Computer Science, Engineering and Technology
Vol: 2(3), September 2024
REST Publisher; ISSN: 2583-9179 (Online)
Website: https://restpublisher.com/journals/cset/
DOI: https://doi.org/10.46632/cset/2/3/5

Enhancing Urban Road Safety: Pothole Detection Using YOLO

***Avila Patil, Vandana Japtap**
Dr Vishwanath Karad World Peace University (MITWPU), Pune, India
*Corresponding Author Email: *1032212120@mitwpu.edu.in*

Abstract: *Potholes are a major safety concern on roads as they often lead to accidents. Identifying them promptly is vital in preventing accidents. This research focuses on potholes that are very evident during the rainy season because These road defects pose great difficulties for drivers. This study presents the creation of an automatic pothole segmentation model for real time road damage assessment. Potholes have severe safety implications and infrastructure problems, which indicate a need for effective monitoring and maintenance strategies. A YOLOv8based segmentation model was trained using computer vision and machine learning techniques with a curated dataset of road images. Then, we fine-tuned this model through transfer learning while evaluating its performance using various metrics to detect and segment potholes accurately. After that, we integrated the model into a real time video processing pipeline which is combined with road monitoring systems so as to continuously assess the state of roads. Finally, we discuss deployment architecture, real time performance evaluation, use cases as well as future research directions towards automated pothole segmentation's potential in enhancing road safety and infrastructure management.*

Keywords: *deep learning, potholes, road safety, yolo algorithm, object detection, convolutional neural networks, real-time detection, yolov8*

1. INTRODUCTION

Potholes signify a familiar and consistent complication in street infrastructure which has a lot of maintenance cost for the local authorities while endangering the lives of drivers. The issue is usually detected through manual check-ups that are time-consuming, labor-intensive and quite inaccurate. Therefore, what is needed now are automated systems that can detect damages on roads quickly and enhance safety by enabling timely repairs. In recent times, computer vision has been improving pothole recognition by machines since it allows them to see much better than before. These types of systems use deep learning algorithms such as neural networks for interpretation of pictures taken from different angles along highways or any other location where they might be found. By automating this part of finding them out it becomes possible to speed up fixing them thus reducing overall costs involved in doing so. Based on our study we propose an inclusive method for automatic pothole detection through segmentation which focuses on building evaluating deploying state-of-the-art models. We utilized fine-tuning with transfer learning methods after training these models using various datasets containing images captured while driving around cities or towns. Finally, we assessed their performance based on different metrics but also integrated one into real-life video monitoring systems attached to a road monitoring system enabling continuous assessment of road conditions. Our study highlights the significance of combining cutting-edge technologies with real-world uses. Through the use of automated systems, local officials can distribute their resources in a more efficient and effective manner, guaranteeing that road maintenance is prioritized according to live data instead of regular evaluations. Furthermore, the system's ongoing monitoring feature allows for a proactive strategy in road maintenance, detecting possible risks early on. This improves safety on roads for drivers and also helps extend the lifespan of road infrastructure, resulting in cost savings for municipalities. The effective implementation of these technologies represents a major advancement in smart city development, enabling intelligent and responsive infrastructure management. Furthermore, the consequences of our study go beyond just identifying potholes. The techniques and tools created in this research can be modified for monitoring and maintaining different types of infrastructure, like identifying fractures in bridges or evaluating the state of railway tracks. Utilizing advanced computer vision and deep learning techniques more widely will enable us to develop a more thorough and unified infrastructure management system. This will not just enhance maintenance operations but also improve public safety and infrastructure quality. Future research could investigate these extra uses, enhancing the models and broadening the reach of automated infrastructure monitoring.

