Research Article Available online www.ijsrr.org ISSN: 2279–0543

International Journal of Scientific Research and Reviews

Pothole Detection on Roads for Autonomous Vehicles Using Deep Learning

S.Kavitha[1], K.R.Baskaran[2] and P.Janani[3]

[1]Department of Information Technology, Kumaraguru College of Technology, Coimbatore, 641049, kavitha.s.it@kct.ac.in
[2]Department of Computer Science and Engineering, Kumaraguru College of Technology
[3]Department of Computer Science and Engineering, Kumaraguru College of Technology

ABSTRACT

Pothole is a large fissure in the earth. Potholes usually lead to underneath caverns and pits. Potholes in the road make the ride difficult. When it is driven by human it can be seen and moved away but when it comes for self-driving or autonomous vehicles, it is to be detected and avoided. Pothole detection is very important task in the self-driving vehicles. To detect potholes, autonomous vehicles use a variety of technologies. In this research various CNN deep learning architectures, AlexNet, LeNet and proposed Manual Network (MaNet) are trained to detect the potholes on the road surface. Based on the training accuracy best model MaNet was selected for deployment in the Arduino board with the motor. During the testing if the plain surface image is given as the input the motor run in the normal speed. If the pothole image is given the motor speed is reduced. The resulting slow speed of the motor proves that the model can be incorporated in the autonomous vehicle system to detect Potholes in an efficient manner.

KEYWORDS: Pothole detection, Autonomous Vehicles, Deep Learning, MaNet, AlexNet, LeNet

*Corresponding author:

S.Kavitha

Assistant Professor,
Department of Information Technology,
Kumaraguru College of Technology, Coimbatore, 641049,

Email Id - kavitha.s.it@kct.ac.in

INTRODUCTION

The state of our road infrastructure significantly influences public safety, vehicular maintenance costs, and overall transportation efficiency. Potholes, a common road defect, pose substantial risks to road users and can result in severe accidents and vehicle damage. Timely detection and repair of potholes are crucial for maintaining road safety and preserving infrastructure integrity.Traditional methods of pothole detection rely heavily on manual inspections, often resulting in delayed identification and response. To address this challenge, emerging technologies such as computer vision and deep learning have paved the way for automated pothole detection systems. These systems leverage advanced image processing techniques and predictive analytics to identify and categorize potholes in real-time.

This research paper delves into the methodologies, challenges, and advancements in pothole detection using a synergistic approach of computer vision and deep learning. By automating the detection process, these systems offer a proactive solution to the perennial issue of potholes in road networks. The integration of data-driven technologies not only enhances the speed and accuracy of identification but also facilitates a more efficient allocation of resources for prompt repairs.

The objectives of this research are twofold: first, to comprehensively review the existing literature on pothole detection methodologies; second, to use a deep learning architectures AlexNet, LeNet and MaNet for extracting features and to decide the best model for deployment in the edge device Through this research, we aim to contribute valuable insights to the field of transportation infrastructure management and advance the state-of-the-art in automated pothole detection.As the world navigates towards smart cities and connected transportation systems, the significance of robust and automated pothole detection cannot be overstated. This research paper endeavours to provide a solid foundation for further advancements in this critical area, ultimately contributing to safer and more sustainable road networks.

LITERATURE REVIEW

Road infrastructure plays a pivotal role in the development and functioning of societies, connecting regions, facilitating commerce, and ensuring the smooth flow of daily life. However, the integrity of roads is constantly challenged by various factors, with potholes being a recurrent and potentially hazardous issue. Detecting potholes promptly is crucial for effective road maintenance, ensuring the safety of commuters and minimizing repair costs. In recent years, the intersection of computer vision and deep learning has emerged as a promising avenue for automating the detection of potholes.

Authors[1] haves systematically investigated methods for spotting road defects. At the outset, the researchers provide a concise overview of the area before moving onto categorise existing approaches. They formulated and examined two techniques utilizing stereo-vision analysis for assessing road conditions ahead of a vehicle. Additionally, they crafted two deep-learning models specifically tailored for pothole detection. This study encompasses an empirical assessment of the four developed methods, and conclusions are derived regarding the distinct advantages offered by each of these approaches.

A robust pothole detection algorithm achieves high accuracy while maintaining computational efficiency[2.] The algorithm begins by transforming a dense disparity map to enhance the differentiation between damaged and undamaged road areas. To enhance the efficiency of this disparity transformation, they employed the golden section search and dynamic programming techniques to estimate transformation parameters effectively. Otsu's thresholding method is applied to extract potential undamaged road areas from the transformed disparity map. The disparities within these extracted regions are modeled using a quadratic surface, employing least squares fitting. Accurate pothole detection is achieved by comparing the disparity maps' actual and modeled differences. Finally, the point clouds corresponding to the detected potholes are extracted from the reconstructed 3D road surface. Experimental results demonstrate a high detection accuracy of approximately 98.7%, with an overall pixel-level accuracy of around 99.6% .

