Becoming Smarter at Characterizing Potholes and Speed Bumps from Smartphone Data — Introducing a Second-Generation Inference Problem

M. Ricardo Carlos, Luis C. González, *Member, IEEE*, Johan Wahlström, Raymundo Cornejo, and Fernando Martínez

Abstract—Much has been said regarding the automatic identification of roadway obstacles by analyzing data collected from inertial sensors either fixed to the vehicle or embedded into the drivers' smartphones. Literature is vast in models to, given a record of sensor readings, determine if the sample corresponds to a pothole or speed bump, even with scores beyond 90% in performance. Acknowledging this advance, this article considers the next-generation version of this problem. Specifically, we investigate questions such as: what physical properties of roadway obstacles could be inferred from the same sensor readings? or, what are the best schemes to model this profile problem? To approach these questions we built the first obstacle-detailed data set that is composed of accelerometer and gyroscope readings of 163 potholes and 101 speed bumps. This data set is the first of its kind, since it specifies ground truth labels that correspond to potholes' depths and also, functional status (OK - Not OK) for speed bumps. We approach this fine-grained characterization using three different learning schemes, as a Regression, Classification and Learning to Rank tasks. Results are encouraging, reporting a RMSE for pothole's depth prediction of up to 1.68 cm and classification performance of 0.89 in AUC score. In summary, after more than 10 years of analysis, struggles and achievements, it is time for the community to become smarter and start profiling roads with real detail.

Index Terms—Road obstacle profiling, bump detection, road quality assessment, smartphone sensing, benchmark data set

1 Introduction

In a smart-city environment we expect any citizen to have the opportunity to actively participate in collecting and sharing data about her surroundings with the aim to improve common spaces. In this regard, data with potential application to road quality assessment is crucial since it impacts concepts such as trip safety and (shorter) transportation times. With thousands of citizens systematically reporting this data, robust platforms could be envisioned to generate meaningful information in terms of routes with high quality roads, or warnings for potential hazards on the roads' surface, e.g., potholes or speed bumps without proper signaling. Having this information will empower the common drivers and simultaneously would provide, in real-time, a list of road segments for authorities to attend.

Two main approaches have been reported in the literature to support a scenario where the quality of the road is permanently and automatically assessed. The first approach consists of equipping vehicles with dedicated devices, such as a Ground Penetrating Radar (GPR) to optically analyze obstacles on the road [1], [2]. The second approach looks to exploit data collected by inertial sensors via mobile computing devices carried by the driver (or embedded within the vehicle), thus aiming to gain some knowledge from vibrational signals. Pros and cons for both approaches do exist. For instance, the field of image processing has matured quickly and algorithms with high discriminative capacity are a reality. However, installing and maintaining the required hardware could lead to high costs. On the other hand, although a high noise-to-signal ratio could be present in sensor data collected with current smartphone technology [3], it is important to note that mobile devices offer convenience of use, affordability, and the ability to carry out practically any processing task, even the on-line execution of learning algorithms [4].

With this analysis in hand, many studies have been inclined to work on evaluating ad-hoc representations and algorithms for data obtained using accelerometers and gyroscopes embedded in drivers' smartphones, two specific sensors to detect inertial forces [5]. Several proposals have been focused on analyzing this type of sensor data in quasi-raw format, others in summarizing it by means of statistical scores (both in the time and frequency domain) [6], [7], [8], or even using second order representations [9]. Simultaneously, more sophisticated approaches have been implemented such as threshold-based heuristics [10] and Machine Learning (ML) algorithms [11], [12].

- *M.R. Carlos, L.C. González, R. Cornejo, and F. Martínez are with Facultad de Ingeniería, Universidad Autónoma de Chihuahua, Chihuahua 31125, México. E-mail: {ricardo.carlos, rcornejo82}@gmail.com, {lcgonzalez, fmartine}@uach.mx.*
- *J. Wahlström is with the Department of Computer Science, University of Oxford, Oxford OX1 2JD, United Kingdom. E-mail: johan.wahlstrom@cs.ox.ac.uk.*

Manuscript received 6 Mar. 2019; revised 15 Aug. 2019; accepted 8 Oct. 2019. Date of publication 21 Oct. 2019; date of current version 7 Jan. 2021.
(Corresponding author: Luis C. González.)
Digital Object Identifier no. 10.1109/TMC.2019.2947443

Naturally, since the appearance of the first proposals to exploit vibration-based sensing to infer road quality to nowadays, the vast body of knowledge has been primarily focused on attending a broad problem formulation. That is, once a sensor signal associated to a road segment was collected, the task is to classify that signal into one of several possible labels, i.e., pothole, speed bump, rough road, good road, etc. Note that this formulation does no consider differentiating the severity level or functional status of road artifacts, limiting the proposed solutions to keep minor detail. Following this historical perspective we call this problem formulation a first-generation problem. Fair to say that this problem has been practically solved, since current computational approaches report scores around 90 percent in performance [5]. It is appropriate then, to acknowledge these advances, and propose the next steps to follow. This is what this study attempts, by formally proposing the second-generation version of this problem.

1.1 A Second-Generation Inference Problem for Road Quality Assessment

While traveling on a vehicle, road segments could be sensed by current vibration-based technology such as accelerometer and gyroscope, and by extension any road anomaly could also have a sensor signal associated to it. For our purposes, we define a road anomaly (or road artifact) as any permanent artifact located on the surface of roads that prevents vehicles from driving in a safely and expedite way. Common examples of such artifacts are potholes and speed bumps, which are both handled in this study. We proceed then, to formally define for the first time the Second-Generation problem as follows.

Let a feature vector fv_i, which was derived from the corresponding list of sensor readings of the ith road anomaly, have a corresponding ground truth value dim_i, which in turn defines a physical measure of the given road artifact.[1] Now, consider a set of statistical models $\phi(\cdot) \in \Phi$ that receive as input fv_i and generates as output an estimated value that could be a real number or a class label. We, then, aim to find a model $\hat{\phi}$ such that

$$\hat{\phi} = \arg\min_{\phi \in \Phi} \sum_i \mathrm{error}(\phi(fv_i), dim_i), \tag{1}$$

where function $\mathrm{error}(\cdot,\cdot)$ is some chosen error metric. Moreover, the model $\hat{\phi}(\cdot)$ is expected to have a generalization property, since it could be applied to examples of road anomalies where the ground truth is not known.

The hypothesis behind this problem formulation is that not only a broad classification of road anomalies could be made, but also a fined-grained prediction. For this computational task it is of interest to know about the nature and physical characteristics of anomalies on roads, that is, a step beyond from what traditionally was inferred. This new perspective looks to add detail to current predictions, since based on the assumption that this information could be shared among drivers, it is more valuable for a unaware driver to know that a pothole with certain dimensions, or a given type, awaits ahead in the road than only knowing that a pothole exists in that area.

