SvedbergOpen
DISSEMINATION OF KNOWLEDGE
International Journal of Artificial Intelligence and Machine Learning
Publisher's Home Page: https://www.svedbergopen.com/

Research Paper Open Access

A Multi-Modal Framework For Road Attribute Detection And Evasive Maneuvering In Autonomous Vehicles

Anoop Thomas[1], Dr.Jobin K Antony[2], Samyuktha Sudheer[3], Yadhu Krishna P[4], R Abhijeet[5], and Sandhra Martin[6]

[1,2,3,4,5,6]Department of Electronics and Communication Engineering, Rajagiri School of Engineering and Technology (Autonomous) , Kochi, Kerala
Affiliated to APJ Abdul Kalam Technological, Thiruvananthapuram, India
anoopt@rajagiritech.edu.in, jobinka@rajagiritech.edu.in
Corresponding author: Anoop Thomas

Abstract

This paper presents a robust multi-modal framework for the real-time detection of road anomalies and subsequent autonomous vehicle response. This paper presents a comprehensive comparison of road attribute detection techniques using advanced machine learning models YOLO v4, v6, v8, VGG-16 and ResNet alongside an analysis of depth estimation for potholes using LiDAR data. The objective is to evaluate the effectiveness of these models in accurately detecting road anomalies such as potholes and humps, and to understand how such detections can influence the response mechanisms of an autonomous vehicle. The models were rigorously tested under controlled simulations to detect various road attributes. YOLO versions, renowned for their object detection capabilities, were compared against the classification-oriented VGG-16 and ResNet-34 models to assess their precision, speed, and utility in real-time scenarios. Depth estimation was performed using LiDAR data to gauge the severity of detected potholes, enhancing the vehicle's decision-making process regarding speed adjustment and maneuvering. Fusion between 2D data using computer vision architectures like YOLOv8, YOLOv4, VGG-16, and ResNet-34 with 1D LiDAR depth profiling, the system identifies and evaluates the severity of potholes and humps. The findings reveal distinct strengths and limitations of each model in the context of autonomous driving. The paper also explores how the detection information is utilized by an autonomous vehicle's control system to execute responsive actions such as deviation from potholes and speed reduction over humps. This integration of detection and responsive action aims to improve navigation safety and passenger comfort, showcasing the potential of combining multiple sensory data and machine learning models in the development of autonomous vehicle technologies. The results contribute valuable insights into the selection of appropriate models and systems for enhanced road attribute detection and response strategies in the evolving landscape of autonomous driving.

Keywords: Object Detection, Classification Models, YOLOv4, YOLOv6, YOLOv8, VGG-16, ResNet, Road Attribute Detection, Pothole Detection, Hump Detection, Autonomous Vehicles, architectural Innovations, Detection Accuracy, Computational Efficiency, Real-world Applicability, Depth Estimation, Solid State LiDAR.

1. Introduction

The rapid evolution of autonomous vehicles represents a paradigm shift in the automotive industry, promising advancements in safety, efficiency, and user convenience. However, this transformative journey is accompanied by complex challenges, with one critical aspect being the accurate detection and response to road attributes [1]. In the dynamic and intricate environments where autonomous vehicles operate, the precise identification of elements such as potholes is essential for ensuring occupant safety and optimizing overall system efficiency.

Autonomous vehicles navigate dynamic road environments fraught with challenges, from uneven surfaces to potholes. Accurate detection of these attributes becomes imperative, enabling vehicles to dynamically adapt their behavior for a safe and efficient journey. Beyond safety, precise detection contributes to optimal route planning, allowing vehicles to navigate around potential hazards and disruptions caused by varying road conditions. This real-time decision making capability is crucial n the fast-paced landscape of autonomous driving, enabling vehicles to make instant adjustments for a safe and controlled drive.

Furthermore, accurate road attribute detection enhances safety measures by facilitating timely and effective responses to potential threats. Potholes, in particular, pose risks to vehicle safety, and the ability to detect and respond to such hazards swiftly can prevent collisions and minimize the risk of damage to the vehicle and its occupants. The data-driven adaptability offered by accurate detection methodologies allows autonomous systems to learn from diverse datasets, improving their ability to recognize and respond to a wide array of road attributes. This adaptability not only contributes to individual vehicle safety but also has implications for optimizing the entire road network as vehicles collectively share information about detected attributes.

The methodology employed for accurate pothole detection, exemplified by the comparison between VGG-16, ResNet-34, YOLOv6 and YOLOv8 models and depth estimation using LiDAR data, plays a crucial role in addressing the aforementioned needs. The success of this methodology lies in its emphasis on data quality, utilizing a large and diverse dataset with labeled images capturing various road conditions and pothole manifestations. Leveraging the Ultralytics library and pre-trained weights, the models undergoes iterative training with varying epochs and batch sizes ensuring effective convergence and optimized parameters for enhanced pothole detection capabilities.

Furthermore, the comprehensive evaluation process, encompassing traditional validation sets and real-world performance testing on diverse hardware devices using YOLOv4, ensures the reliability of the detection system. This dual approach provides a holistic understanding of the model's capabilities, from generalization on validation sets to practical efficiency on different computing platforms.

The methodology doesn't conclude with model training but extends to optimization and further improvements. Iterative training with varying epochs and batch sizes, coupled with techniques like data augmentation and transfer learning, underscores a commitment to enhancing the model's efficiency and accuracy over time. In conclusion, the outlined methodology serves as a robust framework for addressing the intricate challenges of accurate road attribute detection in autonomous vehicles, contributing to the broader goal of making autonomous vehicles safer, more efficient, and adaptable to the dynamic nature of our roadways.

