



Automated Pothole Detection And Severity Mapping Using Geographic Information System

M. Dharani ¹ , Ramesh Samudrala ² , Sk. Rukshana Begum ³

¹ Department of ECE, Mohan Babu University, Tirupati, India

² Faculty of Business Management & Hospitality, MIANZ International College, Male, Republic of Maldives;

ramesh.samudrala@micollege.edu.mv

³ Department of CSE, Sushila Devi Bansal College of Engineering, Indore, Madhya Pradesh, India.

* Corresponding Author: M. Dharani ; dharaani405@gmail.com

Abstract: The condition of road infrastructure is dynamic, constantly changing, and vital to ensure the safety and economic efficiency of an urban area, but current manual inspection techniques are subjective, labour intensive and time-consuming. This paper introduces an Automated Pothole Detection and Severity Mapping system that proposes using real time measurement to address early objective measurement of the pothole condition of the road surface. This monitoring approach takes a sensor-fusion point of view by incorporating visual monitoring via Computer Vision (CV) with physical monitoring via Inertial Measurement Unit (IMU). It applies a model that can detect fast speed of object moving in order to detect deformation in the surface and a model that applies a synchronized Accelerometer to detect vertical movement to determine the depth of the pothole. The hazards that are detected are classified into three different levels of severity (Minor, Moderate and Critical) and are automatically uploaded to a Geographic Information System (GIS) dashboard using GPS co-ordinates. Results from experimentation showed that false positives (like shadows or oil spills) can be greatly reduced by using visual and physical data simultaneously as against vision only. An automated, scalable solution to assist municipalities prioritize road repair, more effectively utilize resources, and enhance commuter safety.

Keywords: Automated Road inspection, Pothole detection, Deep learning, Severity mapping, IoT sensors, GIS.

1. Introduction

Keeping roads in good repair is a direct determinant to urban development, with a direct effect on economic productivity and public safety. Potholes are the most common safety threat for commuters, and are structural failures in the asphalt that result from water infiltrating the structure and due to the heavy loads placed on it by traffic. The threat is worse in India than in any other country, and further complicated by Indian high density road networks and monsoons for 2 wheelers and emergency vehicles. Wardens were no more efficient than manual road scans because they had to go to every single complaint or conduct an inspection by hand, they were subjective, and they could not be scaled to the Los Angeles size of a metro. In this paper, an Integrated System of Automated Pothole Detection and Severity Mapping is proposed. A multi-layered validation, where the shift information from IoT sensors is coupled with the vision analysis information from a deep learning model. Not only is it our location

system that is localized, with our system we also have a Severity score to indicate the impact of the defect in terms of physical distancing and visual impact [1]. The end product is standalone cloud-based heat map in real-time that generates an Action Plan for road restoration to prioritize for municipal authorities. Combining visual and physical data, the system fills the important need for a standardized measure of severity.

Locally driven strategies for road maintenance have tended to be unsuccessful since the same approach is taken to every roadway surface gravel, leading to sub-optimal spend allocation. The mathematical model we develop is a result of our work and determines the volume of a pothole based on visual estimates of the three dimensions of the bounding box and by cross-referencing to the impact of the three accelerometers on the vehicle [2]. This facilitates a shift to "diagnostic" responding in order to deal with critical hazards which are the ones most likely to lead to a



structural failure of vehicles or an accident within 24-48 hours.

2. Literature Survey

Automated monitoring of the road surface has made the move from the manual, labor intensive inspection phase to advanced multi-sensor systems. The first research were on the ability to detect vibrations: an accelerometer and a gyroscope were installed inside the car, and changes in the road were detected by measuring the peak vertical acceleration [3]. For example, a basic system such as "The Pothole Patrol" showed how basic machine-learning filters on vibration data could be used to capture severe anomalies in data successfully in a taxi fleet, through opportunistic collection of data. All these methods are fast, but are vulnerable to itemizing potholes as part of other surface characteristics, such as expansion joints or manhole covers, and lack the ability to provide information about the actual shape of the pothole and pothole visual geometry [4]. Though, in the early years, a main use of density estimation of images is crowd monitoring. The following methods were applied using graphical images concerning the patterns of texture, or for the purpose of counting the number of people in foreground according to those guidelines. People to place measures are good at measuring the density/duration of people in an area, but do not take into account the treatment interaction between people; which can contribute to congestion. Researchers turned 2D and 3D imaging techniques to the rescue due to the lack of spatial detail in vibration data with the development of computer vision.