1. For a pothole the physical measure is the dimension of depth, which could be specified in granular units of cm or inches. For the case of a speed bump we simplify its physical dimensions to two categorical levels representing its functional status: OK or Not OK.

Note that given the novelty of the described problem, there is no clear guide on how to model it. This way, in this study we use a Machine Learning approach, which has shown its strengths in tackling the first generation version of the problem [9], [13]. Accordingly, we will compare prominent learning tasks, that given their mature status, are open to well-proven data representation and effective algorithms. These tasks are: classification, regression and learning to rank, which although they represent interrelated methods, learn different patterns within data. Thereby, some of the algorithms may prove to be more valuable (in terms of error prediction and inference granularity) than the others. Thus, in this study we will report results that could be used as baseline by the community, but also to demonstrate to city officials the advantages of the approach to finely evaluate real-time feedback from drivers smartphones throughout the city. This solution makes it possible to take smarter actions to mitigate deterioration in road infrastructure via a what-to-attend-first list based on the severity of road anomalies.

To approach this new problem, we need to introduce a novel data set composed by accelerometer and gyroscope readings of potholes and speed bumps, carefully characterized by a ground truth value that specifies real-world depths (in cm) and functional status for potholes and speed bumps, respectively. This is the largest data set of its kind, and it is unique in its level of detail. We announce that this data set will be publicly available, then enabling the benchmarking of future methods and thereby lay the foundation for the next generation of smartphone-based roadway characterization.

The organization of this study is as follows: in Section 2 we revise relevant work associated to this problem. In Section 3 we introduce the novel benchmark data set. In Section 4 we describe the learning approaches and the methods that we will use to model the problem. We organized Section 5 in three experiments, these being the ones associated to each type of problem modeling: regression, classification and ranking. Section 6 presents open inquires that could guide future research paths and, finally, Section 7 concludes this study.

2 Related Work

This section surveys earlier work on road anomaly characterization using inertial sensors. As such, it could be read as a summary of proposals for the first generation version of the problem.

One of the first studies was Pothole Patrol (PP) [14], which equipped a taxi fleet with dedicated accelerometer sensors placed in the glove compartment. The authors exploited information from accelerometer measurements in the vehicle's vertical and lateral axes and reported a pothole detection accuracy of 92 percent (the fraction of reported potholes that were actual potholes). Another seminal work was Nericell [15]. One of their main contributions was to use, for the first time, sensors embedded within the smartphone to detect potholes, bumps, and braking events. This work was also the first to propose applying a re-orientation to align the accelerometer axes with the vehicle axes. Using a threshold-based

TABLE 1
Vehicles Used in This Article for Data Collection

Vehicle	Make	Type
Rio 2018	Kia	Hatchback
Pointer 2004	VW	Hatchback
Cruze 2010	Chevrolet	Sedan
Partner 2013	Peugeot	Pickup Truck
Derby 2001	VW	Sedan
Accord 2007	Honda	Sedan
Van 2013	Chevrolet	Passengers Truck
Spark 2013	Chevrolet	Hatchback

heuristic, the best performance that Nericell reported was 8 percent False Positives (FP) and 41 percent False Negatives (FN) in a high speed scenario.

Many studies on smartphone-based pothole characterization followed after Nericell. These include threshold-based methods, learning-based methods using features from the frequency domain, as well as other machine-learning algorithms. For example, Fazeen et al. [10] proposed a multi-label threshold-based strategy to classify speed bumps, potholes, uneven road, and regular road, reporting an average accuracy of 85 percent. Astarita et al. [16] designed a threshold-based heuristic that worked on the vertical acceleration signal for monitoring pavement quality, registering a True Positive Rate (TPR) of 93 percent for pothole detection. Studies using features from the frequency domain include [6], where the authors applied a Fast Fourier Transform to the accelerometer signal and calculated the mel frequency cepstral coefficients to build a feature vector. The feature vector was fed to a Support Vector Machine (SVM). Their best result was a score of 89 percent using the G-means metric (the geometric mean of sensitivity and specificity). Similarly, Seraj et al. [8] built a feature vector with stationary Wavelet decomposition applied over both accelerometer and gyroscope sensors, and using a clustering algorithm reported a G-means score of 89 percent.

Among other methods, Celaya et al. [12] proposed a Genetic Algorithm to evolve logistic functions to classify segments of road having a pothole. To this end the authors used dedicated sensors mounted in a car, reporting that the best model using five statistical features reached an accuracy of 97.14 percent. In the work of Wang et al. [17] the main hypothesis was that Mahalanobis distance is better suited to measure correlations in features derived from accelerometer and gyroscope sensors. Applying a threshold-based heuristic the authors report an error rate for pothole detection within the range of 2.33 and 2.49 percent, whilst for speed bump detection the error rate was below 0.02 percent. Surveys of methods for smartphone-based road anomaly characterization can be found in [18] and [19].

Several studies have contrasted the performance of existing algorithms. For example, Mednis et al. [7] evaluated four threshold-based methods that they named Z-THRESH, G-ZERO, Z-DIFF and STDEV(Z). The best performance was achieved with the Z-DIFF algorithm which represent pothole events using the difference between consecutive accelerometer measurements. Similarly, Carlos et al. [13] made the first public implementation (with available code) of six threshold-based heuristics: Nericell, Pothole Patrol, Z-THRESH, G-ZERO, Z-DIFF and STDEV(Z). The results showed that the best performer was the heuristic called STDEV(Z), which represent events using the standard deviation of accelerometer measurements along the vertical axes. In the same work, the authors took advantage of the way STDEV(Z) worked and proposed a novel feature vector feeding a SVM, finally outperforming STDEV(Z).

One big issue has been the lack of public data sets that can be used to compare and benchmark competing algorithms (most studies have only tested their proposal on proprietary data sets). Hence, in a first attempt to get closer to an objective evaluation of several works, Gonzalez et al. [9] constructed five data sets containing most of the anomalies analyzed in prior works, and using several ML classifiers they contrasted their results against the scores reported previously in [15], [6], [8], and [16].

Beyond basic identification or classification of road anomalies, Xue et al. [20] proposed a smartphone's sensing scheme called P^3 (Perceiving Pothole Profiles) whose aim was to identify the depth and length of only potholes. P^3 reported an estimation error rate (the mean absolute estimation error divided by the true magnitude of the quantity) of about 13 and 16 percent, for the task of inferring pothole's depth and length, respectively. They based their estimation on a degree-of-freedom vibration model, where the vehicle is simplified as a box connected to a suspension system. Some characteristics of their study that contrast with this work are the extremely small data set that they employed, and that their system requires knowing beforehand where the smartphone is placed within the vehicle and the vehicle's instant speed (variables that our methodology does not require).

3 Benchmark Data Set

For this study we exhaustively work on constructing a fine-grained data set unique in its type and freely available.[2] Data was collected while driving in the northern city of Chihuahua, Mexico during the summer of 2018. Eight vehicles, listed in Table 1, were used for this purpose.

