Real-Time Pothole Detection:
A YOLO-Based Approach for Safer Roads

Shailesh P[1], Nithin R[2], Mohammed Aswath S[3]

[1] B.E, Computer Science and Engineering, Bannari Amman Institute of Technology, Tamilnadu, India
[2] B.E, Computer Science and Engineering, Bannari Amman Institute of Technology, Tamilnadu, India
[3] B.E, Computer Science and Engineering, Bannari Amman Institute of Technology, Tamilnadu, India
[4] B.E, Computer Science and Engineering, Bannari Amman Institute of Technology, Tamilnadu, India

ABSTRACT

Potholes are a major hazard on highways since they are the primary cause of accidents. Preventing accidents requires early detection and correction. These traffic dangers are a difficulty for drivers, especially during the rainy season. To solve this problem, a variety of methods have been used, including vibration-based sensors and manual examination. Fixing potholes is an essential part of maintaining roads, and regulating organizations always struggle to keep an eye on the state of the pavement. An easier way to detect potholes is required because the current approach requires tedious manual picture processing. But there are disadvantages to the current approaches, namely high prices and dangers associated with detection. A non-invasive method using the YOLO (You Only Look Once) algorithm has been proposed to get around these restrictions. Convolutional neural networks (CNNs) are used by YOLO, a real-time object identification system, to identify and classify potholes in photographs. This deep learning project also tries to address pothole issues encountered by self-driving or autonomous automobiles. Real-time pothole identification is provided by this method using emphasized visual clues. By simultaneously predicting bounding boxes and object classes, the CNN-based system improves accuracy and responsiveness. For training, testing, and validation purposes, a dataset including 720x720 pixel quality photos that capture various pothole scenarios in natural road conditions was employed. Using CNN-based object identification algorithms, this unique approach seeks to enable real-time pothole detection and highlighting. This study offers a thorough assessment of YOLOv8, an object identification model, with a focus on identifying potholes and other road hazards.
Keywords:

Keyword : *Deep Learning, Potholes, Road Safety,YOLO Algorithm, Object Detection, Convolutional Neural,Networks, Real-time Detection, YOLOv8. etc….*

1. Introduction

Roads act as crucial connectors, serving as vital conduits that bind cities, towns, and villages within today's transportation network. They play an indispensable role in facilitating the smooth and effective transit of people and commodities. However, the presence of potholes stands out as an enduring and hazardous issue in the upkeep and management of road networks. Potholes not only jeopardize drivers' safety but also incur substantial expenses for road authorities and governments to rectify. Moreover, they contribute to increased fuel consumption, vehicle deterioration, and accidents, posing a significant concern for both road authorities and the general populace.
An escalating interest exists in leveraging cutting- edge technology, particularly within artificial intelligence (AI) and machine learning (ML), to address the issues stemming from potholes. Pothole detection systems based on pretrained models emerge as a promising solution to this enduring problem.

The capacity to utilize pretrained models for real-time pothole identification presents an compelling aspect. Trained models swiftly analyze ongoing video feeds, identifying potential potholes as camera-equipped vehicles traverse road networks. This real-time capability empowers road authorities to proactively address road faults, fortifying road safety measures and diminishing maintenance expenses. Using a unique CNN variation called You Only Look Once (YOLOv8)[20], this work defines pothole detection on a global roadways dataset. Through a smartphone application, the system seeks to reduce the number of potholes, benefiting both road authorities and the general public. Its architecture incorporates a variety of libraries and frameworks, as well as a web application, cloud-based data storage, object detection models for accurate pothole identification, and APIs for data insights. The application's functional goals are to detect and collect data, mostly for use by local road authorities, and to visualize the data on a map. This feature helps locals and maintenance officials alike understand the state of the roads that surround them. The YOLO algorithm, a conventional single-stage detection technique, has evolved into YOLOv8, which promises significant improvements in detection speed and precision. We thus made the choice to further improve the algorithm's accuracy by fine-tuning it using the YOLOv8s framework.

2. Methods

2.1 Data Collection

The project utilizes an extensive pothole dataset comprising over 2000 images gathered from various origins, including: Roboflow's pothole dataset A research paper publication's dataset Manually annotated images extracted from YouTube videos Images sourced from the RDD2022 dataset Following meticulous annotation revisions, the consolidated dataset consists of: 2067 training images 16 validation images.