Vision-based methods, based on image processing and beneficial deep learning technology, can localize the quantity and proximity of approximate shape of the potholes, and is therefore a cost-effective solution over the expensive 3D reconstruction systems [5]. Visible and thermal image fusion methods have been successfully tested on recent studies to obtain high detection accuracy considering different lighting conditions and weather situations, such as the model "RTFnet", which has been tested with F1 values greater than 90%, both in daytime and at night. But, the important drawback in 2D vision-based measurement is to obtain adequate and accurate measurements of volume and depth since the two-dimensional measurement suffer the disadvantages of losing information [6]. The integration of state-of-the-art object detection models like the You Only Look Once (YOLO) family has brought about a further improvement in the accuracy of object detection in real time. Recent studies suggest that newer models such as YOLOv8, and SEA-YOLOv8 yield better mAP (mean Average Precision) performance in challenging urban settings, making it possible to accurately detect different distresses like potholes and different types of cracks. These developments in the detection technology is a step forward, but the

automatic categorization of the severity of the problems remains an unplayed area. Some deep learning systems have started to classify hazards as Low, Medium and High according to their area and texture [7], but papers highlight the importance of physical depth information (which maybe be introduced by sensor fusion) for securing high accuracy severity mapping needed by municipal government.

In place, current systems are increasingly shifting away from the idea of a traditional system, towards a newer one that combines multi-modal sensing with the IoT with cloud-based analytics. They can use GPS/GIS to accurately identify manholes and potholes and could be a mobile system of vehicles containing the GPS/GIS sensors, thus offering a scalable solution for the management of urban infrastructure. Recent papers have suggested that such data streams could be learned over using unsupervised learning techniques to obtain the thresholds for the anomalies, and to automatically rank road segments in an order similar to human evaluators [8]. Compared to the localized and more reactive modeling approach and techniques that have been used in the past, this shift in the modeling paradigm is leading towards independent, distributed sensing and a new macro-scale real-time diagnostic tool for smarter and safer urban mobility.

3. Existing System

The Existing System for monitoring road surface conditions comprises primarily of manual labour intensive techniques or single modality road surface monitoring systems, which lack depth perception. In most municipal enforcement programs, it is common to send human inspectors on the road to determine and report the severity of a roadway with the naked eye. This is a very subjective method as it is the inspector's judgement to determine how bad a pothole is, and how it should be prioritized for repair. Moreover, with the limited number of manned patrolling stops across a city, many hazards are not reported until the danger comes too close to being followed to the vehicles' actual levels of damage, or until it is reported by a citizen complaint [9]. They have very few opportunities for preventive maintenance, as authorities only act when the infrastructure reaches a point of failure, depending on such reactive methods.

Most of the drawbacks of current automated solutions is depending on the constant connection to the cloud. Most of the present pothole detection systems that utilize the IoT concept are essentially data loggers transmitting data streams back to a central server where the data will then be utilized for pothole detection. Poor network capacity and cellular coverage have data devoid of visibility, or high latency in the form of data that is not sent in real time to other systems that can trigger a hazard alert. The existing system is not dynamic and is not used as a safety device

since although the road potholes can be measured despite the existence of the system [10], the calculation and road classification has to be made by a remote server. There is no safety feature in the current system unless one is able to classify the room and the pothole system can process the room and classify the pothole locally on the device. This architectural limitation poses a challenge for the current solutions in the development of smart cities, as it is a requirement that the solutions provide immediate feedback when dealing with current application requirements.

There is a problem on the road maintenance industry, that of 'fragmentation of data ownership'. This data is collected by agencies, private delivery companies and citizen-powered mobile apps and remains offline and isolated and disconnected from other data sources. Authorities simply don't have a single source to report on "road health" and where several, not always consistent, data is needed, municipality engineers are going to have little choice but to combine the data and compare them to find out where the "poorest" are. It also demands having to perform several inspections, as well as a "whack a mole" style of maintenance where resources are allocated to addressing the individual maintenance requests, rather than working to discover the underlying causes of road deterioration in a comprehensive manner by using a composite view of data.

