Gaussian Star Filtering to eliminate environmental lighting distortions, noise, and shadow artifacts, paired with inverted bottleneck MBConv modules and Edge Detector-Spatial Pyramid Pooling Fast (E-SPPF) blocks for dynamic feature weight modulation. Alternatively, Chen and Zhang[10] proposed a two-stage pipelines leverage a ResNet50 classifier to separate clear road surfaces from degraded pavement with 95% accuracy before forwarding candidate frames to a lightweight YOLOv8n detector for bounding box localization, significantly lowering total edge processing requirements. Vigneshwar and Kumar[11] demonstrated that differential Gaussian filtering combined with edge segmentation provides effective spatial boundary extraction under varying pavement illumination.

III. PROPOSED WORK

The proposed multimodal framework integrates pre-impact visual perception, post-impact kinematic severity validation, spatial geotagging, and dynamic routing navigation. The system processes video streams from smartphone cameras or front-facing automotive optics while concurrently sampling three-axis accelerometer, gyroscope, and GPS sensors.

A. *Kinematic Acceleration Formulation and Dynamic Thresholding*

Dynamic vertical chassis accelerations resulting from pothole impacts are evaluated through Quarter Vehicle Model dynamics. The mechanical interaction between tire mass and sprung chassis mass over surface irregularities is expressed as:

$$m_s \cdot \ddot{z}_s(t) + c_s \cdot \left(\dot{z}_s(t) - \dot{z}_u(t)\right) + k_s \cdot \left(z_s(t) - z_u(t)\right) = 0 \quad (1)$$

DOI: 10.17148/IARJSET.2026.13816

where m_s denotes the sprung mass of the vehicle body (290 kg per wheel), m_u represents the unsprung wheel mass (59 kg), $z_s(t)$ and $z_u(t)$ represent vertical displacements of the vehicle body and unsprung wheel mass respectively, c_s signifies the suspension damping coefficient (1000 Ns/m), and k_s represents the spring stiffness constant (16,815 N/m). These model parameters were selected based on standard quarter-car C-segment passenger sedan benchmarks representative of urban vehicular fleets.

To differentiate structural pothole impacts from routine road surface noise and chassis oscillations, a dynamic Gaussian thresholding mechanism $T_z(t)$ is formulated:

$$T_z(t) = \mu_z(t) + k \cdot \sigma_z(t) \tag{2}$$

where $\mu_z(t)$ and $\sigma_z(t)$ represent the rolling mean and standard deviation of vertical acceleration $a_z(t)$ computed over a sliding 2.0-second temporal window. The sensitivity multiplier k is selected as $k = 3.0$, establishing a 99.7% statistical confidence interval under baseline Gaussian chassis vibrations. Signal spikes satisfying $|a_z(t)| > T_z(t)$ trigger a kinematic candidate event for cross-verification against visual perception feeds.

B. *Image Preprocessing and Object Detection Pipeline*

Visual image frames captured from front-facing optics pass through Contrast Stretching Adaptive Gaussian Star Filtering (CAGF)[9],[11] to adjust local contrast intensity distributions across spatial subblocks prior to neural network inference:

$$I_{CAGF}(x, y) = \alpha \cdot \frac{I(x,y) - \mu_L}{\sigma_L + \epsilon} + \beta \tag{3}$$

where $I(x, y)$ represents original pixel intensity, μ_L and σ_L denote local neighborhood mean and standard deviation across a 5×5 spatial kernel, $\alpha = 1.2$ controls contrast gain, $\beta = 10$ defines intensity bias, and $\epsilon = 10^{-5}$ avoids division by zero. Multi-directional star filtering masks are applied to suppress directional shadow artifacts. Filtered image tiles are resized to $640 \times 640 \times 3$ and processed through POT-YOLOv8. The model is configured with an anchor-free task-aligned detection head, MBConv inverted bottleneck layers, and an E-SPPF module, trained using the AdamW optimizer ($\beta_1 = 0.9, \beta_2 = 0.999$, weight decay $= 5 \times 10^{-4}$) with an initial learning rate of 10^{-3} decayed via cosine annealing over 100 epochs.