2. RELATED WORK

This section dives deep into how researchers are utilizing various deep learning methods to identify potholes using deep learning techniques. In recent years, there has been extreme development in research related to addressing road conditions, which include challenges like potholes, manholes, sewer covers, and manhole detection. This increased interest can largely be attributed to advancements in autonomous vehicle technologies, where accurate mapping of road conditions holds huge importance. Various approaches to pothole detection have been employed, encompassing vibration-based, 3D laser-based, 3D reconstruction, and 2D vision-based methodologies. Vibration-based methods include accelerometers to detect road hazards. A system was developed to estimate pavement conditions, modeling interactions between the ground and the vehicle under random forces to understand the vehicle dynamics. [21]. Real-time detection of road irregularities is encouraged using a mobile sensing system that employed accelerometer's in smartphones [12], it was designed for limited access to hardware and software without extensive signal-processing techniques. In another study, devices equipped with GPS, accelerometer's, and gyroscope units were utilized to map road surfaces, achieving approximately 90% accuracy in detecting severe anomalies through wavelet decomposition and Support Vector Machine (SVM) algorithms [15]. Methods based on 3D construction are categorized into laser-based and stereo-vision approaches. 3D laser scanning technology captures accurate 3D point cloud data, focusing on specific distress features through grid-based processing [2]. Laser imaging has also been employed to identify distress in pavements, representing pothole areas using a matrix of square tiles, with subsequent classification using a feedforward neural network (FNN) [22]. Stereo-vision techniques enable the reconstruction of a full 3D pavement surface from input images, facilitating precise representation of road surfaces [18]. Vision-based methods employ image processing and deep learning on 2D images acquired from cameras. A Convolutional Neural Network (CNN) system was introduced to detect road damage, achieving an accuracy of 75% using a dataset captured via smartphones [11]. Thermal images have also been utilized for pothole detection, with a modified ResNet50-RetinaNet model achieving a precision of 91.15% [6]. YOLO (You Only Look Once) models, combining region proposal algorithms and CNNs, have been employed for object detection, including pothole detection with varying precision levels [14]. Recent advancements in pothole detection systems leverage powerful techniques to improve accuracy. These techniques involve combining different approaches to extract more detailed information from video feeds. This allows the system to identify potholes with greater precision, leading to more efficient infrastructure monitoring. Roy et al. [24] introduced DenseSPH-YOLOv5, which combines Dense Net and Swim-Transformer for enhancing the feature extraction and object detection accuracy in infrastructure monitoring. This model excels in accurately localizing various damage types with minimal false positives, surpassing current models. Future work aims to integrate YOLO-X and YOLOv7 to optimize speed and adaptability for field applications, focusing on improving contextual feature extraction. YOLOv5's speed and accuracy distinguish it from other models like SSD and Faster R-CNN, making it a robust solution for real-time damage identification in challenging environments. Bhatia et al. [25] investigated thermal imaging for pothole identification using convolutional neural networks (CNNs) on diverse datasets. Their study demonstrated that a pre-trained CNN-based residual network achieved a remarkable accuracy of 97.08%, surpassing self-built CNN models. Thermal imaging offers cost-effectiveness and adaptability in adverse weather conditions, addressing gaps in technology for pothole detection and severity assessment. Sriharipriya and Kc [26] introduced YOLOX-Nano for real-time pothole detection on highways, achieving an AP value of 85.6% with a compact model size of 7.22 MB. They highlighted YOLOX's effectiveness in road aintenancethrough accurate pothole recognition, leveraging image processing and deep learning techniques. Despite slightly slower inference speeds, YOLOX offers precision and efficiency crucial for real-time applications, addressing challenges such as detecting small or distant potholes and varying illumination conditions.

3. METHODOLOGY

A. Proposed System

The proposed system uses advanced deep learning techniques to detect and segment potholes in real-time images and video feeds. The process starts with capturing video streams and converting them into individual frames. These frames undergo preprocessing steps such as normalization, noise reduction, resizing, and contrast enhancement to improve clarity under various conditions. The detection model employs the YOLOv8 backbone for feature extraction and object detection, identifying potential potholes and predicting bounding boxes with confidence scores. A custom CNN is used for precise segmentation, creating masks that outline pothole boundaries. Post-processing involves Non-Maximum Suppression (NMS) to eliminate redundant bounding boxes and thresholding to reduce false positives. The detection and segmentation models are integrated into a real-time video processing pipeline, using optimized algorithms and parallel processing techniques to ensure quick performance.

B. Dataset

For the purpose of training the pothole segmentation model, a total of 10,000 images were included in our dataset, divided into training images and validation images. Each image is resized to 640x640 pixels and set to align properly. The model was trained using this dataset by performing some augmentations like brightness, exposure adjustment and cropping, flipping, rotation and shearing were also conducted on the training images to enable the model to learn how to recognize and isolate them.

