

# Smartphone-Based Pothole Detection Utilizing Artificial Neural Networks

Charalambos Kyriakou<sup>1</sup>; Symeon E. Christodoulou, A.M.ASCE<sup>2</sup>; and Loukas Dimitriou<sup>3</sup>

**Abstract:** Roadway pavement maintenance to the preferred level of serviceability comprises one of the most challenging problems faced by civil and transportation engineers, with regard to transport infrastructure management. This paper presents a study on the detection of roadway pavement anomalies by use of smartphone sensors and on-board diagnostic (OBD-II) devices, which can lead to low-cost roadway infrastructure assessment. The proposed approach, which, in addition to smartphone sensors, also utilizes artificial neural network (ANN) techniques in the analysis, captures a vehicle's interaction with a roadway pavement while the vehicle is moving, and utilizes the observed interaction patterns for the detection of potholes in the pavement. The method utilizes four metrics in the analysis and shows a detection accuracy of about 90%. Preliminary results on the inclusion of additional roadway defects in the analysis and on the ability of the method to distinguish between potholes and other pavement defects (e.g., patches, local upheavals, rutting, and corrugation) have been positive. The study's results confirm the value of smartphone sensors in the low-cost (and eventually crowdsourced) detection of potholes. DOI: 10.1061/(ASCE)IS.1943-555X.0000489. © 2019 American Society of Civil Engineers.

**Author keywords:** Potholes; Smartphones; Artificial neural networks.

## Introduction

The efficient detection and mitigation of roadway anomalies is an important pillar in any roadway assessment and management program, since such defects affect the road condition and the ride quality for travelers. Since the condition of pavement surfaces is highly time dependent and load dependent, their continuous monitoring is typically of utmost difficulty and of high cost. Consequently, the focus of interest in recent years with regard to transport infrastructure has shifted away from the construction of new roadways and to the sustainable management of existing ones.

Within this context of efficiency and sustainability in the operations and maintenance (O&M) of transport infrastructure, the technological advancement and current widespread usage of smartphones has taken a promising role, for they have furnished researchers and practitioners with a noteworthy technology for roadway assessment. The work described herein is based on a low-cost approach to roadway assessment, investigating the applicability of smartphone-based tools and collected data to the efficient condition assessment of roadway networks. The proposed approach relies on the premise that standalone modern smartphones and their built-in sensors can be utilized in capturing sensor data without the need for expensive and proprietary probes and systems. The aforementioned capabilities come at a low cost, and their applications can easily be extended to include the behavior and

performance of vehicles using the roadway infrastructure, by means of on-board diagnostic (OBD) devices on the vehicles and Bluetooth connections to paired smartphones. Smartphone-based hardware and software components enable, for example, the monitoring of a vehicle's forward, lateral, and vertical acceleration; roll and pitch rotations; global positioning system (GPS) latitude and longitude; speed; engine revolutions per minute (RPM); and performance (e.g., gas consumption, emissions).

Furthermore, as Kyriakou et al. (2016b) noted, the use of smartphone technology for roadway anomaly detection gives rise to participatory-sensing implementations and low-cost pavement management systems (PMSs). In such a participatory-sensing scheme, a number of vehicles collecting roadway data could be used in order to highlight rough pavements and potholes within a roadway network, with geocoded events marked at points where a moving car exhibits abnormal behavior. Participatory-sensed event points from a number of vehicles can then be aggregated, allowing engineers to identify where vehicles are collectively experiencing rough riding conditions.

It should be noted, though, that given the heterogeneity of the data sets produced by such a participatory-sensing scheme, it is important that crowdsourced data be collected from a statistically significant number of probe vehicles (Dennis et al. 2014). The use of existing specialized probe vehicles is, however, not an efficient proposition in terms of either technology or cost. On the contrary, probe vehicles equipped with smartphone-based technologies hold the promise of a more effective and cost-efficient approach.

This paper presents ongoing research on pothole detection utilizing smartphones and artificial neural networks. In addition to these introductory paragraphs and a section with a brief outline of related past work and of the motivation for the research work described, the section on methodology presents the proposed data collection process and the data structure of the sensed information. A section on the proposed pothole detection method, obtained results, and the method's efficiency follows. The paper concludes with key findings, limitations of the method, and proposed research directions.

<sup>1</sup>Postdoctoral Researcher, Dept. of Civil and Environmental Engineering, Univ. of Cyprus, Nicosia 1087, Cyprus. Email: kyriakou.charalambos@ucy.ac.cy

<sup>2</sup>Professor, Dept. of Civil and Environmental Engineering, Univ. of Cyprus, Nicosia 1087, Cyprus (corresponding author). ORCID: <https://orcid.org/0000-0002-9859-0381>. Email: schristo@ucy.ac.cy

<sup>3</sup>Assistant Professor, Dept. of Civil and Environmental Engineering, Univ. of Cyprus, Nicosia 1087, Cyprus. Email: lucdimit@ucy.ac.cy

Note. This manuscript was submitted on August 20, 2016; approved on November 9, 2018; published online on May 17, 2019. Discussion period open until October 17, 2019; separate discussions must be submitted for individual papers. This paper is part of the *Journal of Infrastructure Systems*, © ASCE, ISSN 1076-0342.

## State of Knowledge in Roadway Anomaly Detection

As aforementioned, pavements are roadway infrastructure assets of high significance, with operations and management requiring the use of specialized PMSs. PMSs are generally considered to improve the scheduling of pavement works, the analysis of the cost-effectiveness of treatment repairs, and the comparison of various maintenance strategies and budget scenarios (AASHTO 1990). Thus, investment in PMSs is worthwhile, albeit typically at high initial and operating cost since they are specialized hardware and software platforms that are typically hosted on dedicated vehicles.