2. Literature Review – Related works

Autonomous vehicles rely on advanced sensing and detection mechanisms to navigate through dynamic environments safely. In recent research, some notable studies address critical aspects of object detection and pothole classification, contributing valuable in- sights to the field of autonomous vehicle technologies.

2.1 Fusion of 3D LIDAR and Camera Data for Object Detection in Autonomous Vehicle Applications [2]

This study introduces a novel approach for robust object detection in autonomous vehicles by integrating 3D LIDAR sensors and vision cameras. The methodology leverages 3D LIDAR data to generate precise object-region proposals, seamlessly mapping them to the image space. Extracted regions of interest are then fed into a convolutional neural net- work (CNN) for object recognition, emphasizing multi-scale feature extraction. Rigorous evaluation using the KITTI dataset demonstrates the method's effectiveness in candidate proposal accuracy, processing speed, and overall identification precision, particularly in distinguishing between car and pedestrian entities. The findings highlight its potential as a valuable contribution to autonomous vehicle object detection.

2.2 Pothole Classification Model Using Edge Detection in Road Image [3]

This research focuses on a pothole detection model employing edge detection in road images and introduces an innovative approach for pothole classification using the You Only Look Once (YOLO) algorithm. The proposed method begins by converting RGB image data into grayscale, simplifying computational complexity. An object detection algorithm is applied to identify and isolate potholes, with a pixel value of 255 assigned to represent the background. The contour of the pothole is then extracted through edge detection to capture its distinctive characteristics. Potholes are ultimately detected and classified using the YOLO algorithm. Performance evaluation includes assessing distortion rate and restoration rate, validating the model's effectiveness and classification ac- curacy. This research contributes to advancing road damage detection methods, offering a promising solution for accurately classifying potholes in road images through a combination of grayscale conversion, object detection, edge detection, and the YOLO algorithm.

2.3 Camera and Radar Sensor Fusion for Robust Vehicle Localization via Vehicle Part Localization [4]

This study recognizes the challenges posed by individual sensors in delivering precise position information for highly automated driving functions and advanced driver-assistance systems. The research proposes a camera-radar sensor fusion framework for robust vehicle localization, leveraging vehicle part (rear corner) detection and localization. The approach enhances the azimuth angle accuracy of radar information by detecting and localizing the rear corner part of the target vehicle from an image. The proposed methodology involves a deep learning-based model for vehicle part classification, candidate point generation, and particle filter-based tracking of the rear corner part. Experimental results showcase significantly improved localization performance in the lateral direction compared to radar alone, with notable reductions in maximum errors and root mean squared errors. The study demonstrates the potential of the proposed method to enhance the accuracy and robustness of vehicle localization, particularly in scenarios involving occlusions.

2.4 Convolutional Neural Networks Based Potholes Detection Using Thermal Imaging [5]

This research explores the use of thermal imaging for pothole detection. Thermal images of roads are processed using Convolutional Neural Networks (CNNs) to train a model for accurate pothole detection. The developed model is tested against existing models, demonstrating superior accuracy in pothole detection. The application of thermal imaging provides a unique perspective for enhanced pothole detection capabilities.

In summary, these studies contribute significantly to the advancement of autonomous vehicle technologies. However, identified gaps include the need for further research in real-world deployment scenarios, addressing environmental challenges, and exploring the integration of multiple sensor modalities to enhance overall system robustness. Future studies could focus on closing these gaps to facilitate the broader adoption of autonomous vehicle technologies

3. YOLO Architectures Overview

3.1 YOLO v4

YOLOv4 [6] brought notable improvements over previous versions by introducing CSPDarknet53 as its backbone, which leveraged Cross-Stage Partial (CSP) connections for improved feature extraction. It incorporated critical components such as spatial pyramid pooling (SPP) and the path aggregate network (PAN) to optimize feature fusion and facilitate multiscale detection. The model was designed to strike a balance between high detection accuracy and real-time processing efficiency.

3.2 YOLO v6

YOLOv6 [7] was developed with a focus on industrial applications, focusing on computational efficiency while maintaining high precision. It utilized lightweight backbones, including EfficientNet and YOLOv6-tiny, to minimize inference time. Modules such as the Path Aggregation Network (PANet) and Adaptive Spatial Feature Fusion (ASFF) significantly improved object localization. Furthermore, YOLOv6 incorporated advanced loss functions such as focal loss and IoU loss, ensuring robust and effective training.

3.3 YOLO v8

YOLOv8 [8] is the latest version in the series, focusing on multiscale detection and high precision. It introduced a hierarchical feature pyramid network to enhance the detection of small and complex objects. The model incorporated advanced techniques like dynamic anchor assignment and label smoothing, boosting its performance in challenging and diverse scenarios

3.5 VGG -16

VGG-16[9] is a deep convolutional neural network recognized for its simplicity and effectiveness. It consists of a series of convolutional and max-pooling layers that extract features, followed by fully connected layers for classification. Although VGG-16 performs well in classification tasks, it lacks object localization capabilities, which makes it less suitable for detection processes.

3.6 ResNet -34

ResNet-34[10] is a residual network designed to overcome the vanishing gradient problem. By incorporating skip connections, it allows for the training of deeper networks. Although ResNet-34 is highly effective for image classification, similar to VGG-16, it requires additional modifications to be applied effectively for object detection tasks.