Authors[3] proposed an adaptive speed navigation method with a focus on crack detection for autonomous vehicle path planning. This approach automatically adjusts the vehicle's speed in areas where road cracks are detected. Using images of the road environment, they developed an image processing algorithm to accurately locate cracks. Leveraging this information along with obstacle data, they introduced a Bat-Pigeon algorithm (BPA) for the adjustable speed navigation of autonomous vehicles. The image processing algorithm precisely identifies crack locations, enabling the BPA to navigate the vehicle with adjustable speeds. The BPA integrated the global search capability of the Pigeon-inspired optimization (PIO) algorithm with the local search of the Bat algorithm (BA). This combination enhances speed and convergence algorithm performance. The proposed method guides the autonomous vehicle to decelerate around small crack areas while planning collision-free paths with minimal travel time.

The viability and precision of using infrared imaging, specifically convolutional neural networks (CNN), for pothole identification was explored[4]. Traditional methods like vibration-based sensors and manual reporting have limitations. The CNN-based ResNet model developed achieved an impressive 97.08% accuracy in identifying potholes, outperforming pre-trained models. The system, utilizing thermal imagery, offers advantages such as enhanced accuracy, cost-effectiveness, simplicity, night and fog functionality, and safety for drivers. Beyond mere detection, the model can assess pothole severity, aiding in prioritizing areas for immediate repair. This innovation holds promise, especially for resource-constrained nations in Africa and Asia, revolutionizing road maintenance with AI-driven techniques.

Authors[5] introduced a novel method for comprehensive and precise measurement of pothole properties using 3D line laser technology, offering a high-resolution pavement profile. Through a laboratory test on a simulated pothole, the proposed method demonstrated reliability with a relative error ranging from 3.66% to 4.78% and excellent repeatability. The study suggests that the longitudinal spacing between adjacent transverse profiles significantly influences pothole property measurements, advocating for an optimum spacing of less than 2 mm for accurate results. Field tests on interstate way 107 in Xi'an, Shaanxi Province, confirmed the method's practical accuracy, with a relative error of less than 8%.

Authors[6] proposed an automated method for the detection and assessment of these distress types in real-life video clips from Indian highways. Utilizing image processing techniques and heuristic decision logic, the proposed method quantifies potholes, cracks, and patches. Implemented using the OpenCV library, the method demonstrates robustness and efficiency, outperforming previous approaches and current practices. The extracted information can be instrumental in determining maintenance levels and facilitating timely repair and rehabilitation of Indian roads.

The detection and classification of pavement distress types, such as cracks, potholes, patches, and pavement markings, to reduce manual inspection costs was focussed[7]. The methodology employed Convolutional Neural Networks (CNN) and a low-cost video data collection strategy. Distress types are categorized, and the models are trained and tested on an image dataset from Montreal's road pavements. The proposed approach achieves a promising 83.8% detection rate and classification accuracy over the test set. Notably, F1-scores for individual classes are high, with improvements seen by merging certain crack classes. They demonstrated the potential of CNNs for automated pavement distress analysis.

Authors[8] proposed a system that aims to automatically recognize potholes on both muddy and highway roads using deep learning algorithms. Image datasets from muddy roads and highways are collected from internet sources and Kaggle. Pretrained models, including Resnet50, InceptionV2, and VGG19, are utilized for training the model. A web application is implemented to test the model's ability to identify roads with or without potholes. Performance analysis reveals that the VGG19 model achieves the highest accuracy, with 97% for highway roads and 98% for muddy roads, surpassing Resnet50 and InceptionResNetV2 models. This system offers an efficient way to detect potholes and prevent road disaster.

Authors[9] proposed a method for the simultaneous detection and recognition of traffic signs and potholes. Unlike previous approaches that often focus on either pothole detection or traffic sign recognition separately, this work presents a unified model designed specifically for Indian roads. The model utilizes optimized features extracted from road traffic signs, employing Hybrid Features from Accelerated Segment Test and Random Sample Consensus algorithms. Pothole detction involves the use of an improved Canny Edge detector and a bio-inspired Contour detection method. The Support Vector Machine classifier is then employed for the classification of both potholes and traffic signs. Experimental results demonstrate the superior performance of the proposed unified model in terms of accuracy, sensitivity, specificity, Matthew’s correlation coefficient, and F1-Score values compared to existing models.