At present, the methodology used for the collection of pavement distress data utilizes surveys that generally address an evaluation or detail measurement of distress (Walker et al. 2002). The collection of the road surface condition data can be either manual (from a moving vehicle or by “walking” the pavement) or automated (by use of vehicles fitted with specialized cameras and sensors). Two of the most prominent broad categories of data collection are (1) vibration-based methods; and (2) vision-based methods.

Relevant with regard to vibration-based methods for the assessment of roadway pavements, are the research works of De Zoysa et al. (2007), Eriksson et al. (2008), Mohan et al. (2008), Tai et al. (2010), Strazdins et al. (2011), Seraj et al. (2014), Alessandroni et al. (2014), Mohamed et al. (2015), and Kyriakou et al. (2016a, b, 2017).

De Zoysa et al. (2007) described a system (called BusNet) that was based on acceleration sensors placed on public buses, which were utilized for identifying potholes through changes in the vertical acceleration of the buses while the buses were moving. Similarly, Eriksson et al. (2008) developed a mobile sensor system (called Pothole Patrol) hosted on a small fleet of taxis, to gather data and identify potholes in the roadway network by use of GPS, wireless fidelity (Wi-Fi), and three-axis accelerometers. The aforementioned system and similar systems, though, all exhibited a number of problems. The most important of these problems was the inability of the sensors to distinguish between real potholes and road conditions causing sudden vehicular accelerations (such as road bumps and expansion joints) or other vibration-inducing events (such as doors being knocked and unexpected swerves). Furthermore, the sensor-reported values were dependent on the mount and speed of the host car, thus requiring calibration.

A similar system proposed by Mohan et al. (2008), called Nericell, used smartphones that were equipped with an array of sensors (GPS, accelerometer, microphone) in tandem with communication radios, focusing on how these sensors and radios could be used to detect bumps and potholes, and on localizing these anomalies. Tai et al. (2010) used a mobile phone with triaxial accelerometer to collect acceleration data while riding a motorcycle. Strazdins et al. (2011) proposed a method utilizing a smartphone with GPS, three-axis accelerometer and a communication channel (cellular or Wi-Fi), and a system consisting of two application components, one for the data-collection device and one for a data server. Seraj et al. (2014) proposed a system that detected road anomalies using a mobile phone equipped with inertial accelerometers and gyroscopes sensors, coupled with a method for removing the effects of speed, slopes, and drifts from sensor signals. The system proposed by Alessandroni et al. (2014) included both a custom mobile application and a georeferenced database system, with the estimated roughness values for the roadway pavement computed and visualized on a back-end geographical information system. Their approach introduced an integrated system for monitoring applications in a scalable, crowdsourced collaborative sensing environment. Mohamed et al. (2015) proposed a road condition

monitoring system for speed bumps. They suggested that the main indicator for road anomalies is the gyroscope around gravity rotation and that the accelerometer sensor should act as a cross-validation method.

The works of Kyriakou et al. (2016a, b, 2017) discussed the development of a low-cost pavement assessment method and a geographic information system (GIS)-based decision support system (DSS). Data were collected by means of a smartphone and then analyzed for anomalies by use of artificial neural networks, robust regression, and various classification algorithms. The researchers reported pothole-detection accuracy levels of about 90%.

With regard to vision-based methods, relevant work can be found in recent efforts by Koch and Brilakis (2011), Koch et al. (2013), Radopoulou et al. (2016), Radopoulou and Brilakis (2016), and Hadjidemetriou et al. (2016).

In the work by Koch and Brilakis (2011), a vision-based method for automated pothole detection in asphalt pavement was proposed, by which an image was first segmented into defect and nondefect regions using histogram shape-based thresholding, and then the texture inside a potential defect shape was extracted and compared with the texture of the surrounding nondefect pavement in order to determine if the region of interest represented an actual pothole. The aforementioned camera-based pothole-detection method was subsequently extended by Koch et al. (2013) for assessing the severity of potholes by incrementally updating a representative texture template for intact pavement regions and using a vision tracker to reduce the computational effort. In Radopoulou and Brilakis (2016), video data collected from a car's parking camera were utilized to detect defects in frames, coupled it with elevation signals collected from accelerometers attached to the car to reconstruct the profile of the road, and classified detected defects according to their type and severity. The researchers reported that the initial identification of frames including defects produced an accuracy of 96% and approximately 97% precision and that further experiments on such frames, aiming at the detection of potholes, patches, and three different types of cracks resulted in over 84% overall accuracy and over 85% precision.

In terms of analysis methods, common approaches utilize data mining, artificial intelligence, and dynamic models for the extraction and deduction of knowledge on roadway anomalies. Indicative works can be found in the works by Li et al. (2015) and Yi et al. (2015). Li et al. (2015) described a model-based pothole detection algorithm, which empirically breaks down the responses of hitting potholes into three phases governed by simpler dynamic system submodels. Each submodel was based on a rigid-ring tire and quarter-car suspension model, which were then combined into a unified model and a pothole detection algorithm based on unscented Kalman filter (UKF) and Bayesian estimation. Yi et al. (2015) proposed a smartphone probe car (SPC) system (essentially an ordinary vehicle with a mounted smartphone that runs sensing programs to assess bumping caused by road anomalies), which uses a signal processing heuristic for the extraction of the vertical acceleration components from the accelerometer readings. Furthermore, the researchers proposed a road anomaly indexing heuristic that is representable for road anomalies rather than vehicle conditions, and a DENCLUE-like algorithm to mine road anomaly information from reported events.

Finally, Hadjidemetriou et al. (2016) presented a method based on image processing and support vector machine (SVM) classification. The feature vectors used in the classification consisted of an image histogram and two texture descriptors, and the output was a binary image where each image block was classified as *patch* or *no patch*. The performance of their proposed method was rated by a



detection accuracy of 81.97%, a precision of 64.21%, and a recall of 91.21%.