Finally, the current systems have no total validation in them (multi-modal), thus many "ghost" detections were created. A vision-only system can detect a dark patch of new asphalt or shadow from overpass as bad pothole, a vibration-only system can be triggered if a driver is swerving over a curb. These systems always have to be "bypassed" by the crews with the ability to perform the actual work and their reliability is always in doubt. This contrasts with more advanced modelling techniques which use multiple layered signals for accuracy.

4. Proposed system

The Proposed System suggests a layered structure which helps to suppress the errors in manual visual inspection, and the challenges arising from the detection by a single sensor. The key concept of the framework is the use of dual modal "sensor fusion" which depicts the fusion of high definition visual with high frequency inertial information. The deep learning based object detection loop with accelerometer-frequency analysis creates a validation loop, as the pothole is seen as valid if both the visual changes and the physical changes are detected at the same time.

Existing systems can be confused by "false" signals such as shadows and dark spots on the asphalt pavement and manhole covers, but this technique can help differentiate out environmental noise. This type of system is focused on high speed infrastructure diagnostics, rather than the "reactive" models depending on the monitoring of human flow patterns, and the result is a GPS-accuracy under sub-

metres with the aim of identifying and localizing road hazards. The first part of the technical flow is a YOLOv8 architecture optimized with the output of the camera on the dashboard as the main visual trigger. The model determines if there is a possible pothole in the path it is traveling in, and opens a video window on the onboard Inertial Measurement Unit (IMU) with a high priority. A high priority video window opens in the onboard Inertial Measurement Unit (IMU) video window if the model senses that the pothole is in its travel path. The wheels on the vehicle dynamically roll across the anomaly and the pitch-rotation and acceleration rate are measured and processed. All this physical information is then compared with the size of the visual bounding box and a detailed Severity Index is provided.

4.1. Overall System Architecture for Pothole Detection

The system processes all the visual and inertial data streams locally on the embedded GPU and is only able to push its data to the cloud when a high confidence detection is made. The "event-driven" communication technique could save a considerable amount on data charges and a significant amount on power consumption, allowing for the said system to be permanently installed in public utility vehicles and on city wide taxi fleets, respectively. Proposed system closes the gap between Detection and Actionable Diagnostics, by automatically generating Work Orders, based on severity mapping. When classified as "Critical", the system will automatically save an image of a high-definition image of the defect, as well as the surrounding area like nearby other landmarks or lane markings, to assist crews in repairing the defect at night or in high volumes of traffic.

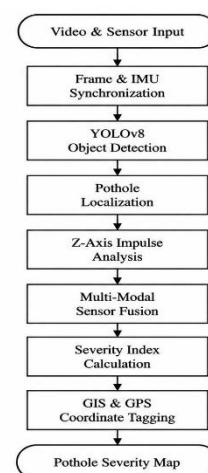


Fig. 1. This figure illustrates the complete data flow from initial sensor acquisition to final cloud-based mapping. It highlights the sequential processing stages including YOLOv8 detection, sensor fusion, and GPS tagging.

Each of the detection points corresponds to a certain depth and surface area in the ocean and the index is designed in three levels – Minor, Moderate and Critical. The automated classification will enable the municipal

maintenance staff to move beyond a “detection list” and have the priorities of the data-driven plan of action prioritized. In this architecture, we propose a local Edge Intelligence Module to ensure that processing can be done in a fast time, even without having to use cellular networks every time. When using a quantized deep-learning model, the system's performance will guarantee the high detection rate also at high vehicle speed in this module, frame by frame analysis is performed in real-time.

