B. TECH PROJECT REPORT

On

ROAD DAMAGE DETECTION USING IMAGERY DATASETS AND ML ALGORITHMS

By

POKALA SAIRITHWIK

190004024

DISCIPLINE OF CIVIL ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY INDORE

November 2022

ROAD DAMAGE DETECTION USING IMAGERY DATASETS AND ML ALGORITHMS

A PROJECT REPORT

Submitted in partial fulfillment of the requirements for the award of the degrees

of

BACHELOR OF TECHNOLOGY

In

CIVIL ENGINEERING

Submitted by:

Pokala Sairithwik

Guided by:

Dr. Guru Prakash

(Assistant Professor, Civil Engineering)

INDIAN INSTITUTE OF TECHNOLOGY INDORE

November 2022

CANDIDATE'S DECLARATION

I hereby declare that the project entitled "**ROAD DAMAGE DETECTION USING IMAGERY DATASETS AND ML ALGORITHMS**" submitted in partial fulfillment for the award of the degree of Bachelor of Technology in 'Civil Engineering' completed under the supervision of Dr. Guru Prakash, Assistant Professor, Civil Engineering, IIT Indore is an authentic work.

Further, I/we declare that I/we have not submitted this work for the award of any other degree elsewhere.

P. Sai Rithwik. 28/11/2022

Signature and name of the student(s) with date

CERTIFICATE by BTP Guide(s)

It is certified that the above statement made by the students is correct to the best of my/our knowledge.

Guru Prakash 28/11/2022

Signature of BTP Guide(s) with dates and their designation

(Assistant Professor)

Preface

This report on "**ROAD DAMAGE DETECTION USING IMAGERY DATASETS AND ML ALGORITHMS**" is prepared under the guidance of Dr. Guru Prakash, Assistant Professor Civil Engineering.

Through this project, I have made some techniques to collect data from a smartphone mounted on a car dashboard and made a simple algorithm to classify the images, which are used to train a supervised machine learning model. Further, I use different machine learning techniques to train the data and choose the best of them as my model based on the obtained results like accuracy score, recall, f1-score, and precision.

I have tried to the best of my abilities and knowledge to explain the content in a lucid manner using tables, source codes, flow charts, etc.

Pokala Sairithwik

190004024

BTech 4th Year

Discipline of Civil Engineering

IIT Indore

Acknowledgments

I wish to thank Prof. Guru Prakash sir for his kind support, expertise and valuable guidance. He provided a perfect environment for critical thinking and research acumen and was always available for discussions, doubt clearance and guidance at every part of the project. He has constantly motivated me to take the project to its very culmination.

I also wish to thank Ph.D. Scholar Revanth Dugalam for his whole support and guidance throughout the project. He provided me with all the resources that are required for the study and was also available for me whenever I was in need of his guidance.

Without their support, this report would not have been possible.

Pokala Sairithwik

190004024

B.Tech. 4th Year

Discipline of Civil Engineering

IIT Indore

TABLE OF CONTENTS

TABLE OF FIGURES

TABLE OF TABLES

1. INTRODUCTION

1.1 INTRODUCTION:

The problem of potholes is not new. Road users are put at risk by potholes, which can severely harm the drivers and their vehicles. Pothole repairs are expensive and require careful budgeting. To commute, drivers seek a smooth, potholes-free road. For efficient traffic management, roads and highways must be maintained. Road and highway maintenance is crucial for a healthy economy and efficient traffic flow. Maintaining such a vast road network need both resources and experience. Over time, experts are utilized to inspect roads, and the responsible authorities fix these roads. These methods are time-consuming, costly, and labor-intensive, significantly increasing these experts' workload.

Additionally, this approach cannot keep up with the rising need for maintaining roads in good shape. As already mentioned, there is a significant delay in road damage repairs. The officials attributed the delay in fixing these potholes to a lack of funding and human resources. Both the number of drivers and the road network are expanding over time. As the advancement in technologies, an increasing number of systems take on autonomy and use artificial intelligence (AI) to provide better services to improve the convenience and safety of the consumer experience. New technology will inevitably be used to maintain the road's quality and safety to keep up with demand. Effective maintenance will be facilitated by the early detection of abnormalities in the road surface. By utilizing Deep Learning models, the most recent technological advancements enable more systems to provide pothole detection and autonomy. According to the recent studies, artificial neural networks has been used to examine and detect road surface degradation.

