Convolution Neural Networks for Pothole Detection of Critical Road Infrastructure

Anup Kumar Pandey[1], Rahat Iqbal[1], Tomasz Maniak[2], Charalampos Karyotis[2], Stephen Akuma[3], Vasile Palade[1]

[1]Coventry University, UK; [2]University of Dubai, UAE [3]Interactive Coventry, UK; [4]Benue State University, Makurdi, Nigeria

pandeya4@coventry.ac.uk, riqbal@ud.ac.ae, tomasz.maniak@interactivecoventry.com, charalampos.karyotis@interactivecoventry.com, sakuma@bsum.edu.ng, ab5839@coventry.ac.uk

Abstract

A well developed and maintained highway infrastructure is essential for the economic and social prosperity of modern societies. Highway maintenance poses significant challenges pertaining to the ever-increasing ongoing traffic, insufficient budget allocations and lack of resources. Road potholes detection and timely repair is a major contributing factor to sustaining a safe and resilient critical road infrastructure. Current pothole detection methods require laborious manual inspection of roads and lack in terms of accuracy and inference speed. This paper proposes a novel application of Convolutional Neural Networks on accelerometer data for pothole detection. Data is collected using an iOS smartphone installed on the dashboard of a car, running a dedicated application. The experimental results show that the proposed CNN approach has a significant advantage over the existing solutions, with respect to accuracy and computational complexity in pothole detection.

Keyword: Pothole detection, convolution neural networks, crowdsource data, accelerometer data, highway maintenance

1 Introduction

Poorly maintained highways and roads impact upon traffic, lead to accidents and hinder economic development. According to the Asphalt Industry Alliance report, a one-time repair of the road surface damage is estimated to take more than 10 years and cost over £10 billon. The report mentions that every 19 seconds one pothole is being filled and every year nearly 1.5 million potholes are being filled in England alone. The cost to fill these potholes is nearly £100 million. A survey conducted by Kwikfit suggested that nearly 10 million drivers had their car damaged due to poor road surface conditions over a year, spending over £1.25 billion in repairs. According to RAC, the UK economy lost £8 billion in 2016 because of traffic jams. Therefore, better road networks are needed for smooth traffic flow and for protecting driver's health and vehicles from damage.

Maintaining a vast network of carriageways requires a significant financial investment by the relevant authorities. However, the authorities managing the road network are usually facing lack of human and financial resources. Consequently, there is a severe delay in the repair works. Over the years road inspections were completed by experts in the field, and later this work was transferred into specialised departments of the government. These processes are human dependent, expensive, and time-consuming. This has led to an increased workload and need for experts in the field. Moreover, inspections carried out by experts may not be consistent, as they are based on human visual perception. As technologies are improving, more and more systems are becoming available, which can support pothole detection and provide autonomy. Modern systems rely on Artificial Intelligence based models for detection, such as deep neural networks to perform their task.

The accuracy of any deep learning model depends on the quality of the training dataset. A variety of sensors have been used to support road pothole detection. Recently, smartphones have been utilised to collect GPS and accelerometer data, which can be used to train deep learning models. In [1], the authors discussed participatory sensing crowd data collection. This approach has become a popular method to collect data by using modern smartphones that are widespread and have great sensing features. In [2], the researchers used sensor data collected by smartphones to map road surface roughness based on the international roughness index. Several other studies were also based on the use of Neural Networks to identify road surface and road conditions [3]. Similarly other machine learning methods were used to classify road surface anomalies and have also proven to be viable and cost-effective [4]. The above facts demonstrate that it is possible to automate the process of road surface inspection by using state of the art machine learning and sensory technologies.

This research proposes a crowd data collection approach that utilizes standard smartphone sensors and Convolution Neural Network (1D-CNN) inference to detect road defects. An overview of the proposed approach is illustrated in Figure 1 below.

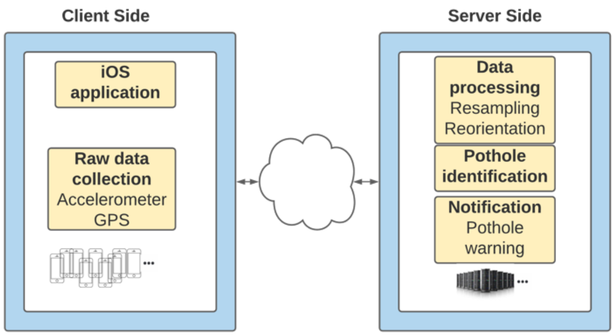

Figure 1: Proposed architecture of pothole detection system.

The rest of the paper is organised as follows. Section 2 presents a literature review focused on the state-of-the-art road pothole detection methods using sensor data. Sections 3 discusses the methodology used in this study and how the data was procured. Section 4 discusses how data is pre-processed in order to be used by the 1D-CNN classifier. Section 5 presents the proposed model and training. Section 6 discusses the experimental results and performs a comparative analysis with existing solutions. Section 7 concludes the findings of the paper and discusses limitations and possible future directions.

2 Previous Work

Road surface anomalies detection has been studied for the last many years. The methods used to detect road surface anomalies and potholes have changed over time. Traditionally, there are three approaches to monitor road surface: 3D reconstruction, computer vision-based and vibration-based. The 3D reconstruction approach requires a 3D laser scanner, which scans the surface and creates an accurate representation of the current state of the road surface that is compared with the original state to detect anomalies. However, such laser scanners are very costly, and the methods depend on the local accuracy of the 2D scan [5]. [6] proposed the polarization method to calculate the difference between horizontal and vertical polarisation. However, the polarisation filters may affect the quality of the images, hence reducing the detection accuracy. In addition, such laser scanners are expensive and these methods depend on the local accuracy of the 2D scan [5].

