https://doi.org/10.65025/ICAIC25222s

Vision-Based Pothole Detection and Reporting System for Smart Cities (Leveraging YOLOv8-Seg and Predictive Maintenance Models)

Paras Sawale
Department of Information Technology JSPM's Rajarshi Shahu College of Engineering
Pune, India
parassawale02@gmail.com

Kusumlata Pawar
Department of Information technology
JSPM's Rajarshi Shahu College of Engineering
Pune, India
kusumlata00pawar@gmail.com

Nihar Ranjan
Department of Information technology
JSPM's Rajarshi Shahu College of Engineering
Pune, India
hod_it@jspmrscoe.edu.in

Omkar Adsare
Department of Information technology
JSPM's Rajarshi Shahu College of Engineering
Pune, India
omkaradsare2003@gmail.com

Gayatri Shelke
Department of Information technology
JSPM's Rajarshi Shahu College of Engineering
Pune, India
shelkegayatri73@gmail.com

Shrutika Jagdale
Department of Information Technology JSPM's Rajarshi Shahu College of Engineering
Pune, India
shrutikajagdale8@gmail.com

Abstract

Maintenance of roads is an important part of city infrastructure and development, whereby potholes play a major role in causing accidents, vehicle damage, and financial losses. Traditional pothole detection involves manual inspection that is time- consuming, inefficient, and manpower-intensive. This paper reviews available techniques for pothole detection and proposes an AI-based framework intended for real-time detection, severity estimation, and report generation. The suggested system uses modern computer vision models like YOLOv8-Seg for the identification and segmentation of potholes along with MiDaS/DPT for estimating depth and XGBoost to predict severity. The solution incorporates GPS- based geo-tagging along with the Google Maps API and cloud platforms for reporting to government authorities automatically. Unique features are a centralized pothole ranking dashboard according to priority, cost estimation for repair, and a citizen app for crowdsourced verification. The system provides a smart and scalable solution for improving road safety.

Keywords

Pothole Detection, YOLOv8-Seg, MiDaS, DPT, Predictive Maintenance, Smart City Infrastructure, Road Safety, Computer Vision.

1. Introduction

Urbanization and the expansion of road networks have been significant challenges to cities worldwide regarding the upkeep of roads. Potholes are some of the most prevalent and hazardous road surface imperfections caused by these challenges. They lead to congestion in traffic, accidents, vehicle damage, and economic losses for municipal governments. Recent studies has shows that accidents caused by potholes account for a major percentage of road accidents and cost millions of rupees in repair costs annually. Traditional pothole detection methods are founded on visual observation and manual surveys, which are labor intensive, time-consuming, and not appropriate for large road networks. These methods also do not provide real- time feedback, thereby limiting their ability to enable proactive maintenance. The vast size of urban road infrastructure and limited finances highlight the need for automated, standardized, and scalable detection mechanisms. Computer vision, artificial intelligence, and intelligent transportation systems developments bring with them new opportunities for road condition monitoring. The solutions are capable of detecting potholes in real time, rating severity, and predicting future patterns of damage.

This paper presents a review and implementation proposal for a Vision based AI - pothole detection and reporting system. The solution uses object detection (YOLOv8-Seg), depth estimation (MiDaS/DPT), and prediction models (ARIMA, XGBoost) to provide real- time detection, automatic reporting to urban municipal governments, and predictive repair planning.

2. Related Work

Pothole detection has been done by using different methods, including traditional image processing and deep learning object detection models. Safyari et al. [1] categorized these methods into 2D image processing, 3D modeling, and hybrid approaches. They emphasized the importance of high-quality annotated data. Frnda et al. [2] compared YOLOv7 and Faster R-CNN in low light.