The accuracy of any deep learning model depends on the standard of the information set accustomed to train the model. Information assortment has been difficult within the past, requiring an ardent recording device to gather the data to develop Machine Learning Models. However, within the event of technologies, smartphones are present and might record smart imagery data. Machine Learning strategies are viable and efficient for paved surface damage detection. Hence, its imperative to automatise road surface damage detection using progressive Machine Learning techniques.

This thesis suggests a deep learning model that utilizes a fusion model based on Convolution Neural Networks to handle the problem of road potholes (CNN). Road potholes will be located using visual data from mobile phones as inputs to the fusion model. The server will employ a trained CNN model at the back end to spot potholes in the road and alert users who have signed up for this service. New potholes will be added to the road dataset in real-time as they are discovered. Road authorities can use the real time datasets of road damages and work on these damages to keep the road in good condition.

1.2 MOTIVATION:

The goal of the research is to improve our understanding of Deep Learning models and their practical applications that benefit society. Potholes cause many accidents and deaths in our country Fig 1 shows the roads with potholes. Various Deep Learning and Machine Learning approaches were investigated in support of this project, along with several real-world implementations. For the research effort, a well-known unresolved issue with pothole recognition was discovered. With the help of cutting-edge deep learning technology, this initiative intends to eliminate potholes by automating the monitoring of road surfaces. The initial study was done to find out what methods are currently used to monitor the state of roads in terms of road damage detection. The study found that several methods for identifying road irregularities had been researched throughout the years.

https://www.livemint.com/news/india/3564-accidents-in-india-due-to-potholes-in-2020-11639651532979.h

https://thefederal.com/news/potholes-caused-over-5000-deaths-shows-2018-2020-road-accidents-data/

Figure 1:Images showing potholes

1.3 LITERATURE REVIEW:

In (Bouilloud et al., 2009), To forecast France's road conditions, Bouilloud employed the reactions between Soil, Biosphere, and Atmosphere "Crocus" model. This model depends on long-term surface condition simulation utilizing spatialized meteorological data and short-term meteorological forecasts. The most recent advances in computer vision and Complex computing were made possible by the computational capability of modern computers and GPU. Deep learning that automatically finds the surface damages on the road. In DNN(Depth neural networks) are employed in many different contexts and have grown in prominence in the industry. A vast amount of data is necessary to develop models using deep neural networks. Using a CPU to train these models could be time-consuming and sometimes unprofitable.

In (Huidrom & Das 2013), Huidrom employed a vehicle-mounted charge-coupled device (CCD) camera to identify a road surface flay quickly. This technique uses a downward camera to record photos of the road surface from a moving vehicle. The essay advises utilizing the gloss of the road surface and computing the absolute deviation in light of the low luminance level. They assumed that the low brightness signal corresponded to the actual road surface. Higher levels, however, imply introspection. The study computes the luminance deviation and anticipates the state of the road.

(Steinkraus et al., 2006), On the ImageNet dataset, the researchers showed how they had trained one of the biggest Convolution Neural Networks to improve performance. The paper created a 2D convolution neural network implementation that was optimized. In this study, the scientists produced image data sets by scaling them to 256x256 and trained neural networks using the unprocessed RGB values of the pixels. The significance of activation functions and the time required for their training were also covered in the paper. Before describing how the author trained networks on multiple GPUs to shorten training time.

Shaoqing (2015) has talked about how region-based convolution neural networks (R-CNN) and the region proposal technique have helped with object detection. The region proposal algorithm, however, could be time- and money-consuming. Given the expense of the detection network, the study suggested a tweak in the algorithm where proposal computation is almost free. The approach described in the research has convolutional layers in common with cutting-edge object detection networks. The quicker R-CNN object detection system, which consists of two modules, is presented in the study. The fast-CNN detector is the second module, and the first is fully coupled to convolution neural networks that suggest regions.