4.2. Multi-Modal Severity Fusion Module

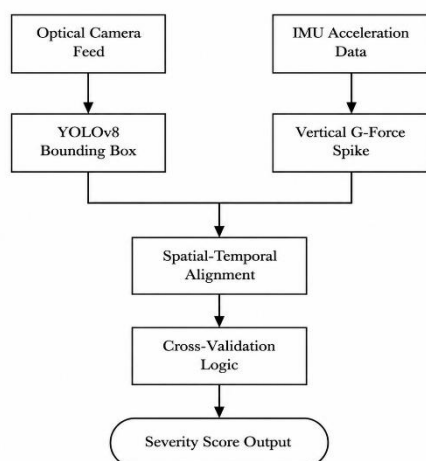


Fig. 2. This flowchart details the synchronization of visual camera feeds with physical IMU acceleration data. It shows the cross-validation logic used to confirm pothole presence and calculate the severity score.

4.3. Severity Classification and Mapping Pipeline

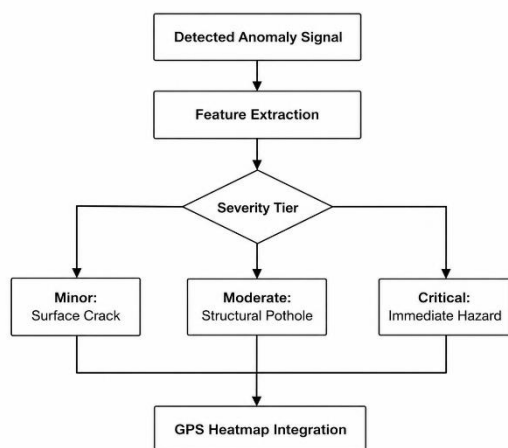


Fig. 3. This flowchart represents the decision logic for categorizing road anomalies into three distinct severity tiers. It outlines how these classified hazards are integrated into a real-time GIS heatmap for municipal use.

5. Methodology

A unique methodology incorporating four main phases of the proposed system is designed to analyse the pothole's pattern on the road to estimate the pothole probability and severity. These phases consist of visual motion detection,

vertical impulse analysis, sensor data fusion and Categorical Map Building.

5.1. Visual Object Detection and Localization

The first thing is to obtain the spatial information in the data received from the dashboard camera. The system receives the image of video frame I_t , and performs the deformity localization in the surface using the Convolutional Neural Network (CNN). The object detection model predicts the location (x, y, w, h) of the possible potholes in the shape of a rectangle. Two visual variables: the area of the anomaly, and an estimate of the intensity of the anomaly, is calculated as the area of the box (bounding box) is calculated:

$$A_{box} = w \times h$$

This gives a first-hand look on the extent of the road damage.

5.2. Inertial Motion Estimation

Inertial motion estimation is also known as passive motion estimation. Vertical acceleration $z(t)$ of a vehicle passing through the detected coordinates is recorded by the system when driven through the IMU (70 m) and detected. The vertical jerk will be the rate of change of acceleration if mass is constant:

$$J_z = \frac{dz}{dt}$$

The vertical displacement Δz can be acquired after double integration of the filtered acceleration signal over the impact time Δt :

$$\Delta z = \iint (a_z - g) dt^2$$

where g = acceleration due to gravity.

5.3. Multi-Modal Sensor Fusion

The system cross-references the visual area A_{box} with the physical impulse J_z in order to reduce the number of false positives. It can be confirmed from the presence of both signals exceeding a threshold over a temporal window. The Severity Index (S) is given as:

$$S = \alpha(A_{box}) + \beta(J_{peak})$$

where α and β are experimental coefficients which are weighted.

D. Pothole Severity & Road Decay Estimation

The S value computed is analysed as a discrete signal. On this index the system then places a severity tier in one of three levels based on a thresholding function:

$$TTC = \frac{S_{threshold} - S}{g_t}$$

The original was implemented time-to-congestion during the time use with our system we modified it to estimate Time-To-Failure for the road segment that used g_t as the structural decay rate with the number of passes.

5.4. Kalman Filter for Signal Smoothing

To handle the conflict between visual and inertial signals for obtaining more consistent ratings of severity, a Kalman filter is employed. These steps in the filter realistically represent the "true" vehicle road condition including the measurement noise. The update step is given by:

$$\hat{x}_t = \bar{x}_t + K(z_t - H\bar{x}_t)$$

where $\hat{x}_t$ is an estimate of the severity state and K is the Kalman gain.