The computer vision-based methods need sophisticated image processing algorithms to extract relevant texture. This is then compared with the normal texture to find anomalies on the surface. In these methods, a camera is installed in the vehicle, facing downwards, to capture the image of the road surface. The collected images also contain metadata including geographical information (i.e., longitude and latitude). Then, edge detection methods, such as Canny edge detection, are applied to detect anomalies on the road surface [7]. The image-based methods are cost-effective in comparison to the 3D laser scan method. However, an image-based method is sensitive to environmental factors such as light, shadow, rain, etc. Also, the image-based methods are demanding in terms of computational resources and in some cases the use of outdated computer vision algorithms leads to lower pothole detection accuracy and hinders the real time delivery of results.

In order to overcome the beforementioned limitations several, previous studies proposed the use of Deep Leaning (DL) approach to identify road pothole. Several researchers have also proposed vibration-based methods as the most suitable approach for pothole detection. Vibration-based methods rely on accelerometer data to detect potholes. These methods use accelerometer sensors to collect data which are later processed using machine learning and statistical approaches to detect anomalies. The use of deep learning for pothole detection and vibration- based methods are presented in detail in the following sections.

2.1 Deep Learning for Pothole Detection

Deep Learning is a popular Artificial Intelligence (AI) method that mimics the human brain in processing data and creating patterns. Deep learning algorithms rely on multiple levels of interconnected neurons. It has recently formed the computational backbone of many innovative solutions across different scientific domains and applications, from traffic management to digital health [8]. As demonstrated by various studies, deep learning can achieve superior classification and forecasting performance in a variety of machine learning tasks. Its unique characteristics make it an excellent candidate solution for computationally intensive tasks, such as image recognition, big data analytics and real time processing of sensory data [8]. Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN) are some of the most popular deep neural network architectures. CNNs and RNNs, as typical deep neural networks, are structured in consecutive layers. Each of the layers learn more complex representations of hidden features in the data under investigation. Training is conducted by iteratively adjusting the weights of the network. The aim is to achieve the optimal value for a cost function that represents the fitness of the model to the data [8]. RNNs and CNNs have been used successfully in many machine learning tasks. Demonstrative examples are the work by [9], where an 1D CNN was used to extract features, which were later fed into an SVM classifier. Another example is the work by Lee and Cho, where an 1D-CNN that utilized accelerometer data was proposed for human activity recognition [9]. In these works, the applied deep neural networks were able to outperform other machine learning techniques, and effectively support crucial machine learning tasks such as feature space reduction [9].

Deep learning's potential in object detection, classification and other machine learning tasks has been used successfully in recent research efforts towards pothole detection purposes. [10] presented a solution for pothole detection that utilized convolutional neural networks. The solution used image data from different places and under different weather conditions and demonstrated excellent performance [10]. In [11], deep learning-based approaches were used to monitor the road surface and identify potholes. The researchers investigated the application of different deep neural networks, including CNNs and Long Short-Term Memory (LSTM) networks. Their results showcased the ability of deep learning approaches to achieve excellent classification accuracy in the tasks [11]. [12] utilized different deep neural network architectures to support automatic assessment of the road surface and detect potholes. Their experiments resulted in reasonably high accuracy for the deep learning models and low computational times to deliver the results [12]. In a recent work [13], a system that utilized ultrasonic sensory data was proposed for identifying humps. The researchers were able to overcome the considerable challenges of real time analysis by utilizing a CNN-based approach. Their approach achieved better performance compared to other conventional machine learning techniques [13].

2.2 Vibration based Methods for Pothole Detection

The vibration-based pothole detection method uses accelerometer data to detect potholes. Data are being collected using a variety of IoT devices currently present in our everyday lives. In recent times, smartphone and Internet of Things (IoT) technologies have emerged rapidly and have become an integral part of day-to-day life. According to a report by Ericsson, there will be 31.4 billion connected IoT devices by 2023. These intelligent devices are generating vast amounts of valuable data, which can be used to solve many important issues such as road pothole detection. Vibration based methods exploit these data and are cost-effective, require small storage and can be used in real-time [5]. The vibration-based methods can broadly be divided into three categories: (1) Threshold-based, (2) Dynamic Time Warping (DTW) and (3) Machine Learning methods. The threshold-based methods detect anomalies when there is a difference in amplitude or in some of the signal's other properties compared to a specified threshold value. Dynamic Time Wrapping (DTW) measures similarities between two sequences, which may vary in space and time. Machine-learning methods utilize various data sources to automatically identify anomalies on the road surface and optimize the performance of pothole detection applications.

In the recent work of [13], a system that utilized ultrasonic sensory data was proposed for identifying humps. In [14] , the authors discussed the vibration-based method to detect potholes in their work. Data samples were collected using a customised application, later detection algorithms Z- thresh and

Z-peak were applied to find potholes. In [15], Chen and Lu designed and developed a device to collect accelerometer data. The device collected accelerometer data on three axes: X, Y and Z, along with GPS information. The study discussed the limitations of the Z-peak method and why a uniform threshold cannot be used to classify road surface. The study proposed the i-Gaussian Mixture Model (GMM), which accommodated the variability of speed.

In [16] the authors have proposed an accelerometer data and filter-based method to classify road surface. The data set for the study was prepared by driving a car to collect accelerometer data. A threshold was applied on the z-axis data to classify road surface. [17] proposed crowdsourcing using a smartphone to collect sensor data. The researchers used a clustering algorithm – Denclue- to extract road anomalies and index these anomalies. [18] created a specially designed device to collect images and sensor data of the road surface. They have used the Iterative Closest Point (ICP) method to identify potholes. The accuracy of the model is in the range of 90%.