Tsung (2017) talks about using feature pyramid networks to find objects. The feature pyramids are scale-invariant, which makes it difficult to distinguish things at various scales.FPN is memory- and time-intensive. Hence it has recently been avoided in Deep Learning techniques. The paper explains the drawbacks and practical impracticality of showing each level of a picture pyramid in an actual application. The impossibility of training end-to-end deep networks using an image pyramid is discussed. It creates uneven movement, tests time references, and necessitates extensive memory and time. The sliding window thrives the main topic of the paper.

Different deep neural network topologies and photos were used by Ukhwah et al. (2019) to facilitate automatic evaluation of the road surface and pothole detection. Their research showed that the Deep Learning networks delivered results quickly and with a fair amount of accuracy.

2. OBJECTIVE

Roads are constantly exposed to different weather conditions. Extreme weather events have a right away impact on our roads. Road surfacing will soften and expand under the sun's extreme heat in heat waves. In high rainwater floods and times of serious rainfall, pavements in need of maintenance can fall victim to the look of damages, potholes and different surface damages. so, we need to inspect roads frequently, which is impossible by traditional methods for a wide range of road networks in countries like India.

Project objective:

1. To build the best machine learning model to detect and classify different types of road surface defects based on Image data collected through available smartphones.
2. The project's goal is to gather enough high-quality and quantity imaging data to identify potholes. using unique deep learning-based methods to automatically recognise and locate potholes.
3. To investigate the problem of pothole detection and emphasize its importance for the economy, safety, and health.
4. To research cutting-edge machine learning methods for pothole identification.To locate various pothole detection-related existing data sets.

Figure 2:Flowchart of methodology

3. METHODOLOGY

- The first stage was to collect imagery data.
- Following the completion of the imaging data set, a Deep Learning model was created. To create a suitable model, the simulations were ran using various hidden layer and hyperparameter combinations.
- Prior to feeding the imaging data into the trained 2D-CNN model, the area of interest was chosen. The model's output was imagery data with pinpointed road potholes and their GPS coordinates.
- Fig 3 describes the methodology to train the 2D-CNN model.