4. Proposed Methodology

4.1 Integrated LiDAR-Vision Fusion

The proposed system employs a late-fusion strategy to combine 2D object detection with 1D depth profiling[11]. The vision subsystem is powered by the YOLO algorithm, identifies the Region of Interest (RoI) for road anomalies within the RGB camera frame. Once an anomaly (pothole or hump) is localized with a bounding box (x,y.w,h),the system maps these coordinates to the LiDAR's spatial scan area. The LiDAR sensor performs a horizontal sweep across the road surface. Depth estimation is calculated by comparing the baseline road distance (d_{base}) with the instantaneous distance measured within the detected RoI (d_{roi}). The severity of the anomaly is quantified as: $d_{delta}= | d_{base}- d_{roi} |$. This fusion [15] allows the system to filter out 'false positives' (such as shadows or oil spills) that a camera might misidentify as a pothole but which lack the physical depth confirmed by the LiDAR.

4.2 Comparative study of deep learning models for pothole and speed bump detection

To address the critical issue of road anomaly detection and classification in road images, we conducted an extensive investigation leveraging deep learning techniques. Our primary goal was to develop robust models capable of accurately detecting road anomalies like potholes and humps within road images and effectively classifying images based on the presence or absence of them. For this, we employed state-of-the-art object detection frameworks, with a focus on architectures optimized for real-time performance and high accuracy. Among these, we explored variants of the YOLO (You Only Look Once) architecture, including YOLOv4, YOLOv6, and YOLOv8,[16] due to their efficiency and effectiveness in detecting objects within images. We fine-tuned these models on annotated datasets containing images labeled with pothole bounding boxes to enable them to identify and localize potholes accurately.

In parallel, we addressed the classification aspect by training convolutional neural networks (CNNs) on a dataset comprising road images annotated with binary labels indicating the presence or absence of road anomalies. Leveraging architectures such as VGG16 and ResNet, we trained classification models to distinguish between images containing potholes and those devoid of them. Through extensive experimentation with different model architectures, training strategies, and data augmentation techniques, we aimed to optimize the classification performance and generalization capabilities of our models.

Following model training and validation, we rigorously evaluated the performance of both the detection and classification models on a separate test dataset comprising unseen road images. Performance metrics such as precision, recall, F1-score, and accuracy were computed to assess the models effectiveness in accurately detecting potholes and classifying images. Additionally, qualitative analysis was conducted to scrutinize the models behavior o real-world road images, ensuring their practical applicability and reliability. The gathered results provide valuable insights into the efficacy of our developed models in addressing the challenge of pothole detection and classification, paving the way for future enhancements and deployment in real-world scenarios to enhance road safety infrastructure maintenance.

4.2.1 Dataset Preparation

To create a robust dataset for training a road attribute detection system, we followed a meticulous methodology involving several key steps. Initially, we gathered images from multiple sources, notably Kaggle and roboflow, ensuring a wide representation of road conditions such as potholes, speed humps, and regular road surfaces. These images formed the foundation of our dataset, reflecting the diverse scenarios encountered on roads. To facilitate effective training, each image underwent manual classification into distinct categories, specifically potholes and speed humps. This classification process resulted in approximately 2000 images per category, offering a substantial dataset for training purposes.

Moreover, to enhance the dataset's utility, detailed annotations were created for each image, including precise labels and bounding boxes, providing essential information for the road attribute detection system to learn and accurately identify various road features during inference. Through these systematic steps, we ensured the creation of a comprehensive and well-prepared dataset, laying the groundwork for the development of a robust road attribute detection system.

4.3 Depth Estimation and segmentation using LiDAR dataset

In this section, we delve into the fascinating realm of depth estimation and segmentation using LiDAR (Light Detection and Ranging) datasets, a pivotal aspect of modern sensing technologies in autonomous systems and environmental monitoring. Our exploration begins with the acquisition of LiDAR Matrix datasets, obtained from sensors meticulously aligned using checkerboard method [16]. These LiDAR Matrix datasets encapsulate precise distance values, providing a rich source of spatial information crucial for depth estimation tasks.

The LiDAR Matrix datasets undergo a transformative journey as they are fed into the High-Level Synthesis (HLS) tool, a powerful software component adept at translating high-level descriptions into hardware implementations. Within the HLS environment, intricate logic unfolds as the matrices undergo operations such as subtraction, thresholding and averaging, orchestrated to unveil the depth characteristics of the environment, particularly focusing on the identification and delineation of potholes.

The culmination of this process yields not only insights into the depth of potholes but also tangible outcomes in the form of Register Transfer Level (RTL) descriptions. These RTL descriptions encapsulate the synthesized hardware logic, representing a blueprint for the hardware realization of the depth estimation algorithm. Furthermore, detailed analysis of hardware utilization metrics provide invaluable insights into resource allocation, enabling optimization and refinement of the implemented algorithms.

Through this comprehensive approach, we aim to harness the power of LiDAR datasets and HLS methodologies to advance the frontier of depth estimation and segmentation, with a specific focus on enhancing road safety through the precise identification and characterization of potholes. By integrating cutting-edge sensing technologies with hardware acceleration techniques, we strive to pave the way for more efficient and accurate solutions in autonomous navigation and infrastructure maintenance domains.

4.3.1 Dataset

The dataset created is a crucial component in our endeavor to accurately estimate depth and segment potholes using sensor fusion techniques, specifically leveraging data from Solid State LiDAR and camera sensors. This meticulously curated dataset comprises a 660x24 matrix, meticulously obtained after aligning the LiDAR and camera sensors using checkerboard calibration technique.[16]

Each entry in this matrix encapsulates vital distance measurements captured by the LiDAR sensor, precisely mapped with corresponding high-resolution images obtained from the camera sensor. This alignment process ensures that the data from both sensors are accurately synchronized and co-registered, facilitating seamless integration and fusion during subsequent analysis.