They discovered that Faster R-CNN provided better accuracy, while YOLO was faster. Gujar et al. [3] trained YOLOv8 on custom datasets, achieving a mean average precision (mAP) of 95.82% and demonstrating strong performance in real-time detection. Other researchers focused on improving model efficiency. Addanki and Lin [4] pointed out that YOLOv8m is compact at 6.3 MB but still offers good detection results. Ahmed [5] reviewed several architectures and found that Faster R-CNN with ResNet50 provided the best accuracy, although it required more computing power. Similarly, Chavan et al.[6] showed that YOLOv4 outperformed CNN models in both speed and accuracy for real-time applications. Many suggestions have been made to enhance robustness. Rout et al. [7] combined YOLOv7 with ESRGAN to handle low-resolution inputs. Meanwhile, Bhavana et al. [9] introduced POT-YOLOv8 with preprocessing filters, achieving 99.1% accuracy. Researchers have also explored lightweight models like YOLOX-Nano [8] for devices with limited resources, aiming to balance accuracy and speed. Dharneeshkar et al. [10] highlighted the need for region-specific datasets and IoT-enabled deployment using Raspberry Pi. Overall, research shows that YOLO-based architecture performs well in both accuracy and speed. However, challenges remain in dealing with poor lighting, varied road conditions, and scaling for larger urban projects.

S/No.	**Author & Year**	**Methodology /Algorithm Used**	**Performance Metrics**	**Conclusion**
1	Yashar Safyari, Masoud Mahdianpari and Hodjat Shiri [1].	U-Net	Accuracy - 97%, Mean Intersection over Union (mIoU) - 0.86	Advanced computer vision and machine learning techniques like U-Net enable accurate (97%) and reliable pothole detection (mIoU 0.86)
2	Jaroslav Frnda, Srijita Bandyopadhyay, Michal Pavlicko, Marek Durica, Mihails Savrasovs and Soumen Banerjee [2].	YOLOv7, FRCNN	FRCNN surpasses YOLO in precision (0.92 vs 0.91) and mAP@ [0.5:0.95] (0.79 vs. 0.63), while both have equal recall at 0.91.	FRCNN's superior performance in detecting potholes under low-light conditions, YOLOv7's efficiency in terms of speed and model size.
3	Dr. S. N Gujar, Prathmesh Shinde,Shlok Parihar, Aditya Pawar, Saikrishna Yemul [3].	YOLO V8, CNN	mAP - 95.82%, Precision - 0.95, Recall - 0.97	YOLOv8 demonstrates robust real-time pothole detection, addressing critical road safety challenges effectively.
4	Bhavana N, Mallikarjun M Kodabagi, Muthu Kumar B,Ajay P, Muthukumaran N and Ahilan A [9].	POT- YOLOv8	POT-YOLOv8 model achieves outstanding performance with ACU:99.10%, PRE:97.65%, RCL:93.52%, And F1S: 90.2%.	The POT-YOLOv8 model demonstrates exceptional accuracy (99.10%) in pothole detection, outperforming other methods.
5	Nirmal Kumar Rout, Gyanateet Dutta,Varun Sinha, Arghadeep Dey, Subhrangshu Mukherjee, Gopal Gupta [7].	YOLOv7x + ESRGAN, YOLOv7 multi +ESRGAN, YOLOv7 tiny multi ESRGAN	YOLOv7x+ ESRGAN: Precision 1.00 YOLOv7 multi + ESRGAN: Precision 0.984 YOLOv7 tiny multi + ESRGAN: Precision 0.947	YOLOv7x excels in precision, YOLOv7 multi balances speed and accuracy, YOLOv7 tiny is fastest.

CASE STUDY: POTHOLE DETECTION IN SMART CITIES: STATE OF ART

Introduction:

Potholes have long been an issue for urban road networks, leading to accidents, vehicle damage and pricey repairs. Many possible solutions have been considered in the literature– from traditional image processing methods to deep learning through actual smart city algorithms, including YOLOv7, YOLOv8 and U-Net, but hardly any studies consider real implementation of these techniques in a smart city environment. Since it is a case study, we can see how research insights can be used and incorporated into a well-established, end-to-end pothole treatment, management system. This practical application bridges the gap between theoretical innovation.

Objectives:

The studies show that the YOLO-based systems are always better than the CNNs, with the precision with YOLOv8 near to 95% on-line.Depth Prediction: MiDaS and DPT projects show us the possibility of more intelligent severity estimation than classical classification. Reporting: very few systems have integrated automated reporting and reporting systems: Very few systems have reporting (reporting automation in reporting pipeline or municipal dashboards because they have reporting systems focused just for detection. Predictive Analytics: Based on previous works it is clear that predicting pothole formation is a large research gap. Only few studies with ARIMA or ML based models have been published in this field at all.