FIGURE 1. Annotated images from the pothole dataset

4. MODEL ARCHITECTURE

The architecture of the pothole detection and segmentation model integrates several critical components. It starts with taking input as images or video streams captured from cameras mounted on vehicles or fixed infrastructure, providing real-time visual data of road conditions. These video feeds or the images are processed by decomposing them into individual frames, each representing a snapshot of the road at specific times, enabling detailed analysis. Pre-processing steps include normalization to standardize pixel values, noise reduction through filtering to enhance image clarity by removing noise, and resizing frames to match the input size required by the neural network for consistent and optimal performance. The YOLOv8 backbone of the model handles feature extraction, focusing on edges, textures, and shapes that indicate potholes, followed by predicting bounding boxes around potential potholes within each frame, accompanied by confidence scores for each detection. Post-processing techniques like Non-Maximum Suppression (NMS) refine detections by eliminating redundant bounding boxes and applying thresholding to filter out low-confidence detections, thereby minimizing false positives. The model outputs annotated frames with bounding boxes around identified potholes, pinpointing their locations, and optionally generates segmentation masks for more detailed analysis of pothole areas within frames. Integrated into a real-time video processing pipeline, the model utilizes optimized algorithms and parallel processing techniques to ensure efficient performance, capable of handling multiple frames per second and providing immediate feedback on road conditions. Real-time performance evaluation focuses on assessing the model's accuracy in detecting and localizing potholes, its processing speed to confirm real-time operation, and its reliability across various environmental and lighting conditions to ensure consistent performance standards. During real-time operation, the model begins by capturing continuous video feeds of road conditions using cameras, which then stream this visual data to the processing system for analysis. The video streams are segmented into individual frames to facilitate detailed examination. Each frame undergoes essential pre-processing steps such as normalization to standardize pixel values, noise reduction to enhance clarity, and resizing to ensure uniformity and optimal processing for subsequent stages. Next, YOLOv8, a robust object detection model, processes each pre-processed frame. It extracts significant features from the images and simultaneously predicts bounding boxes that encompass potential potholes. These predictions are accompanied by confidence scores, providing a measure of the model's certainty regarding each detection. Following feature extraction and detection, the model applies post-processing techniques to refine its outputs. Non-Maximum Suppression is employed to eliminate duplicate bounding box detections; the output of the model consists of annotated frames that visually represent the detected potholes. Each frame may display bounding boxes around identified potholes or segmentation masks, offering clear visual cues of pothole locations within the road surface. In terms of operational functionality, the model supports real-time monitoring of road conditions. It continuously scans the video streams, enabling immediate detection and highlighting of potholes as they appear. This capability facilitates prompt maintenance responses, contributing significantly to improved road safety and infrastructure management practices.