Through this meticulous alignment and mapping process, we have created a comprehensive dataset that serves as the foundation for our depth estimation and segmentation tasks. The synchronized data from LiDAR and camera sensors enable us to leverage the strengths of each modality, combining precise distance measurements from LiDAR with rich visual information from the camera to enhance the accuracy and reliability of pothole detection and segmentation algorithms.

4.3.2 Depth Estimation

In the methodology for depth estimation of potholes using LiDAR (Light Detection and Ranging) datasets, a systematic approach unfolds, encompassing data preprocessing, thresholding, depth calculation, RTL generation and hardware resource utilization analysis[17].

The process begins with data preprocessing, wherein the LiDAR dataset, consisting of matrices of distance values, is read into the computational environment provided by Vitis HLS software. Each matrix represents a snapshot of the environment captured by the LiDAR sensor. Subsequently, a matrix subtraction operation is performed, subtracting the height of the LiDAR sensor placement from the road surface from each distance value in the matrix. This normalization step ensures consistent reference frames for depth estimation. Following data preprocessing, thresholding is applied to identify potential pothole regions. A minimal depth value indicative of a pothole is defined empirically, based on the characteristics of the LiDAR dataset and the expected depth of potholes. Distance values below this threshold are retained as potential pothole regions, while others are set to zero.

With the pothole regions isolated through thresholding, depth calculation ensues. Various approaches can be employed, such as averaging the depth values within each pothole region or applying specific algorithms to estimate the depth profile accurately.Concurrently, Register Transfer Level (RTL) descriptions are generated using the Vitis HLS software. These descriptions capture the hardware logic required to implement the ndepth estimation algorithm efficiently on hardware platforms. Furthermore, an analysis of hardware resource utilization is conducted to assess the hardware requirements of the depth estimation algorithm. This analysis provides insights into the utilization of FPGA resources such as logic cells and memory blocks.

4.3.3 Depth First Search Algorithm

The Depth First Search (DFS) algorithm is a fundamental graph traversal technique that explores as far as possible along each branch before backtracking. It starts at a selected node and explores as far as possible along each branch before backtracking. In essence, DFS prioritizes depth over breadth when traversing a graph or tree structure. This algorithm is particularly useful in various applications, such as finding connected components, determining paths, and detecting cycles within graphs. It operates efficiently in situations where the depth of the search tree is relatively small compared to the number of vertices or edges. In the context of LiDAR data analysis for pothole detection, DFS can be utilized effectively to determine if potholes are large enough based on the depth of the pothole. By representing the LiDAR data as a graph, with each point as a node and the connections between points as edges, DFS can traverse through the data to identify regions with significant depth variations indicative of potential potholes. Through a recursive exploration process, DFS can efficiently locate and analyze these regions, providing valuable insights into the size and distribution of potholes within the surveyed area. This approach allows for the automated detection and assessment of potholes from LiDAR data, enabling efficient road maintenance and infrastructure management strategies.

4.4 Response System

The second phase of our work focused on the development of a response system for an autonomous vehicle prototype, which reacts dynamically to road attributes such as potholes and humps. For this phase, we utilized the YOLOv4 model, leveraging its trained weights to detect potholes and humps from video feeds simulating real road scenarios. The video input, which included various instances of potholes and humps, was processed by the detection system implemented on a Jetson Nano. The choice of YOLOv4 was driven by its efficiency in real-time object detection, allowing our system to identify road anomalies accurately and promptly.

Upon detection, the response actions were executed through a hardware setup involving an Arduino controller. The Jetson Nano communicated with the Arduino via serial communication, sending commands that were contingent on the nature and location of the detected road attribute. For instance, if a pothole was detected on the left side of the video frame, the Arduino was instructed to maneuver the vehicle prototype to the right, effectively avoiding the pothole. Conversely, detection of a hump would trigger a command to slow down the vehicle, mitigating the impact of the hump on the vehicle's stability and the comfort of its passengers.

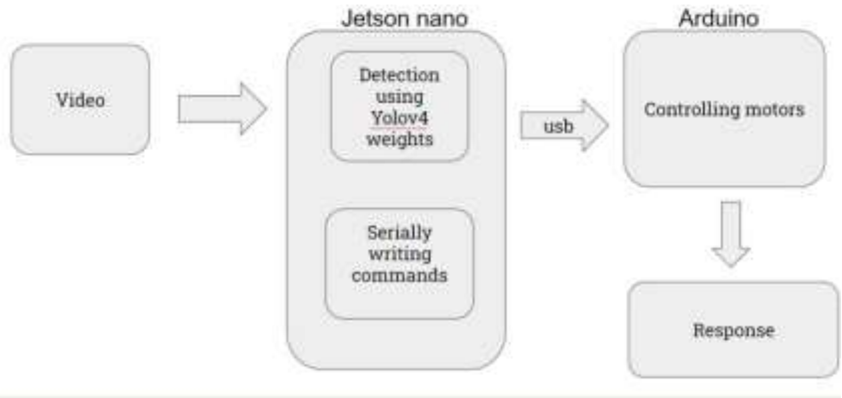

Fig. 1. Response Part Block Diagram

5. Implementation & Results

5.1 Comparitive study of deep learning models for pothole and speed bump detection

5.1.1 Dataset Overview

The custom pothole and speed bump dataset comprises 4967 images meticulously gathered for training VGG-16, ResNet-34, and YOLO models. Annotation of these images was manually conducted using the CVAT online annotation platform, with annotations stored in text format. This annotated dataset serves as a robust resource for developing advanced computer vision algorithms, providing precise ground truth data for accurate pothole and speed bump detection.