In [19] , the authors also proposed a method in which accelerometer data is collected using a smartphone and then a Gaussian model and x-z ratio filter are applied to detect road anomalies. The data was collected when the vehicle was running at a speed of 15-20 km/h. The estimated error was in the range of 5.08% to 61.93%. In [20] the authors used accelerometer sensor data to detect anomalies using the DTW method. The method produced an accuracy in the range of 88% and was not sensitive to speed. [21] proposed a two detectors-based method to detect bumps and potholes. The proposed method is sensitive to the speed of the vehicle and was conducted at 25 km/h.

In [22], the authors proposed the use of the embedded sensor in a smartphone to collect gyro rotation data. They used variability in the gyro rotation data to detect abnormality in the road surface. The Google maps application was used to find the location of abnormalities and validate the results. Then, they used Dynamic Time Warping (DTM) to find road abnormalities. In [23], the authors discussed their motivation to help people during a natural calamity. They proposed collecting accelerometer data from pedestrians' smartphones and used Support Vector Machine (SVM) to decide whether the surface is flat or not. Before using the SVM classifier, another statistical method was used to extract features from the sensor data. After classification of the road surface, nodes were generated to suggest a map of the safest route.

In [24],the authors proposed a sensor data and Machine Learning based method to classify road surface. Originally, the authors had developed a smartphone application to capture sensory data. The study used Best First, Ranker and Greedy Stepwise to optimise feature selection. Following this, the study utilised Support Vector Machine, Random Tree and Random Forest to perform the surface classification. The results were discussed with respect to the feature optimisation method. Random Forest approach generated the best results.

In [11], the authors used a Deep Learning method and sensor data to monitor the road surface and identify potholes. The authors investigated the application of different deep neural networks, including convolutional and Long Short-Term Memory (LSTM) networks. Their results showcased that Deep Learning approaches were able to achieve excellent classification accuracy in the mentioned tasks. In [25] authors also discussed the use of a Machine Learning approach for road pothole detection using smartphones. The paper used data processing and Machine Learning classification methods, such as logistics regression, Support Vector Machine (SVM), and random forest to detect potholes. In the proposed method, features from collected data were first extracted and then Machine Learning classification was used to detect potholes. This paper did not use neural networks, arguing that it requires a large set of data and does not guarantee higher efficiency.

In the recent work of [13], a system that utilized ultrasonic sensory data was proposed for identifying bumps. The authors were able to overcome the considerable challenges of real-time analysis by utilizing a Convolution Neural Networks-based approach. Their approach achieved better performance than other conventional Machine Learning methods [13].

As demonstrated by our literature review, the vibration methods are cost-effective, require small storage and can be used in real-time [5].This paper focuses on vibration-based methods in combination with a 1D-CNN.

The contributions of this paper are as follows:

- The creation of a training dataset containing sensory data (accelerometer and GPS) captured by an application available for iOS smartphones.
- The development of an optimal 1D-CNN model for pothole detection after evaluating the effect of different hyperparameters on the detection accuracy.

3 Methodology

The proposed methodology is comprised of four steps as described below: (1) data collection, (2) data processing, (3) training an 1D CNN deep learning model to classify potholes, and (4) optimisation of the hyperparameters to obtain a more efficient model (as shown in Figure 2). The first step of the experiment was to collect data. The data for the experiment was collected using a dedicated application installed on a smartphone. This smartphone was mounted on the windshield of the vehicle. In the second stage, the data was processed to ensure that the sampling rate was 100 Hz, it was labelled, and divided into “pothole” or “no pothole” classes. The third stage of the experiment covered the design of a 1-DCNN model and training the model using the collected dataset. The results of the model were studied and analysed. Subsequently, the hyperparameters of the models were changed and more experiments were run until a satisfactory result was achieved. At the end of the fourth stage, the optimal model was obtained in order to be used for road surface classification purposes. The steps of the methodology are discussed in detail in the following sections.

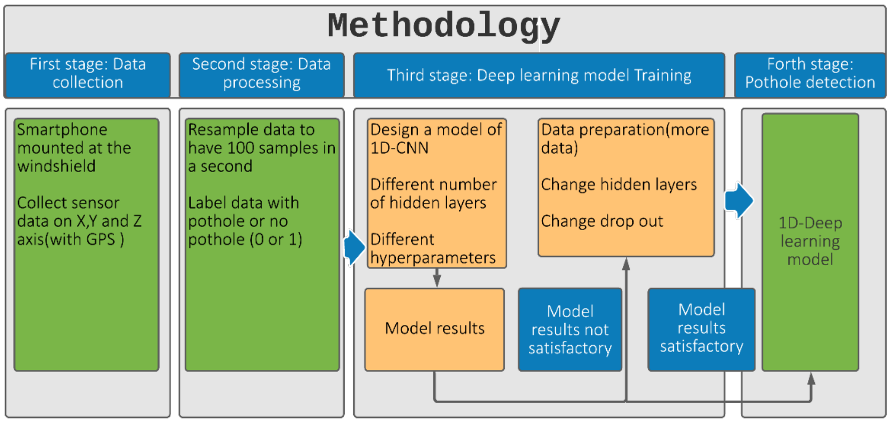

Figure 2:Proposed methodology for pothole detection

4 Data Collection

For data collection, an application (Sensors Pro by Philip Broder) from the iOS store was used to record 3-dimensional (X, Y and Z) accelerometer data, along with the timestamp and GPS information. The sampling rate of the accelerometer was set at 100 Hz. However, due to hardware limitations, the GPS sampling rate was set to 1 Hz. The data was stored in the iOS smartphone (iPhone XR Max) and later was downloaded to a desktop computer (MacOS:11.1, RAM: 24 GB, Processor:3.2 GHz Quad-Core Intel Core i5, graphic AMD Radeon R9 M390 2 GB) for data processing, modelling and offline analysis. The smartphone was securely mounted on the windscreen to record accelerometer data. The data for this study was collected while driving on the roads mentioned in Table 1, which shows the distance travelled to collect data and the types of the road surface. The motorways and A-roads are maintained very well and have very few potholes. However, B roads and the roads in town do have potholes. It took nearly 2 hours and 50 minutes (10,200 seconds) to cover the mentioned distances.