We made each vehicle deliberately pass over potholes or speed bumps at a speed in the range between 20 and 40 km/h (5.5-11 m/s), which is common in urban areas [3], [12]. We randomly drove through different sections of the city and chose a pothole (or a speed bump) at the time the driver or copilot spotted it. In some cases, the vehicle passed between three and five times over the same anomaly to capture subtle differences in how the anomaly is recorded when influenced by speed, direction and anomaly's shape, aiming to simulate how different drivers would approach the obstacle. Consider Fig. 1, which shows the case of the same pothole by passing over it from different directions on the same vehicle.

3.1 Smartphone Specifications

Where does a regular driver or passenger place her smartphone when traveling? In an attempt to analyze sensor readings simulating a naturalistic scenario up to five smartphones were simultaneously used for data collection,[3] placed at different

2. This data set can be downloaded from http://www.accelerometer.xyz/datasets/
3. Smartphones models include Motorola Moto G, Samsung Galaxy S7, Huawei P20, and Xiaomi Redmi.

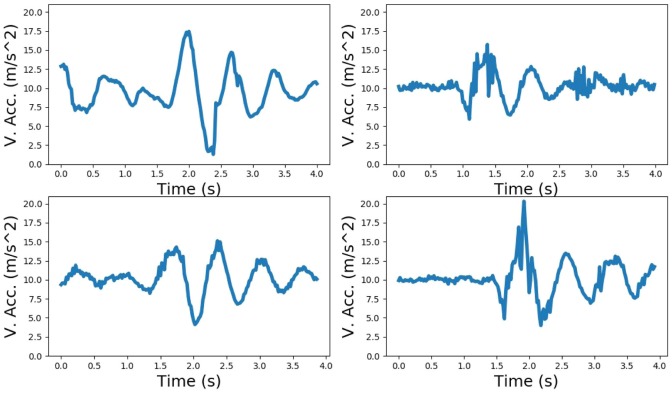


Fig. 1. Vertical acceleration signals for a pothole, with the vehicle approaching from different directions. Top left: reverse (direction), top right: forward, bottom left: reverse, and bottom right: forward.

locations inside the vehicle: the copilot's door, glove compartment, central console (or cup holder), etc. We placed smartphones in ten different locations in total since not all vehicles provide the same locations. The copilot controlled the beginning and end of the data collection using an in-house software application installed in a tablet, and this application controlled the smartphones used as sensors. All the smartphones recorded accelerometer data with a sampling rate of 50 Hz, such as some other works [9], [13].

3.2 Potholes

Fig. 2 illustrates the distribution of magnitudes of potholes' depths that compose the data set. Fig. 3 schematically shows how pothole's depth was determined, that is, as the maximum difference between the pavement level and the lowest point of its depression. To determine the exact value of depth (ground truth) a measuring tape was used. The hook was positioned at the lowest part of the pothole, and the reading was taken at the pavement level.

3.3 Speed Bumps

Although both asphalt and metal bumps can be considered the same type of artifact given their effect on the vehicle's suspension, a hypothesis is that possible differences could exist in the sensor signal based on the material that they are made of. For the specific case of asphalt bumps, all events that we considered correspond to *"...sudden and sharp 3 to 4 inch rises with a 1 to 3 foot base width..."* as described in [21]. Further, two qualitative labels were used to describe them

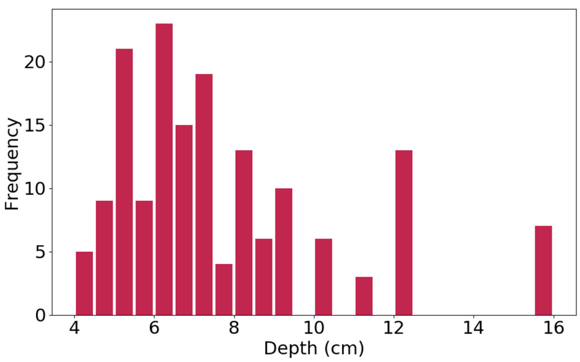


Fig. 2. Histogram for potholes' depths considered in this article.

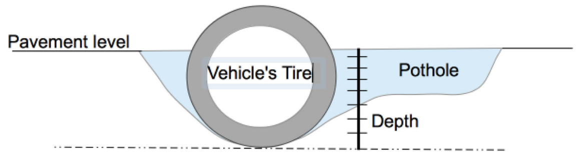


Fig. 3. Schematic representation of the depth of a pothole.

in terms of functionality (ground truth): OK or acceptable condition, and Not OK, for those found with some damage. To label a speed bump, a consensus had to be reached between the driver and the copilot of the vehicle. They both had to agree after visually inspecting the speed bump that the bump itself does not have any damage, missing part, or significant wear to assign the OK label. Consider Fig. 4 for an illustration of the concept of damage in speed bumps.

Table 2 summarizes the number of road artifacts that are present in the data set. Observe that an event corresponds to one pass of the car over the artifact. The third column indicates how many records we have of all events given all the smartphones placed at different vehicle's compartments.

3.4 Signal Preprocessing

The orientation of the smartphones was not recorded during the data collection stage. That is, the devices were placed without any consideration with respect to vehicle. As prepossessing step all signals had to be reoriented to obtain vertical acceleration (parallel to gravity) following the procedure described in [16]. Since only information from vertical axis is needed no reorientation was performed on the longitudinal/lateral smartphone's axes.

4 Experimental Setup

Since one of our aims is to contrast several learning tasks and initiate a discussion about what would be the best approach to follow, we will then, model the novel inference problem using three specific learning schemes. To properly define these tasks we use the following notation: n_{train} refers to the subset of training examples that are used to train a given model, and n_{test} refers to the subset of testing examples that are used to evaluate if an already trained model can be generalized. Observe that $n_{train} \cap n_{test} = \emptyset$.

4.1 Learning Schemes

1) Regression task for Pothole Profiling. This task consists of finding a model ϕ capable of approximating a physical quantity (such as the depth of a pothole) from a number of input variables. The output of this model is by definition continuous. Specifically, the problem is stated as follows: given a vector $X_i^T = [x_{i,1}, x_{i,2}, x_{i,3}, \ldots, x_{i,m}]$ of m features to represent example i, where $1 < i \leq n_{train}$, find a vector $W^T = [w_1, w_2, w_3, \ldots, w_m]$ of weights so that $\sum_{i=1}^{n_{test}} Error(\phi(X_i) \cdot W, target_i)$ is minimum. This paradigm is used to evaluate how accurate the depth of a pothole can be exactly inferred.