Table 1. Hump and pothole dataset split

Category	Count	
	Train	Test
Hump	2241	83
Pothole	2563	80
Total	**4804**	**163**

5.1.2 Performance Metrics of different models

Table 2. Performance Metrics of VGG -16 Model

Class	Precision	Recall	F1 Score	Accuracy
All	0.94	0.94	0.94	93.8%
Pothole	0.91	0.97	0.94	
Hump	0.97	0.90	0.94	

Table 3. Performance Metrics of ResNet -3

Class	Precision	Recall	F1 Score	Accuracy
All	0.93	0.93	0.93	92.6%
Pothole	0.89	0.97	0.93	
Hump	0.97	0.88	0.92	

Table 4. Performance Metrics of YOLOv4

Class	Precision	Recall	F1 Score	Accuracy
All	0.88	0.63	0.74	87%
Pothole	0.89	0.68	0.77	
Hump	0.85	0.62	0.71	

Table 5. Performance Metrics of YOLOv6

Class	Precision	Recall	F1 Score	Accuracy
All	0.83	0.63	0.77	55%
Pothole	0.84	0.51	0.77	
Hump	0.82	0.76	0.71	

Table 6. Performance Metrics of YOLOv8

Class	Precision	Recall	F1 Score	Accuracy
All	0.88	0.63	0.83	63%
Pothole	0.84	0.63	0.78	
Hump	0.91	0.72	0.87	

The YOLO models displayed diverse performance across different metrics. YOLOv8 achieved the highest precision of 0.877, demonstrating its ability to correctly identify objects with minimal false positives. However, YOLOv4 showed a slightly better recall at 0.63, indicating its proficiency in identifying the majority of true positives but with a trade-off of higher false positives. Overall, the YOLOv8 model stood out for its high precision and recall balance, making it particularly effective for complex detection tasks with small and overlapping objects. Meanwhile, the VGG-16 and ResNet-34 models demonstrated strong classification performance but faced challenges in object localization and dynamic detection tasks.

Table 7. Different models of YOLO

Metric	YOLOv4	YOLOv6	YOLOv8
Precision (P)	0.88	0.83	0.877
Recall (R)	0.63	0.633	0.677
mAP50	0.74	0.77	0.826
mAP50-95	N/A	0.444	0.503

5.2 Depth Estimation and Segmentation of road anomalies using LiDAR Data

5.2.1 Depth Estimation.

The process outlined entails a systematic approach to analyze LiDAR data, primarily focusing on the detection and characterization of surface irregularities, particular potholes.Initially, the LiDAR distance matrix is meticulously loaded into an HLS tool, enabling subsequent processing steps. Through matrix subtraction, height differences are calculated, providing valuable insights into variations in elevation indicative of potholes or other surface anomalies. To ensure data integrity, noise filtering techniques are applied via thresholding, allowing for the selective retention of relevant depth information while mitigating the impact of extraneous noise. Following this, averaging methodologies are employed to estimate the average depth within regions of interest, thereby facilitating accurate depth assessments crucial for identifying and evaluating potholes. Post-synthesis, a thorough analysis of resource utilization is conducted to ascertain the compatibility of the design with the targeted hardware specifications, ensuring optimal performance and efficiency. It’s important to note, however, that the current implementation lacks the capability to address scenarios involving multiple potholes or accurately locate specific regions of potholes within the LiDAR data. This limitation underscores the potential for further refinement and enhancement of the methodology to better address real-world scenarios involving complex road conditions.

5.2.2 Implementation of segmentation at the logic level.

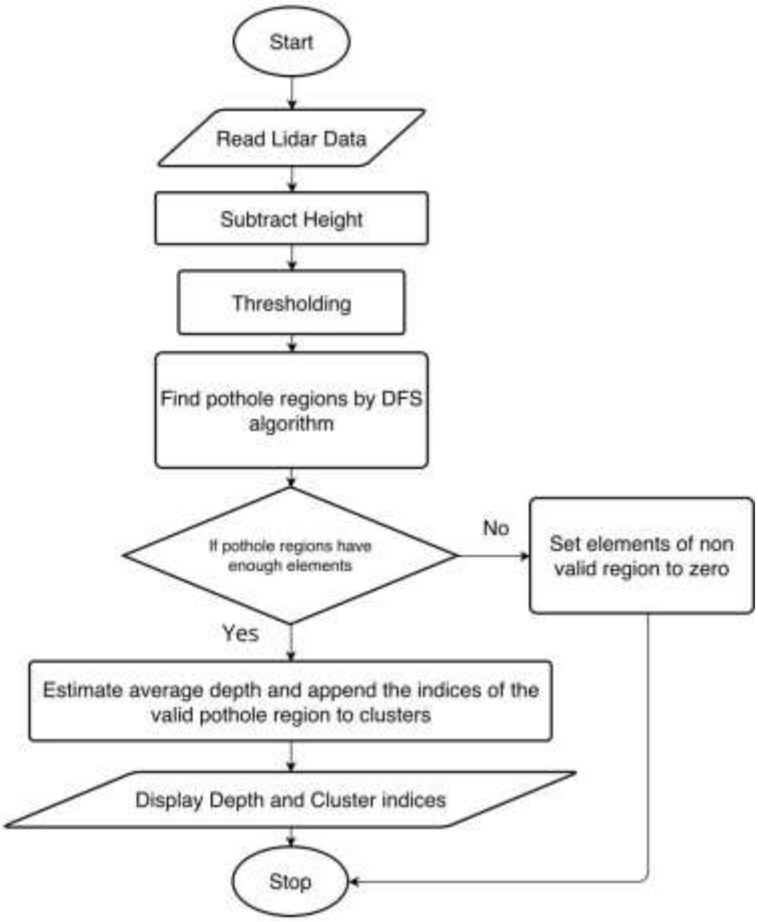

Fig. 32. Flowchart for logic level implementation of segmentation

The initial step in the pothole detection process from LIDAR data involves standardizing the dataset by subtracting the height from the road surface from each depth value. This adjustment ensures that subsequent analyses are conducted relative to the baseline road level, enhancing the accuracy of pothole detection. Following this preprocessing step, thresholding mechanisms are employed to discern significant irregularities in the road surface from noise or minor fluctuations. By defining predetermined depth thresholds, the algorithm can effectively filter out insignificant variations, focusing solely on potential pothole formations.