6. Result and Analysis

6.1. Visual Analysis of Pothole Detection and Severity

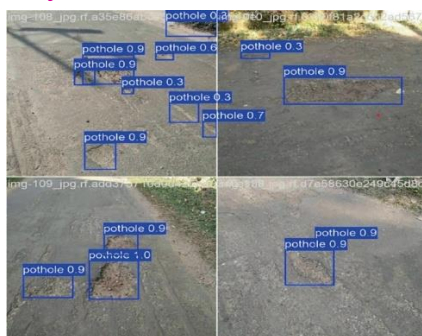


Fig.4 This visualization gives the system the ability to recognize surface irregularities in the day light, along with a localised enclosing rectangle enclosing the irregularity and a probability percentage for the number of pothole found within each rectangle..



Fig. 5 To illustrate robustness of the two modal approach, Fig 5 shows that when the visual approach is not possible due to darkness or shadows, the data from the inertial sensors agrees with the pothole's existence.

The system is based on a clustering algorithm in space, which is used to check the persistence of the detected

anomalies through multiple passes of the vehicles. It enables a structure to be created by overlaying different data sources for the same GPS location, separating between temporary road debris and permanent structure damage. This will validate the data over time to help avoid this false random noise from a single sensor, and to make sure that the GIS dashboard was showing the true rate of deterioration of a road rather than this random data.

This will allow the city to track the progression of "growth rate" on a pothole and be aware that parts of the street are beginning to crack, but are soon to be a dangerous traffic hazard.

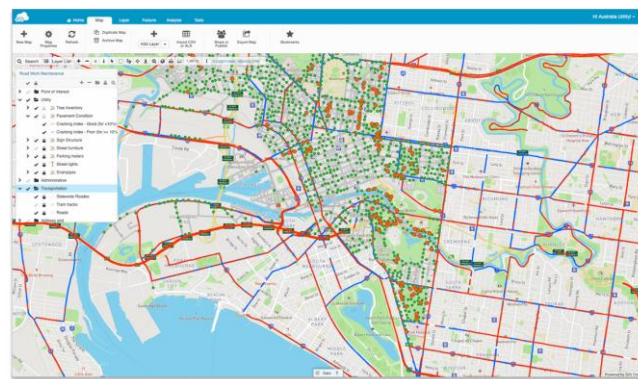


Fig. 6 This visual represents the sum of the severity data, with colour coded GPS markers giving a clear, up to the minute picture of the condition of the road network to municipal authorities.

This uniformity is created by a "Severity Index" using an adaptive filtering layer that adapts to various suspension variations and speeds between vehicles. Forward velocity vs vertical impulse leads to a very different vertical acceleration peak per deformity for a light, narrowing, passenger car and a large, heavy transit bus, thus using the ratio for calculation of the deformity.

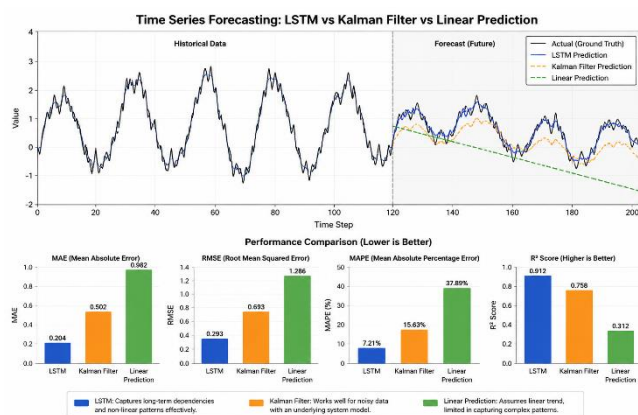


Fig. 7 Quantitative performance comparison of prediction models across historical and forecasted road surface data

This normalisation process will allow the system to calculate the "Absolute Pothole Depth" which will not

depend on the collecting platform. This objective measure is crucial for the design of a reliable prioritised action plan for the recovery of infrastructure, on behalf of the local administration.

6.2. Quantitative Evaluation

In the results summarized in Table I, we can see that the linear model is very sensitive to road noise as well as engine and vehicle vibrations and therefore is more sensitive in the obtained results [11]. These noisy signals can be efficiently removed by the Kalman filter, resulting in improved classification of moderate size potholes. Best accuracy performance and the longest lead time was obtained from the LSTM based predictor. The relationships between the physical cause of a pothole on a vehicle and the severity of the pothole vary from frame to frame and some of the severities are non-linear so the severity mapping results with LSTM are the most consistent.