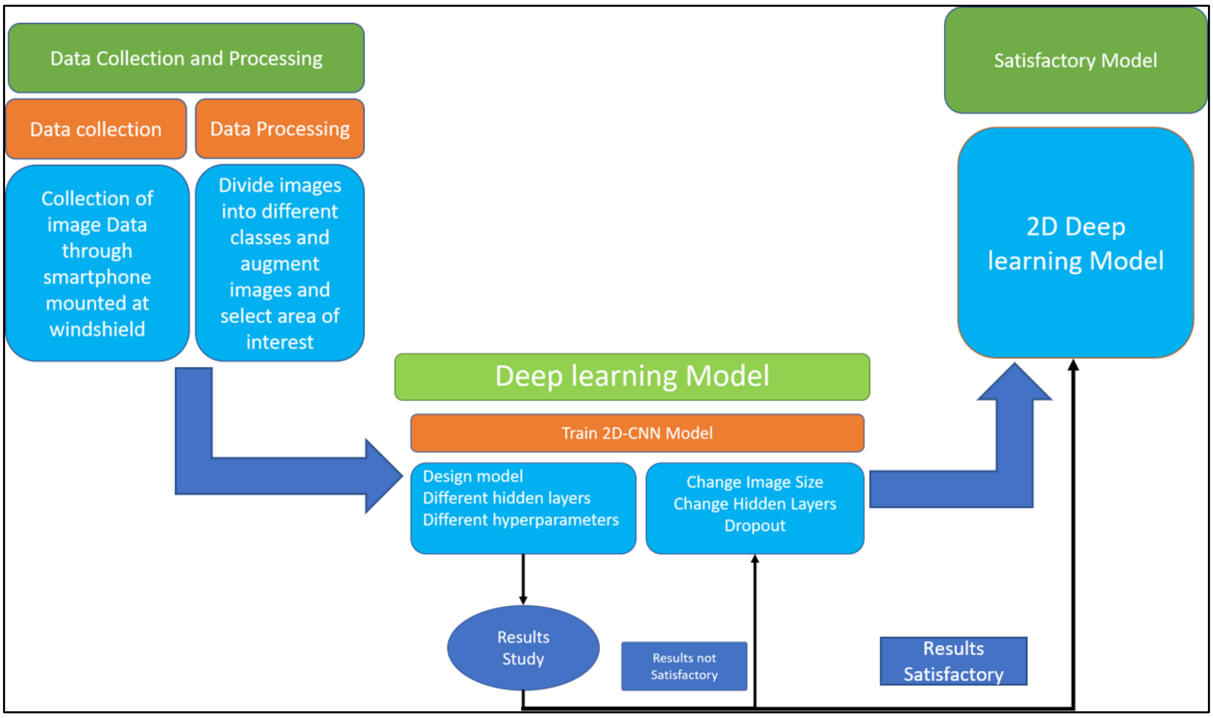


Figure 3:Methodology to train the 2d-CNN model on imagery data

4.FIELD STUDY

4.1 VEHICULAR MODEL:

For the field study we choose the car model of Maruti Suzuki Ertiga, uses MacPherson Struts suspension system in front that work in tandem with a Torsion Beam suspension at the rear. The suspension setup feels the defects of pavement when we ride on speed below 50 km/hr. The long wheel base contributes to the good stability and wide tyres also cushion the ride well and offer decent grip. The smartphone would be mounted on vehicle dashboard which captures the video of the road through the moving vehicle. The dashboard is strong enough that won't produce unwanted movements, since the smartphone is mounted on dashboard. The detailed information about car and its suspension system is mentioned in Table 1.

Table 1:Model design and suspension system of vehicle selected for study.

Engine Type	K15C Smart Hybrid
Length (mm)	4395
Width (mm)	1735
Wheel Base (mm)	2740
Gross Weight (Kg)	1785
Power Socket	12V
Dashboard	Sculpted with metallic Teak-Wooden finish
Front suspension	Mac pherson strut & coil spring
Rear suspension	Torsion beam & coil spring

4.2 DATA COLLECTION:

The Imagery data will be collected using a camera of a smartphone of high resolution, and the collected data are stored in the memory of the phone used and later transferred to the desktop of required specification that can run the selected model to perform the training. As seen in Fig 5, the mobile was carefully positioned on the dashboard of the car. The mobile's camera was employed to record videos, which were then converted into photos and utilised in our research. To achieve high accuracy in road damage detection, the location of the smartphone on the dashboard is essential. In order to have a clear view of the road, the smartphone was safely mounted in the centre of the dashboard's breadth with a clear vision of capture, as illustrated in Fig 5. The study's smartphone can capture 4K footage at 30 frames per second. In this investigation, potholes on the road will be located using GPS after being identified using a dashboard camera. Therefore, pothole data that was used in this project was gathered while the car was being driven at an average speed of 30 to 40 KMPH. But all kinds of roads were used to capture standard photographs of the road surface. The view of road from mounted smartphone is as shown in the Fig 6 .

Figure 4:Position of smartphone on the dashboard

Figure 5: The view of road from the phone.

4.3 DATA PRE-PROCESSING:

Depending on the model and makeup, the smartphone camera's captured photos and video frames have varying sizes. Images of 128x128 pixels were used as input in the 2-D convolution neural network in the study. The subsequent actions were taken to get the dataset ready. The images used in the dataset came from a variety of sources. As a result, the image sizes vary. Making all of these images the same size and removing any that wouldn't work for the experiment was the first step. The initial tests were carried out using photos that are 256x256 in size. Later, for research, the image size was modified to 128x128.To analyze the outcomes using the models. The result of adjusting the size of an image has been discussed in the results section. Images that were larger than the required size (256x256) were divided into two images, and then each image was manually reviewed to determine whether it was appropriate for the experiments. The next step was to accurately name these images after they had all been scaled down to 256x256. A key component of supervised learning is accurate data labelling. How accurately data is labelled impacts how accurate detection is. The photos were manually reviewed to filter out any high-motion blur or undesirable elements.