Table 1:Road type, quality and distance travelled

Road Type	Road Quality	Distance (miles)
Motorway	Very Good	80
A Road	Good	40
B road and in town	Not good	30

4.1 Smartphone placement

Another challenge in detecting potholes with the use of a smartphone is to determine the optimal position of the smartphone in the vehicle. In most research efforts, the smartphones were mounted on either the windshield or the dashboard. However, a few studies investigated the performance when the smartphone was kept in the glove box or the driver's pocket. In these cases, [27] confirmed that lower detection rates have been found. For this study, the smartphone was mounted on the windshield in the upright position, as shown in Figure 5. This is the standard mounting position to use a smartphone as a satnav device or as a dashboard camera. The data collected with the smartphone placed in this position will not require data rotation. The data rotation operation has not been covered in this paper. Figure 4 shows the axes of the iOS smartphone in the upright position.

4.2 The orientation of the smartphone

Pothole detection is sensitive to the orientation of the sensors. The smartphone in this study was placed in an upright position (Figure 5) at the windscreen to make sure that accelerometer data for all three axes are in synch with the axes of the vehicle (Figure 3) [27].

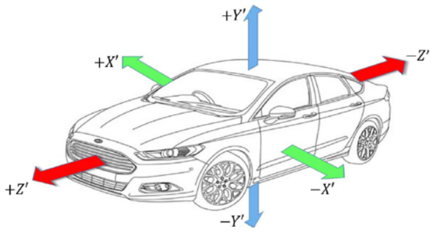

Figure 3: Axes of the vehicle

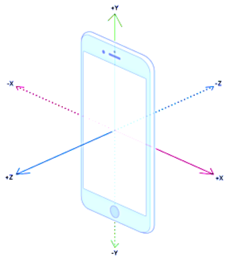

Figure 4: Cartesian coordinate axes of the iOS smartphone accelerometer

Figure 5 shows how the smartphone was placed. Figure 6 shows how a pothole looks from the smartphone perspective.

Figure 5:Placement of the smartphone

Figure 6:A pothole seen from smartphone

4.3 Speed Dependency

The speed of the vehicle influences the performance of the road anomaly detection when using data from a smartphone. [27] showed that average speed plays an important role when estimating road roughness. When a vehicle passes over a pothole at different speeds, the amplitude of the signal collected by the accelerometer will be different. For this study, data was collected from all kinds of roads, as mentioned in Table 1. However, most potholes (95%) were recorded on a B road and in town. The potholes on the motorway and on A roads were not included in the development of the model because, at the speed of 60 miles/hour, a vehicle will cover over 26 meters in a second. Practically it will not be common to find such a big pothole. Also, the current sampling rate is restricted to 100 Hz, which will not be sufficient to capture a pattern in accelerometer data. A higher sampling rate will provide more data in a second and over a certain distance will help to improve the detection rate.

5 Data pre processing

The pre-processing of data is a prerequisite step for obtaining satisfactory and meaningful research results. The data processing for this study was broken down into three stages: resampling, augmentation and labelling. This study aimed to use convolution neural networks to process raw data as they are obtained from the smartphone, without applying many data processing methods. This study is undertaken using 1D-CNNs, which can learn all such internal features and variations to produce an acceptable result.

5.1 Resampling

It was observed that the smartphone was not able to sample the accelerometer data at the fixed frequency uniformly. The sampling rate in the iOS application's accelerometer was set at 100Hz, but it was noted that the accelerometer sampled in the range of 70-100Hz. The 1D-CNN takes fixed-length data and, hence, to have a consistent sampling rate at 100 Hz, the data with lower sampling rates would have to be deleted. However, for this study, the data were resampled at 100 Hz, and missing values were filled by interpolating the data uniformly. The inbuilt library of SciPy (spline interpolation) was used for this purpose.

5.2 Data augmentation

To increase the size of the dataset, data augmentation methods such as permutation and scaling were applied. The scaling factors +/-5% were used to accommodate variation in the data. The permutation on the dataset was applied to increase the number of potholes. In the permutation operation, data for a given time stamp were rearranged to imitate position variation of a pothole, for example data at time T1 on axis X=[X1,X2,X3,…X100] were rearranged to X=[X2,X5,X10..X100].

5.3 Labelling

A supervised learning model's performance depends upon the quality of data and data labelling [27]. In this study, the images with potholes were identified manually. Later, the sensory data was labelled using a software script to match location and time from the photos and videos to the sensor’s data location and time. GPS location and timestamp were used to tag and produce tuples of 100 samples (sampling rate). The data were recorded at the sampling rate of 100 Hz on all three axes (X, Y, Z). A one-second data sample has 100x3 timestamps, 100 timestamps on each X, Y and Z axis. As shown in Figure 7, the accelerometer reading on the Y-axis presents a significant movement in the case of a pothole.

Figure 7 shows a 4-second data sample with pothole detection in the 3rd sec (200-300). It can be noticed that the accelerometer reading on the Y axis had a significant dip, at stamps 200-300, when the pothole was detected. The whole 1-sec window was labelled as a pothole.