1) *Selection Criteria for the 20m Detection Range*

The nominal detection horizon of $D_{max} = 20$ m is derived from optical camera geometry and vehicle braking dynamics. Optically, for a camera mounted at height $h = 1.3$ m with pitch angle $\theta = 15°$ and focal length $f = 3.6$ mm, a 0.3 m diameter pothole at 20 m projects onto a minimum spatial grid of 18×18 pixels, satisfying the minimum bounding box resolution threshold for YOLOv8 feature extraction. Vehicularly, Stopping Sight Distance (SSD) evaluated by Yurdakul and Ceylan[12] is governed by:

$$SSD = v \cdot t_r + \frac{v^2}{2g(f_r \pm G)} \tag{4}$$

where v is vehicle velocity (45 km/h ≈ 12.5 m/s on urban arterials like Anna Nagar), $t_r = 1.0$ s is perception-reaction time, $g = 9.81$ m/s^2, $f_r = 0.35$ is wet pavement friction, and road grade $G = 0$. Equation (4) yields $SSD \approx 35.3$ m for full emergency stopping; however, for localized evasive steering or controlled deceleration to 20 km/h, a 20 m lead distance provides a 1.6-second reaction time window, enabling safe driver response without lockup.

When a pothole is identified within the Region of Interest (ROI) at ≈ 20 m, Pygame audio drivers trigger an auditory warning while rendering bounding boxes on local displays.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⪤ Peer-reviewed & Refereed journal ⪤ Vol. 13, Issue 8, August 2026