2) Classification task for Pothole and Speed Bump Profiling. For this task the first step is to train a model $\phi()$ over the set of feature vectors $X_i^T = [x_{i,1}, x_{i,2}, x_{i,3}, \ldots, x_{i,m}]$, where $1 < i \leq n_{train}$, to match the output between $\phi(x_i)$ and the true label y_i of the ith example. Afterwards, $\phi()$ is applied over

(a) Metal bump with ”Not-Ok” label

(b) Metal bump with ”Not-Ok” label

(c) Asphalt bump with ”Not-Ok” label

(d) Asphalt bump with ”Not-Ok” label

(e) Metal bump with ”Ok” label

(f) Asphalt bump with ”Ok” label

Fig. 4. Examples of speed bumps considered in this article.

the testing partition to predict the right label of each unseen example. Note that in this case the output is categorical: discrete levels that describe a pothole's depth (e.g., *shallow*, *of medium depth*, or *very deep*). In the case of speed bumps, this study will focus on identifying two classes: *OK* and *Not OK*.

3) Learning to Rank (L2R) task for Pothole Profiling. This task is slightly different and more challenging than the previous two, since it takes as input a list of objects and returns an ordered permutation of such objects, so any incorrect prediction for a given object can be highly penalized when the whole ranking is evaluated. For the case of potholes, the order is expected to be from the most shallow one to the deepest one. Thus, the model is expected to learn a comparison function that will then allow to automatically sort previously unseen examples, yielding a ranked list. We evaluate this task since it could be useful to offer an immediate feedback to the city, under a what-to-fix-first principle, bypassing the regression and classification steps.

4.2 Feature Extraction

A key component of a successful model is the data representation. In this regard we evaluate a novel set of 52 features (shown in Table 3) to describe the signals in the time and frequency domains.[4] Feature extraction is performed in two stages. Three new time series are obtained in the first stage: the first derivative of the signal (jerk in the vertical axis), and the first and second integrals (vertical speed and displacement, respectively). The derivative is calculated as the difference between consecutive acceleration samples ($\Delta x = x_i - x_{i-1}$), while the integrals are calculated by means of omega arithmetic [23]. Now, with four available time series representing each example, the thirteen features listed in Table 3 (under the column Descriptor) are calculated for each of these time series, producing a total of 52 features. All features are standardized by removing the mean and scaling to unit variance before being fed into the regressors and classifiers.

4. The longitudinal speed of the vehicle was also evaluated as a feature but, in accordance with the results reported in [22], it did not produce any gain in prediction, so it was decided to keep it out of the final feature vector.

4.3 Algorithmic Machinery

Table 4 presents the algorithm-task pairings that will be evaluated. The rationale behind only choosing machine learning algorithms is that these algorithms have outperformed other approaches in the first generation version of this problem [13]. All implementations were the ones provided by the python-based scikit-learn library [24].

In addition to these algorithms, we added three baseline procedures: A Linear Regression model (LR), and two dummy functions MEAN and RAND. For a given physical dimension of an anomaly, MEAN always returns its mean value in the data set as the prediction, whereas RAND returns a random number in the range between the minimum and maximum values.

5 Experiments and Results

As a visual aid for the organization of the experiments we refer the reader to Fig. 5. A 10-fold cross-validation scheme was used to obtain the training and testing partitions. We performed the computations for thirty runs. All splits are also available at http://www.accelerometer.xyz/datasets/.

TABLE 2
Overview of the Data Set for Pothole and Speed Bump Characterization

Anomaly	Events	W/all smartphones
Potholes	163	857
Asphalt Speed Bumps	59 (Ok: 43, Not Ok: 16)	308
Metal Speed Bumps	42 (Ok: 26, Not Ok: 16)	211
Total	264	1,376

TABLE 3
Features Considered in This Article

Descriptor	1st Deriv.	Original	1st Int.	2nd Int.
Mean	0 (*)	13	26 (*)	39(*)
Std. Dev.	1	14 (*)	27 (*)	40(*)
Max.	2 (*)	15 (*)	28 (*)	41(*)
Min.	3 (*)	16 (*)	29 (*)	42(*)
Definite integral	4	17 (*)	30 (*)	43(*)
25th percentile	5 (*)	18	31 (*)	44(*)
50th percentile	6	19 (*)	32 (*)	45(*)
75th percentile	7 (*)	20 (*)	33 (*)	46(*)
Zero crossing rate	8 (*)	21 (*)	34 (*)	47(*)
Spectral centroid	9 (*)	22 (*)	35 (*)	48(*)
Spectral flatness	10	23	36 (*)	49(*)
Spectral roll-off	11	24	37 (*)	50(*)
Spectral bandwith	12	25 (*)	38	51(*)

TABLE 4
Algorithm-Task Pairings to be Evaluated

Algorithm	Learning task
Random Forest (RF), 150 trees, Gini crit	Regress, Class, L2R
Gradient Boosting (GB), 150 trees, 3 nod/tree	Regress, Class, L2R
k Nearest Neighbor (kNN), K=5, Mink dist	Regress, Class, L2R
Bayesian Ridge (BR), $\alpha_1, \alpha_2, \lambda_1, \lambda_2 = 1e-6$	Regress, L2R
Decision Tree (DT), Gini crit	Regression
Multi-Layer Perceptron (MLP), 26 units, Relu	Class, L2R
Support Vector Machine (SVM), Linear, C=1	Classification
Nave Bayes Classifier (NB)	Classification

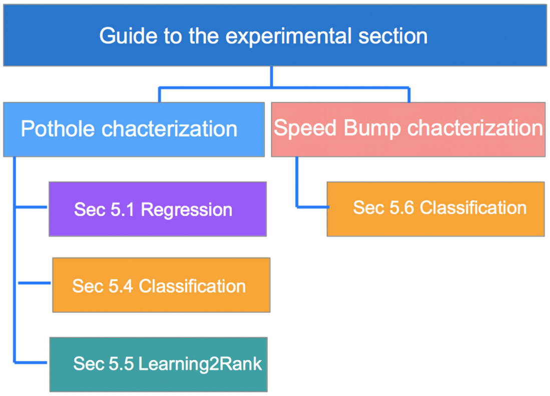


Fig. 5. Organization of the learning tasks approached in this study.

TABLE 5
Average Performance Metrics for Pothole Depth Regression

Algorithm	ϵ	R^2	RMSE
RF	**0.22**	**0.45**	**1.95**
RF (41 Feat - F test)	0.22	0.45	1.95
GB	0.23	0.41	2.02
RF* (5 Feat - Gini)	0.23	0.40	2.03
BASELINE 1 (LR)	0.24	0.34	2.13
KNN	0.24	0.33	2.15
BR	0.24	0.33	2.15
DT	0.29	-0.1	2.78
BASELINE 2 (MEAN)	0.29	-0.02	2.8
BASELINE 3 (RAND)	0.65	-2.76	5.14

5.1 Pothole Characterization by Regression

For the regression task we consider three performance metrics. The relative error (ϵ), which presents the average of errors as a ratio of the actual measured values

$$\epsilon = \frac{1}{n}\Sigma_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right|, \tag{2}$$

where $\hat{y}_i$ corresponds to the prediction made by a model for an example i, y_i is the real (observed) value of the example, and n the number of examples.