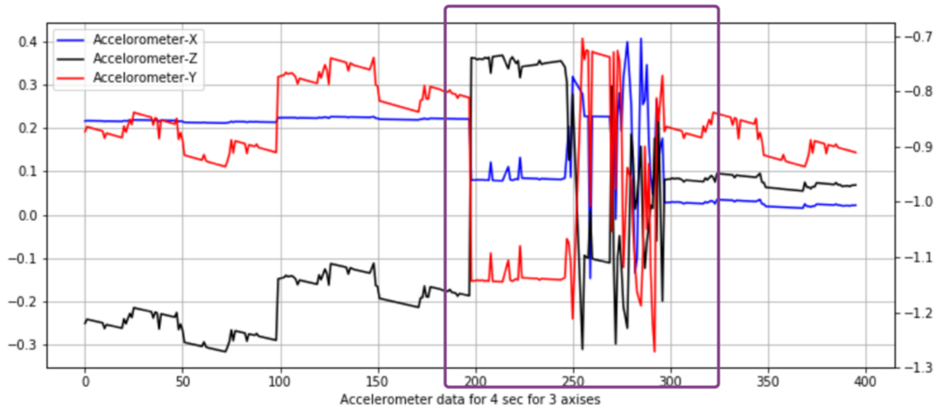

Figure 7: Labelling of a pothole in accelerometer data

6 Training the 1D-CNN Model for Classification

As demonstrated in [9], 1D-CNN can effectively recognise human activities based on accelerometer data collected using a smartphone [9]. The promising results in the study inspired us to use 1D-CNN to classify potholes. Initially the data were labelled. Later a script was used to extract time and GPS information to match the data with the accelerometer data and label the potholes. There was no requirement for domain expertise to process the data before feeding them into the model manually. The 1D-CNN model can automatically learn the relevant features of the time series data and produce a good model from the dataset. 1D-CNN models with multiple layers of non-linear transformation ranging from two to seven layers were used to obtain the classification results. Figure 8 shows the network with two hidden layers. The number of hidden layers was selected randomly to experiment with. The result of each model was analysed. The best result was for the model that had two hidden layers. A kernel of size three and five, and dropouts .25 and .5 were used during the experiment. The dropout technique is used to avoid the overfitting of the model. Theoretically, any value between 0 and 1 can be used. However, practically a random value between .2 and .5 is used while training the model.

The binary cross-entropy loss function was used (Eq.1), where $\hat{y_{i-p}}$ and $\hat{y_{i-n}}$ are the predicted positive and negative probabilities respectively, and N is the sample size. The stochastic gradient descent (Eq.2) optimisers were used.

$$L(x,y,\theta) = -\frac{1}{N}\sum_1^N(-y_p.\log(y_{i-p}^{\wedge}) - y_n.\log(y_{i-n}^{\wedge}))$$ ---------------------Eq.1

$$\theta = \theta - \eta.\nabla_\theta L(x^i, y^i, \theta)$$ ---Eq.2

In Eq (2), $\boxed{\theta}$ is the weight parameter, which starts with a random value and in each iteration the gradient of loss function Eq(1) is calculated to adjust the weight parameter for the next iteration. The training aims to minimise loss (Eq.1) and get the optimal weight parameter $\boxed{\theta}$. For the sample $\boxed{x_i}$, the predicted negative probability is denoted by $\boxed{y_{i-n}^{\wedge}}$ and positive probability by $\boxed{y_{i-p}^{\wedge}}$. The SoftMax function was used to classify between "pothole" and "no pothole" cases.

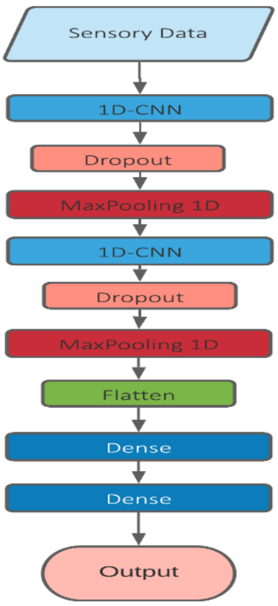

Figure 8: 1D-CNN Model with 2 Layers

Table 2: Algorithm to train the 1D-CNN model

Algorithm 1: Train the 1D-CNN model
Input: Labelled dataset: { X,Y}
Output: Optimal θ^*;
Initial $\boxed{\theta}$, epoch = 0, learning rate α
repeat
1. Sampling labelled data batch $\{x_i, y_i\}$ from $\{X, Y\}$
2. Performing forward propagation of the network and compute $[y_{i-n}^{\wedge}, y_{i-p}^{\wedge}]$
3. Compute loss L by
4. Compute adaptive gradient by SDG (Eq:2)
5. Update parameter $\theta \leftarrow \theta - \alpha\frac{\partial L}{\partial \theta}$
6. epoch = epoch + 1
until (epoch > Epochs)

Table 3: Pothole prediction using the 1D-CNN model

Algorithm 2: Pothole prediction using the 1D-CNN model

Input: Sensory data

Output: Pothole or No pothole$\boxed{;}$ GPS location

1. Capture sensor data using an accelerometer sensor app of the smartphone
2. Upload data on a cloud server
3. Download data from server to a computer
4. Pre-process data to make that one-sec data sample has 100 timestamps (on all three axis X, Y, Z), if not do resampling
5. Feed data from the previous step to the classifier model
6. If $\boxed{y_p > y_n}$ $y_p > y_n$ then
7. The fed sample represents a pothole
8. Else
9. No pothole (normal road surface)
10. End if
11. Tag the sample with pothole and GPS details.
12. Upload the classifier on the cloud server
13. Send notification to other vehicles (who have opted for) that are in the proximity.

Table 2 gives details of the training process. Table 3 provides details of pothole detection using the trained model. Algorithm 1, Table 2, and Algorithm 2,Table 3, shows the steps executed inside the software libraries.