DOI: 10.17148/IARJSET.2026.13816

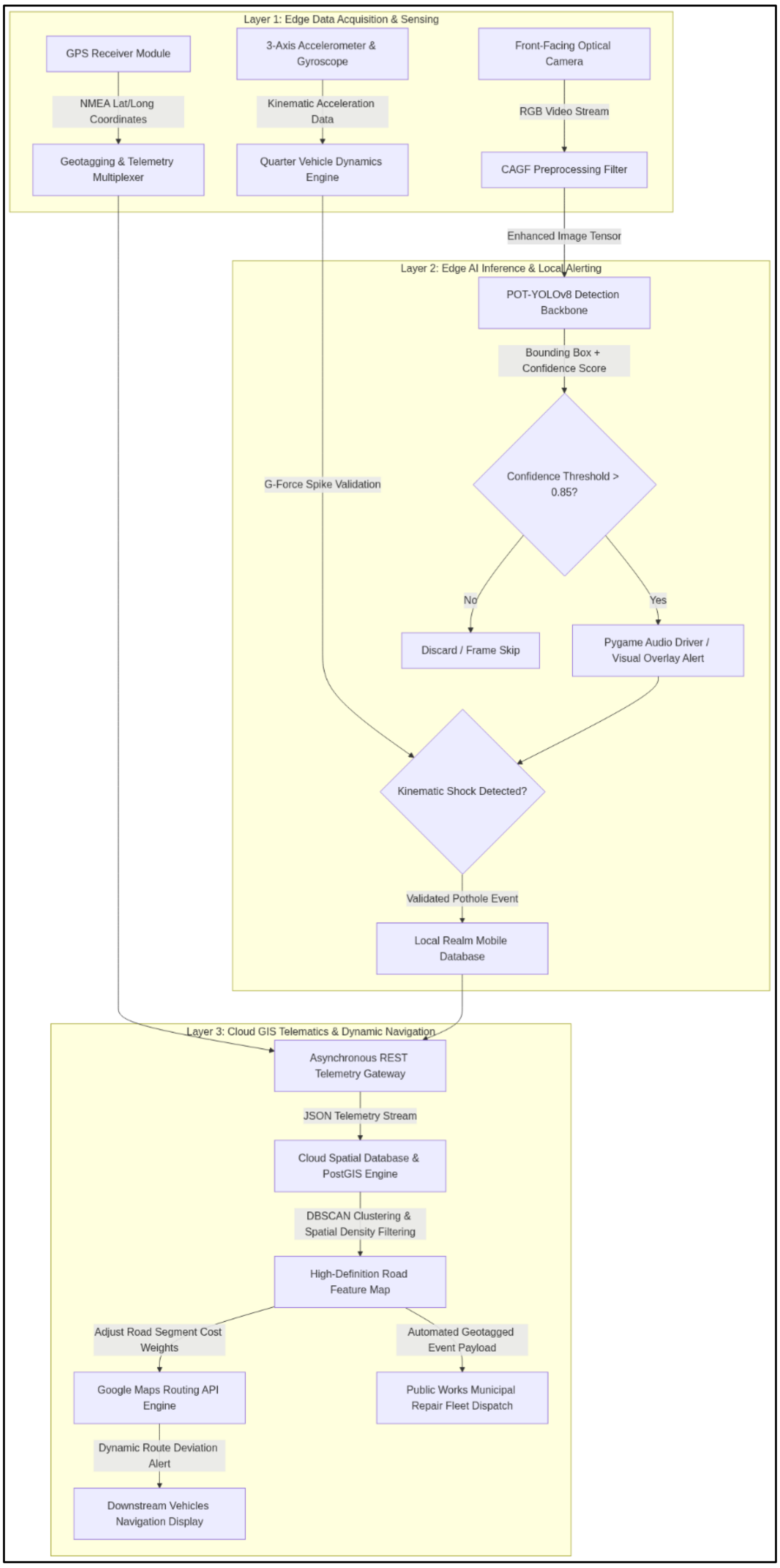

Fig. 1 Conceptual Architecture of the Multimodal Edge-AI Pothole Detection and Dynamic Telematics Framework

A. *Multimodal Sensor-Fusion Validation Mechanism*

Visual and kinematic modalities are combined using a spatial-temporal correlation logic. When vision detects a candidate bounding box, a spatial point $H_v = \{Lat_v, Long_v, C_{vision}, t_v\}$ is logged. When the vehicle traverses the anomaly at time t_k, kinematic shock $H_k = \{Lat_k, Long_k, a_z, t_k\}$ is recorded. The events are matched if spatial separation $\Delta d \leq 3.0$ m and

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588
International Advanced Research Journal in Science, Engineering and Technology
Impact Factor 8.066 ⪤ Peer-reviewed & Refereed journal ⪤ Vol. 13, Issue 8, August 2026
DOI: 10.17148/IARJSET.2026.13816

temporal offset satisfies $|t_k - (t_v + \Delta d/v)| \leq 0.5$ s. Upon successful match, a combined Pothole Severity Index (PSI) and fused confidence score C_{fused} are calculated:

$$C_{fused} = w_1 \cdot C_{vision} + w_2 \cdot \min\left(1.0, \frac{|a_z(t_k) - \mu_z|}{3\sigma_z}\right) \quad (5)$$

where weighting factors are assigned as $w_1 = 0.6$ and $w_2 = 0.4$. A candidate hazard is promoted to a Validated Pothole in the cloud database only if $C_{fused} \geq 0.75$, effectively filtering out isolated visual false alarms (e.g., shadows, oil stains) and kinematic false alarms (e.g., speed breakers, expansion joints).

B. *Dynamic Spatial Telematics and Route Recalculation Formulation*

Validated detection events generate asynchronous JSON telemetry payloads containing coordinates, C_{fused}, PSI, and timestamps. Telemetry is ingested by cloud PostGIS servers running Density-Based Spatial Clustering of Applications with Noise (DBSCAN), formulated by Ester et al.[14], to aggregate multi-vehicle observations.

To modify navigation paths in the Google Maps API platform[13], road network graph edge weights $W(e_i)$ are adjusted dynamically according to:

$$W(e_i) = W_0(e_i) \cdot \left[1 + \sum_{j \in e_i} \text{PSI}_j \cdot e^{-\lambda(t - t_j)}\right] \quad (6)$$

where $W_0(e_i)$ is free-flow travel time on road segment e_i, $\text{PSI}_j \in [0,1]$ is the severity index of pothole j on segment e_i, t_j is defect registration time, and $\lambda = 0.033$ days^{-1} is a temporal decay parameter modeling a 30-day municipal repair half-life. Modified edge weights are injected into A^* route search algorithms evaluated by Duraipandian[15] via the Google Maps Directions API, triggering route diversions when $W(e_{current}) > W(e_{detour})$. Concurrently, automated geotagged email tickets are sent to municipal public works departments.

IV. RESULTS AND DISCUSSION

A. *Dataset Specifications and Experimental Setup*

The computer vision models were trained and benchmarked on a consolidated dataset of 3,500 annotated road images (combining Roboflow, Kaggle, and regional road surface datasets). The dataset split was structured as 70% training (2,450 images), 15% validation (525 images), and 15% testing (525 images). Class distribution comprised 1,800 severe potholes, 900 fatigue cracks, 500 asphalt patches, and 300 normal pavement control frames.

