

LETTER

DeepBus: Machine Learning based Real Time Pothole Detection System for Smart Transportation using IoT

Kashish Bansal¹ | Kashish Mittal¹ | Gautam Ahuja¹ | Ashima Singh¹ | Sukhpal Singh Gill*²

¹Computer Science and Engineering Department, Thapar Institute of Engineering and Technology, Patiala, Punjab, India

²School of Electronic Engineering and Computer Science, Queen Mary University of London, London, UK

Correspondence

*Sukhpal Singh Gill, School of Electronic Engineering and Computer Science, Queen Mary University of London, Mile End Rd, Bethnal Green, London E1 4NS, UK.
Email: s.s.gill@qmul.ac.uk

Summary

Road related accidents have always been a nuisance to drivers and pedestrians alike. Every year countless accidents and deaths occur due to potholes which could have been preventable if there had been a prior warning or if the civic authorities were able to repair these potholes in time. This paper proposes a machine learning based pothole detection system called *DeepBus* for real time identification of surface irregularities on roads using Internet of Things (IoT). DeepBus uses IoT sensors to detect potholes in real time while an end user is driving vehicles on the road. The location of these potholes would be available on a centrally hosted map which can be accessed by both end users and civic authorities. Thus, it would serve as a warning system to all users as well as a database of potholes with their locations to the authorities for quick repair and action. We have compared the performance of various machine learning models (Logistic Regression, Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Naive Bayes, Decision Tree, Random Forest and Ensemble Voting) based on different parameters (Accuracy, F-score, Precision and Recall) and identified that Random Forest is the best model for pothole detection.

KEYWORDS:

Machine Learning, Pothole Detection System, Surface Irregularities, Smart Transportation, IoT

1 | INTRODUCTION

Due to beginning of digital India, the current state of roads in India is improving. However, natural disasters such as heavy rains and floods destroy road infrastructure continuously. There are millions of potholes are existing on the roads which cause various road related deaths and injuries every year¹. A lot of people are losing their lives or getting injuries in India every year due to potholes. In 2017, 3597 people were died in pothole related incidents and 803 were injured². There is no such mechanism is existing to notify concerned authorities and citizens about potholes or worsening road conditions, especially over the Internet. There is no mobile application is available in India to identify road conditions/potholes and warn users in real-time. With the increase in world's population, there has been an increasing load on the infrastructure of the country. Roads have been flooded with the vehicular traffic. It has become very tough to manage such an amount of traffic. This is the primary reason behind developing a system for a vehicle which is an intelligent enough to help drivers in various fields to provide them comfortable ride. There are many reasons of road deterioration such as rainfalls, oil spills on the road, road accidents or wear and tear which damage the road and make it difficult to drive on it. Further, it leads to road accidents and increases the consumption of fuel by vehicles, which effects the environment³. There is a need of machine learning based pothole detection system to identify the road irregularities using Internet of Things (IoT) sensors to avoid future accidents.

TABLE 1 Comparison of DeepBus with existing pothole detection systems*

System	Implementation			MA	Performance Parameters				Machine Learning Models						
	AR	G	RPi		AY	F	P	R	KNN	SVM	LR	NB	DT	RF	EV
Nericell ⁴	✓	✓													
CECE ⁵	✓	✓			✓										
MSS ⁶	✓	✓			✓										
Pothole Patrol ⁷	✓	✓			✓										
DeepBus (our work)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

1.1 | Our Contributions

The motivation behind this research work is to propose a machine learning based smart pothole detection called *DeepBus* for real time identification of surface irregularities on roads using IoT sensors. We have designed an application for smartphones with IoT sensors (accelerometer, gyroscope, GPS) to identify if a vehicle is moving over a pothole or a speed bump in real time and update a common map for all the users with the precise location of road conditions. The main *objectives* of this research work are: 1) to reduce pothole related deaths and injuries, 2) to make drivers aware of the location of potholes before they actually drive over them, 3) to share the exact location of potholes among the Governments and civic authorities for quick repair and 4) to build a real time map which is constantly updating with latest road conditions (presence of potholes). Further, We have compared various machine learning models (Logistic Regression, Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Naive Bayes, Decision Tree, Random Forest and Ensemble Voting) based on various performance parameters (Accuracy, F-score, Precision and Recall) and identified that Random Forest is the best model for pothole detection.