6.1 Model evaluation

To evaluate the 1D-CNN classifier described in section 6, the confusion matrix (error matrix) was used. A confusion matrix is a tabular representation of the performance of a classification algorithm. Classification performance measures, such as accuracy, precision, recall and loss, are derived from the confusion matrix. Here TP-True positive, TN-True negative, FN-False negative, FP-False positive and F1 score were calculated. The F1 score gives an overall measure of a model's accuracy by combining precision and recall.

7 Results and Discussion

The data set (Table 4) was used to run simulations with various combinations of hyperparameters. The Rectified Linear Unit (Relu) activation function was used during the experiments. Each training was conducted with batch size 100 and number of epochs 500. The analysis involved a comparison on test accuracy, precision and recall for various models.

Table 4 : Dataset Details

Dataset for 1D-CNN Model		
Pothole	No-Pothole	

11016	22344	
Training	Validation	Test
23352(70%)	5004(15%)	5004(15%)

Table 5: 1D-CNN Implementation (for Kernel Size=3)

Number of hidden layers	Model Accuracy	Average Precision Rates	Average Recall Rates	F1-Score
7	93.66%	99.00%	73.00%	84.00%
6	93.54%	97.00%	74.00%	84.00%
5	94.40%	99.00%	77.00%	86.00%
3	95.78%	99.00%	80.00%	89.00%
2	95.98%	94.00%	90.00%	92.00%
1	95.10%	85.00%	92.00%	88.00%

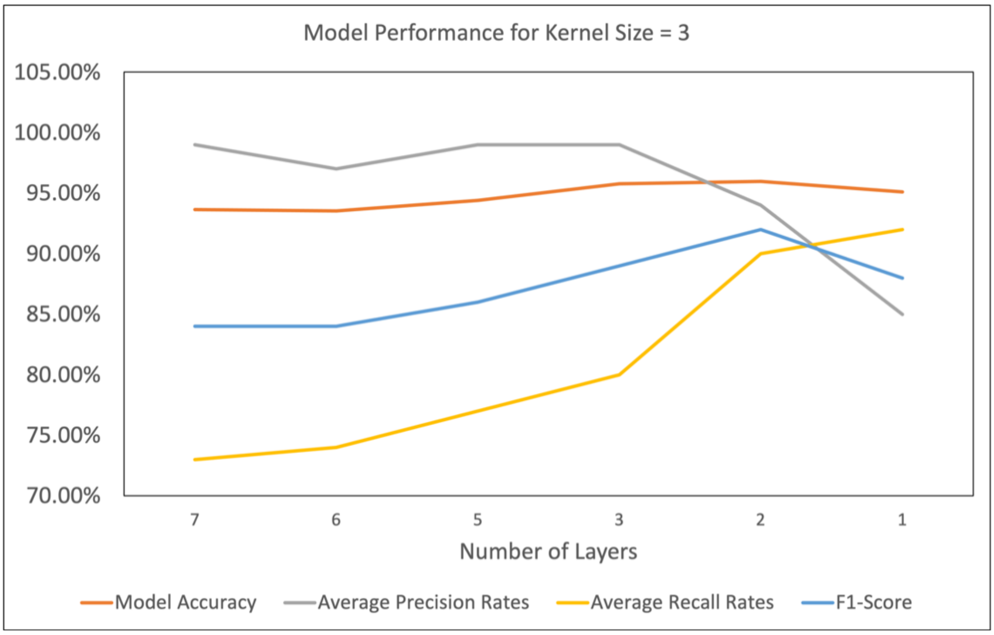

Figure 9: Model Performance for Kernel Size 3

Table 5 and Figure 9 show the simulation results of the 1D-CNN models with Kernel size 3 and hidden layers ranging from 1 to 7. The model's accuracy varied in the range of 95%. However, the recall rate and F1 scores were better for the models with fewer hidden layers. The recall rate was the highest for the model with just one hidden layer, and F1 score was the highest for the model that had two hidden layers. It should be noted that the time to train and classify goes down as the number of hidden layers goes down. Looking at all the performance factors discussed above, the model with two hidden layers is the best model when a kernel size 3 is chosen.

Table 6: 1D-CNN Results for Kernel Size 5

Number of hidden layers	Model Accuracy	Average Precision Rates	Average Recall Rates	F1-Score
7	94.87%	99.00%	79.00%	88.00%
6	95.43%	97.00%	83.00%	89.00%
5	95.63%	98.00%	81.00%	89.00%
3	96.29%	98.00%	84.00%	91.00%
2	95.74%	89.00%	93.00%	91.00%
1	95.10%	98.00%	81.00%	89.00%