The aforementioned systems show that pothole detection by use of smartphone sensors is technologically feasible. The current research aims to generate additional insight into roadway anomaly detection and the key factors affecting detection, and to provide an alternative method for achieving it by use of low-cost sensor technology. Furthermore, the goal is to address some of the limitations of existing methods, such as the ones reported by De Zoysa et al. (2007) and Mohamed et al. (2015).

## Methodology

The technology required to widely implement pavement condition monitoring by use of smartphones is established and has already partially been proven workable (De Zoysa et al. 2007; Mohan et al. 2008; Strazdins et al. 2011; Kyriakou et al. 2016b; Seraj et al. 2014; Kyriakou et al. 2017). However, smartphone-based data are fundamentally different in form and function from traditional pavement condition data; thus, additional considerations need to be made. For example, a model similar to today's practice in which one vehicle is capable of capturing a specific metric in a single pass is not likely to be workable with smartphone-derived crowdsourced data. It is more likely that a multitude of data sources (probe vehicles) will need to participate in data collection to arrive at a usable metric.

There are four main disadvantages of collecting data from vehicles and smartphones, which need to be overcome. First, there is no method for collecting sensor data from a vehicle or from smartphones without the contribution and authorization of their owners. Second, accelerometer data from a controller area network (CAN) cannot be collected without the cooperation of the vehicle manufacturer. Third, accelerometers are not standard across vehicles, and they demand individual assessment and calibration for each vehicle. Finally, vertical acceleration data captured from vehicles are not at a sufficiently high sample rate to produce usable roughness data for traditional metrics such as the International Roughness Index (IRI) and the Pavement Condition Index (PCI).

The roadway anomaly type examined in this case study is potholes (and by extension, manholes) [Figs. 1(a–c)], and related data were collected by use of a car fitted with a smartphone (mounted on the car's windshield) with its camera, GPS accelerometers, and gyroscopes activated. Furthermore, the smartphone was connected to an OBD-II reader and an OBD-II interface for recording and exporting georeferenced and timestamped sensor readings. Data readings were taken at time intervals of 0.1 s per reading, resulting in a high spatial resolution.

In addition to  $x$ ,  $y$ ,  $z$  accelerations and vehicle speed data, the sensors also recorded the smartphone's angular (roll and pitch) movements [Fig. 2(a)], which could then be related to the host vehicle's roll and pitch movements. The roll relates to a car's acceleration difference between its left and right front wheels, while the pitch relates to a car's acceleration difference between its front and rear wheels. In tandem, roll and pitch describe the manner in which the host car is off balance both sideways and front/back.

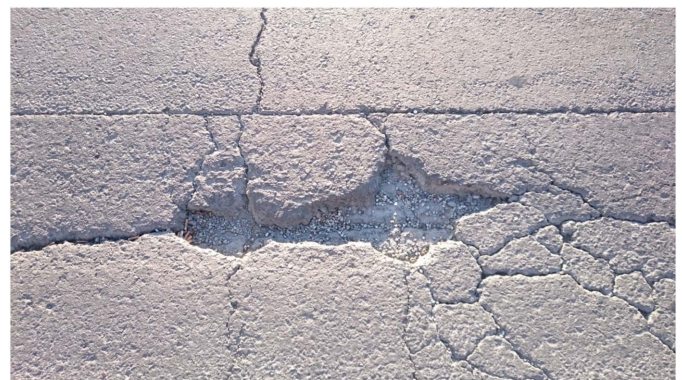
The field investigation included the collection of sensor data for 31 parameters (e.g., speed,  $x$ ,  $y$ ,  $z$  accelerations,  $x$ ,  $y$ ,  $z$  rotations, GPS coordinates, time, and engine revolutions per minute) and regression techniques (multiple regression and robust regression) were used in the analysis of the data for investigating the statistical significance of the factors.



(a)



(b)



(c)

Fig. 1. Potholes examined for detection.

Multiple regression estimates the coefficients  $\beta_i$  of each factor  $x_i$ , and the residual error  $\varepsilon_j$  for each observation  $j$ , in the equation  $y_j = \beta_0 + \beta_1 x_{1j} + \beta_2 x_{2j} + \dots + \beta_p x_{pj} + \varepsilon_j$ , which is assumed to relate the dependent variable  $y_i$  to the independent variables  $x_i$ . The independent variables are the 31 parameters used in this analysis, as aforementioned, and the dependent variable relates to the detected status of the pavement (1 for no defect, 2 for potholes). The subscript  $j$  represents the observation (row) number (approximately 70,000 timestamped data rows were used in the analysis, of which approximately 1,450 were pothole-related observations). Although the regression analysis problem might be addressed by a number of different techniques, the most-used method is the least-squares method, in which the values of  $\beta_i$  are selected so as to minimize the sum of the squared residuals. This set of  $\beta_i$  is not necessarily the right set since it may be distorted by outliers (i.e., points that are not representative of the underlying data pattern). Outliers violate the assumption of normally distributed residuals in least-squares regression, and thus they need to be filtered from the analysis. Robust regression provides an alternative to least-squares regression.