The average root-mean-square error (RMSE)

$$RMSE = \sqrt{\frac{1}{n}\Sigma_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2} \tag{3}$$

captures the magnitude of the average error in the same units of the measurement. For both metrics, zero indicates perfect performance, and higher values reflect the amount of errors. And finally, we compute the coefficient of determination

$$R^2(y_i, \hat{y}_i) = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}, \tag{4}$$

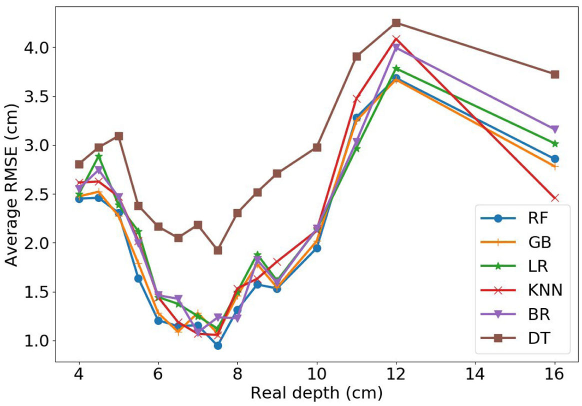
to measure the extent to which the variance of the variable of interest, i.e., depth, is predictable. For this metric a value between 0 and 1 indicates the extent to which unknown examples will be correctly predicted; zero and negative values indicate poor prediction.

Table 5 summarizes the performance of all methods for the regression task to predict pothole's depth when only using accelerometer data. From this results Random Forest (RF) emerges as the most effective regressor, with a relative error (ϵ) of about 0.22. These results also suggest that ensemble-based methods (RF and GB) provide a robust modeling alternative for this problem. Interestingly, Linear Regression, a baseline method, yields better results than other three ML approaches. The output of the other two (dummy) baseline functions could help contextualize to what extent knowledge inference is being achieved by the other methods. The Kruskal Wallis H test (with a significance level of 0.05) confirms the differences in the performance of the algorithms, and a post-hoc Mann-Whitney U tests validates that RF's ε is superior to the rest of the classifiers.

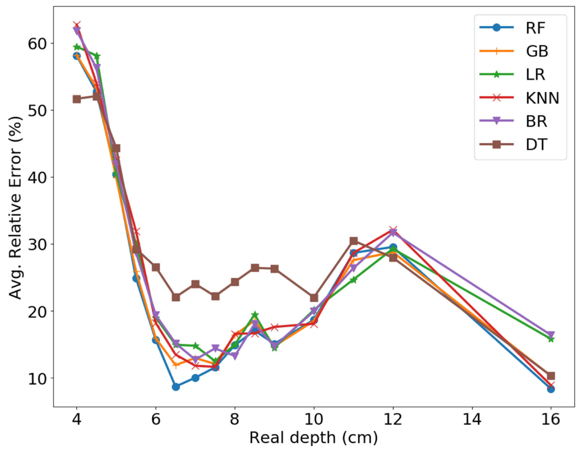
As we observed in Fig. 2, the data set is composed by examples of potholes having different depths, so now, we investigate how depth affects prediction error. Fig. 6 shows the performance of all regressors in function of depth. We observe that there exists a depth range for which most of the regressors report their minimum error, these samples correspond to medium-sized potholes (5-8 cm deep), even reaching up to 1 cm in RMSE.

5.2 Feature Selection

To analyze the impact that the number of features has in the performance of the algorithms we applied two strategies. First, we aimed to identify the minimum number of features that could be used to achieve the best performance. Accordingly, we applied the F-test, which gave as a result a subset of 41 features (shown with "*" in Table 3). Then, we evaluated how a method would perform if we just kept the most relevant features according to certain metric. To do this we computed the Gini importance over the complete set of



(a) Average RMSE by real depth for the evaluated algorithms.



(b) Average ϵ by real depth for the evaluated algorithms

Fig. 6. Error scores for pothole's depth prediction as a function of depth.

52 features. This analysis, shown in Fig. 7, clearly highlights five relevant features: the definite integral of vertical acceleration, the maximum value of vertical speed (the first integral of our original signal), the standard deviation of vertical acceleration, the minimum value of vertical acceleration, and the maximum value of the first derivative of the signal (jerk). Using these two strategies for feature selection, we built new feature vectors and fed the best regressor (RF). Results, presented in Table 5, suggest that it is possible to reduce the number of features and still be competitive. More important, if RF only uses 5 features, a reduction of 90 percent over all features, only loses a 4.5 percent performance in ϵ.

5.3 Analysis of Gyroscope Data

As explained in the experimental setup, the vast majority of the sensor signals that were collected come from the accelerometer sensor, present in all the smartphones that were employed. Even when there was only one smartphone that provided gyroscope, we include in this study the analysis of gyro information for pothole characterization as motivated by [25], [26]. To this end, we applied the same feature extraction procedure over the gyroscope signals that we did previously, that is, only using the five most important features and feed all the algorithms. Table 6 summarizes the error rates for the regression task using only gyro information. We observe that the performance of the classifiers is consistent (similar ordering) with those only using acceleration data and that most of the classifiers improve their error rate. More specifically, the gyroscope data improves the RMSE and ϵ by about 8 percent in comparison to the best performance using only accelerometer data (shown with an "*" in Table 6). Based on this improvement, we decided to test sensor fusion by concatenating both sensor data within the same feature vector. The complete set of features computed originally for accelerometer were also obtained for gyroscope. This characterization was able to obtain the same error metrics as only using gyro data, but it improved R^2 from 0.42 to 0.47 (not shown in Table 6). This finding could suggest the convenience to exploit angular velocity for pothole fine-grained characterization.

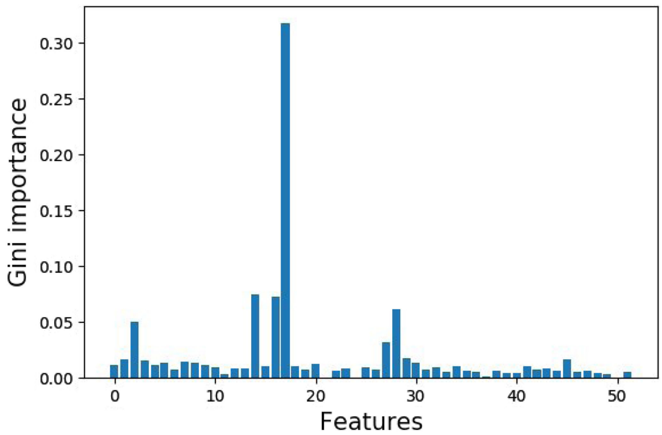


Fig. 7. Gini importance for each feature as computed by the Random Forest. The features are represented by their ordinal position, shown in Table 3.