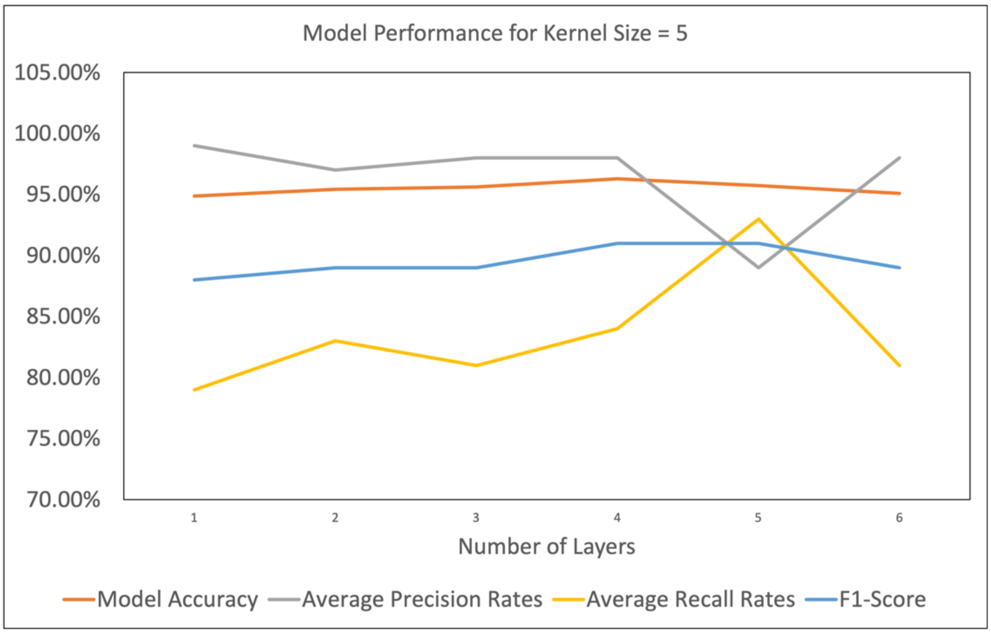

Figure 10: Model Performance for Kernel Size 5

Table 6 and Figure 10 show the simulations results of the 1D-CNN models with Kernel size 5 and hidden layers ranging from 1 to 7. As discussed above for the models with kernel size 3, the accuracy of the model with kernel size 5 ranged from 94.87% to 96.29%. The recall rate and F1 scores were better for the models with the smaller number of hidden layers. The recall rate was the highest for the model with two hidden layers, and then it went down for the model with one hidden layer. The F1 score is in the range of 88%-91% and is the highest for the models that have two and three hidden layers. Considering all the performance factors discussed above, the model with two hidden layers might be the best model with kernel size 3. So, after reviewing the obtained results, it can be concluded that 1D-CNN with two layers and kernel size of 5 is sufficient for the dataset described in Table 4. This study shows that 1D-CNN can be used to effectively detect potholes on the road. The overall average accuracy of the models ranges from 93.66% to 96.29%.

Using a deep neural network-based approach enabled our study to achieve better accuracy compared to the work of [33]. In [33],the researchers used sensory data and applied filters to isolate potholes and achieved a pothole detection rate of 65%. Moreover, the recall rate, precision and F1-Score in this study were better, compared to the study in [25], which utilized simpler machine learning approaches such as Linear Regression(LR), Support Vector Machine(SVM) and Random Forests (RF) to perform the classification tasks and obtained the testing accuracy 95.2%, 94.8% and 95.7 % for LR , SVM and RF methods respectively.

8 Conclusions and Future Work

This paper presented a CNN-based pothole detection method that utilizes accelerometer data. For the purposes of this research, we conducted extensive experiments on a variety of roads and collected data reflecting different weather and road conditions by using an iOS smartphone. The data were used to train and test different network architectures.

Our experiments demonstrated that 1D-CNNs can be used to effectively detect potholes based on accelerometer data collected from standard smartphones. Our results showed that a 2-hidden layer 1-D CNN with kernel size 3 is able to achieve excellent classification results, i.e., a precision of 98%, recall of 84%, and F1-score 91%. These scores correspond to state-of-the-art performance for pothole detection that utilizes sensory data.

The presented method is easy to adopt, efficient and cost-effective. While its accuracy rate is considerably high, it does not require any special equipment to collect and process the data. Moreover, the raw data that is collected by the smartphone sensor can be used to train the 1D-CNN model without the need of excessive data processing. This fact reduces the computational complexity and the time needed for delivering the classification results. This study was conducted offline. However, given the advantages of the method, the data can be collected and processed in real time. Data collected by different vehicles can be uploaded to a cloud server. The cloud server data can be classified in real time using the proposed model. Given the classification results, notification could then be sent to the road users based on their location to alert them of road hazards. Our future work will include further refinement of the proposed approach and its application to a larger dataset to increase the detection accuracy. We will also implement the proposed methodology to work in real-time and explore the development of a hybrid system that utilizes both sensory and imagery data using CNNs.

9 References

[1] J. A. E. Burke and M. B. Srivastava, “Participatory sensing,” UCLA: Center for Embedded Network Sensing, 2006.

[2] K. Zang and S. Jie , “Assessing and Mapping of Road Surface Roughness based on GPS and Accelerometer Sensors on Bicycle-Mounted Smartphones,” Sensors, vol. 18, no. 3, 2018.

[3] L. Huidrom, L. Das and S. Sud, “Method for automated assessment of potholes, cracks and patches from road surface video clips,” Procedia Soc. Behav. Sci., vol. 104, pp. 213-321, 2013.

[4] . M. Artis, S. Girts and Z. Reinholds , “Real Time Pothole Detection using AndroidSmartphones with Accelerometers,” 2011 International Conference on Distributed Computing in Sensor Systems and Workshops (DCOSS), pp. 1-6, 2011.

[5] T. Kim and S.-K. Ryu, “Review and Analysis of Pothole Detection Methods,” Journal of Emerging Trends in Computing and Information Sciences, vol. 5, no. 8, pp. 603-605, 2014.

[6] T. ItakuraY, “Polarizationon- board sensor for detection of wet road surface, using polarization characteristics of reflection,” IEEE, p. 66, 1982.

[7] E. Buza and S. Omanovic, “A. Pothole detection with image processing and spectral clustering.,” In Proceedings of the 2nd International Conference on Information Technology and Computer Networks, Athens, Greece, pp. 48-53, 2013.

[8] R. Iqbal, R, T. Maniak and C. Karyotis, “Intelligent remote monitoring of parking spaces using licensed and unlicensed wireless technologies,” IEEE Network, pp. 23-29, 2019.