Authors[10] focused on developing a fissure detection algorithm for concrete surfaces using transfer learning. Addressing challenges such as limited labeled data, feature transferability, domain shift, model bias, and complexity, the dataset consists of 40,000 evenly distributed images with and without cracks. Comprehensive deep learning models, built upon existing pre-trained models like EfficientNetB0, ResNet50, InceptionResNetV2, VGG19, Xception, DenseNet201, MobileNetV2, and InceptionV3, are trained through transfer learning. The proposed InceptionV3 model, augmented with three additional dense layers and one pooling layer, achieves an impressive accuracy score of 98.78%. Incorporating Explainable AI to interpret model results, this research offers improved explanation and classification accuracy.

Feature Fusion Network (FFN) is introduced, enhancing traditional deep network models. Transfer learning models, including VGG16, InceptionV3, and ResNet50, are employed on the Northeastern University-DETection (NEU-DET) Dataset[11], significantly reducing training time. Data augmentation, utilizing a Generative Adversarial Network, enhances input images. An Explainable Artificial Intelligence (XAI) classifier aids in understanding surface defect classifications. The proposed Hybrid FFN (HFFN), combining features from pre-trained networks, achieves high precision, recall, f-score, and accuracy of 98.65%, 98.42%, 98.51%, and 98.54%, respectively, on the NEU-DET dataset. The Particle Swarm Optimization (PSO) algorithm is applied to reduce features in the HFFN, resulting in a highly effective flaw classification system.

PROPOSED MODEL

Fig1shows the flow diagram of the proposed system.

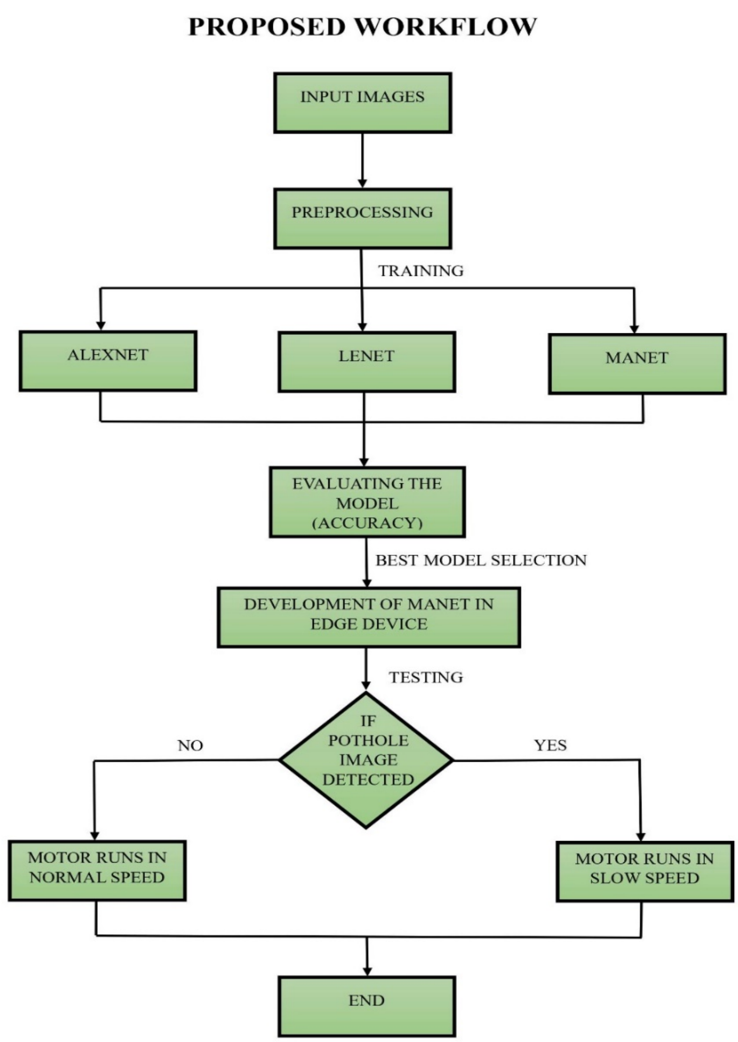

Figure 1. Proposed Workflow

Data Set

To detect potholes in the road surface, the data utilized in this study is sourced from Kaggle and the given link[12] (provides access to the dataset. Precisely, the Kaggle dataset is divided into two sets: a training folder and a valid folder, each comprising two classes – plain and pothole This dataset encompasses approximately 370 images. An illustration of plain and pothole images is presented in Fig 2. The dataset is partitioned into training and validation sets at an 70/30 ratio for the experiment, which is conducted in Google Colab.