The crux of the detection process lies in the application of Depth-First Search (DFS) algorithms. DFS systematically traverses the LIDAR data matrix, meticulously probing neighboring points to identify contiguous clusters indicative of potential potholes. Through recursive exploration, DFS efficiently navigates the matrix, evaluating depth values against predefined thresholds to discern clusters that exhibit characteristics synonymous with potholes. This algorithmic approach allows for the systematic identification of regions of interest, laying the groundwork for subsequent analysis.

Within the DFS logic, each point in the LIDAR matrix undergoes rigorous scrutiny to ascertain its eligibility to contribute to a potential pothole cluster. Points exceeding specified depth thresholds are earmarked as part of potential pothole clusters, guiding the algorithm towards regions warranting further investigation. Moreover, the algorithm incorporates checks to ensure that detected potholes are substantial enough to warrant attention. By evaluating the size and depth of potential potholes against predefined criteria, the algorithm can differentiate between minor irregularities and significant road surface anomalies. Upon the detection of potential pothole clusters, the C++ implementation undertakes further refinement to generate actionable insights. This refinement involves computing essential metrics such as average depth and spatial distribution within detected clusters, providing a comprehensive understanding of the identified irregularities.

Table 8. Resource Utilization

Name	FF	LUT
Expression	0	18
Instance	539	614
Register	72	-
Multiplexer	-	190
Total	611	822
Available	106400	53200
Utilization(%)	~0	~1

Fig. 43. Output

5.3 Response

The response system was implemented as an integrated control loop where detection triggers specific serial commands ('h' for hump, 'r' for right turn) to the motor driver for real-time maneuvering. The response mechanism utilizes a serial communication link between the Jetson Nano and an Arduino microcontroller. Upon detecting a 'hump' (Class ID 1), the system transmits a specific control byte (b'h') to the Arduino, which executes a Pulse Width Modulation (PWM) duty cycle reduction to decelerate the vehicle. For lateral anomalies (potholes), the system calculates the object's centroid relative to the frame width to initiate an evasive steering maneuver (b'r' for a right turn if the object is in the left hemisphere). For implementing the response system Ubuntu 20.04.3 LTS was flashed to a SD card which was then used to boot the jetson nano. Then installed required depedencies like Opencv and Tensorflow.The detection model was trained on google colab and corresponding weights file was downloaded into local directory. This weights file was used for detection. In the response part there are 4 dc motors which are controlled using the arduino. The commands are given in such a way that if a pothole is detected to the left of the frame a deviation must be made to the opposite side thatThe vehicle is to the right and vise versa.If a hump is detetcted correspondingly the speed is reduced otherwise it will move forwardequipped with normal speed. The power supply for motors are given usinga 4-wheel drive system powered by a 12V DC battery with help of aand an L298N motor driver and the jetson is powered by 20000mah power bank.