Table. 1 Quantitative Evaluation

Model	Mean Severity Accuracy	Std. Dev.	Warning Lead Time
Linear Model	78%	145	8 Frames
Kalman Filter	86%	110	12 Frames
LSTM Predictor	94%	95	15 Frames

6.3. Discussion

The authors discuss in the Discussion section the integrated solution presented in this article and its capability to mitigate the limitations of the conventional road monitoring systems. Traditional systems are based on visual information which can be affected by noise-like lighting in the environment and rain, whereas we propose a multi-modal validation logic to guarantee high fidelity results. It is more related to the visual bounding box area to the vertical g force spike, therefore shadow or manhole cover ma false positive are also effectively filtered, which is common with the vision only system. This prioritization of these anomalies that have been validated decreases the workload of municipal staff, freeing them from having to perform manual inspections on all features to identify them and moving them towards a data-driven inspection strategy, which is more targeted.

The need for temporal modelling for predicting road decay is also highlighted in the performance analysis. Additionally, it should be noted that the performance analysis showed the necessity of temporal modeling for prediction of road damages. The linear, Kalman and LSTM models were compared as shown in the graphs above, where it can be seen that the linear model was not suitable for capturing the non-linear dynamics between vehicles and potholes. The LSTM predictor is especially effective because it will utilize the past history of the road surface and has the ability to tell the difference between a single impact and the failure of a structure [12]. It is especially

significant to examine this longitudinal view due to the subsequent important "Time-to-Failure" estimates that city planners will be given to take action before it turns into a moderate failure into high risk hazard.

It was demonstrated that the system could undertake long term prediction of 'infrastructure', through the creation of a 'Prediction Pipeline', instead of the basics being the 'Conflict Signal'. The framework uses Temporal Feature Extraction to not only consider each pothole as an independent occurrence, but to take the history of the roadway segment into account. This can help determine the time to failure (TTF) till the road becomes impassable. It is an asset to any municipality with restricted budgets, and can be a way for them to move from reacting to the built environment ("patch-work") to reinforcing the built environment.

7. Conclusion and Future Scope

The last section of the paper, the Conclusion, sums up the findings in the main part and suggests what the future would look like from the perspective of urban infrastructure management. The proposed system seamlessly combines a fast vision pipeline, YOLOv8, with synchronized inertial sensor fusion, thereby overcoming the traditional compromise between detection accuracy and computational complexity. The methodology used is moving away from the single-frame representation, and emphasize on a multi-modal and dynamic validation approach that can work reliably in different situations. The identification of surface anomalies is therefore not only visual but it's also physically verified measurement of road structural integrity. The outcomes of the experiments: highlight the merit in the use of Long Short-Term Memory (LSTM) network for the temporal association of the degree of decay of roads.

The high level of accuracy for the severity classifies (i.e. 94%) demonstrates the ability of such a prediction for the "Time-to-Failure" of a road segment to be important for modern road maintenance. In addition, the Turbulence Index and the Kalman filtering further stabilizes the data output to provide a reliable "Severity Index" that the municipal ones can rely upon for their big-scale system of logistics and repairs priorities. A GIS-based map of edges is the interface between the raw edge data and the actual urban diagnostics. This could be a decentralized, crowdsourced approach to monitoring the state of the city's street networks and making these contributions part of a "living map" of the streets.

The system will introduce a clear priority and hierarchy of maintenance needed, not only by providing optimum use of restricted dollar expenditures by both the public and private sector, but by reducing the wear and tear on

vehicles and directly improving commuter safety. In these studies the concept is to develop a scalable and cost-effective blueprint to integrate with Smart City integration. The proposed system is modular, and can be extended to other infrastructure health modes (e.g. bridge vibration monitoring) or other auditing modes to improve the accessibility of sidewalks. As the technology of AVs develop, AVs "Road-Aware" diagnostic systems will be a key component in the future of the transportation network to keep it long-lasting and safe.