Figure 2.Sample Plain and Pothole Images

AlexNet

AlexNet is a convolutional neural network (CNN) architecture designed for image classification. It was developed by Alex Krizhevsky, Ilya Sutskever, and Geoffrey Hinton, and it won the Image Net Large Scale Visual Recognition Challenge (ILSVRC) in 2012. AlexNet played a pivotal role in popularizing deep learning, demonstrating the effectiveness of deep neural networks in image recognition tasks. AlexNet architecture as shown in Fig 3 consists of eight layers, including five convolutional layers followed by max-pooling layers and three fully connected layers. It employs the rectified linear unit (ReLU) activation function. The convolutional layers in AlexNet use filters of various sizes to capture different levels of abstraction. These layers are designed to detect features at different scales, from simple edges and textures to complex pattern.

AlexNet incorporates LRN, a normalization technique that enhances the model's ability to generalize by normalizing responses across adjacent channels-pooling layers are employed to down sample the spatial dimensions of the feature maps, reducing computation, and providing a form of translation invariance. The final layers of AlexNet are fully connected, leading to a softmax output for image classification. These layers contribute to learning high-level features and making predictions. To prevent overfitting, AlexNet uses dropout regularization in the fully connected layers during training. Dropout randomly drops some neurons during each training iteration.

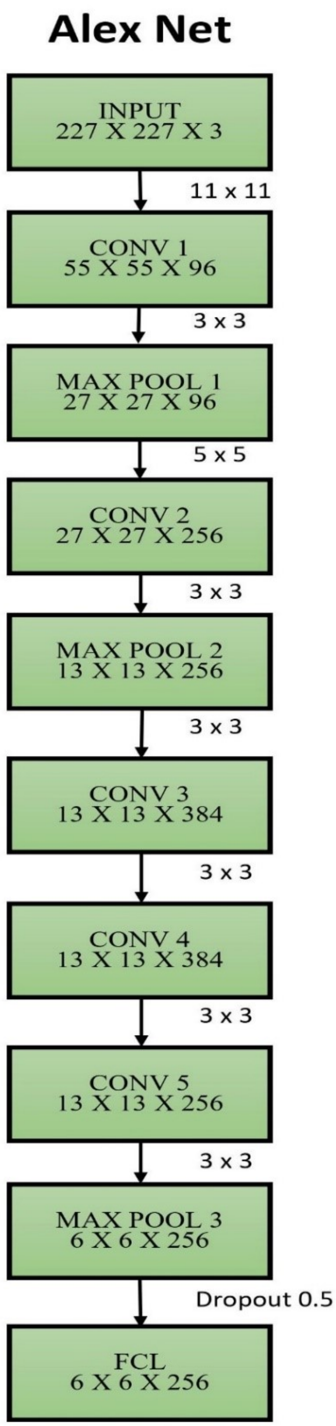

Figure 3. AlexNet Architecture

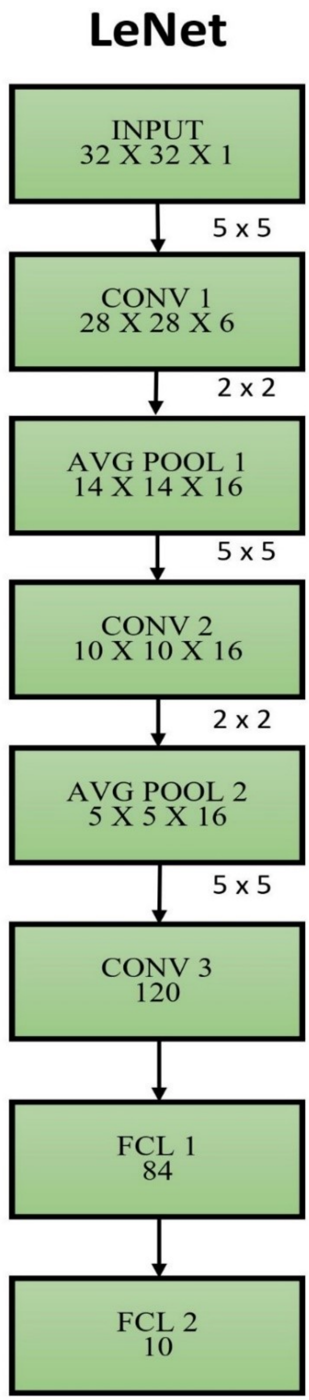

Figure4. LeNet Architecture

LeNet

LeNet, short for LeNet-5, is a convolutional neural network (CNN) architecture designed for handwritten digit recognition. It was developed by Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner and was introduced in 1998. LeNet is one of the pioneering CNN architectures and played a crucial role in the development of deep learning for image recognition.