2 | RELATED WORK

Literature reported that very limited research work has been done using accelerometer or gyroscope sensors for classification of pothole data using various machine learning algorithms. Prashant et al.⁴ proposed a system called Nericell for rich monitoring of road and traffic conditions using smartphones equipped with multiple sensors (GPS, accelerometer, microphone) and communication radios. Song et al.⁵ proposed a Cost Effective and Computationally Efficient (CECE) model to categorize potholes using smartphone sensors by processing the data. They utilized Transfer Learning and Inception V3 to achieve flexible ways of application. The success of classification was based on the data collected which can vary according to vehicle type, bump shape and pothole shape. Mednis et al.⁶ proposed Mobile Sensing System (MSS) for road irregularity detection using Android OS based smart-phones and explored that that an accelerometer can be used to classify potholes. They devised various algorithms for deployment on devices with low computational ability and their evaluation on the data acquired using different Android based smartphones. Using a scheme of identifying when the sensor readings reached a certain threshold value, they achieved true positive rates as high as 90% in the context of classifying different road irregularities. Eriksson et al.⁷ proposed Pothole Patrol, which contains a crowd sourced fleet of taxis using their Accelerometer, Gyroscope, and GPS data to classify potholes and other road anomalies with misidentification rate less than 0.2%. Table 1 compares the proposed system, DeepBus, with existing pothole detection systems.

***Abbreviations used in Table 1 are** - LR: Logistic Regression, NB: Naive Bayes, DT: Decision Tree, RF: Random Forest, EV: Ensemble Voting, AY: Accuracy, F: F-score, P: Precision, R: Recall, AR: Accelerometer, G: Gyroscope, RPi: Raspberry Pi and MA: Mobile App

3 | PROPOSED MODEL

To achieve our goals, we have designed a pothole detection system called DeepBus which is capable of classifying potholes along with their coordinates (latitude and longitude) whenever a vehicle drives over it. The classifying device will be installed on a strategic position on the vehicle such as to give maximum reliability and accuracy. By collecting data manually and using machine learning algorithms, the device will improve in accuracy over time. There will also be a companion app which will display real time data of the potholes. As soon as a driver drives over a pothole, it will be classified and displayed to all other users who are driving. The next time someone travels on the same road, they will be aware that their route contains potholes,

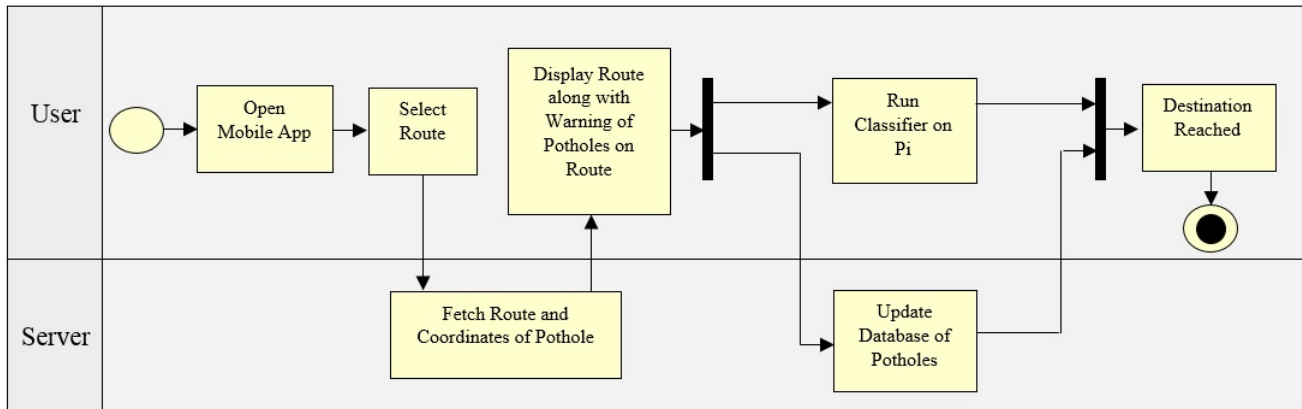


FIGURE 1 System Model of DeepBus

giving them the choice to avoid it, or drive slower and more carefully in its vicinity. This will also enable us to share these exact locations with municipal authorities which will allow them to quickly repair these potholes. DeepBus marks the potholes on the map and alert the user beforehand about the upcoming potholes on its paths so that the driver can take appropriate decisions and prevent any major accidents/damages. The working of DeepBus is shown in Figure 1 and performs various functions, which classifies into two different classes such as User and Server. Server fetches the route and coordinates of the pothole and updates the database. User can find routes with information about potholes and selects the best route for destination by running the classifier on Raspberry Pi. The overall process is divided into three main phases:

- *Data Collection and Preprocessing*: Using the sensors available to collect data on road information. This will involve manually labelling timestamps with presence of potholes and bumps. This can be done by using a smartphone or a Raspberry Pi with appropriate sensors. After data is collected, it will need to be preprocessed which includes removal of missing data, outlier detection, grouping data etc. It is to be noted that the suspension, steering, braking, acceleration of all vehicles are different. So, it is possible but not necessary that multiple models will have to be trained in the future for different vehicles. For our domain, we have done this research work on a single vehicle.
- *Training*: Training various machine learning models on the obtained datasets. This poses questions like which model will help us achieve the best Accuracy, and least False Positive and False Negative rate.
- *Classification/Testing*: After the first two phases, the model is deployed on phones and collects data from the smartphone sensors in real time and classify the presence of a pothole/bump. Misclassifications can be minimized by methods like Weighted Polling, Majority Voting etc. The locations will then be marked on a centrally hosted map which will be available to all users.

4 | PERFORMANCE EVALUATION

4.1 | Experimental Setup

We have created testbed using Raspberry, Accelerometer and Gyroscope sensor to generate the dataset which would be fed into various machine learning models for training⁸. Further, experimental setup consists of Raspberry Pi 3B+ microcomputer which has a Broadcom BCM2837B0 quad-core A53 (ARMv8) 64-bit @ 1.4GHz SoC, a Broadcom Videocore-IV GPU and is equipped with 1GB LPDDR2 SDRAM. Figure 2 shows the Raspberry Pi device is connected with GPS module and IoT sensor such as Accelerometer and Gyroscope. We have detected potholes using Accelerometer and Gyroscope sensors. Accelerometer measures total specific applied force on the sensor. For example, if the device is not moving, it will show some value corresponding to the earth's gravitational force. An accelerometer under free fall in vacuum will not show any value. It can measure acceleration on three perpendicularly placed axis i.e. X, Y and Z. It can calculate acceleration up to 6g. Gyroscope measures change in orientation of the device in terms of angular velocity or angular displacement per second. Like the Accelerometer, it senses values along the X, Y and Z axis. Pothole vibrations are generally vertical and can be measured on the vertical axis (Z) readings.

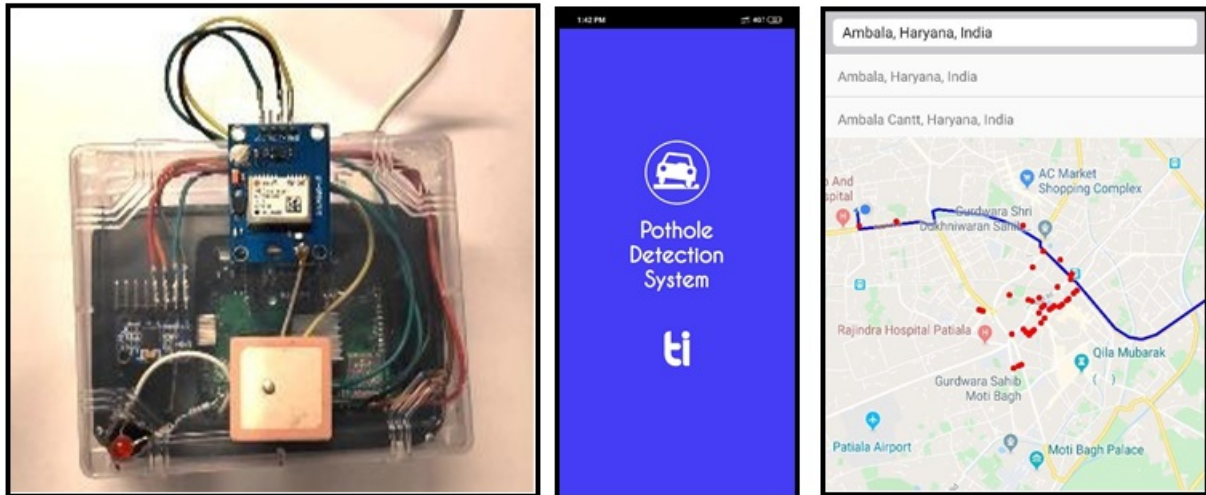


FIGURE 2 Testbed and Mobile App for DeepBus (Red Circles are Potholes)