5.4 Pothole Characterization by Classification

For this task we model the pothole characterization problem as a binary and multi-class classification problem (with three and four labels). These categories cover equally spaced intervals between 4 and 16 cm of depth for potholes (e.g., 4-10 cm and 10-16 cm as limits to assign a binary tag). By considering several quantizations, we can evaluate the performance at different levels of granularity. Thus, for the binary problem we are only interested in assigning labels that can be mapped to descriptions such as *deep* or *shallow*, whereas for three classes the labels could be *deep*, *medium*, and *shallow*, and so on.

We measured classification performance by employing the Area under the curve (AUC) score, which may be interpreted as the probability of classifying a random positive instance higher than a random negative instance. A perfect AUC score would be 1.0. The AUC is calculated as

$$AUC = \frac{1}{2}\left(\frac{TP}{TP+FN} + \frac{TN}{TN+FP}\right), \tag{5}$$

where TP, FP, TN, and FN denote True positives, False positives, True negatives, and False negatives, outputs respectively. In order to use this binary classification metric for multi-class cases, we transformed the problem into binary classification with one-versus-the-rest (OvR) approach [27].

Since there is some degree of imbalance in the data we also calculate the G-means score [28] as

TABLE 6
Average Performance Metrics for Pothole Depth Regression from Gyroscope Data

Algorithm	RMSE (cm)	ϵ
RF (Gyro+Acc)	**1.68**	**0.19**
RF	**1.68**	**0.19**
GB	1.73	0.19
KNN	1.57	0.18
BR	1.72	0.20
RF* (Accel)	1.95	0.22
DT	2.24	0.23
BASELINE 1 (LR)	3.29	0.33
BASELINE 2 (MEAN)	3.12	0.31
BASELINE 3 (RAND)	5.31	0.68

$$G = \sqrt{\frac{TP}{(TP+FN)} \times \frac{TN}{(TN+FP)},} \tag{6}$$

which penalizes results when the classifier is better at correctly identifying positive over negative examples, and vice versa.

Table 7 presents the classification results for different levels of granularity (n categories). As it could be expected when the number of labels increases the problem becomes harder, but a good trade-off between granularity and performance seems to be with $n = 3$. In most levels the best performers continue being RF and GB, as in the regression task. These differences are confirmed with the same statistical procedure explained previously. These results suggest that it is possible to properly classify potholes' depth in discrete levels, then being possible to signal potholes that need urgent repair.

5.5 Pothole Characterization by Learning to Rank

For this task we evaluate two different ways to generate a ranking for a collection of potholes:

- A point-wise approach, in which we use the estimation made by depth regression (see Section 5.1), as an input to a sorting algorithm to produce the ordered list for the testing data; a naive approach for which we need the regression task to be solved beforehand.
- A pair-wise approach, in which we train a binary classifier on all the permutations for a pair of potholes in the training set to learn an ordering based on their actual depths. Afterwards, the learned model is used by a sorting algorithm as a comparison function to produce an ordered list of potholes in the testing set.

TABLE 7
Average Performance Metrics for Pothole Categorical Depth Classification

	$n = 2$		$n = 3$		$n = 4$	
Algorithm	AUC	G	AUC	G	AUC	G
GB	**0.89**	0.6	**0.84**	0.61	0.74	0.47
RF	0.87	0.47	0.83	0.49	**0.75**	0.31
MLP	0.84	**0.64**	0.80	**0.62**	0.71	0.48
KNN	0.78	0.52	0.76	0.54	0.69	0.46
SVM	0.83	0.37	0.82	0.36	0.74	0.28
NB	0.76	0.57	0.69	0.55	0.64	**0.51**

TABLE 8
Average Kendall's τ over the 10-Fold Cross-Validation for the Point-Wise Approach

Algorithm	τ training	τ testing
RF	**0.9**	**0.34**
GB	0.79	0.31
KNN	0.47	0.24
LR	0.36	0.28
BR	0.34	0.27
MEAN	−0.08	−0.08
RAND	0.0004	−0.02

To evaluate the output of both approaches we use the Kendall rank correlation coefficient (τ) [29]. This score measures the similarity of two orderings of ranked data based on the number of different pairs between ordered sets. In this study, the two orderings correspond to the output of each classifier (prediction) and the list of the testing anomalies ordered by their true known depths (the ground truth). This correlation coefficient can take values within the continuous range between -1 (total agreement but in inverse order) and 1 (total agreement), with 0 representing no agreement at all. The metric is defined as

$$\tau = \frac{P - Q}{\sqrt{(P+Q+T) * (P+Q+U)}}, \tag{7}$$

where P is the number of concordant pairs, Q is the number of discordant pairs, T is the number of ties only in the first list, and U is the number of ties only in the second list (a tie is not counted if it occurs in both lists).

5.5.1 Point-Wise Ranking

Table 8 presents the results for this task using the point-wise approach. The contrast between the results for the learning and testing phases suggests the degree of complexity of this task. However, there is some learning that is being achieved, since both baseline methods (MEAN and RAND) are farther away by more than 200 percent in τ score only considering the performance in the testing phase. For this approach, the best performance is obtained by Random Forest.

5.5.2 Pair-Wise Ranking

The results of this task are presented in Table 9. For this case, the baseline approaches are RAND (randomly choose between pothole 1 or pothole 2) and FIRST (always returning the first pothole in the pair-wise comparison). Similar to the point-wise approach, these results suggest that the best option to offer a ranked list of potholes is the one produced by Random Forest.

Note that to obtain a perfect τ score for this task there must not be any error in the ordered position of any pothole within the output list. But, how hard is this? To get an idea of this task's complexity we created a slopegraph (Fig. 8), where the elements of the correct list are joined by line with the position that they receive in the predicted list. We observe that although there are some predictions that are acceptable (red lines) the majority of potholes were not properly ranked. Interestingly, the potholes that show less error are located at

TABLE 9
Average Kendall's τ Over the 10-Fold Cross-Validation
for the Pair-Wise Approach

Algorithm	τ training	τ testing
RF	0.**95**	**0.29**
GB	0.52	0.29
MLP	0.71	0.22
RAND	−0.003	−0.002
FIRST	−0.08	−0.09

the bottom of the ranking, corresponding to those that are the worst in terms of depth.

5.6 Speed Bump Characterization

Table 10 presents classification results for each type of speed bump, and also when both types are combined in the same data set.

These results suggest that if the goal is to recognize functional status, it is better to treat asphalt and metal bumps separately. This could be done given previous literature findings, for instance check González et al. [9, pp. 2921], who claimed that: "*... sub classification of bumps is possible,* [even beyond 90% accuracy when differentiating metal from asphalt bumps] *which may open new research directions to fine-grained classification for all events*". To exploit this, a novel pipeline could apply a binary classifier to identify if a bump is either asphalt or metal, and then characterize it with a material-specific model as the ones proposed in this study.

Regarding the comparison of the models, Gradient Boosting (GB) and Random Forest are the two top options for this task, both statistically validated with Kruskal-Wallis H and Mann-Whitney U tests, with a significance level of 0.05.