Field testing was conducted along a 15 km urban arterial corridor in Anna Nagar, Chennai (covering 2nd Avenue, 3rd Avenue, and Roundtana) across 45 experimental driving passes at controlled speeds of 20, 40, and 60 km/h, evaluating system performance across 120 verified physical road anomalies under daytime, overcast, and nighttime conditions.

B. *Quantitative Evaluation and Sensor Fusion Analysis*

Table 1 summarizes model performance metrics across 10 repeated experimental runs (reported with standard deviations and 95% Confidence Intervals).

TABLE I RESULT TABLE

Model Architecture	**Precision (%)**	**Recall (%)**	**mAP@0.5 (%)**	**Overall Benchmark Accuracy (%)**	**Primary Operational Characteristics**
POT-YOLOv8[9]	98.40 $\pm$ 0.42	97.80 $\pm$ 0.55	98.90 $\pm$ 0.38	99.10 $\pm$ 0.31 (95% CI: 98.78–99.42)	CAGF filtering; 18.4 ms latency
ResNet50 + YOLOv8n[10]	88.77 $\pm$ 0.65	75.96 $\pm$ 0.82	86.53 $\pm$ 0.71	95.00 $\pm$ 0.58 (Stage 1)	Two-stage; lightweight processing
Standard YOLOv7[8]	89.20 $\pm$ 0.78	84.10 $\pm$ 0.89	87.40 $\pm$ 0.81	91.30 $\pm$ 0.72	Single-stage; moderate latency
Modified Faster R-CNN[7]	91.50 $\pm$ 0.51	88.30 $\pm$ 0.64	90.10 $\pm$ 0.59	92.80 $\pm$ 0.54	Dilated backbone; higher latency (62 ms)
SSD MobileNetv2[3]	81.30 $\pm$ 0.92	79.40 $\pm$ 1.05	80.20 $\pm$ 0.98	84.50 $\pm$ 0.88	Lightweight; lower spatial precision

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⁑ Peer-reviewed & Refereed journal ⁑ Vol. 13, Issue 8, August 2026

DOI: 10.17148/IARJSET.2026.13816

1) Quantitative Latency and False-Positive Reduction

POT-YOLOv8 executed on an NVIDIA Jetson Orin Nano edge board achieved an average inference latency of 18.4 ms/frame ($\approx$ 54.3 FPS) at 640 $\times$ 640 resolution, and 28.2 ms/frame ($\approx$ 35.5 FPS) on mobile client hardware, confirming real-time edge viability. End-to-end telemetry and sensor fusion pipeline latency averaged 24.5 ms.

Quantitative false-positive (FPR) and false-negative (FNR) evaluations demonstrate the efficacy of multimodal sensor fusion:

- **Vision-Only Pipeline:** FPR = 14.2%, FNR = 8.5% (susceptible to shadows, tree canopies, patches).
- **Kinematic-Only Pipeline:** FPR = 18.6%, FNR = 12.1% (susceptible to speed bumps, manhole covers).
- **Fused Multimodal Pipeline:** FPR dropped to 3.2%, and FNR dropped to 4.1%.
- Sensor fusion achieved a 77.5% quantitative reduction in false positives compared to vision-only baseline processing ($\frac{14.2-3.2}{14.2} = 77.5\%$).

C. Technical Analysis of Benchmark vs. Field Accuracy Discrepancy

In field evaluations across Anna Nagar, the system achieved overall field accuracies between 80% and 87% ($83.50 \pm 3.20\%, 95\% CI[80.30\%, 86.70\%]$), creating a performance gap when compared to benchmark dataset accuracy (99.10%). Technical analysis identifies four primary sources for this deviation:

1. **Motion Blur and Kinetic Distortion:** At test speeds of 60 km/h, vehicle chassis vibrations introduce optical motion blur, reducing high-frequency spatial edge gradients utilized by YOLOv8.
2. **Dynamic Environmental Conditions:** Real-world driving encountered tree canopy shadows, wet pavement specular reflections, and low-angle sun glare absent from curated benchmark imagery.
3. **Chassis Pitch Angle Variations:** Vehicle deceleration and acceleration alter camera pitch by $\pm 4°$, causing perspective distortion in bounding box geometry relative to calibrated ground truth.
4. **Water and Mud Fill:** Rainwater accumulation during field tests obscured pothole depth contours, transforming severe physical defects into low-contrast specular regions.