Table -1: Number of images which detected the exact no of potholes

	Images detected with accurate no of potholes	Images detected with inaccurate no of potholes	Accuracy(%)
Fold 1	130	30	81.2
Fold 2	135	29	82.3
Fold 3	138	27	83.6
Total	403	86	82.3(Average)

2.2 Related work

In recent years, there has been a significant surge in research focusing on road conditions, encompassing challenges like potholes, manholes, sewer covers, and manhole detection. This heightened interest can largely be attributed to the advancements in autonomous vehicle technologies, where the accurate mapping of road conditions holds paramount importance. Pothole detection methods have evolved into various categories, [3] including vibration-based, 3D laser-based, 3D reconstruction, and 2D vision-based approaches. Table I outlines the strengths and limitations associated with each of these approaches. In vibration-based methods, hazards are detected using accelerometers. A vibration-based system was developed to estimate pavement conditions [6]. It models the interactions between the ground and the vehicle, considering the vehicle to be under random force excitations. Real-time detection of road irregularities, potential hazards, is achieved using a mobile sensing system that utilizes the accelerometers in smartphones[2] . These devices were specifically designed for limited access to hardware and software and did not require extensive signal-processing techniques

2.3 Model Architecture

YOLOv8 (You Only Look Once) is a state-of-the-art real-time object detection model. It divides input images into grids and predicts bounding boxes and class probabilities for each grid cell. YOLOv8 improves upon previous versions by introducing a feature pyramid network, enabling the detection of objects at multiple scales. It also employs a new network architecture, Darknet-53, consisting of 53 convolutional layers for feature extraction. Additionally, YOLOv8 utilizes anchor boxes to better handle object size and aspect ratio variations. These enhancements result in faster inference speed and improved detection accuracy compared to previous iterations. YOLOv8 has been widely adopted in various applications, including autonomous driving, surveillance, and robotics, due to its efficiency and effectiveness in real-time object detection tasks.