6 Open Inquires

We have provided some evidence that tackling a second-generation problem with the proposed models is attainable,

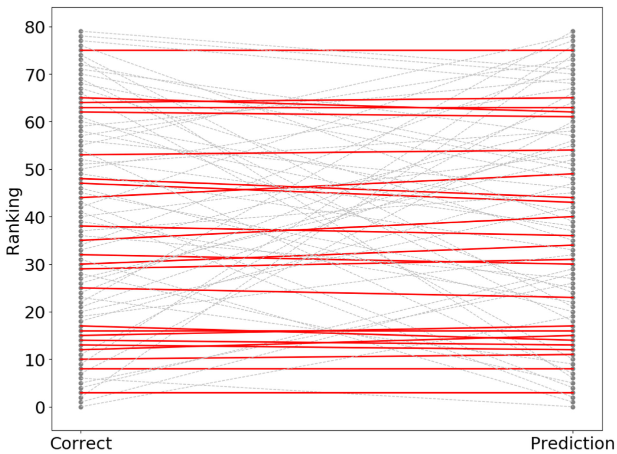


Fig. 8. A slopegraph that contrasts the correct rank of a list of potholes (left) with the actually predicted rank by the point-wise approach (right). A line connects the same pothole on both ranked lists. Solid red lines indicate a position error of less than 5 places, and dotted gray ones shows a larger error in position.

TABLE 10
Average Performance Metrics for Binary Speed
Bump Condition Classification

	Asphalt		Metal		Combined	
Algorithm	AUC	G	AUC	G	AUC	G
GB	0.75	0.57	**0.72**	**0.61**	0.69	**0.52**
RF	0.**76**	0.54	0.71	0.55	**0.70**	0.48
MLP	0.75	0.57	0.66	0.53	0.62	0.49
SVM	0.75	0.60	0.59	0.47	0.63	0.48
KNN	0.67	0.48	0.60	0.38	0.59	0.45
NB	0.69	**0.61**	0.62	0.48	0.59	0.41

however, we have detected some open questions that deserve more attention and need to be further investigated:

- Regarding appropriate source of information. Can it be confirmed that gyro data has more advantages to perform a fine-grained characterization than accelerometer data?
- Regarding sensor fusion. What other ways can be envisioned to effectively incorporate data from different sensors?[5]
- Regarding problem formulation. Are there other ways to model this problem that could offer better results than the ones obtained by regression or classification?
- Based on the assumption that different damage levels of speed bumps may vary greatly, is it possible to have even a finer representation of such cases?
- Regarding the performance of the Learning to Rank task. Based on the assumption that no overfitting exists, what pipeline could provide a better ranking of road anomalies?
- Regarding the size of the Benchmark Data Set. Is it possible to apply a data augmentation strategy to generate meaningful examples to better train current models? (Even opening the possibility to find better data representation by deep models).

7 Conclusion

This study proposes to go beyond a simple classification scheme to recognize road anomalies, and instead attempt to find physical characteristics for these anomalies from accelerometer and gyroscope time series. Hence, this work focused on inferring the depth of potholes and also the functional status of asphalt and metal bumps. To accomplish this, we built a new data set that is composed of the aforementioned sensor readings that correspond to 163 potholes and 101 speed bumps, all of them with ground truth labels, and publicly available.

This second-generation inference problem was modeled in three different ways: as regression, classification and learning to rank tasks. For each task we evaluated several Machine Learning algorithms. In the case of pothole depth regression, the best model was built with a Random Forest fed with a feature vector containing both data from accelerometer and

5. Consider that data representation plays a crucial role to improve classification performance, consider the recent related work of [30], where the authors show that treating sensor information by means of a second-order representation improved classification accuracy.

gyroscope and yielding an average RMSE of 1.68 cm, clearly outperforming a Gradient Boosting (GB) strategy, Bayesian Regression and Decision Tree. For pothole classification, we evaluated different levels of granularity with respect to the number of depth categories, considering 2, 3 and 4 classes. For this task, GB and RF were the top performers in terms of AUC score, surpassing a Multi-layer perceptron and Support Vector Machine, among other popular algorithms. For the third task we were interested in evaluating the possibility of devising a method to create a ranked list of potholes, based on their severity (which we assume is directly correlated to their respective depth). We evaluated two approaches for ranking: point-wise and a pair-wise, being the former approach the one that yielded the best results, with a Kendall's τ of 0.34 using the RF as a learning model. If gyroscope data is used for this task, then Kendall's τ improves, obtaining a value of 0.41 in contrast to 0.34 (obtained when using only accelerometer data).

After all this analysis, we suggest that the best way to model the problem of pothole characterization is given by the Classification task. This, since a competitive score of 0.84 in AUC was achieved using up to three levels of granularity, i.e., shallow, medium and deep, suggesting then a good trade-off between performance and detail.

For speed bump profiling, "OK" and "not OK" labels were used to represent functional status. Experimental results show that GB and RF can deem the status of asphalt and metal bumps with an AUC score of ~ 0.75 when treating them separately. Altogether, the introduction of this second-generation problem opens the possibility to profile roads and avenues with real detail, alerting drivers of the severity of hazards in their route, and also helping city managers to better focus state resources to improve transportation infrastructure.

Acknowledgment

The authors would like to acknowledge the support received from Google by means of the 2017 Latin American Research Award for the project "Learning Roadway Surface Disruption Patters to Improve Transportation Infrastructure." The first author acknowledges the graduate student scholarship granted by the National Council of Science and Technology of Mexico (Conacyt). All authors are grateful to the anonymous reviewers that significantly help improve the quality of the manuscript.

References

[1] E. Buza, S. Omanovic, and A. Huseinnovic, "A pothole detection with image processing and spectral clustering," in *Proc. 2nd Int. Conf. Inf. Technol. Comput. Netw.*, 2013, pp. 48–53.

[2] Y. Jo and S. Ryu, "Pothole detection system using a black-box camera," *Sens.*, vol. 15, no. 11, pp. 29316–29331, 2015.

[3] S. G. Christopoulos, S. Kanarachos, and A. Chroneos, "Learning driver braking behavior using smartphones, neural networks and the sliding correlation coefficient: Road anomaly case study," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 1, pp. 65–74, Jan. 2019.

[4] M. R. Carlos, F. Martínez, R. Cornejo, and L. C. González, "Are android smartphones ready to locally execute intelligent algorithms?" in *Proc. 15th Mexican Int. Conf. Artif. Intell. Part II*, 2017, pp. 15–25.

[5] S. Sattar, S. Li, and M. Chapman, "Road surface monitoring using smartphone sensors: A review," *Sens.*, vol. 18, no. 11, 2018.

[6] M. Perttunen et al., "Distributed road surface condition monitoring using mobile phones," in *Proc. Int. Conf. Ubiquitous Intell. Comput.*, 2011, pp. 64–78.