Two classifications were created out of the images with damages and no damages. After labelling the images accurately, certain image processing techniques were used to add to these images in order to increase the size of the dataset. The dataset was made carefully in which images from completely different weather conditions, together with a dry road and a hole full of water, likewise as different times of day to make sure for brightness variation. However, given the time And geographic constraints, data augmentation was applied to the dataset images so the dataset has more images for model evaluation. To get additional images, within a range of 0.2 to 1, the brightness of the photos in the dataset was changed. The image's brightness was adjusted to copy images taken at many points in time with variable brightness.

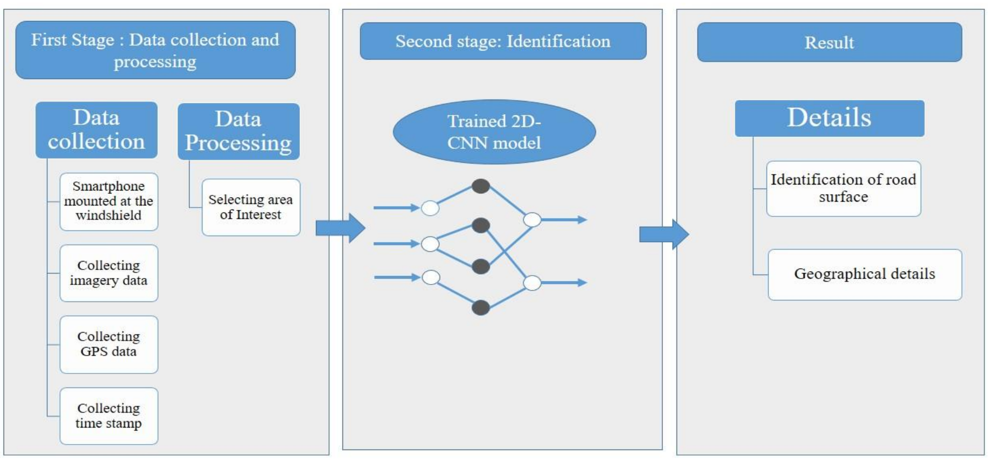


Figure 6: Methodology to detect road damages using image data with the 2D-CNN model.

Figure 7:Road surface with no damage.

Figure 8: Road surface with damage.

4.4 WHAT ARE CNN'S?

Convolutional Neural Networks, also known as CNNs, are a common type of artificial neural network used in Deep Learning for object/image recognition and categorization. Thus, Deep Learning uses a CNN to identify items in a picture. CNNs are important for a variety of activities and applications, including speech recognition in natural language processing, localization and segmentation tasks for computer vision, video analysis, and self-driving car obstacle detection.

The CNN algorithm is composed of the following components, which are organised in a particular workflow:

- Input Image
- Convolution Layer (Kernel)
- Pooling Layer
- Classification—Fully Connected Layer
- Architectures