We classify the pothole on the basis of the change in value of accelerometer and gyroscope combined. MPU6050 Gyroscope and Accelerometer module are connected which are equipped with a 3-axis accelerometer and a 3-axis gyroscope integrated on the same chip. The accelerometer is capable of measuring the gravitational acceleration of the device along 3 axes namely A_x , A_y , A_z . Its output is in meters per second. The gyroscope measures the rotational velocity or the rate of change of angular position over time along the same 3 axes: G_x , G_y , and G_z . Its output is in degrees per second. A push button is also connected to the Pi which can record binary states 0 and 1 (label 1 -> pothole, label 0 -> no pothole). A script is written using Python and loaded on to the Raspberry Pi to collect these data points and the state of the button (0 or 1), which is saved in a .csv file.

4.2 | Dataset

We have used Motorbike, which is integrated with the Raspberry Pi component on its tank and is ridden by 2 people where each instance of a pothole occurring was noted by pressing the push button by the person sitting at the back. The corresponding 6 features (A_x , A_y , A_z , G_x , G_y , G_z) are also noted in the .csv file. We used a 125 CC CBF Honda Stunner bike as the vehicle which has a top speed of 101 Km/h. The data collection has been done in Patiala, Punjab (India). After a couple of hours of driving and recording data, we have collected around 50,000 data records (around 5000 seconds or 83.3 minutes of data) out of which around 1000 records are labelled 1 (Pothole). Since the frequency of the MPU6050 chip is approx. 10 Hz, we are able to obtain 10 data points per second with each sample containing 6 features and a single label. We of course did not drive over 1000 potholes but instead drove on each pothole multiple times to augment the dataset. After running a script on the raw data, we obtained the final dataset which was in .csv format with the following 8 columns: [Time, A_x , A_y , A_z , G_x , G_y , G_z , Pothole]. In this research work, label 1 denotes “pothole” and label 0 denotes “no pothole”. As observed by⁷, a greater accuracy was achieved by combining the data records into 2 second segments as compared to a single data point which is not always accurate due to variation in the exact time the button was pressed each time a pothole was encountered. So, we aggregated the 10 rows above label 1 and 10 rows below label 1 to obtain 2 second readings, and then we have done same for label 0 as well. The aggregation is done using various aggregation techniques such as mean, standard deviation, maximum, minimum for A_x , A_y , A_z , G_x , G_y , G_z each. Therefore, each row of 20 labels will have an added feature of the above-mentioned measures. As a result, the input data that would be fed into our Machine Learning model will contain 24 input features and give one output feature which identifies the presence of pothole. Figure 2 shows the user interface of DeepBus and identification of potholes on Google map using our mobile app.

4.3 | Experimental Results

To test the performance of DeepBus, various machine learning models such as Logistic Regression, SVM, KNN, Naive Bayes, Decision Tree, Random Forest and Ensemble Voting are compared using various performance parameters such Accuracy, F-score, Precision and Recall to identify the best machine learning model for pothole detection. The formulas for performance parameters are given below:

- *Accuracy (a)*: It is the percentage of correct predictions over all the predictions on the dataset. Basically, it identifies how correct a classifier actually is. The formula is given by:

$$a = \text{CorrectPredictions}/\text{TotalPredictions} = (TP + TN)/(TP + TN + FP + FN) \quad (1)$$

Here TP is True Positive, TN is True Negative, FP is False Positive, and FN is False Negative.