3. Methodology:

The adopted approach represents a vision-based system for such system, which real-time pothole detection, reporting, and prediction for smart city infrastructure are made possible using advanced deep learning algorithms and predictive maintenance, in intelligent city infrastructure. This method is established on a multi-step pipeline with embedded data, computer vision, geo-tagging, cloud reporting, and predictive analytics.

3.1. Data Acquisition

Road condition observations can be gathered on the vehicles through mobile cameras, dashcams, and employ Raspberry Pi modules attached to the vehicles themselves. Each snapshot has GPS information that will be utilized to create geo-referenced inputs into input forms from these snapshots for the incorporation of recorded material. Additionally, the dataset is enriched further with the introduction of a citizen-lead app-enabled mobile app for pothole images submission by the participants for faster road maintenance reporting.

3.2. Pothole Detection and Segmentation and Assessment

Detected are also constructed with deep-learning model YOLOv8-Seg to perform real-time object detection and segmentation. And all this is done to ensure that the potholes get positioned exactly in the video frames, regardless of whatever traffic is passing and the weather. In low-light or low-quality images, image- enhancing techniques can be applied like ESRGAN to enhance the quality of detection

3.3. Determination of depth and severity

Pothole severity prediction with monocular depth estimation tools MiDaS and Dense Prediction Transformer (DPT) are examples. A severity index can be calculated using depth, width, and traffic exposure from one image for depth, size, and volume estimations. Severity scoring is performed in this scenario by way of a supervised learning model, XGBoost, and enables the prioritization of instant corrections by government authorities.

3.4. Geo-Tagging and Cloud-based Reporting

The Google Maps API geo-tags automatically each pothole detected. The information such as images, severity scores, coordinates of the location is stored in a cloud database (say Firebase/AWS). The information is stored and tabulated using a top-level dashboard (a single, aggregated view that categorizes pothole severity according to their urgency and offers municipal officials a clear-cut view of response alternatives.

3.5. Predictive Maintenance

From historical road data, weather, and traffic flow patterns to train predictive models like ARIMA and XGBoost. These models forecast the probable occurrence of potholes by predicting when scheduling maintenance can be done (preventive maintenance). These predictive layers reduce future repair expenses and the road safety threat by avoiding the dangerous condition created by the presence of potholes.

3.6. Dashboard in Software and The Public Interface

The centralized dashboard offers real-time monitoring, which categorizes the kind of potholes and the cost of pothole repairs estimates available to the authorities. Furthermore, the citizen app enables citizens to report potholes, be notified when potholes develop on hazardous parts of the road and monitor repair potholes.

Conclusion

Based on previous integrated pothole management (using those techniques) and surveying methods combined, high-detection accuracy (YOLOv8), robustness to harsh environment(GAN-augmented,YOLOv7),depth estimation (MiDaS/ DPT), predictive modeling (ARIMA, XGBoost), we are able to build an integrated smart city solution. The resultant theoretical model demonstrates the bridge between theory and practice; providing the bridge from academia to engineering in its practical applications.

4. Proposed System

System architecture of the proposed Vision-Based Pothole Detection and Reporting System for Smart Cities has been conceptualized as a multi-layered system with the use of computer vision, machine learning, cloud computing, and smart city infrastructure. The system begins with data acquisition and proceeds through detection, severity estimation, geo-tagging, cloud reporting, predictive analytics, and visualization on a centralized dashboard. As shown in Fig. 1 (System Architecture Diagram) and Fig. 2 (Workflow Diagram), the system operates the following way: Data Collection Layer: There is continuous video taken on the road by vehicle-mounted cameras, mobile phones, or Raspberry Pi devices. GPS information accompanies each frame to offer geo-referenced mapping of detected potholes. Detection and Segmentation Layer: The frames are detected and segmented by YOLOv8-Seg, doing real-time detection and segmentation of potholes. For clarity improvement under poor conditions, ESRGAN is utilized. Depth and Severity Analysis Layer: Depth estimation is done using MiDaS and Dense Prediction Transformer (DPT). They estimate the size of the pothole, and a severity score is produced by XGBoost from depth, width, and contextual factors such as traffic density. Geo-Tagging and Cloud Integration Layer: Detected potholes are geo-tagged using the Google Maps API and uploaded to cloud storage sites (Firebase/AWS).This provides them with accessibility, scalability,and integrability with municipal monitoring systems. Predictive Maintenance: Layer Historical weather trends, traffic flow, and road conditions are modeled using ARIMA and machine learning algorithms to predict potential pothole hotspots.