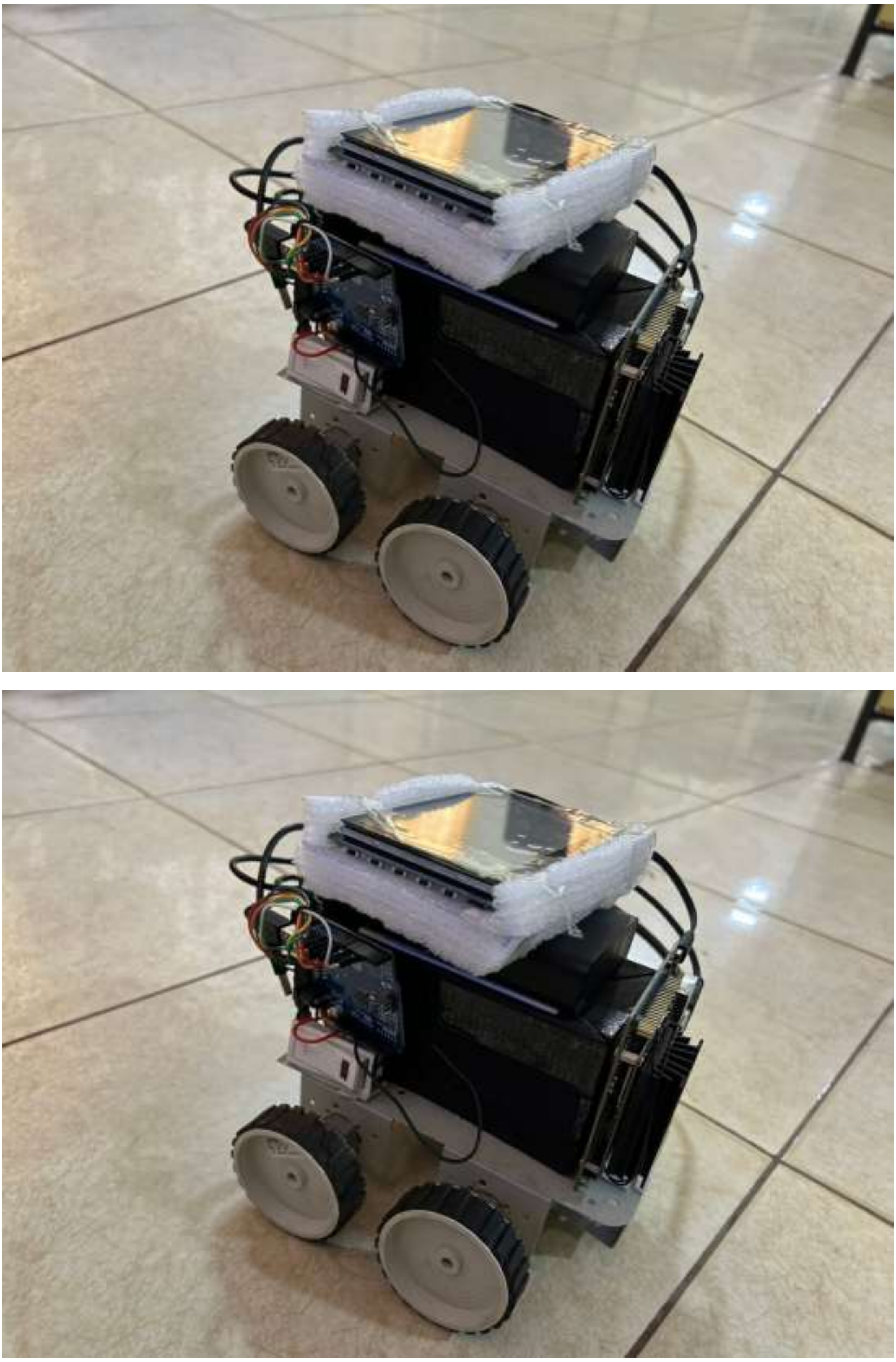

Fig. 5. Response Prototype

In evaluating the response capabilities of our autonomous vehicle prototype, the YOLO v4 model was employed to detect road attributes such as potholes and humps from video inputs. The model achieved a frame rate of 4.65 frames per second (FPS) during these detection tasks. This rate, although functional, presents certain limitations in terms of the vehicle's real-time Response efficacy in rapidly changing road conditions.Detection rate of 4.65 FPS is sufficient for the current prototype's low-speed indoor/controlled environment testing, future iterations will implement TensorRT optimization or model pruning to reach 15-20 FPS required for real-world road speeds [18]. Upon detection of a pothole on either side of the lane, the autonomous vehicle was programmed to take deviation with a speed of 50 RPM towards the opposite direction. Similarly, detection of a hump triggered a reduction in vehicle speed to 78.4 RPM and upon no detection the vehicle will move with a speed of 200 RPM. These responses were facilitated through a controlled interface between the Jetson Nano and an Arduino unit, which commanded the vehicle to take deviation and slow down systems based on the detection inputs.

Fig. 35: Delay per Frame

6. Conclusion

In conclusion, the comparative analysis of the YOLOv4, YOLOv6, YOLOv8, YOLOv10, VGG-16, and ResNet-34 models highlights their varied strengths and limitations in object detection and classification tasks. YOLOv8 proved to be the most effective for complex detection scenarios involving small and overlapping objects, offering a strong balance between precision and recall. YOLOv10 showed robustness in challenging environments, particularly under low-light conditions, as reflected by its mAP50-95 score. Conversely, VGG-16 and ResNet-34 excelled in classification tasks, achieving high accuracy, but faced challenges in object localization and dynamic detection.

When considering LiDAR data, localizing potholes presents a critical challenge in road safety and maintenance. Advanced models like YOLOv8 and YOLOv10 offer significant potential in detecting and localizing potholes within LiDAR point clouds, especially when dealing with small or overlapping potholes in complex road conditions. These models' strong performance can enhance traditional detection methods, improving both speed and accuracy. However, further improvements are needed to better handle LiDAR-specific characteristics, Such as point cloud density and spatial features to optimize pothole localization.

Overall, these results emphasize the importance of choosing the right model based on the specific task and environment. YOLO models, particularly YOLOv8 and YOLOv10, show great potential for real-time object detection applications. Meanwhile, VGG-16 and ResNet-34 provide excellent classification performance in more controlled settings. Future research should focus on refining these models to improve their detection capabilities and adaptability, particularly for LiDAR-fusion based pothole localization.

While the YOLOv4 model achieved a processing speed of 4.65 FPS on the Jetson Nano, which is sufficient for low-speed prototype validation, future work will focus on utilizing NVIDIA TensorRT and FP16 quantization to optimize the inference pipeline for higher-speed navigation.

References

[1] Anoop Thomas, Juergen Trost ,A Study on Implementing Autonomous Intra City Public Transport System in Developing Countries – India, Procedia Computer Science. Elsiever; p.375-382, 2017, https://doi.org/10.1016/j.procs.2017.09.093.

[2] X. Zhao, P. Sun, Z. Xu, H. Min and H. Yu, Fusion of 3D LIDAR and Camera Data for Object Detection in Autonomous Vehicle Applications,IEEE Sensors Journal, vol. 20, no. 9, pp. 4901-4913, May1, 2020, https:// doi.org/10.1109/JSEN.2020.2966034.

[3] Baek, J.-W.; Chung, K. Pothole Classification Model Using Edge Detection in Road Image. Appl. Sci. 2020, 10, 6662. https://doi.org/10.3390/app10196662

[4] Daejun Kang, Dongsuk Kum, Camera and Radar Sensor Fusion for Robust Vehicle Localization via Vehicle Part Localization, IEEE Access PP(99):1-1, April 2020, https://doi.org/10.1109/ACCESS.2020.2985075

[5] Aparna, Yukti Bhatia, et al. , Convolutional neural networks based potholes detection using thermal imaging, Journal of King Saud University - Computer and Information Sciences, Volume 34, Issue 3,2022, Pages 578-588, https://doi.org/10.1016/j.jksuci.2019.02.004.