LeNet-5 as shown in Fig 4 consists of seven layers, including three convolutional layers, two subsampling (pooling) layers, and two fully connected layers. The first layer of LeNet is a convolutional layer that applies convolutional operations to input images. The convolutional layers use small filter sizes, such as 5x5 or 3x3, to capture local features. After each convolutional layer, LeNet incorporates subsampling layers, which perform spatial Down sampling by taking the maximum or average value in a local region. This helps reduce the spatial dimensions and the computational load.

LeNet uses the sigmoid activation function in the convolutional and fully connected layers. The choice of the sigmoid function was common in neural network architectures of that era. The final layers of LeNet are fully connected layers, like a traditional neural network. These layers contribute to high-level feature learning and classification. The output layer of LeNet uses the SoftMax activation function, allowing the network to output probabilities for different classes.

MaNet

The proposed Manual Net Architectureis shown in Fig 5.

Layers:

Convolutional Layers (Conv Layers): There are 6 convolutional layers in the architecture. Each convolutional layer involves applying convolutional filters to the input data. Maxpooling Layers There are 5 maxpooling layers. Maxpooling is a downsampling operation that reduces the spatial dimensions of the input volume. There are 2 fully connected layers. Fully connected layers connect every neuron in one layer to every neuron in the next layer. The architecture includes a softmax layer, typically used in the output layer for multi-class classification problems. Softmax converts the raw output scores of the network into probability distributions over multiple classes.

Activation Function:

ReLU (Rectified Linear Unit): ReLU is the activation function used in each convolutional layer. It introduces non-linearity to the model by replacing all negative values in the feature map with zero.

Optimizer:

Adam Optimizer: The Adam optimizer is chosen for updating the weights during the training process. Adam is an adaptive optimization algorithm that adjusts the learning rates of each parameter individually, leading to faster convergence and better performance in many cases.

Overall Workflow:

The neural network follows a sequential architecture where the input data passes through multiple layers, including convolutional layers for feature extraction, maxpooling layers for downsampling, and fully connected layers for classification.

ReLU is applied after each convolutional layer to introduce non-linearity.

The Softmax layer is used for the final classification, providing probability distributions over the output classes.

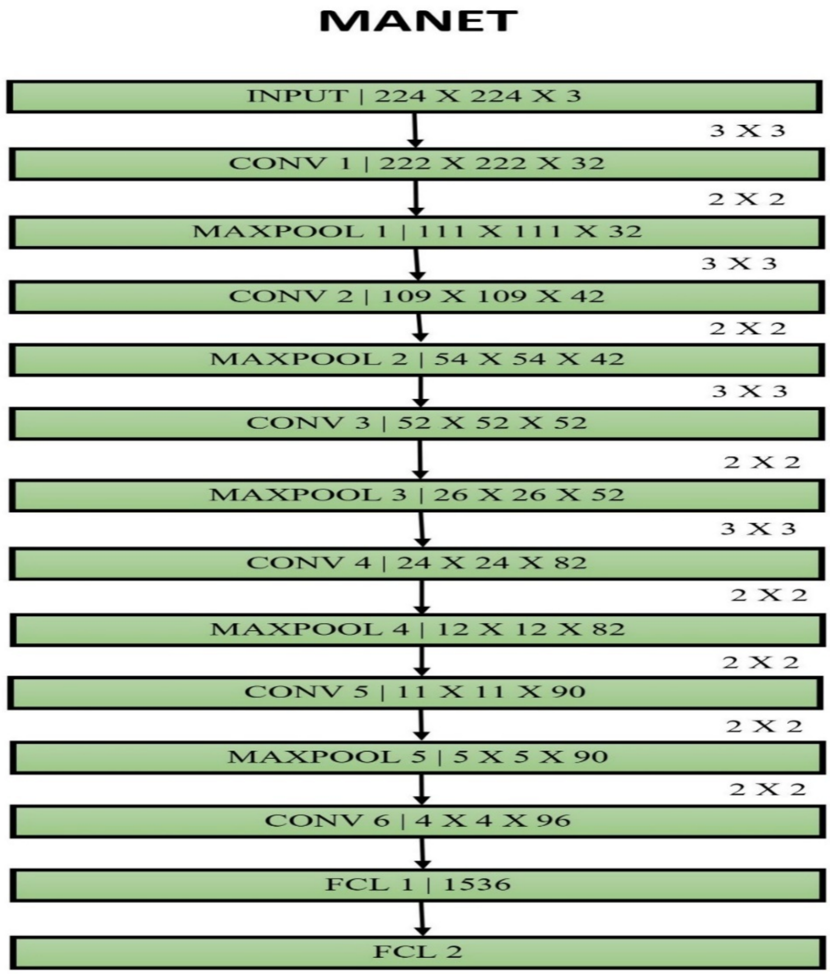

Figure5. MaNet Architecture

RESULTS AND DISCUSSION

In this work the pothole images are trained using three different architectures AlexNet, LeNet and MaNet (proposed model). Three models are trained for 60epochs with Adam optimizer and loss is calculated using categorical cross entropy . The performance of the models is evaluated with accuracy.