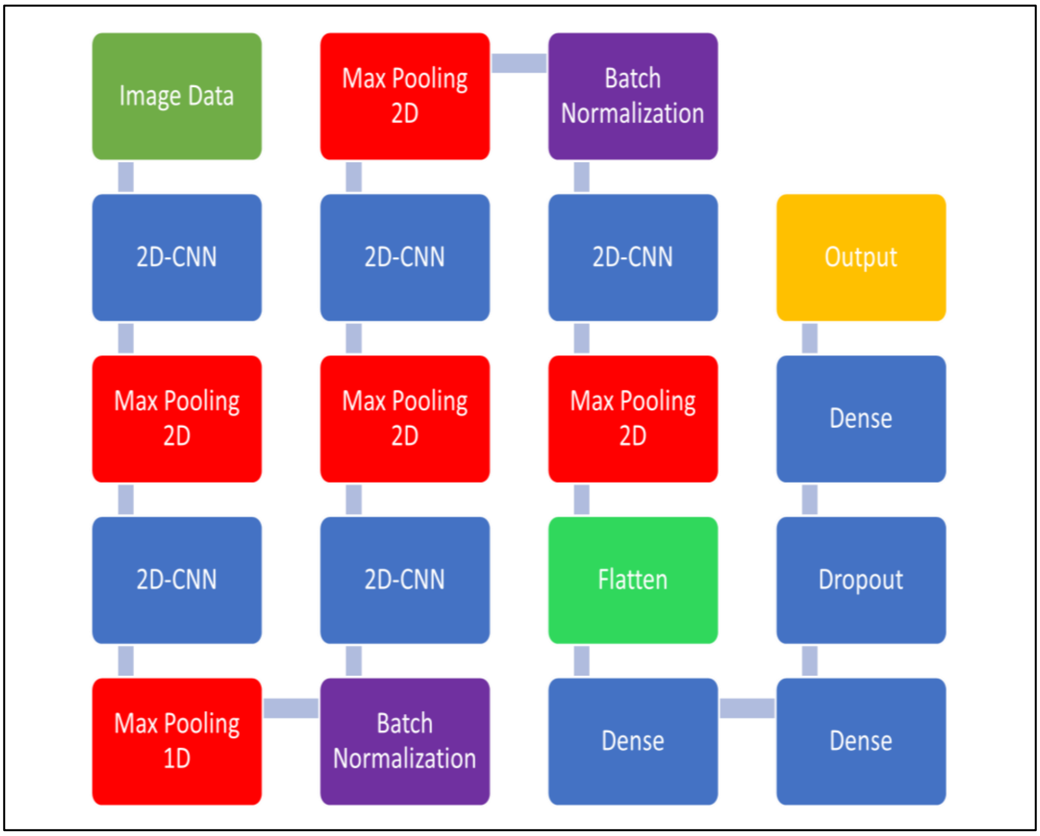


Figure 9: Convolution neural network model with five layers.

2D CNN Model:

Table 2: Table showing different layers in the CNN model.

Layer (type)	Output Shape	Param #
input_3 (Input Layer)	[(None, 120, 120, 3)]	0
input_3 (Input Layer)	(None, 118, 118, 16)	448
max_pooling2d_4 (MaxPooling2D)	(None, 59, 59, 16)	0
conv2d_5 (Conv2D)	(None, 57, 57, 32)	4640
max_pooling2d_5 (MaxPooling2D)	(None, 28, 28, 32)	0
global_average_pooling2d_2	(None, 32)	0
dense_2 (Dense)	(None, 1)	33

Total params: 5,121
Trainable params: 5,121
Non-trainable params: 0

4.5 TRAINING THE MODEL:

Firstly, we take the images collected in 2 different files damages as positive file and no damages as negative file.

we load in the file paths to the files into a data frame along with the labels and then we pass in that data frame to the generator and it will flow in the images through those file paths and give the directory names for the positive class and negative class naming positive dir and negative dir . now data frames have been created, for this we a function called generate data frame generate (df) it's going to take in a directory. Individual data frames of positive and negative directories are created and the both data frames are concated together into a single A sample of 600 photos selected from the available images is used to evaluate the model. These 600 images are then divided into test and train datasets. 30% of the data goes to the test module, 70% goes to the train module. We use some of the train data for validation purposes, taking a validation split of 0.2.

The final data split for the experiment is now 56% train, 14% validation, and 30% test.

this is referring to how many photos are in each dataset.

- 336 images in train dataset
- 84 images in validation dataset
- 180 images in test dataset.

The 600-image dataset was split into 3 parts: training (56%), testing (30%), and validation (14%), after which the CNN model was trained using a variety of hyperparameter combinations. A model with two layers classifies damage and no-damage was first examined for the investigation. A new model with three classes and five hidden layers was also taken into consideration. On hidden layers, the Rectified Linear Unit activation function was utilised, and on the output layer, SoftMax. A 32-person batch size was chosen, and training is done separately for each batch using 100 epochs and a dropout of 0.5. Through precision and recall, the model test accuracy was evaluated during the analysis.

Table 3: System configuration.

Processor	9th Gen Intel® Core™ i7-955U
Memory	16GB, 2x8GB, DDR4, 3200MHz
Graphics	Intel® Iris® Xe Graphics

Table 4: Deep learning model with 5 hidden layers and 2 classes.