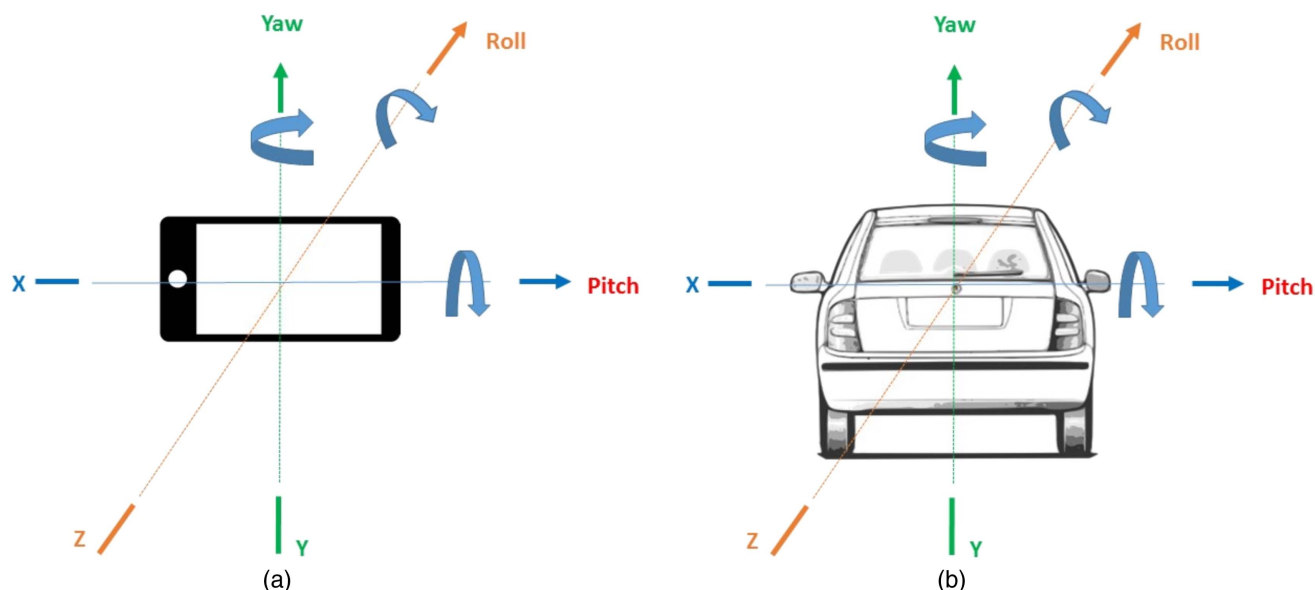


Fig. 2. Smartphone's (a) roll and pitch directions; and (b) relation to car wheels' differential.

| Robust Multiple Regression Report  |   |                            |                           |                                      |            |                   |  |
|--|---|----------------------------|---------------------------|--------------------------------------|------------|-------------------|--|
| Dataset  | C:\...17.Robust\4 1406 pothole detection.xlsx |                            |                           |                                      |            |                   |  |
| Dependent  | C7  |                            |                           |                                      |            |                   |  |
| Run Summary Report   |   |                            |                           |                                      |            |                   |  |
| Item   | Value   | Rows                       |                           |                                      |            | Value             |  |
| Dependent Variable   | C7  | Number Processed           |                           |                                      |            | 1405              |  |
| Number Ind. Variables  | 4   | Number Used in Estimation  |                           |                                      |            | 1405              |  |
| Weight Variable  | None  | Number Filtered Out        |                           |                                      |            | 0                 |  |
| Robust Method  | Tukey's Biweight                              | Number with X's Missing    |                           |                                      |            | 0                 |  |
| Tuning Constant  | 4.685   | Number with Weight Missing |                           |                                      |            | 0                 |  |
| MAD Scale Factor   | 0.6745  | Number with Y Missing      |                           |                                      |            | 0                 |  |
|  |   | Sum of Robust Weights      |                           |                                      |            | 1106.207          |  |
| Run Information  | Value   |                            |                           |                                      |            |                   |  |
| Iterations   | 23  |                            |                           |                                      |            |                   |  |
| Max % Change in any Coef   | 0.001   |                            |                           |                                      |            |                   |  |
| R <sup>2</sup> after Robust Weighting  | 0.9754  |                            |                           |                                      |            |                   |  |
| S using MAD  | 0.07  |                            |                           |                                      |            |                   |  |
| S using MSE  | 0.05  |                            |                           |                                      |            |                   |  |
| Completion Status  | Normal Completion                             |                            |                           |                                      |            |                   |  |
| Regression Coefficients T-Tests Assuming Fixed Weights   |   |                            |                           |                                      |            |                   |  |
| Independent Variable   | Regression Coefficient b(i)                   | Standard Error Sb(i)       | Standard-ized Coefficient | T-Statistic to Test H0: $\beta(i)=0$ | Prob Level | Reject H0 at 5% ? |  |
| Intercept  | 1.45474                                       | 0.003631                   | 0.0000                    | 400.653                              | 0.0000     | Yes               |  |
| AUX_ACCEL_FORWARD_Gs   | 7.61014                                       | 0.037703                   | 1.6698                    | 201.844                              | 0.0000     | Yes               |  |
| AUX_ACCEL_LATERAL_Gs   | -2.74957                                      | 0.081288                   | -0.7562                   | -33.825                              | 0.0000     | Yes               |  |
| AUX_ROTATION_PITCH_Ä   | -0.12678                                      | 0.000563                   | -1.9389                   | -225.228                             | 0.0000     | Yes               |  |
| AUX_ROTATION_ROLL_Ä  | -0.04393                                      | 0.001408                   | -0.7168                   | -31.210                              | 0.0000     | Yes               |  |
| Estimated Equation   |   |                            |                           |                                      |            |                   |  |
| C7 =   |   |                            |                           |                                      |            |                   |  |
| 1.45473581036334 + 7.61014286269618 * AUX_ACCEL_FORWARD_Gs - 2.74956802902307 * AUX_ACCEL_LATERAL_Gs - 0.126777462130123 * AUX_ROTATION_PITCH_Ä - 0.0439314311674655 * AUX_ROTATION_ROLL_Ä |   |                            |                           |                                      |            |                   |  |

Fig. 3. Robust regression.