AlexNet

AlexNet was appertained network trained on the Image Net dataset, which consists of millions of labelled images across 1000 categories. Training was performed using stochastic gradient descent (SGD) with momentum. Data parallelism was applied, splitting the training data across multiple GPUs. To train the pothole image dataset, the weights of AlexNet is used to extract the features. The model's performance was evaluated with the accuracy measure. For the pothole image dataset, the accuracy obtained was 49%.The accuracy and loss of the AlexNet during training and testing is shown in the Fig 6.

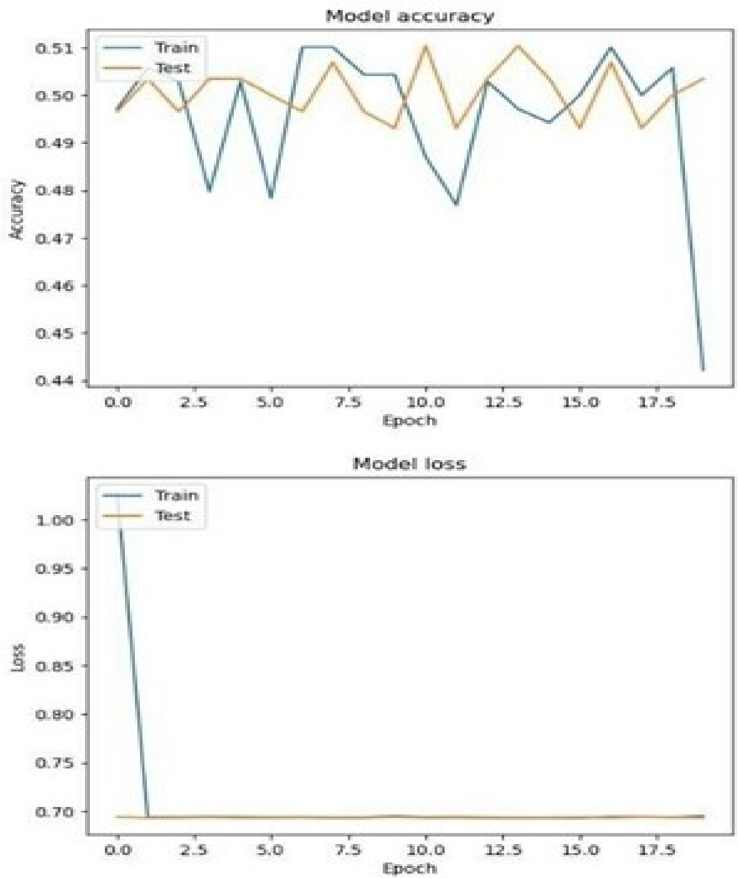

Figure6. Accuracy and Loss of AlexNet Architecture

LeNet

LeNet was designed for digit recognition, its principles influenced the development of subsequent CNN architectures for broader image classification tasks. Sigmoid and tanh activation functions were used in the LeNet architecture. It is trained using backpropagation and gradient descent to minimize the classification error. To train the pothole image dataset, the weights of LeNet is used to extract the features. The model's performance was evaluated with the accuracy measure. For the pothole image dataset, the accuracy obtained was 89%.The accuracy and loss of the LeNet during training and testing is shown in the Fig 7.