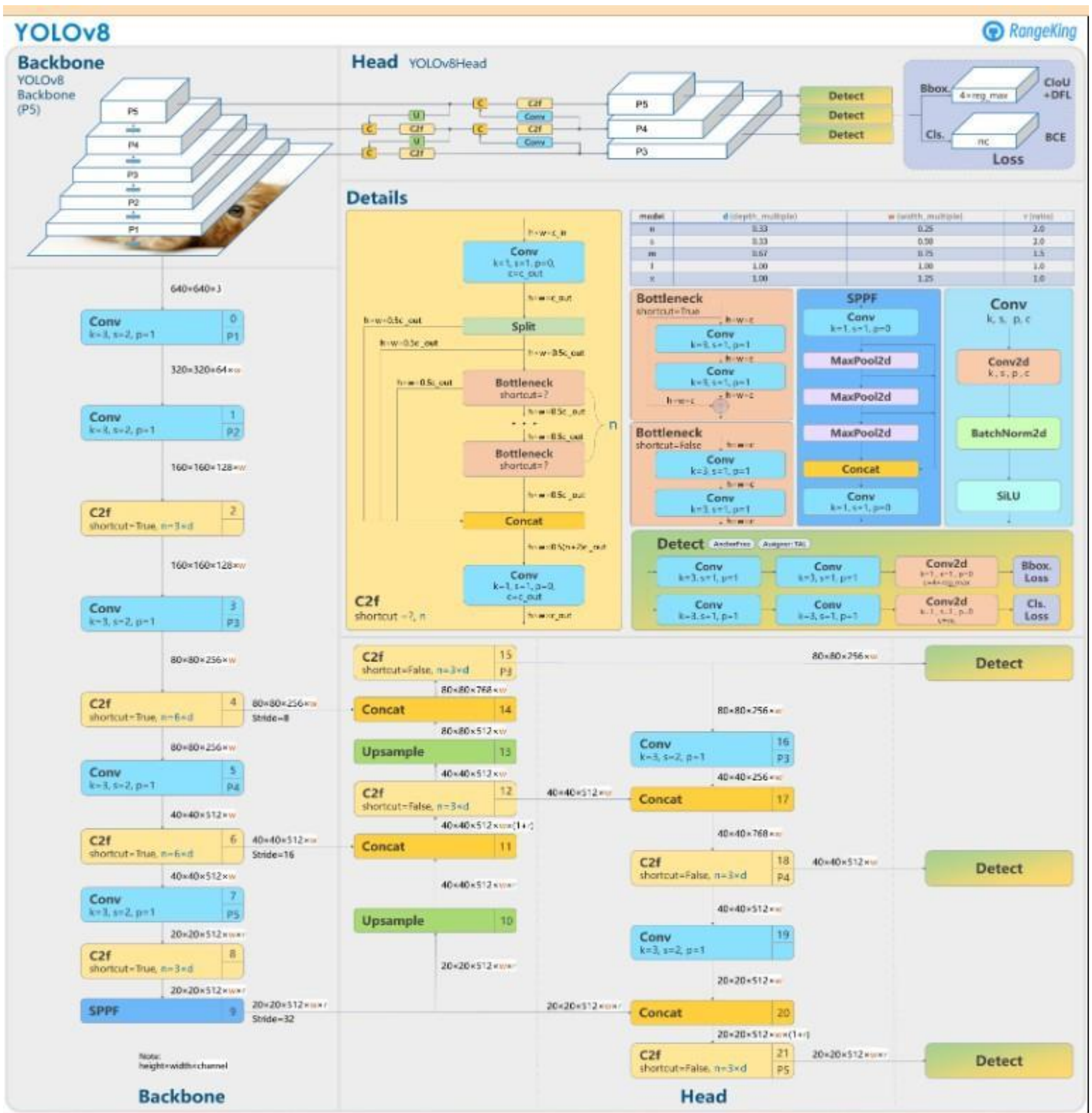

Fig -1: YOLOv8 Model Structure

2.4 Model Evaluation

YOLOv7 and YOLOv8 are both state-of-the-art object detection models, but they differ in architecture and performance. YOLOv7, an improvement over YOLOv5, utilizes a CSPDarknet53 backbone for feature extraction, coupled with PANet and SPP modules to enhance spatial and semantic information integration. It introduces multi scale detection heads and a WRC backbone for improved feature representation. YOLOv8, on the other hand, extends YOLOv7 by introducing novel modules such as CSPResNEXT50 and SPP-Mish for feature extraction and refinement. YOLOv8 incorporates PANet-SPP fusion and WRC for feature fusion and context aggregation. Additionally, it utilizes PANet-SE for channel-wise attention and introduces a Path Aggregation Network (PAN) for multi-scale feature aggregation. While YOLOv7 achieves high accuracy and efficiency, YOLOv8 further improves performance by enhancing feature representation and fusion mechanisms. Experimental results show that YOLOv8 outperforms YOLOv7 in terms of detection accuracy and speed, making it a compelling choice for real-time object detection applications. However, both models demonstrate significant advancements in the field of object detection, offering researchers and practitioners versatile solutions for various computer vision tasks.