**Table 1.** Data collected and variables used for classifying roadway anomalies

| Variable | Name  | Description          |
|----------|---|----------------------|
| VAR_1    | AUX.ACCEL.FORWARD (Gs)                        | Forward acceleration |
| VAR_2    | AUX.ACCEL.LATERAL (Gs)                        | Lateral acceleration |
| VAR_3    | AUX.ROTATION.ROLL ( $\hat{\text{A}}^\circ$ )  | Roll                 |
| VAR_5    | AUX.ROTATION.PITCH ( $\hat{\text{A}}^\circ$ ) | Pitch                |

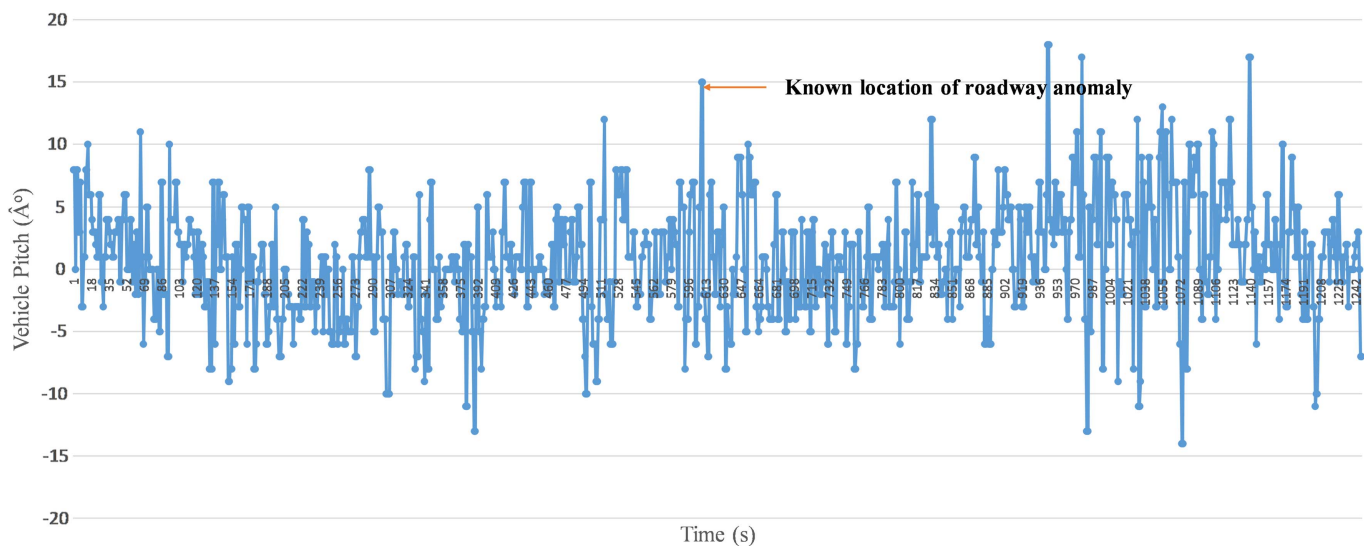
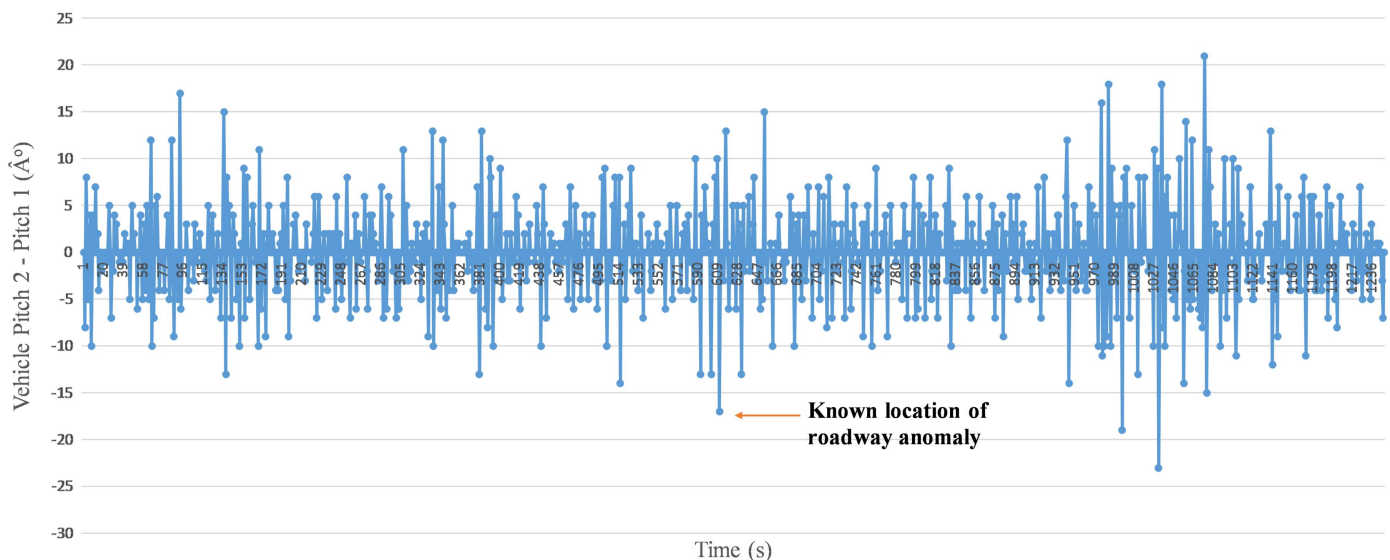
which works with less restrictive assumptions and provides much better regression estimates when outliers are present in the data.