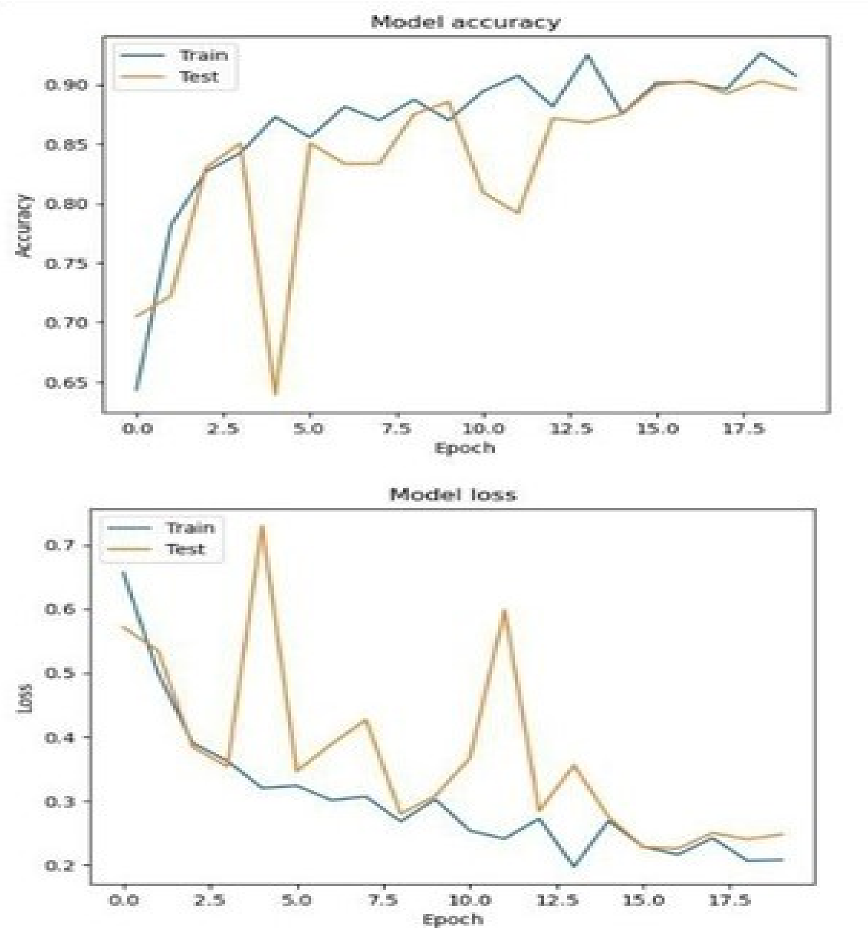

Figure7. Accuracy and Loss of LeNet Architecture

MaNet

The Manual Net Architecture comprises six convolutional layers, five max-pooling layers, two fully connected layers, and one SoftMax layer. Each convolutional layer is composed of convolutional filters and employs the ReLU non-linear activation function. The process of max pooling is implemented through dedicated pooling layers. The chosen optimizer, known for its optimal performance and higher accuracy, is the ADAM optimizer. Increase in the number convolutional layers in the proposed model increases the performance. The accuracy obtained by the MaNet is 95% and the model lite to deploy in the edge device. The accuracy and loss obtained by the model was shown in the Fig 8.

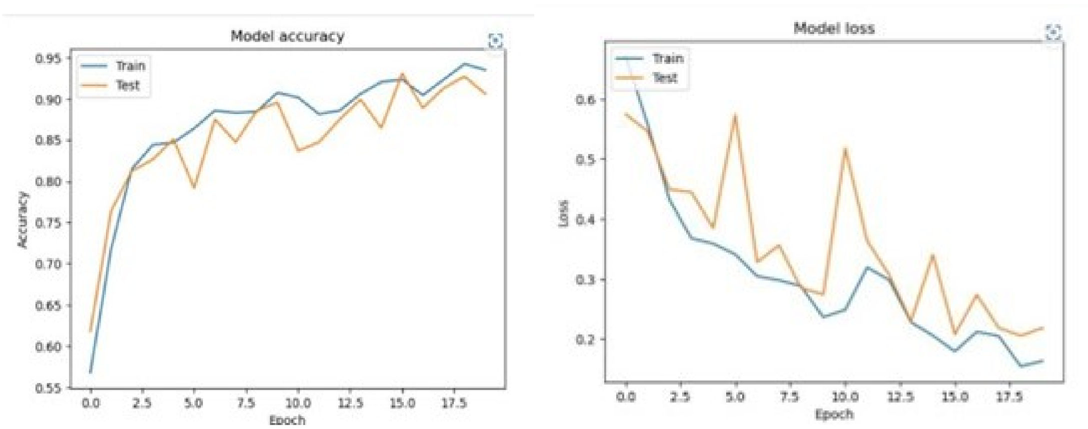

Figure8. Accuracy and Loss of MaNet Architecture

Table 1 shows the comparative analysis of architecture based on the Convolution layers used in the deep learning models.

Table 1. Comparative Study of the Architectures

Architecture	Layers	Accuracy
AlexNet	5convolutionallayers,3max-poolinglayers, 2normalisationlayers,2fullyconnectedlayers and 1 soft max layer	**49.83**
LeNet	3convolutionlayers,2averagepoolinglayers, 2 fully connected layers and 1 softmaxclassifier.	**89.50**
ManualNet	6ConvolutionalLayers,5Max-pooling layers, 2 Fully connected layers and 1Softmaxlayer.	**95.00**

Deployment of the Model

To deploy the selected proposed MaNet deep learning model, the model is first saved as hierarchical data format file(.h5). The model is deployed in the Arduino UNO connected with the Motor and LCD display. The deployment setup is shown in the Fig 9.