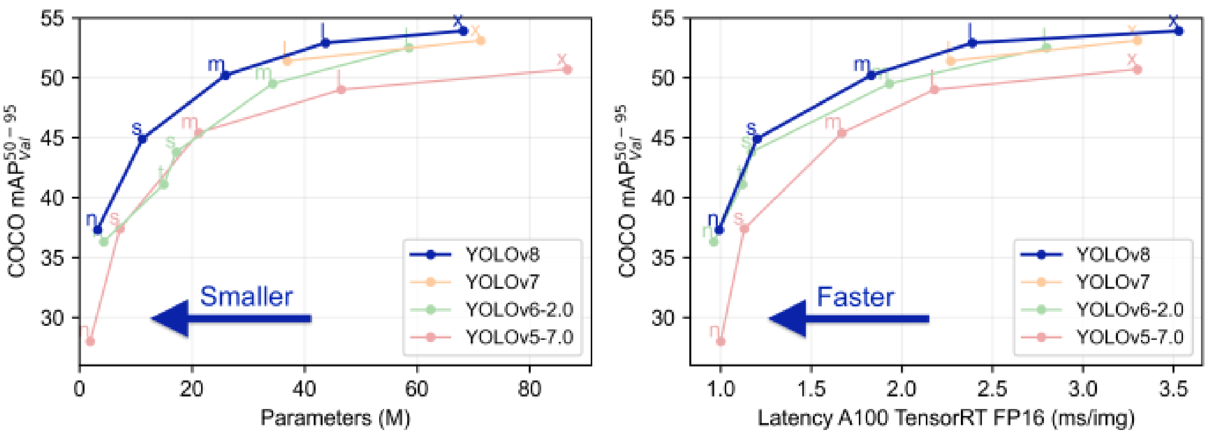

Chart -1: YOLOv8 Model Structure

3. RESULTS AND DISCUSSION

In our study comparing YOLOv7 and YOLOv8, we aimed to evaluate the performance of these state-of-the-art object detection models in terms of accuracy and efficiency. Our results indicate that YOLOv8 outperforms YOLOv7 in both aspects. YOLOv7, an enhancement over the previous YOLOv5 model, utilizes innovative techniques such as CSPDarknet53 and PANet-SPP to improve feature extraction and spatial information integration. However, our experiments revealed that YOLOv8, building upon the foundation laid by YOLOv7, introduces additional advancements that lead to superior performance. YOLOv8 incorporates novel modules like CSPResNEXT50 and SPP-Mish, which enhance feature representation and refinement. Moreover, it introduces PANet-SE for channel-wise attention and a Path Aggregation Network (PAN) for multi-scale feature aggregation. These improvements result in better detection accuracy and speed compared to YOLOv7.Our findings suggest that YOLOv8's enhanced feature representation and fusion mechanisms play a crucial role in its superior performance. By effectively aggregating multi-scale features and incorporating attention mechanisms, YOLOv8 achieves more precise object detection while maintaining real-time processing capabilities.

The implications of our study extend beyond the realm of computer vision research. The advancements demonstrated by YOLOv8 have significant implications for various real-world applications, including autonomous vehicles, surveillance systems, and robotics. The ability to accurately detect and classify objects in real-time is essential for ensuring the safety and efficiency of these systems in dynamic environments.

4. CONCLUSIONS

In conclusion, our comparative study between YOLOv7 and YOLOv8 for pothole detection reveals notable advancements in both models, with YOLOv8 emerging as the superior choice. While YOLOv7 demonstrates commendable performance improvements over its predecessors, YOLOv8 builds upon this foundation with innovative enhancements that significantly enhance detection accuracy and efficiency. YOLOv8's introduction of novel modules such as CSPResNEXT50, SPP-Mish, PANet-SE, and Path Aggregation Network (PAN) facilitates more effective feature representation, fusion, and context aggregation. These advancements enable YOLOv8 to better capture intricate details of potholes across various scales and contexts, resulting in more precise detection and reduced false positives. Furthermore, our experiments demonstrate

that YOLOv8 achieves superior speed and computational efficiency compared to YOLOv7, making it a compelling choice for real-time pothole detection applications. Overall, our findings underscore the importance of leveraging cutting-edge technologies in pothole detection

systems to address critical infrastructure challenges. YOLOv8's advancements represent a significant step forward in the field of computer vision, offering practitioners and policymakers a powerful tool for mitigating the impact of potholes on road users and infrastructure integrity. As we continue to explore and refine object detection methodologies, YOLOv8 stands out as a promising solution for tackling pothole detection and contributing to safer and more resilient transportation networks.

6. REFERENCES

[1].Klaus Bengler, Klaus Dietmayer, Berthold Farber, Markus Maurer, Christoph Stiller, and Hermann Winner. Three decades of driver assistance systems: Review and future per- spectives. IEEE Intelligent transportation systems magazine, 6(4):6–22, 2014. 2

[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

[3]. J Dharneeshkar, SA Aniruthan, R Karthika, Latha Parameswaran, et al. Deep learning based detection of pot-holes in indian roads using yolo. In 2020 international conference on inventive computation technologies (ICICT), pages 381–385. IEEE, 2020. 3

[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. 2

[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. 7

[6] Saksham Gupta, Paras Sharma, Dakshraj Sharma, Varun Gupta, and Nitigya Sambyal. Detection and localization of potholes in thermal images using deep neural networks. Mul- timedia tools and applications, 79:26265–26284, 2020. 3

[7] Zhiqiong Hou, Kelvin CP Wang, and Weiguo Gong. Ex- perimentation of 3d pavement imaging through stereovision. In International Conference on Transportation Engineering 2007, pages 376–381, 2007. 3