Number of Layers in Model	5
Number of Classes	2(damage and no-damage)
Train Accuracy	87.62%
Test Accuracy	90.00%

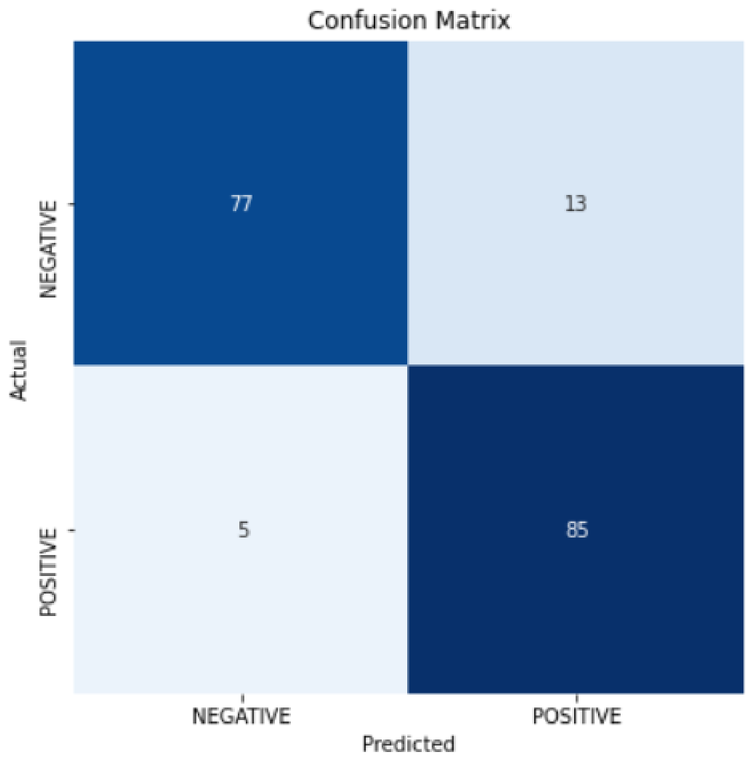
5. RESULTS

In order to train the identical dataset of two variables (damage and no damage), a convolution neural network model with five hidden layers was utilised. The training and testing outcomes for the Convolution Neural Networks model with hidden layers five are displayed in the Figur e. The test accuracy of this model ranged between 66% and 100%, with a median of 90.00%, whereas the training accuracy ranged between 57% and 93%, with a median of 87.62%.

Figure 10: Training and Validation Loss Over time.

Test Loss: 0.27871
Test Accuracy: 90.00%

Figure displays the model's confusion matrix for various combinations. With five hidden laye rs and two classes, the Convolution Neural Networks model has 90% F1-Score, 90% recall, a nd 90% precision. In light of the dataset's size and variety in image types, the outcome is resp ectable.



CLASSIFICATION REPORT:

The below Table 5 represents the classification report.

Table 5:Classification Report of the tested images using DL model.

	precision	recall	f-1 score	support
Negative	0.94	0.86	0.90	90
Positive	0.87	0.94	0.90	90
Accuracy			0.90	180
Macro _ average	0.90	0.90	0.90	180
Weighted _ average	0.90	0.90	0.90	180

The size of the dataset and the amount of variation in pothole type size, shape, and depth will affect how accurate the 2D-CNN model is.

6. CONCLUSION

In this study, a deep learning model has been trained for the detection of road anomalies. The performance of the proposed model has been compared with other models. During testing, the deep learning 2D CNN model has shown better accuracy for prediction of. In both the cases, the performance of deep learning 2D CNN model is higher than the other models. The prepared model was applied on field data, it gave better and almost precise number of road surface defects. Hence it is clear that deep learning 2D CNN model with the proposed features can be used for the health monitoring of longer road networks with semi-skilled labour and without procrastination. Further we can make proceedings like alerting about upcoming pothole in rainy season, since this project detected position of damages on roads precisely.