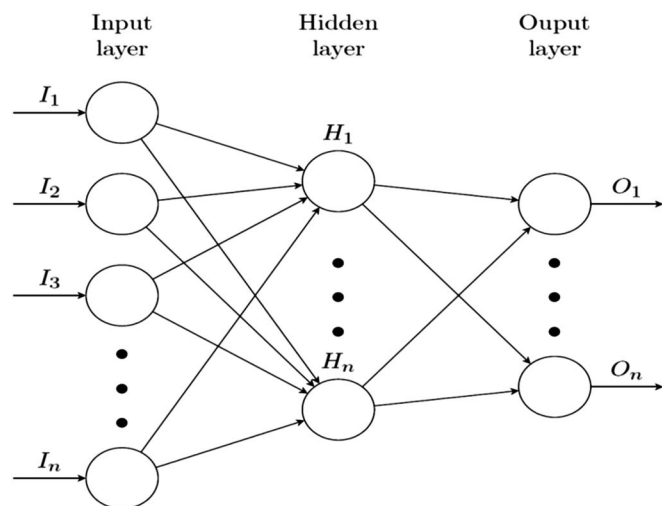
The data set was thus analyzed using the robust regression method, aiming for the identification of the most important independent variables and the resulting regression equation. The analysis results are shown in Fig. 3 and Table 1, listing the most significant variables (with  $p$ -values  $\leq 0.05$ ).

## Results and Discussion

The signals of the raw sensor data (Figs. 4 and 5) depict the complexity of the detection problem at hand. Variables such as the vehicle roll and vehicle pitch are not by themselves as conclusive as initially presumed. Neither the signal of the vehicle's roll (Fig. 4) or pitch, nor the signal of the incremental differences in vehicular roll (or pitch) between subsequent locations (Fig. 5), show any patterns that can serve as detectors. The vertical variability is random, and even at a point of a known roadway anomaly, the variability in the signal does not mean that the existence of the anomaly is definite.

Even though the aforementioned analyses may indicate possible signal anomalies such as deviations from the signal's moving average or unusually high or low values, these signal anomalies are not reliable indicators of pavement anomalies. Furthermore, the figures do not provide any additional information on parameters such as the forward and lateral acceleration of the host vehicle, which may affect the recorded signals.

**Fig. 4.** Vehicle pitch over time.**Fig. 5.** Point-to-point vehicle pitch variation over time.



**Fig. 6.** ANN model schematic architecture.

**Table 2.** ANN training and validation statistics for pothole detection

| Classes    | Training accuracy (%) | Validation accuracy (%) | Testing accuracy (%) | Overall accuracy (%) | Average accuracy (%) |
|------------|-----------------------|-------------------------|----------------------|----------------------|----------------------|
| (1) Type_1 | 98.0                  | 100                     | 98.6                 | 98.4                 | 98.8                 |
| (2) Type_2 | 100                   | 100                     | 100                  | 100                  |                      |

Note: See Fig. 1 for classes.

**Table 3.** Training confusion matrix

| Output class | Target class |               | Overall (%)  |
|--------------|--------------|---------------|--------------|
|              | 1            | 2             |              |
| 1            | 342<br>71.1% | 7<br>1.5%     | 98.0<br>2.0  |
| 2            | 0<br>0.0%    | 132<br>27.4%  | 100.0<br>0.0 |
| Overall      | 100%<br>0.0% | 95.0%<br>5.0% | 98.5<br>1.5  |

Note: Class 1 = no defect; and Class 2 = potholes/manholes (Fig. 1).

In an effort to strengthen the detection of hidden patterns in the sensed data, the data sets are processed by means of an artificial neural network (ANN) consisting of 4 inputs ( $I_1$ ,  $I_2$ ,  $I_3$ , and  $I_4$ ), 10 hidden neurons ( $H_1$ ,  $H_2$ , ...,  $H_{10}$ ), and 2 outputs ( $O_1$  and  $O_2$ ). The ANN inputs are the top four parameters identified by the regression analysis (listed in Table 1) and the ANN outputs, which are binary in nature (1 for no defect, 2 for defect), are used to classify data readings into *no defect* and *potholes* (Fig. 6).

Roadway pothole anomalies at known locations are used in training the ANN, with *ClassType\_1* referring to the no-defect case and *ClassType\_2* referring to potholes/manholes. Approximately 1,450 pothole-related timestamped data were utilized in training, validating, and testing the ANN (summary results are shown in Table 2) split in two batches for validation and testing purposes (Tables 3–7). In training and validating the ANN, raw time-series sensor data were used (no filtering or preprocessing was utilized), and the pothole-free data samples were taken from defect-free pavement sections. In each case, 70% of the data were used for training, 15% for testing, and 15% for validating the ANN.

**Table 4.** Validation confusion matrix

| Output class | Target class |              | Overall (%)  |
|--------------|--------------|--------------|--------------|
|              | 1            | 2            |              |
| 1            | 74<br>71.8%  | 0<br>0.0%    | 100<br>0.0   |
| 2            | 0<br>0.0%    | 29<br>28.2%  | 100.0<br>0.0 |
| Overall      | 100%<br>0.0% | 100%<br>0.0% | 100<br>0.0   |

Note: Class 1 = no defect; and Class 2 = potholes/manholes (Fig. 1).

**Table 5.** Test confusion matrix

| Output class | Target class |               | Overall (%)  |
|--------------|--------------|---------------|--------------|
|              | 1            | 2             |              |
| 1            | 72<br>69.9%  | 1<br>1.0%     | 98.6<br>1.4  |
| 2            | 0<br>0.0%    | 30<br>29.1%   | 100.0<br>0.0 |
| Overall      | 100%<br>0.0% | 96.8%<br>3.2% | 99.0<br>1.0  |

Note: Class 1 = no defect; and Class 2 = potholes/manholes (Fig. 1).

**Table 6.** All confusion matrix

| Output class | Target class |               | Overall (%)  |
|--------------|--------------|---------------|--------------|
|              | 1            | 2             |              |
| 1            | 488<br>71.0% | 8<br>1.2%     | 98.4<br>1.6  |
| 2            | 0<br>0.0%    | 191<br>27.8%  | 100.0<br>0.0 |
| Overall      | 100%<br>0.0% | 96.0%<br>4.0% | 98.8<br>1.2  |

Note: Class 1 = no defect; and Class 2 = potholes/manholes (Fig. 1).