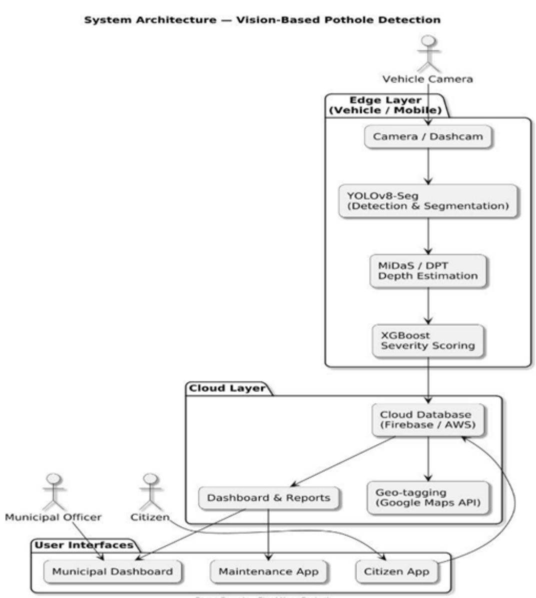

Fig 1. System Architecture Diagram

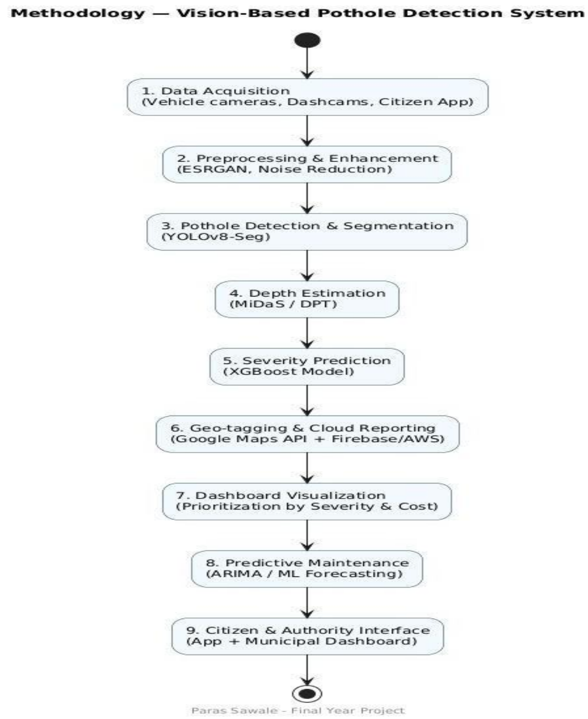

Fig 2. Workflow Diagram

Comparison of average of precision:

The bar chart plots the Mean Average Precision (mAP@0.50) of some of the YOLO-based models. YOLOX-Nano is the most accurate with a perfect 100%, beating everyone else by quite a margin. To everyone's surprise, it outperforms YOLOv4-Tiny by 21.3% and that speaks volumes about its object-detecting power. Of the rest, YOLOv5s, YOLOv4-Tiny, and YOLOv4 are the runner-ups, while YOLOv3 and YOLOv3-Tiny lag behind. This comparison states why YOLOX-Nano is more accurate, and it is suitable for accuracy-critical applications.