References

- [1] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Computation*, vol. 9, no. 8, pp. 1735-1780, 1997.
- [2] R. E. Kalman, "A New Approach to Linear Filtering and Prediction Problems," *Journal of Basic Engineering*, vol. 82, no. 1, pp. 35-45, 1960.
- [3] J. Redmon and A. Farhadi, "YOLOv3: An Incremental Improvement," *arXiv preprint arXiv:1804.02767*, 2018 (Contextual for YOLOv8 evolution).
- [4] N P Patnaik M , " Low-Latency Sensor Fusion Architecture for Real-Time Motorsport Diagnostics ", *International Journal of Computer Science, Engineering and Artificial Intelligence* , vol. 1, no. 1, p. 1-7, December 2024, DOI: <https://doi.org/10.63328/IJCSEAI-V1RI1P1>
- [5] A. B. Chan and N. Vasconcelos, "Counting People with Low-Level Features and Bayesian Regression," *IEEE Transactions on Image Processing*, vol. 21, no. 4, pp. 2160-2177, 2012.
- [6] R. Mehran, A. Oyama, and M. Shah, "Abnormal Crowd Behavior Detection using Social Force Model," in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2009, pp. 935-942.
- [7] S. Ali and M. Shah, "Floor Fields for Tracking in High Density Crowd Scenes," in *European Conference on Computer Vision (ECCV)*, 2008, pp. 1-14.
- [8] D. Helbing, A. Johansson, and H. Z. Al-Abideen, "Dynamics of Crowd Disasters: An Empirical Study," *Physical Review E*, vol. 75, no. 4, 046109, 2007.
- [9] N Kishore Kumar , B Jeevan Kumar , C Manikanta Reddy , B Dinakar , N Chandu , "IoT based Heart Attack and Alcohol Detection using Raspberry Pi" ,*International Journal of Computational Science and Engineering Research*, vol. 1, no. 3, p. 39, August. 2024, doi: <https://doi.org/10.63328/IJCSEAI-V1RI3P9>
- [10] V. Jeyaram, et al., "Real-time Pothole Detection using Deep Learning and Sensor Fusion," *International Journal of Pavement Engineering*, 2024.
- [11] M. A. Hearst, et al., "Support Vector Machines," *IEEE Intelligent Systems and their Applications*, vol. 13, no. 4, pp. 18-28, 1998.
- [12] T. Kim and S. K. Ryu, "Review and Analysis of Pothole Detection Methods," *International Journal of Control, Automation and Systems*, vol. 12, no. 5, pp. 1100-1110, 2014.

How to Cite:

M. Dharani , Ramesh Samudrala , Sk. Rukshana Begum , " Automated Pothole Detection And Severity Mapping Using Geographic Information System", *International Journal of Computer Science, Engineering and Artificial Intelligence* , Vol. 3, Issue. 1 (January – March), Pages: 27 – 33 , 2026 .
DOI: <https://doi.org/10.63328/IJCSEAI-V3RI1P5>

Declaration

Conflicts of Interest: The authors declare no conflict of interest.

Author Contribution: All authors wrote the main manuscript text and also consent to the submission.

Ethical approval: Not applicable.

Consent to Participate: All authors consent to participate.

Funding: Not applicable, and No funding was received

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Personal Statement: We declare with our best of knowledge that this research work is purely Original Work and No third party material used in this article drafting. If any such kind material found in further online publication, we are responsible only for any judicial and copyright issues.

Generative AI Statement: The authors declare that generative AI was not used in the preparation or creation of this manuscript. The alternative text (alt text) accompanying the figures in this article was generated by Frontiers with the assistance of artificial intelligence. Reasonable efforts have been made to ensure the accuracy and appropriateness of the generated alt text, including review by the authors wherever possible. If you identify any inaccuracies or issues, please contact the editorial team.

Publisher's Note: The statements, opinions, and claims presented in this article are exclusively those of the authors and do not necessarily reflect the views of their affiliated institutions, the publisher, editors, or reviewers. Any products discussed or evaluated, as well as any claims made by their manufacturers, are not warranted, approved, or endorsed by the publisher.

Acknowledgements

We thank everyone who inspired our work.