**Table 7.** Tested on new data sets

| Output class | Target class |               | Overall (%)  |
|--------------|--------------|---------------|--------------|
|              | 1            | 2             |              |
| 1            | 488<br>68.1% | 7<br>1.0%     | 98.6<br>1.4  |
| 2            | 0<br>0.0%    | 222<br>31.0%  | 100.0<br>0.0 |
| Overall      | 100%<br>0.0% | 96.9%<br>3.1% | 99.0<br>1.0  |

Note: Class 1 = no defect; and Class 2 = potholes/manholes (Fig. 1).

Table 2 shows a synopsis of the obtained detection results for the original data set, whereas Table 7 shows a synopsis of the results from the application of the proposed model on new data sets.

As the tables and the corresponding ANN confusion matrices depict, the ANN classification arrives at a high degree of accuracy, not only when investigating for the existence of a pothole, but also when investigating for the no-defect case. The ANN confusion matrices tabulate the numbers (and percentages) of not only the accurate classifications (i.e., perfect matches between target and output classes), but also of erroneous and of false-positive classifications.

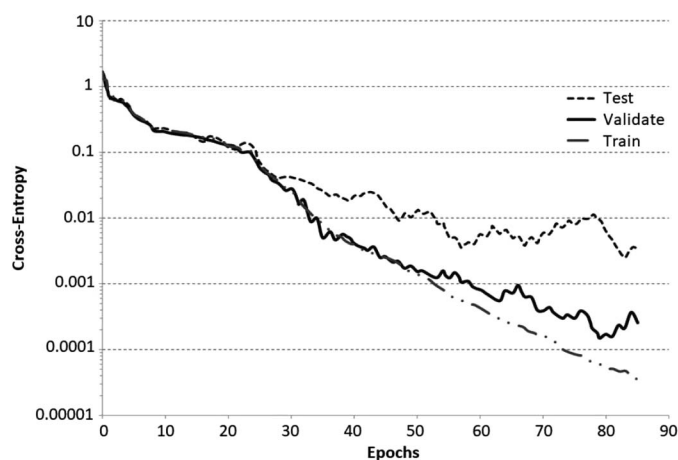


Fig. 7. ANN training, testing, and validation performance.

The horizontal axis in each confusion matrix indicates the target class (the expected result of the ANN classification), and the vertical axis indicates the output class (the actual result of the ANN classification). It should be noted that a good classifier yields a confusion matrix that is dominantly diagonal, as all off-diagonal elements on the confusion matrix represent misclassified data.

As shown in Tables 3–6, the ANN correctly classifies roadway anomalies (the diagonal elements for each of the confusion matrices) with a high degree of accuracy. Furthermore, convergence to these classifications is achieved in short ANN training/validation cycles (epochs) (Fig. 7), with the classification accuracy (indicated by lower cross-entropy values) being higher during the training and validation phases and lower during the test phase.

In essence, the ANN classifier spots and precisely detects potholes (Target class 2) while also distinguishing the no-defect condition (Target class 1), thus separating normal and abnormal roadway pavement conditions.

## Limitations

One significant way in which vehicle data are different from current roadway condition assessment data is that the former do not directly refer to pavement condition, and thus the data require further processing. Smartphone-based data monitor the performance of the vehicle in response to pavement conditions, and the assessment of a pavement's condition is inferred from data regarding interactions between the driver, the vehicle, and the pavement. This approach creates limitations on the accuracy of smartphone data collection. Furthermore, it is also expected that sensor data will vary by vehicle, smartphone properties, and behavior of the driver. Differences in vehicle dynamics, vehicle condition, smartphone properties, and sensor condition might create different readings. Furthermore, even if with the same vehicle, the same driver, multiple passes from the same segment of roadway can create significantly different results for the pavement condition. However, if it can be shown that one vehicle can repeatedly collect similar data results from multiple passes on a segment road, this would also show the ability to provide an accurate model of that vehicle's interaction with the pavement. Similarly, since vibration-based detection methods rely on vibrations caused by tires hitting a roadway anomaly, a question that usually arises is how accurate the method can be if it cannot be guaranteed that a traveling car will hit an existing roadway anomaly.

Counteracting the mentioned weaknesses, the primary strength of smartphone-based data is that it can be crowdsourced. The use of

crowdsourced (or participatory) sensing, in which multiple cars are used as agents of detection with the collective sensing aggregated into single detections, allows for the collection of volumes of data by obtaining pavement condition information from smartphone-based sensors. The power of crowdsourced data is that large data sets, which are collected through multiple data sources, negate the limitations in generalizability of data collected from a single data source. Even though multiple vehicles might provide conflicting data relating to pavement condition, the total effect and knowledge inherent in the data provides an accurate model of the roadway condition in relation to how an average user experiences the pavement condition.

## Conclusions and Future Work

Transportation agencies can improve the condition and operation of their roadway networks by implementing a pavement management system that utilizes smartphone-based data collection, OBD connections, and decision support software. The popularity of smartphone technology in vehicles provides an opportunity to efficiently collect vehicle data and process it by use of connected and distributed systems. Even though vehicle data is not likely to directly provide traditional assessment metrics (such as IRI and PCI), new metrics might supplement and eventually supplant traditional metrics. This paper presented a study on the utilization of smartphones and artificial neural networks for the detection of potholes. The applied methodology is readily available, low-cost, and accurate and can be utilized in participatory-sensing applications leading to automated roadway assessment and pavement management systems. The presented study documents the detection of potholes, exhibiting accuracy levels higher than 90%. The proposed methodology is currently field-tested with a higher number of roadway defect types and with larger data sets, as well as a GIS-based DSS for pavement management.