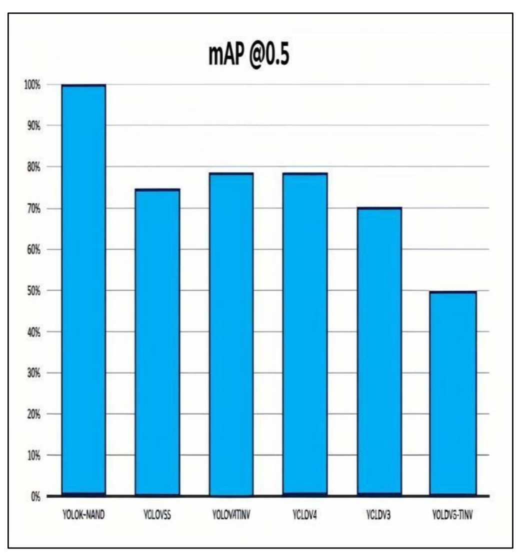

Fig 3. Precision Comparison

The results determine performance improvements in model architecture and indicate that even smaller models will be better than earlier, bigger models in accuracy. This shift highlights the growing efficiency of newer training techniques and optimization strategies. As a result, developers can achieve high accuracy with reduced computational costs, making advanced AI more accessible.

The model size comparison is obvious from the bar graph. YOLOX-Nano being the smallest having the least computing requirements stands at only 7.22 MB. YOLOv5s, twice as large standing at 14.8 MB, and YOLOv5m, standing at 43.3 MB, are relatively heavier in comparison. Even more substantial models like YOLORP6 (291.8 MB) and YOLOR-W6 (624.84 MB) are quite resource-intensive in storage and computation and thus only for cloud installation. The above comparison indicates the advantage of YOLOX-Nano in scenarios where there is a requirement for swift and efficient detection with less hardware constraint.

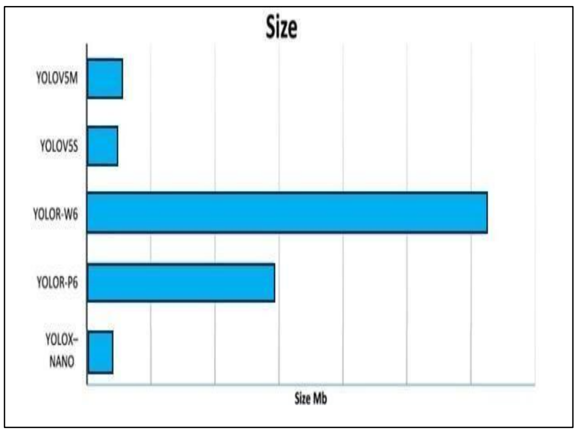

Fig 4. Data Size Comparison FUTURE SCOPE:

A future research and development directions:

- Enhancing the sensitivity of our model to the various environmental parameters (lighting parameters, weather, road conditions and other related features of our model.
- Real time detection on a Raspberry Pi or mobile camera is optimized for resource-limiting devices.
- It is necessary to extend predictive analytics and prediction of the potholes from just a conventional prediction of predictions, to the prediction based on traffic, weather, past repairing records, and repair history.
- Providing citizen engagement features that allow people to use mobile applications to report potholes, and verify them.
- Smart city systems integration, automating maintenance prioritization and resource management for a multitude of tasks.
- Our Vision-Based Smart Pothole Detection & Reporting System which focuses on smart urban infrastructure management, encompassing multiple key enablers such as automation, predictive maintenance, and citizen engagement at an urban, suburban or other level.

5. Conclusion:

Artificial intelligence based pothole detection and reporting systems using computer vision, machine learning, and predictive maintenance techniques were all studied in this survey. Legacy pothole inspection methods are time-consuming, laborious, and prone to human errors. The vision-based smart pothole detection & reporting system is introduced for these limitations, by modern approaches such as YOLOv8-Seg for real-time pothole detection, MiDaS/DPT for depth estimation, and XGBoost for predictive maintenance. Through GPS based geotagging, real time local authority reporting plus centralized dashboard for ordering pothole and monitoring them, this system increases efficiency, reduces repair cost and enhances public safety. The study highlights that although the precision and predictability of AI-driven systems increases, several difficulties have been encountered like data quality, model robustness, real time operation in various environmental conditions, and interfacing with current city infrastructure. These challenges underscore the need for continuous refinement and adaptive deployment strategies.

References:

[1] Yashar Safyari, Masoud Mahdianpari, and Hodjat Shiri, “A Review of Vision-Based Pothole Detection Methods Using Computer Vision and Machine Learning,” Sensors, 2024, 24, 5652. doi: 10.3390/s24175652.