- *Precision*: It is the ratio of positive predictions that were actually positive. In our case, it means that whenever the classifier identified a pothole, was it actually a pothole? This is crucial, because mis-identifications have to be reduced. False locations of potholes will cause a waste of time and provide misleading information to both users and civic authorities. The formula is given by:

$$\text{Precision} = TP/(TP + FP) \quad (2)$$

- *Recall*: It is the percentage of actual positives that were classified correctly. Although less “damaging” than a low precision, a high recall is necessary to ensure reliability and consistency of the system. The formula is given by:

$$\text{Recall} = TP/(TP + FN) \quad (3)$$

- *F1-score*: It is the harmonic mean of precision and recall of the model. The formula is given by:

$$F1 = 2 * ((\text{Precision} * \text{Recall})/(\text{Precision} + \text{Recall})) \quad (4)$$

Figure 3 shows the comparison of various machine learning algorithms. As per Figure 3, Random Forest model performs better than other models and it achieves highest Accuracy (86.8%), highest Precision (83.7%), highest Recall (84.4%) and highest F1-score (83.7%) identified as a best model for pothole detection. The reason of better performance are: 1) Random Forest algorithm avoids over-fitting problem in classification problems and 2) with a single training pass, the same Random Forest can be used for classification as well as regression tasks. We used this classifier to classify potholes existing on roads along with their coordinates on mobile app.

5 | CONCLUSIONS AND FUTURE SCOPE

In a developing country like India, it is difficult for the government to maintain a regular surveillance on the road conditions and thus sometimes these small potholes result in large accidents leading to injuries and loss of lives. In this research work, we proposed machine learning based pothole detection system called DeepBus, to pinpoint the location of potholes present on roads. The live data of these potholes is made available through a real time map for all users to enable smart transportation. With this data, warnings can be given to drivers and their locations shared with civic authorities for quick repair. Further, we have compared various machine learning models based on different performance parameters and identified that Random Forest classifier has achieved 86.8% accuracy on the collected dataset for pothole detection.

Various future improvements can definitely be made to improve and expand the scope of DeepBus. If the system can differentiate and classify speed bumps it would further add functionality to the DeepBus. Apart from marking potholes, developing a system to map road conditions would help drivers make more informed choices. Next feature can be added to classify the severity of a pothole. Differentiating a deep pothole with a shallow one will enable Governments to assign priorities while fixing potholes. Another important future direction is reducing misclassifications. A misclassified pothole is detrimental and a waste of time and money to authorities as well as users.



FIGURE 3 Performance comparison of various machine learning algorithms

References

1. More deadly than terrorism': potholes responsible for killing 10 people a day in India, [online] Available at: <https://www.theguardian.com/world/2018/jul/25/more-deadly-than-terrorism-potholes-responsible-for-killing-10-people-a-day-in-india> [Accessed 15 Nov. 2019].
2. Over 9300 deaths, 25000 injured in 3 years due to potholes, [online] Available at: <https://www.indiatoday.in/india/story/over-9300-deaths-25000-injured-in-3-years-due-to-potholes-1294147-2018-07-24>. [Accessed 15 Nov. 2019].
3. Gill, Sukhpal Singh, Shreshth Tuli, Minxian Xu, Inderpreet Singh, Karan Vijay Singh, Dominic Lindsay, Shikhar Tuli et al. Transformative effects of IoT, Blockchain and artificial intelligence on cloud computing: evolution, vision, trends and open challenges. *Internet of Things*. 2019; 8:100118.
4. Mohan, Prashanth, Venkata N. Padmanabhan, and Ramachandran Ramjee. "Nericell: rich monitoring of road and traffic conditions using mobile smartphones." In *Proceedings of the 6th ACM conference on Embedded network sensor systems*, pp. 323-336. 2008.
5. Song, Hyunwoo, Kihoon Baek, and Yungcheol Byun. "Pothole detection using machine learning." *Advanced Science and Technology* (2018): 151-155.
6. Mednis, Artis, Girts Strazdins, Reinholds Zviedris, Georgijs Kanonirs, and Leo Selavo. "Real time pothole detection using android smartphones with accelerometers." In *2011 International conference on distributed computing in sensor systems and workshops (DCOSS)*, pp. 1-6. IEEE, 2011.
7. Eriksson, Jakob, Lewis Girod, Bret Hull, Ryan Newton, Samuel Madden, and Hari Balakrishnan. "The pothole patrol: using a mobile sensor network for road surface monitoring." In *Proceedings of the 6th international conference on Mobile systems, applications, and services*, pp. 29-39. 2008.
8. Chandana, R., S. Jilani, and Mr S. Javeed Hussain. "Smart surveillance system using thing speak and Raspberry Pi." *International Journal of Advanced Research in Computer and Communication Engineering* 4, no. 7 (2015): 214-218.