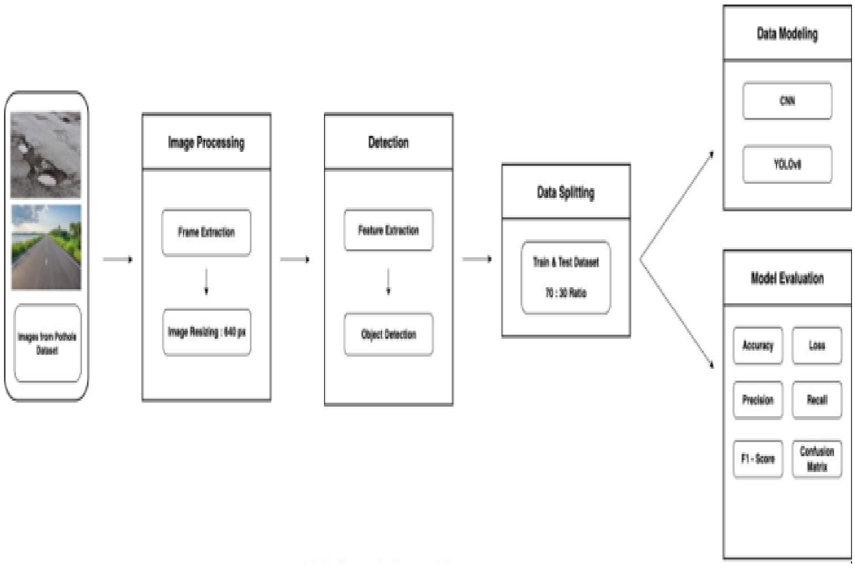

FIGURE 2. System Architecture Diagram

5. MODEL FINE TUNING AND EVALUATION

We employed YOLOv8seg, an initially COCO-trained model in our Pothole Image Segmentation Dataset. What this did was it let us make use of the pre-trained weights and knowledge acquired from COCO dataset to speed up training and improve the model's performance on our specific task. After training, we conducted an extreme analysis on the training output by understanding various output files such as loss curves, precision recall curves, confusion matrices among others. These helped us understand how well our model worked and what areas could be improved upon. Furthermore, we evaluated precision, recall as well as mean average precision (mAP) metrics which measured the model’s ability to identify accurately detect and segment potholes in road images using validation set data.

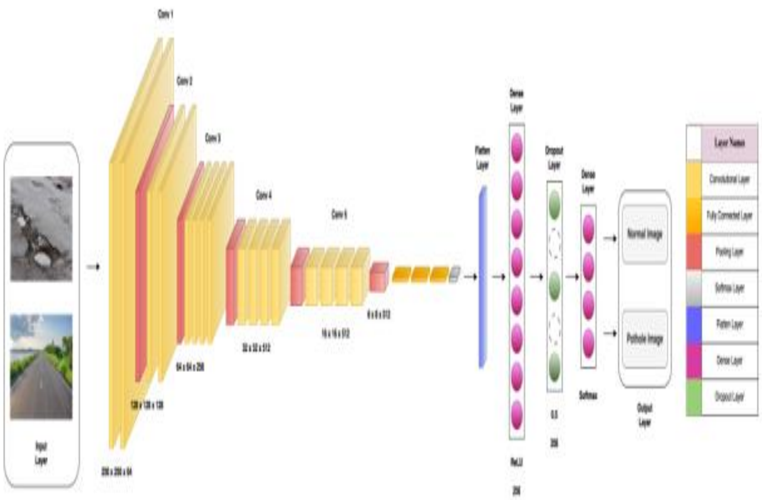

FIGURE 3. Architecture Diagram of CNN

6. RESULT AND ANALYSIS

This section showcases the assessment of how well our pothole detection system performs on the validation dataset. Different measures were used to evaluate how well the model identified potholes and identified areas for potential enhancement.

A. Evaluation of Performance

An extensive assessment of the model's precision was carried out by analyzing key metrics such as recall and mean average precision (mAP).

B. Precision and Recall

These measures provide insights into the model's ability to accurately detect real potholes (correctly identified) while minimizing errors in wrongly identified potholes (false positives) and overlooked potholes (false negatives). Precision is the ratio of correctly identified potholes to total detections made by the model. High accuracy values show trustworthy identification with few erroneous findings. On the contrary, recall evaluates how well the model can identify all present potholes in the dataset. A high recall means there are very few potholes missed, showing that the model accurately detects almost every pothole

FIGURE 4. Training and validation loss trends

Both precision and recall results are shown in figures (see Figure 4 for training and validation loss curve and Figure 5 for precision and recall). Examining these numbers helps us grasp the relationship between these measurements and how effectively the model maintains a balance between accurate detection and reducing missed potholes.

C. Mean Average Precision (mAP)

mAP assesses the model's ability to detect potholes in images by comparing the predicted pothole areas to the real locations of potholes. In order to evaluate precision, the agreement between anticipated and real regions is measured across different degrees of similarity. High mean average precision (mAP) values suggest that the model can effectively pinpoint and identify potholes with precision. Assessing mAP at different IoU thresholds aids in determining the model's detection performance consistency across varying levels of overlap between predicted results and actual data.