## References

- AASHTO. 1990. *Guidelines for pavement management systems*. Washington, DC: AASHTO.
- Alessandroni, G., L. Klopfenstein, S. Delpriori, M. Dromedari, G. Luchetti, B. Paolini, A. Seraghi, E. Lattanzi, V. Freschi, and A. Carini. 2014. "SmartRoadSense: Collaborative road surface condition monitoring." In *Proc., UBIComm-2014*. Wilmington, DE: IARIA XPS.
- Dennis, E., Q. Hong, R. Wallace, W. Tansil, and M. Smith. 2014. "Pavement condition monitoring with crowdsourced connected vehicle data." *Transp. Res. Rec.* 2460: 31–38. <https://doi.org/10.3141/2460-04>.
- De Zoysa, K., C. Keppitiyagama, G. P. Seneviratne, and W. Shihan. 2007. "A public transport system based sensor network for road surface condition monitoring." In *Proc., 2007 Workshop on Networked Systems for Developing Regions*, 9. New York: Association for Computing Machinery.
- Eriksson, J., L. Girod, B. Hull, R. Newton, S. Madden, and H. Balakrishnan. 2008. "The pothole patrol: Using a mobile sensor network for road surface monitoring." In *Proc., 6th Int. Conf. on Mobile Systems, Applications, and Services*, 29–39. New York: Association for Computing Machinery.
- Hadjidemetriou, G. M., S. E. Christodoulou, and P. A. Vela. 2016. "Automated detection of pavement patches utilizing support vector machine classification." In *Proc., Electrotechnical Conf. (MELECON), 2016 18th Mediterranean*, 1–5. Piscataway, NJ: IEEE.
- Koch, C., and I. Brilakis. 2011. "Pothole detection in asphalt pavement images." *Adv. Eng. Inf.* 25 (3): 507–515. <https://doi.org/10.1016/j.aei.2011.01.002>.

- Koch, C., G. Jog, and I. Brilakis. 2013. "Automated pothole distress assessment using asphalt pavement video data." *J. Comput. Civ. Eng.* 27 (4): 370–378. [https://doi.org/10.1061/\(ASCE\)CP.1943-5487.0000232](https://doi.org/10.1061/(ASCE)CP.1943-5487.0000232).
- Kyriakou, C., S. E. Christodoulou, and L. Dimitriou. 2016a. "Road anomaly detection and classification using smartphones and artificial neural networks." In *Proc., Transportation Research Board 95th Annual Meeting*. Washington, DC: Transportation Research Board.
- Kyriakou, C., S. E. Christodoulou, and L. Dimitriou. 2016b. "Roadway pavement anomaly classification utilizing smartphones and artificial intelligence." In *Proc., Electrotechnical Conf. (MELECON), 2016 18th Mediterranean*, 1–6. Piscataway, NJ: IEEE.
- Kyriakou, C., S. E. Dimitriou, and L. Dimitriou. 2017. "Detecting and classifying roadway pavement anomalies utilizing smartphones, on-board diagnostic devices, and classification models." In *Proc., Transportation Research Board 96th Annual Meeting*. Washington, DC: Transportation Research Board.
- Li, Z., I. Kolmanovsky, E. Atkins, J. Lu, and D. Filev. 2015. "Road anomaly estimation: Model based pothole detection." In *Proc., American Control Conf.*, 1315–1320. Piscataway, NJ: IEEE.
- Mohamed, A., M. M. M. Fouad, E. Elhariri, N. El-Bendary, H. M. Zawbaa, M. Tahoun, and A. E. Hassanien. 2015. "RoadMonitor: An intelligent road surface condition monitoring system." In *Intelligent systems'2014*, 377–387. Berlin: Springer.
- Mohan, P., V. N. Padmanabhan, and R. Ramjee. 2008. "Nericell: Rich monitoring of road and traffic conditions using mobile smartphones." In *Proc., 6th ACM Conf. on Embedded Network Sensor Systems*, 323–336. New York: Association for Computing Machinery.
- Radopoulou, S. C., and I. Brilakis. 2016. "Automated detection of multiple pavement defects." *J. Comput. Civ. Eng.* 31 (2): 04016057. [https://doi.org/10.1061/\(ASCE\)CP.1943-5487.0000623](https://doi.org/10.1061/(ASCE)CP.1943-5487.0000623).
- Radopoulou, S. C., I. Brilakis, K. Doycheva, and C. Koch. 2016. "A framework for automated pavement condition monitoring." In *Proc., Construction Research Congress 2016*, 770–779. Reston, VA: ASCE.
- Seraj, F., B. J. van der Zwaag, A. Dilo, T. Luarasi, and P. Havinga. 2014. "RoADS: A road pavement monitoring system for anomaly detection using smart phones." In *Proc., Int. Workshop on Modeling Social Media*, 128–146. New York: Springer.
- Strazdins, G., A. Mednis, G. Kanonirs, R. Zviedris, and L. Selavo. 2011. "Towards vehicular sensor networks with android smartphones for road surface monitoring." In *Proc., 2nd Int. Workshop on Networks of Cooperating Objects*. Chicago: European Commission.
- Tai, Y., C. Chan, and J. Y. Hsu. 2010. "Automatic road anomaly detection using smart mobile device." In *Proc., Conf. on Technologies and Applications of Artificial Intelligence*. Piscataway, NJ: IEEE.
- Walker, D., L. Entine, and S. Kummer. 2002. *Pavement surface evaluation and rating: PASER manual*. Madison, WI: Univ. of Wisconsin and Transportation Information Center.
- Yi, C.-W., Y.-T. Chuang, and C.-S. Nian. 2015. "Toward crowdsourcing-based road pavement monitoring by mobile sensing technologies." *IEEE Trans. Intell. Transp. Syst.* 16 (4): 1905–1917. <https://doi.org/10.1109/TITS.2014.2378511>.