7. FUTURE WORK

The model has been trained on the training data set and has given satisfactory accuracy. The CNN model, which contains five layers and two classes, has been trained successfully. This model is trained only to identify whether there is a pothole in the given image. This model now has to be trained further on different data sets to classify the potholes into different categories, i.e., small and big ones. This classification of potholes is required to assess the road condition that needs repairs properly. More pictures, videos, and diverse data can be added to the database as a result of future research. Future research can investigate optimization techniques to identify the best model parameters and enhance the accuracy of the proposed deep learning model.

8. REFERENCES

1. Pandey, A.K., 2021. *Deep Learning Approach for Road Pothole Detection Using Accelerometer and Imagery Data* (Doctoral dissertation, Coventry University).
2. Bouilloud, L., Martin, E., Habets, F., Boone, A., Le Moigne, P., Livet, J., Marchetti, M., Foidart, A., Franchisteguy, L., Morel, S. and Noilhan, J., 2009. Road surface condition forecasting in france. *Journal of Applied Meteorology and Climatology*, *48*(12), pp.2513-2527.
3. Al-Saffar, A.A.M., Tao, H. and Talab, M.A., 2017, October. Review of deep convolution neural network in image classification. In *2017 International conference on radar, antenna, microwave, electronics, and telecommunications (ICRAMET)* (pp. 26-31). IEEE.
4. Gao, Q., Liu, J., Zhang, H., Hou, J. and Yang, X., 2012. Enhanced fisher discriminant criterion for image recognition. *Pattern Recognition*, *45*(10), pp.3717-3724.
5. Huidrom, L., Das, L.K. and Sud, S.K., 2013. Method for automated assessment of potholes, damages and patches from road surface video clips. *Procedia-Social and Behavioral Sciences*, *104*, pp.312-321.
6. Krizhevsky, A., Sutskever, I. and Hinton, G.E., 2017. Imagenet classification with deep convolutional neural networks. *Communications of the ACM*, *60*(6), pp.84-90.
7. Carr, T.A., Jenkins, M.D., Iglesias, M.I., Buggy, T. and Morison, G., 2018, July. Road damage detection using a single stage detector based deep neural network. In *2018 IEEE Workshop on Environmental, Energy, and Structural Monitoring Systems (EESMS)* (pp. 1-5). IEEE.
8. Ren, S., He, K., Girshick, R. and Sun, J., 2015. Faster r-cnn: Towards real-time object detection with region proposal networks. *Advances in neural information processing systems*, *28*.
9. Lin, T.Y., Dollár, P., Girshick, R., He, K., Hariharan, B. and Belongie, S., 2017. Feature pyramid networks for object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2117-2125).
10. Pandey, A.K., 2021. *Deep Learning Approach for Road Pothole Detection Using Accelerometer and Imagery Data* (Doctoral dissertation, Coventry University).
11. Xia, W., 2018, July. An approach for extracting road pavement disease from HD camera videos by deep convolutional networks. In *2018 International Conference on Audio, Language and Image Processing (ICALIP)* (pp. 418-422). IEEE.
12. Roychowdhury, S., Zhao, M., Wallin, A., Ohlsson, N. and Jonasson, M., 2018, July. Machine learning models for road surface and friction estimation using front-camera images. In *2018 International Joint Conference on Neural Networks (IJCNN)* (pp. 1-8). IEEE.
13. Mandal, V., Uong, L. and Adu-Gyamfi, Y., 2018, December. Automated road damage detection using deep convolutional neural networks. In *2018 IEEE International Conference on Big Data (Big Data)* (pp. 5212-5215). IEEE.

14. Krizhevsky, A., Sutskever, I. and Hinton, G.E., 2017. Imagenet classification with deep convolutional neural networks. *Communications of the ACM*, *60*(6), pp.84-90.
15. He, K., Zhang, X., Ren, S. and Sun, J., 2015. Spatial pyramid pooling in deep convolutional networks for visual recognition. *IEEE transactions on pattern analysis and machine intelligence*, *37*(9), pp.1904-1916.
16. Huidrom, L., Das, L.K. and Sud, S.K., 2013. Method for automated assessment of potholes, damages and patches from road surface video clips. *Procedia-Social and Behavioral Sciences*, *104*, pp.312-321.