[6] Bochkovskiy, A., Wang, C.-Y., Liao, H.-Y. YOLOv4: Optimal speed and accuracy of object detection., Apr 2020. , https://doi.org/10.48550/arXiv.2004.10934

[7] Li, C., Li, L., Geng, Y., Jiang, H., Cheng, M., Zhang, B., Ke, Z., Xu, X., Chu, X. ,YOLOv6 v3.0: Ab full-scale reloading., 2023. https://doi.org/10.48550/arXiv.2301.05586

[8] Jocher, G., Chaurasia, A., Qiu, J. ,Ultralytics YOLOv8.,2023.https://github.com/ultralytics/ultralytics

[9] Simonyan, K., Zisserman, A., Very deep convolutional networks for large-scale image recognition., arXiv 1409.1556 ,Sep 2014 https://doi.org/10.48550/arXiv.1409.1556

[2] [10] He, K., Zhang, X., Ren, S., Sun, J. ",Deep residual learning for image recognition.". Proceedings of Of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2016): 770-778. https://doi.org/10.48550/arXiv.1512.03385

[3] [11] D. J. Bochkovskiy, A., Wang, C.-Y., Liao, H.-Y. "YOLOv4: Optimal speed and accuracy of object detection." (Apr 2020).

[4] Li, C., Li, L., Geng, Y., Jiang, H., Cheng, M., Zhang, B., Ke, Z., Xu, X., Chu, X. "YOLOv6 v3.0: A full-scale reloading." (2023).

[5] Jocher, G., Chaurasia, A., Qiu, J. "Ultralytics YOLOv8." (2023).

[6] Molina, R. S., Gil-Costa, V., Crespo, M. L., Ramponi, G. "High-level synthesis hardware design for FPGA-based accelerators: Models, methodologies, and frameworks." IEEE Access 10 (2022): 90429-90455.

[7] Prasad, V., Kumari, S., Suryatej, I. "Pothole detection using LiDAR." Advances in Automobile Engineering 10 (2021): 1-3.

[8] Yeong, D. J., G. Velasco-Hernandez, G., J. Barry, and J.,. W a l s h , J. "Sensor and sensor f u s i o n technology in autonomous vehicles: A review.", Sensors, vol. 21, no. 6 (2021):, p. 2140., 2021.

[9] Carballo, A., Lambert, J., Monrroy, A., Wong, D., Narksri, P., Kitsukawa, Y., Takeuchi, E., Kato, S., Takeda, K. "LIBRE: The multiple 3D LiDAR dataset." Proceedings of the IEEE Intelligent Vehicles Symposium (IV) (2020): 1094-1101.

[10] Kosinski, Robert J. “A literature review on reaction time.” Clemson University 10, no. 1(2008): 337-344

[11] S. https://doi.org/10.3390/s21062140 [12] S. A. Talha, D. Manasreh, and M. D. Nazzal, "The use of LiDAR and artificial Intelligence algorithms for Detection and size of potholes”, Buildings, vol. 14, no. 4, p. 1078,(,2024), , https://doi: .org/10.3390/buildings14041078.

[12] Z. [13] Z. Duan, J. Shao, M. Zhang, J. Zhang, and Z. Zhai, "A small-object-detection algorithm based on LiDAR point-cloud clustering for autonomous vehicles" , Sensors, vol. 24, no. 16,p. 5423, (2024) https://doi:.org/10.3390/s24165423

[14] Hasanujjaman, Muhammad, Mostafa Zaman Chowdhury, and Yeong Min Jang. Sensor Fusion in Autonomous Vehicle with Traffic Surveillance Camera System: Detection, Localization, and AI Networking[13] D. J. Yeong, G. Velasco-Hernandez, J. Barry, and J. Walsh, “Sensor and sensor fusion technology in autonomous vehicles: A review,”, Sensors, vol. 21 , 2023, no. 6, p. 2140,(2021).: 3335. https://doi.org/10.3390/s23063335

[14] Anoop Thomas, Juergen Trost ,”A Study on Implementing Autonomous Intra City Public Transport System in Developing Countries – India”, 7th International Conference on Advances in Computing & Communications, ICACC (2017), Procedia Computer Science. Elsiever; p.375-382. [15] Anoop Thomas , Jobin K Antony , Ananthu Aji , Anna Koshy , Bhavana Krishna, Caroline Jaibee, “A Novel Method to Detect Potholes Using Yolo for Autonomous Vehicles”,, In Proceedings of the International conference on Sustainable technology and Innovation for Industry-icstii (2022)ICSTII, April 21-22, 2022, CUSAT, Kochi, Kerala, India. ISBN No 978-81-954804-2-5.

[16] Thomas, A., Antony, J.K., Isaac, A.V. et al. ”A novel road attribute detection system for autonomous vehicles using sensor fusion.”. Int. j. inf. tecnol. 2024).. https://doi.org/10.1007/ s41870-024-02255-5

[17] Anoop Thomas, Jobin K Antony, Sarath Chandran S, “Detection of Road Attributes using Solid-State Solid-State LiDAR”,, ICACC, (2024) IEEEExplorer https://doi: .org/10.1109/ICACC63692.2024.10845297

[18] Kosinski, Robert J. ,A literature review on reaction time., Clemson University 10, no. 1(2008): 337-344 Localization, and AI Networking" Sensors (2023), no. 6: 3335. https://doi.org/10.3390/s23063335