[9] W. Huijuan and . C. Jiping, “One-Dimensional CNN-Based Intelligent Recognition of Vibrations in Pipeline Monitoring With DAS,” Journal of Lightwave Technology, vol. 37, no. 17, 2019.

[10] V. Pereira, S. Tamura, S. Hayamizu and H. Fukai, “A deep learning-based approach for road pothole detection in timor leste,” IEEE International Conference on Service Operations and Logistics, and Informatics , pp. 279-284, 2018.

[11] B. Varona, A. Monteserin and A. Teyseyre, “A deep learning approach to automatic road surface monitoring and pothole detection.,” Personal and Ubiquitous Computing, pp. 1-16, 2019.

[12] E. Ukhwah, E. Yuniarno and Y. K. Suprapto, “Asphalt pavement pothole detection using deep learning method based on yolo neural network,” International Seminar on Intelligent Technology and Its Applications , pp. 35-40, 2019.

[13] C. Chellaswamy, M. Saravanan and E. Kanchana, “Deep Learning Based Pothole Detection and Reporting System. I,” 2020.

[14] M. Artis, S. Girts and Z. Reinholds, “Real Time Pothole Detection using AndroidSmartphones with Accelerometers,” 2011 International Conference on Distributed Computing in Sensor Systems and Workshops (DCOSS), pp. 1-6, 2011.

[15] K. Chen and M. Lu, “CRSM: Crowdsourcing based Road Surface Monitoring,” in 2013 IEEE International Conference on High Performance Computing and Communications, China, 2013.

[16] F. E. Gunawan, Yanfi and B. Soewito, “A Vibratory-based Method for Road Damage Classification,” 2015 International Seminar on Intelligent Technology and Its Applications, 2015.

[17] C. Yi, Y. Chuang and C. Nian, “Toward Crowdsourcing-Based Road Pavement Monitoring by Mobile Sensing Technologies,” IEEE Transactions on Intelligent Transportation Systems, pp. 1905-1917, 2015.

[18] Y. L. Chen and et al., “Inexpensive Multimodal Sensor Fusion System for Autonomous Data Acquisition of Road Surface Conditions,” IEEE Sensors Journal, pp. 7731-7743, 2016.

[19] P. Harikrishnan and V. Gopi, “Vehicle Vibration Signal Processing for Road Surface Monitoring,” IEEE Sensors Journal,, pp. 5192-5197, 2017.
[20] G. Singh, D. Bansal and N. Aggarwal, “Smart patrolling: An efficient road surface monitoring using smartphone sensors and crowdsourcing,” Elsevier, vol. 40, pp. 71-88, 2017.
[21] P. Mohan, V. Ramjee and R. Nericell, “Rich monitoring of road and traffic conditions using mobile smartphones,” In Proceedings of the 6th ACM conference on Embedded network sensor systems, Raleigh, NC, USA, pp. 357-358, 2008.
[22] Y. Alqudah and B. H. Sababha, “ Road Surface Conditions Using Embedded Smartphone Sensors,” in 8th International Conference on Information and Communication Systems (ICICS), Jordan, 2017.
[23] Y. Ikeda and M. Inoue, “An Estimation of Road Surface Conditions Using Participatory Sensing,” International Conference on Electronics, Information, and Communication (ICEIC), pp. 1-3, 2018.
[24] M. Dey, U. Satapathy and P. Bhanse, “MagTrack: Detecting Road Surface Condition using Smartphone Sensors and Machine Learning,” IEEE Region 10 Conference (TENCON), pp. 2485-2489, 2019.
[25] W. Chao , W. Zhen , H. Simon and S. Marc , “An Automated Machine-Learning Approach for Road Pothole Detection Using Smartphone Sensor Data,” Sensors, vol. 20, 2020.
[26] V. Douangphachanh and H. Oneyama, “Estimation of road roughness condition from smartphones under realistic settings,” In Proceedings of the 2013 13th International Conference on ITS Telecommunications (ITST), Tampere, Finland,, 2013.
[27] V. Astarita and M. Vittoria, “A mobile application for road surface quality control:UNIquALroad,” Procedia Social and Behavioural Sceince, vol. 34, 2012.

Anup Pandey received the PhD in the Deep Learning and Computer vision from Coventry University in 2021.He bring over 15 years of Industries experience. He has received three master degrees including MBA from Oxford university. He had worked on large scale IT project in the USA and the UK. He has authored few papers in journals and conferences proceedings.

Rahat Iqbal is a Senior Academic, Director, Inventor and Consultant with 20 years’ experience in industrial and academic roles. His main research interests lie in Big Data Analytics, Internet of Things (IoT), Connected and Autonomous Vehicle, Machine Learning and Artificial Intelligence. Prior to joining the University of Dubai, he was Leading the Connected Mobility Research Group at Coventry University, UK.

Tomasz Maniak has over 14 years of commercial experience in applied research, machine learning, software development, IoT and automation. He graduated with a PhD in Computer Science and Machine Learning from Coventry University. Dr Maniak has published a number of papers in peer reviewed international journals.

Charalampos Karyotis is a senior data scientist at Interactive Coventry Ltd. He has a Ph.D. degree in computer science from Coventry University and over 8 years of experience in data-driven technologies and industrial projects. His research interests include machine learning, smart city systems and big data analytics.

Stephen Akuma is a deputy director at Center for Open and Distance Learning in Benue State University. He is an experienced research scientist and a university Lecturer with demonstrated history of working in data science, human-computer interaction and information retrieval, focusing on user modelling and personalization.

Vasile Palade is currently a Professor of Artificial Intelligence and Data Science at Coventry University, UK. He joined Coventry University in 2013, after working for several years as a Lecturer with the Department of Computer Science of the University of Oxford, UK. He has authored 200+ articles in reputed journals and conference proceedings on machine learning.