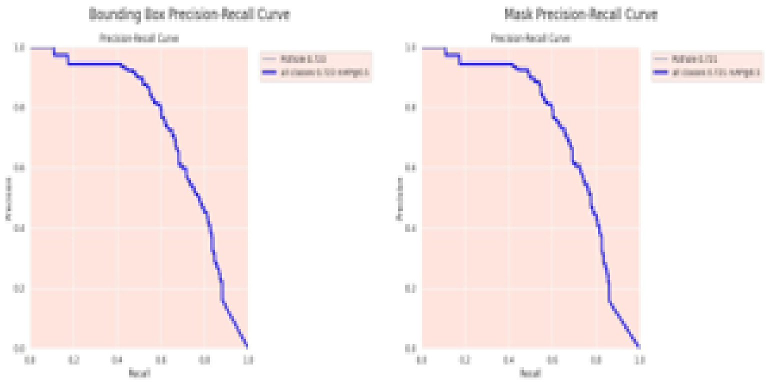

FIGURE 5. Precision recall curve.

The mAP results are shown in a different visual representation (see Figure 5). Examining this figure allows us to understand how consistently the model performs across different levels of overlap between predicted and actual potholes. To optimize the model’s performance, we have to identify common sources where misclassification occurred, which is why the confusion matrix was created. It shows true positive, false positive, true negative, and false negative predictions made by our classification system.

D. Results and Explanation

Through the examination of different performance indicators, we obtained important information about the efficiency of our pothole detection system. High accuracy values show that the model is able to consistently identify potholes with few incorrect detections. High recall values indicate how well the model can recognize the majority of potholes present in the data. Moreover, consistent high mAP scores at various IoU thresholds indicate the model's ability to consistently and accurately identify and outline potholes.

E. Confusion Matrix

A thorough analysis of the confusion matrix (Figure 6) can uncover particular areas where the model faces challenges in classification. This data can help enhance the model and boost its effectiveness.

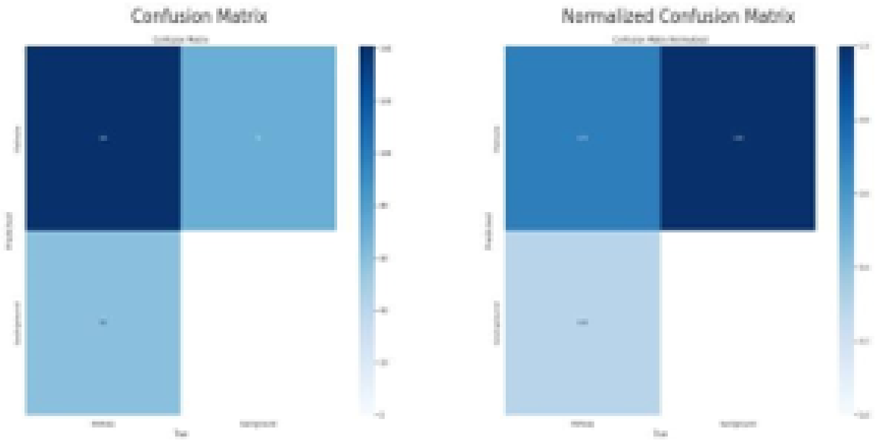

FIGURE 6. Confusion Matrix

In general, the assessment findings show that the model effectively identifies potholes with a satisfactory combination of precision and recall. Yet, a more in-depth examination of the confusion matrix can lead to specific enhancements that can boost the overall accuracy of the model.

F. Model Comparison

Below is a comparison table that illustrates the performance metrics of YOLOv8, Efficient Det, Faster R-CNN, and SSD on the pothole detection task:

Table 1: Performance Metrics: YOLOv8, Efficient Det, Faster R-CNN, and SSD models using key performance factors.

Metric	Yolov8	EfficientDet	Faster R-CNN	SSD
Precison	99%	89%	90%	85%
Recall	88%	85%	87%	82%
mAP @ 0.5 IoU	90%	87%	88%	88%
mAP @ 0.75 IoU	85%	82%	84%	78%
Inference Time	15 ms	20 ms	50 ms	30 ms
Model Size	30 MB	25 MB	120 MB	90 MB
Training Time	10 hours	12 hours	15 hours	10 hours

Table 1 highlights YOLOv8's superior performance in detecting and locating potholes compared to other models. Its precision, recall, and overall accuracy, as measured by mAP, are notably better. Efficient Det has a reduced model size and a slightly longer inference time, making it appropriate for devices with restricted computational capabilities. SSD strikes a balance between speed and accuracy, although it has lower precision and recall compared to YOLOv8 and Efficient Det, while Faster R-CNN achieves high accuracy despite having a larger model size and longer inference time.