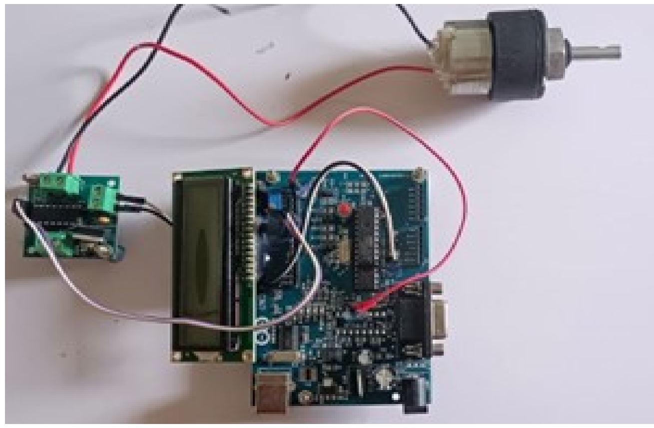

Figure 9. Deployment Setup

The model is tested in the device by processing the test image. If the input image is plain the motor runs in the normal speed. If the input image is pothole, then the motor runs in the slow speed. This model can be tested for the autonomous vehicles.

CONCLUSION

This research paper addresses the critical issue of pothole detection in the context of self-driving or autonomous vehicles. Potholes pose challenges for these vehicles as they require efficient detection and avoidance mechanisms. The study explores various deep learning architectures, including AlexNet, LeNet, and a proposed Manual Network (MaNet), for pothole detection using convolutional neural networks (CNN). The proposed MaNet architecture, with its six convolutional layers and five max-pooling layers, demonstrates a high accuracy of 95% in detecting potholes. The comparative analysis with AlexNet and LeNet highlights the efficacy of the MaNet model. The deployment of MaNet on an Arduino board with a motor suggests practical applicability for real-time scenarios, such as autonomous vehicles. The motor's response to plain and pothole images further validates the model's capability in detecting road surface conditions.

REFERENCES

1. Dhiman A, Klette R. "*Pothole detection using computer vision and learning*". IEEE Transactions on Intelligent Transportation Systems. 2019 Aug 5;21(8):3536-50.
2. Fan R, Ozgunalp U, Hosking B, Liu M, Pitas I. "*Pothole detection based on disparity transformation and road surface modelling*". IEEE Transactions on Image Processing. 2019 Aug 22;29:897-908.

3. Lei T, Luo C, Sellers T, Rahimi S. "*A bat-pigeon algorithm to crack detection-enabled autonomous vehicle navigation and mapping*". Intelligent Systems with Applications. 2021 Nov 1;12:200053.
4. Bhatia Y, Rai R, Gupta V, Aggarwal N, Akula A. "*Convolutional neural networks based potholes detection using thermal imaging*". Journal of King Saud University-Computer and Information Sciences. 2022 Mar 1;34(3):578-88.
5. She X, Hongwei Z, Wang Z, Yan J. "*Feasibility study of asphalt pavement pothole properties measurement using 3D line laser technology*". International Journal of Transportation Science and Technology. 2021 Mar 1;10(1):83-92.
6. Huidrom L, Das LK, Sud SK. "*Method for automated assessment of potholes, cracks and patches from road surface video clips*". Procedia-Social and Behavioral Sciences. 2013 Dec 2;104:312-21.
7. Zhang C, Nateghinia E, Miranda-Moreno LF, Sun L. "*Pavement distress detection using convolutional neural network (CNN): A case study in Montreal*", Canada. International Journal of Transportation Science and Technology. 2022 Jun 1;11(2):298-309.
8. Saisree C, Kumaran U. "*Pothole detection using deep learning classification method*". Procedia Computer Science. 2023 Jan 1;218:2143-52.
9. Lee SY, Le TH, Kim YM. "*Prediction and detection of potholes in urban roads: Machine learning and deep learning based image segmentation approaches*". Developments in the Built Environment. 2023 Mar 1;13:100109.
10. Kavitha S, Baskaran KR, Dhanapriya B. "*Explainable AI for Detecting Fissures on Concrete Surfaces Using Transfer Learning*". In2023 International Conference on Inventive Computation Technologies (ICICT) 2023 Apr 26 (pp. 376-384). IEEE.
11. Sundarrajan K, Rajendran BK. "*Explainable efficient and optimized feature fusion network for surface defect detection*". The International Journal of Advanced Manufacturing Technology. 2023 Jun 21:1-8.
12. https://www.kaggle.com/datasets/virenbr11/pothole-and-plain-rode-images