Chapter 10
Enhanced Autonomous Driving:
Various YOLO Models for Pothole Detection

Abderrahim Waga
Moulay Ismail University, Morocco

Said Benhlima
Moulay Ismail University, Morocco

Ali Bekri
Moulay Ismail University, Morocco

ABSTRACT

The process of repairing potholes entails significant financial and temporal resources. This article presents cutting-edge approaches to pothole detection employing convolutional neural network techniques, focusing exclusively on RGB input images. The primary objective of this research is to assess the performance of three iterations of the You Only Look Once model, namely, YOLOv8 Nano, Small, and Medium, for pothole detection. Additionally, the proposal suggests a system to alert drivers about identified potholes and recommend alternative routes. The design of the solution enables it to function effectively in real-time in various lighting conditions, including low-light scenarios. Evaluation metrics include inference speed and detection accuracy. We merged diverse datasets from multiple sources to facilitate model training and accelerate convergence. The mean average precision at IoU threshold 0.5 (mAP@0.5) for YOLOv8 n, s, and m models is reported as 66%, 71%, and 51%, respectively.

DOI: 10.4018/979-8-3373-1399-3.ch010

INTRODUCTION

In recent years, there has been a significant focus on the advancement of autonomous vehicles (Parekh et al., 2022), driven by their potential to significantly transform transportation and mobility. These vehicles, integrating intelligent transportation systems technologies, promise to enhance road safety, reduce traffic congestion, and provide sustainable urban mobility solutions. However, realizing these benefits requires addressing numerous challenges, including sophisticated perception and decision-making systems and seamless integration with non-autonomous vehicles. Overcoming these challenges is crucial for successfully deploying and integrating autonomous vehicles into our transportation systems. Pothole detection is crucial not only for the safety of autonomous vehicles but also for overall road safety, infrastructure maintenance, and transportation efficiency. By addressing pothole detection, we can significantly reduce vehicle damage, enhance passenger safety, and improve the longevity and quality of road infrastructure. The YOLO V8 model, a robust computer vision algorithm, has emerged as a leading tool for real-time pothole detection to address this challenge. YOLO is a family of real-time object detection models that treat object detection as a single regression problem. This model divides the image into a grid and predicts bounding boxes and probabilities for each region. By leveraging deep neural networks and one-stage object detection approaches, the YOLO V8 model shows promise and instills confidence in its ability to identify potholes and enhance autonomous driving systems effectively.

Several methods exist for detecting potholes and road damage using various techniques such as lasers (She et al., 2021), vibration sensors (Wang et al., 2015), and imaging (Li et al., 2019), but the method that combines both low cost and ease of maintenance is imagery-based. Hence, harnessing the power of deep learning and, precisely, computer vision in pothole detection is crucial. These technologies have been at the forefront of numerous research efforts, including our own, to address the challenges autonomous vehicles face. (Wang et al., 2022) This study attempted to improve driving efficiency by training a Long Short-Term Memory (LSTM) network on the highD dataset and combining it with Vehicle-to-Vehicle (V2V) technology. Another study (Kavas-Torris et al., 2021) aimed to increase perception robustness by merging multiple sources of information. (Woo et al., 2021) These advancements are not just theoretical but practical solutions to the problem of making the right decision within a reasonable time while adhering to predefined standards. Figure 1 presents our overall system, which will be the subject of study in this research. Our overall system is composed of three components. The front camera of the autonomous vehicle captures the scene, utilizing the first component for this purpose. The second component then uses the captured image to perform inference on all three