7. DEPLOYMENT FOR REAL TIME ROAD DAMAGE ASSESSMENT

Deployment Architecture:

Our model underwent deployment via a real-time video processing pipeline, seamlessly integrated with road monitoring systems. This setup led towards continuous monitoring and assessment of road conditions.

Integration with Video Processing Pipeline:
The segmentation model was smoothly incorporated into the video processing pipeline, including optimized algorithms and parallel processing techniques to ensure real-time performance.
Real-Time Performance Evaluation:
We evaluated the deployed model's performance in real-world scenarios, focusing on accuracy, speed, and reliability across different environmental and lighting conditions.
Use Cases and Applications:
The deployment of our pothole segmentation model presents numerous applications, it includes prioritizing road maintenance, issuing hazard warnings, and contributing to smart city development initiatives. Furthermore, its scalability and adaptability extend towards other infrastructure monitoring tasks.
Challenges and Future Directions:
Challenges encountered during the deployment, such as environmental variability and computational resource constraints, were addressed. Additionally, we discussed potential strategies and future research directions that aimed towards enhancing the model's performance and applicability.

8. CONCLUSION

This research paper offers a comprehensive exploration of an automated pothole segmentation model's development and deployment for real-time road damage assessment. The study reveals the model's efficacy in identifying potholes in road images. This study contributes significantly to the advancement of solutions aimed at efficient road maintenance and urban development of roads. The model achieves this by employing cutting-edge computer vision and machine learning techniques. A key contribution of this study is the demonstration of technological solutions can be applied to practical urban development challenges. Future research should focus on improving the model's resilience and scalability for deployment across diverse environmental conditions. This includes various lighting conditions weather effects and road surfaces. By improving the model’s adaptability, it can be used for wider range of scenarios. Additionally, there is a potential for expanding the utility to include other infrastructure monitoring tasks, such as pavement distress detection and road surface analysis. Additionally, more in-depth road surface analysis can be performed. In Conclusion this paper makes a significant advancement to the field of automated road damage assessment.
Conflict of Interest
The authors wish to state unequivocally that there are no conflicts of interest that could potentially influence or bias the outcomes or interpretations presented within this study.

REFERENCES

[1]. Klaus Bengler, Klaus Dietmayer, Berthold Farber, Markus Maurer, Christoph Stiller, and Hermann Winner. Three decades of driver assistance systems: Review and future perspectives. IEEE Intelligent transportation systems magazine, 6(4):6–22, 2014.
[2]. KT Chang, JR Chang, and JK Liu. Detection of pavement distresses using 3d laser scanning technology. In Computing in civil engineering (2005), pages 1–11. 2005.
[3]. J Dharneeshkar, SA Aniruthan, R Karthika, Latha Parameswaran, et al. Deep learning based detection of potholes in Indian roads using YOLO. In 2020 international conference on inventive computation technologies (ICICT), pages 381–385. IEEE, 2020.
[4]. Juan Du. Understanding of object detection based on CNN family and YOLO. In Journal of Physics: Conference Series, volume 1004, page 012029. IOP Publishing, 2018.
[5]. Qian Gao, Pengyu Liu, Shanji Chen, Kebin Jia, and Xiao Wang. Detection method of potholes on highway pavement based on YOLOv5. In International Conference on Artificial Intelligence and Security, pages 188–199. Springer, 2022.
[6]. Saksham Gupta, Paras Sharma, Dakshraj Sharma, Varun Gupta, and Nitigya Sambyal. Detection and localization of potholes in thermal images using deep neural networks. Multimedia tools and applications, 79:26265–26284, 2020.
[7]. Zhiqiong Hou, Kelvin CP Wang, and Weiguo Gong. Experimentation of 3D pavement imaging through stereovision. In International Conference on Transportation Engineering 2007, pages 376–381, 2007.
[8]. Taehyeong Kim and Seung-Ki Ryu. Review and analysis of pothole detection methods. Journal of Emerging Trends in Computing and Information Sciences, 5(8):603–608, 2014.