[2] Jaroslav Frnda, Srijita Bandyopadhyay, Michal Pavlicko, Marek Durica, Mihails Savrasovs, and Soumen Banerjee, “Analysis of Pothole Detection Accuracy of Selected Object Detection Models Under Adverse Conditions,” Journal of Physics: Conference Series, Vol. 2, pp. 209–217.

[3] Dr. S.N. Gujar, Prathmesh Shinde, Shlok Parihar, Aditya Pawar, and Saikrishna Yemul, “Pothole Detection System Using YOLOv8,” International Research Journal of Modernization in Engineering Technology and Science, Vol. 6, Issue 4, April 2024. e-ISSN: 2582-5208.

[4] Ashur Raju Addanki, Jianlin Lin, Yeshiva University, “Pothole Detection with YOLOv8,” ResearchGate, December 2023. Available: https://www.researchgate.net/publication/376685010

[5] Khaled R. Ahmed, “Smart Pothole Detection Using Deep Learning Based on Dilated Convolution,” Sensors, 2021, 21, 8406. doi:10.3390/s2124840.

[6] Kshitija Chavan, Chinmay Chawathe, and Vatsal, "Detection System Using YOLOv4 Algorithm," International Journal of Engineering Research & Technology, Vol. 9, Issue 4, April 2022. e-ISSN: 2395-0056.

[7] Nirmal Kumar Rout, Gyanateet Dutta, Varun Sinha, Arghadeep Dey, Subhrangshu Mukherjee, and Gopal Gupta, "Improved Pothole Detection Using YOLOv7 and ESRGAN," February 2024. DOI: 10.48550/arXiv.2401.0858.

[8] Mohan Prakash B. and Sriharipriya K.C., "Enhanced Pothole Detection System Using YOLOX Algorithm," Autonomous Intelligent Systems, Vol. 2(1), August 2022. DOI: 10.1007/s43684-022-00037-z.

[9] Bhavana N., Mallikarjun M. Kodabagi, Muthu Kumar B., Ajay P., Muthukumaran N., and Ahilan A., "POT-YOLO: Real-Time Road Potholes Detection Using Edge Segmentation Based YOLOv8 Network," IEEE Sensors Journal, August 2024. DOI: 10.1109/JSEN.2024.3399008.

[10] Dharneeshkar J., Soban Dhakshana V., Aniruthan S.A., Karthika R., and Latha Parameswaran, "Deep Learning Based Detection of Potholes in Indian Roads Using YOLO," 2020. DOI: 10.1109/ICICT48043.2020.9112424.

[11] A. Singh and M. Randhawa, "Pothole Sensing and Depth Estimation System Using Deep Learning Technique," Conference Paper, June 2024. [Online]. Available: https://www.researchgate.net/publication/366904739_Pothole_Sensing_and_Depth_Estimation_System_using_Deep_Learning_Technique

[12] . A. Jindal, S. Kumar, S. Verma, and S. Kumar, "A Hybrid Method for Pothole Depth Estimation: Combining MiDaS with Point Cloud and RANSAC," Conference Paper, May 2025. [Online]. Available: https://www.researchgate.net/publication/393336581_A_Hybrid_Method_for_Pothole_Depth_Estimation_Combining_MiDaS_with_Point_Cloud_and_RANSAC

[13] "BER-XGBoost: Pothole Detection Based on Feature Extraction and Optimized XGBoost Using BER Metaheuristic Algorithm," ResearchGate, [Online]. Available: https://www.researchgate.net/publication/376802594_BERXGBoost_Pothole_Detection_based_on_Feature_Extraction_and_Optimized_XGBoost_using_BER_Metaheuristic_Algorithm

[14] "Enhanced Pothole Detection in Road Condition Assessment Using YOLOv8," ResearchGate, [Online]. Available: https://www.researchgate.net/publication/380047987_Enhanced_Pothole_Detection_in_Road_Condition_Assessment_Using_YOLOv8

[15] "Enhanced Real-Time Pothole Detection Using YOLOv8 with EfficientNet Backbone: A Deep Learning Approach for Road Safety," IEEE, March 2025. [Online]. Available: https://doi.org/10.1109/ICVADV63329.2025.10961898