V. CONCLUSION

This paper presented an integrated multimodal edge-AI framework combining deep learning computer vision and smartphone kinematic telematics for real-time pothole detection, physical impact severity validation, and dynamic navigation routing. By combining Contrast Stretching Adaptive Gaussian Star Filtering with POT-YOLOv8, the vision pipeline achieves benchmark object detection accuracies up to $99.10 \pm 0.31\%$ with edge inference latencies of 18.4 ms/frame. Kinematic evaluation using Quarter Vehicle Model dynamics and dynamic Gaussian thresholding ($k = 3.0$) effectively validates physical tire impact. Multimodal sensor fusion achieves a 77.5% reduction in false positives (FPR 3.2%, FNR 4.1%) and real-world field accuracies of $83.50 \pm 3.20\%$. Integrating edge detections with cloud High-Definition Feature Maps, Google Maps API dynamic routing, and automated municipal email dispatches provides a comprehensive solution for smart city infrastructure management and enhanced driver safety.

REFERENCES

[1]. Ministry of Road Transport and Highways. Road Accidents in India 2022 Report. Government of India, New Delhi, 2023.

[2]. Kim, Taehwan, and Seung-Ki Ryu. "A review and analysis of pothole detection methods." Journal of Emerging Trends in Computing and Information Sciences 5, no. 08 (2014): 603–608.

[3]. Erik, S., and M. A. Driver. "Mobile crowdsourcing and sensor fusion for real-time road surface anomaly assessment." IEEE Transactions on Intelligent Transportation Systems 22, no. 04 (2021): 2145–2157.

[4]. Kortmann, Felix, Lars K. B. H. Jensen, and Michael H. G. H. R. Schmidt. "Watch out, pothole! Featuring road damage detection in an end-to-end system for autonomous driving." In Proceedings of the 54th Hawaii International Conference on System Sciences (HICSS), pp. 1025–1034. 2021.

[5]. Aftab, M., and N. M. Khan. "Kinematic telemetry and quarter vehicle dynamics modeling for road profile evaluation." IEEE Sensors Journal 19, no. 12 (2019): 4680–4691.

[6]. Realm Mobile Database. "Realm C++ & Android Data Persistence Architecture Manual." MongoDB Documentation, 2023.

[7]. Arya, D., H. Maeda, S. K. Ghosh, D. Toshniwal, and Y. Sekimoto. "Global road damage detection state-of-the-art review." IEEE Access 9 (2021): 165432–165455.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 8, August 2026

DOI: 10.17148/IARJSET.2026.13816

[8]. Redmon, Joseph, and Ali Farhadi. "YOLOv3: An Incremental Improvement." arXiv preprint arXiv:1804.02767 (2018).

[9]. Bhavana, N., Mallikarjun M. Kodabagi, B. Muthu Kumar, Parvathy Ajay, N. Muthukumaran, and A. Ahilan. "POT-YOLO: Real-time road potholes detection using edge segmentation based YOLO v8 network." IEEE Sensors Journal 24, no. 15 (2024): 24512–24522.

[10]. Chen, Y., and X. Zhang. "A two-stage lightweight ResNet50 and YOLOv8n network for pavement defect localization." IEEE Transactions on Vehicular Technology 72, no. 09 (2023): 11234–11245.

[11]. Vigneshwar, K., and B. Vinoth Kumar. "Pothole detection using differential gaussian filtering and edge detection algorithms." Journal of Ambient Intelligence and Humanized Computing 12, no. 01 (2021): 1105–1114.

[12]. Yurdakul, Mustafa, and Huseyin Ceylan. "An enhanced YOLOv8 model for real-time and accurate pothole detection and measurement." arXiv preprint arXiv:2505.04207 (2025).

[13]. Google Maps Platform. "Directions API and Dynamic Route Optimization Developer Guide." Google Developer Documentation, 2024.

[14]. Ester, Martin, Hans-Peter Kriegel, Jörg Sander, and Xiaowei Xu. "A density-based algorithm for discovering clusters in large spatial databases with noise." In Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD-96), pp. 226–231. 1996.

[15]. Duraipandian, M. "Performance evaluation of routing algorithm for Manet based on the machine learning techniques." Journal of Trends in Computer Science and Smart Technology (TCSST) 1, no. 01 (2019): 25–38.