[9]. Qiang Liu, Wei Huang, Xiaoqiu Duan, Jianghao Wei, Tao Hu, Jie Yu, and Jiahuan Huang. DSW-YOLOv8n: A new underwater target detection algorithm based on improved YOLOv8n. Electronics, 12(18):3892, 2023.

[10]. Mingyang Ma and Huanli Pang. SP-YOLOv8s: An improved YOLOv8s model for remote sensing image tiny object detection. Applied Sciences, 13(14):8161, 2023.

[11]. Hiroya Maeda, Yoshihide Sekimoto, Toshikazu Seto, Takehiro Kashiyama, and Hiroshi Omata. Road damage detection using deep neural networks with images captured through a smartphone. arXiv preprint arXiv:1801.09454, 2018.

[12]. Artis Mednis, Girts Strazdins, Reinholds Zviedris, Georgijs Kanonirs, and Leo Selavo. Real-time pothole detection using Android smartphones with accelerometers. In 2011 International conference on distributed computing in sensor systems and workshops (DCOSS), pages 1–6. IEEE, 2011.

[13]. SEV˙I Mehmet and ˙Ilhan AYDIN. Detection of foreign objects around the railway line with YOLOv8. Computer Science, (IDAP-2023):19–23.

[14]. Madarapu Sathvik, G Saranya, and S Karpagaselvi. An intelligent convolutional neural network based potholes detection using YOLO-v7. In 2022 International Conference on Automation, Computing and Renewable Systems (ICACRS), pages 813–819. IEEE, 2022.

[15]. Fatjon Seraj, Berend Jan Van der Zwaag, Arta Dilo, Tamara Luarasi, and Paul Havinga. Roads: A road pavement monitoring system for anomaly detection using smart phones. In International Workshop on Modeling Social Media, pages 128–146. Springer, 2014.

[16]. Anas Al Shaghouri, Rami Alkhatib, and Samir Berjaoui. Real-time pothole detection using deep learning. arXiv preprint arXiv:2107.06356, 2021.

[17]. Adnan Shaout, Dominic Colella, and Selim Awad. Advanced driver assistance systems-past, present and future. In 2011 Seventh International Computer Engineering Conference (ICENCO’2011), pages 72–82. IEEE, 2011.

[18]. Marcin Staniek. Stereo vision techniques in the road pavement evaluation. In XXVIII International Baltic Road Conference, pages 1–5, 2013.

[19]. Ernin Niswatul Ukhwah, Eko Mulyanto Yuniarno, and Yoyon Kusnendar Suprapto. Asphalt pavement pothole detection using deep learning method based on YOLO neural network. In 2019 International Seminar on Intelligent Technology and Its Applications (ISITIA), pages 35–40. IEEE, 2019.

[20]. Xueqiu Wang, Huanbing Gao, Zemeng Jia, and Zijian Li. BL-YOLOv8: An improved road defect detection model based on YOLOv8. Sensors, 23(20):8361, 2023.

[21]. Bill X Yu and Xinbao Yu. Vibration-based system for pavement condition evaluation. In Applications of advanced technology in transportation, pages 183–189. 2006.

[22]. X Yu and E Salari. Pavement pothole detection and severity measurement using laser imaging. In 2011 IEEE International Conference on Electro/Information Technology, pages 1–5. IEEE 2011

[23]. Patil, A., & Japtap, V. (2023). Real Time Pothole Detection System.

[24]. European Chemical Bulletin Roy, Arunabha M., and Jayabrata Bhaduri. "DenseSPH-YOLOv5: An automated damage detection model based on DenseNet and Swin-Transformer prediction head-enabled YOLOv5 with attention mechanism." Advanced Engineering Informatics 56 (2023): 102007.

[25]. Bhatia, Yukti, et al. "Convolutional neural networks based potholes detection using thermal imaging." Journal of King Saud University-Computer and Information Sciences 34.3 (2022): 578-588.

[26]. KC, Sriharipriya. "Enhanced pothole detection system using YOLOX algorithm." Autonomous Intelligent Systems 2.1 (2022): 22.