[7] A. Mednis, G. Strazdins, R. Zviedris, G. Kanonirs, and L. Selavo, "Real time pothole detection using Android smartphones with accelerometers," in *Proc. Int. Conf. Distrib. Comput. Sensor Syst. Workshops*, Jun. 2011, pp. 1–6.

[8] F. Seraj, B. J. van der Zwaag, A. Dilo, T. Luarasi, and P. Havinga, "RoADS: A road pavement monitoring system for anomaly detection using smart phones," in *Big Data Analytics in the Social and Ubiquitous Context*. New York, NY, USA: Springer International Publishing, 2016, pp. 128–146.

[9] L. C. González, R. Moreno, H. J. Escalante, F. Martínez, and M. R. Carlos, "Learning roadway surface disruption patterns using the bag of words representation," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 11, pp. 2916–2928, Nov. 2017.

[10] M. Fazeen, B. Gozick, R. Dantu, M. Bhukhiya, and M. C. Gonzlez, "Safe driving using mobile phones," *IEEE Trans. Intell. Transp. Syst.*, vol. 13, no. 3, pp. 1462–1468, Sep. 2012.

[11] F. Martinez, L. C. Gonzalez, and M. R. Carlos, "Identifying roadway surface disruptions based on accelerometer patterns," *IEEE Latin America Trans.*, vol. 12, no. 3, pp. 455–461, May 2014.

[12] J. M. Celaya-Padilla et al., "Speed bump detection using accelerometric features: A genetic algorithm approach," *Sens.*, vol. 18, no. 2, 2018, Art. no. 443.

[13] M. R. Carlos, M. E. Aragón, L. C. González, H. J. Escalante, and F. Martínez, "Evaluation of detection approaches for road anomalies based on accelerometer readings—addressing whos who," *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 10, pp. 3334–3343, Oct. 2018.

[14] J. Eriksson, L. Girod, B. Hull, R. Newton, S. Madden, and H. Balakrishnan, "The pothole patrol: Using a mobile sensor network for road surface monitoring," in *Proc. 6th Int. Conf. Mobile Syst., Appl., Services*, 2008, pp. 29–39.

[15] P. Mohan, V. N. Padmanabhan, and R. Ramjee, "Nericell: rich monitoring of road and traffic conditions using mobile smartphones," in *Proc. 6th ACM Conf. Embedded Netw. Sensor Syst.*, 2008, pp. 323–336.

[16] V. Astarita et al., "A mobile application for road surface quality control: UNIquALroad," *Procedia-Social Behavioral Sci.*, vol. 54, pp. 1135–1144, 2012.

[17] H. Wang, N. Huo, J. Li, K. Wang, and Z. Wang, "A road quality detection method based on the mahalanobis-taguchi system," *IEEE Access*, vol. 6, pp. 29078–29087, 2018. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8362926&isnumber=8274985, doi: 10.1109/ACCESS.2018.2839765.

[18] J. Wahlstrm, I. Skog, and P. Hndel, "Smartphone-based vehicle telematics: A ten-year anniversary," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 10, pp. 2802–2825, Oct. 2017.

[19] G. Chugh, D. Bansal, and S. Sofat, "Road condition detection using smartphone sensors: A survey," *Int. J. Electron. Elect. Eng.*, vol. 7, no. 6, pp. 595–602, 2014.

[20] G. Xue, H. Zhu, Z. Hu, J. Yu, Y. Zhu, and Y. Luo, "Pothole in the dark: Perceiving pothole profiles with participatory urban vehicles," *IEEE Trans. Mobile Comput.*, vol. 16, no. 5, pp. 1408–1419, May 2017.

[21] J. Clement, "Speed humps and the thousand oak experience," *Inst. Traffic Eng. J.*, vol. 53, pp. 35–39, 1983.

[22] A. Efthymiou, E. N. Barmpounakis, D. Efthymiou, and E. I. Vlahogianni, "Transportation mode detection from low-power smartphone sensors using tree-based ensembles," *J. Big Data Analytics Transp.*, vol. 1, pp. 57–69, Jun. 2019.

[23] C. Kupfer, "Investigation and measurement of oscillations of the H.E.S.S. telescopes," Master's thesis, Erlangen Centre Astroparticle Phys., Friedrich-Alexander Universitt Erlangen-Nrnberg, Erlangen, Germany, Feb. 2018.

[24] F. Pedregosa et al., "Scikit-learn: Machine learning in python," *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, Nov. 2011.

[25] A. Mohamed et al., "Roadmonitor: An intelligent road surface condition monitoring system," in *Intelligent Systems'2014*, D. Filev, J. Jabłkowski, J. Kacprzyk, M. Krawczak, I. Popchev, L. Rutkowski, V. Sgurev, E. Sotirova, P. Szynkarczyk, and S. Zadrozny, Eds., Cham, Switzerland: Springer International Publishing, 2015, pp. 377–387.

[26] Y. A. Alqudah and B. H. Sababha, "On the analysis of road surface conditions using embedded smartphone sensors," in *Proc. 8th Int. Conf. Inf. Commun. Syst.*, Apr. 2017, pp. 177–181.

[27] P. Gehler and S. Nowozin, "On feature combination for multiclass object classification," in *Proc. IEEE 12th Int. Conf. Comput. Vis.*, Sep. 2009, pp. 221–228.

[28] M. Kubat and S. Matwin, "Addressing the curse of imbalanced training sets: One-sided selection," in *Proc. 14th Int. Conf. Mach. Learn.*, 1997, pp. 179–186.
[29] N. J. Salkind, *Encyclopedia of Measurement and Statistics*, vol. 1. Newbury Park, CA, USA: Sage, 2007.
[30] M. R. Carlos, L. C. González, J. Wahlström, G. Ramírez, F. Martínez, and G. Runger, "How smartphone accelerometers reveal aggressive driving behavior?–the key is the representation," *IEEE Trans. Intell. Transp. Syst.*, pp. 1–11, 2019, early access, Jul. 16, 2019. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8764567&isnumber=4358928, doi: 10.1109/TITS.2019.2926639

M. Ricardo Carlos received the BEng degree in computer systems engineering from the Instituto Tecnológico de Chihuahua II, Chihuahua, Mexico, in 2006 and the MEng degree from the Universidad Autónoma de Chihuahua, in Chihuahua, Mexico, in 2015. He is currently working toward the PhD degree in engineering at Universidad Autónoma de Chihuahua, Chihuahua, Mexico. His current research interests include pattern recognition and mobile computing.

Luis C. González received the PhD degree from The University of North Carolina at Charlotte, Charlotte, North Carolina, in 2011. That same year, he joined the Universidad Autónoma de Chihuahua, in the north of Mexico, where he was a founder (and now a professor) of the graduate program of computer engineering. Currently, he is a member of the National System of Researchers of the National Council of Science and Technology of México (CONACyT). His current research interests include pattern recognition and combinatorial optimization problems. In 2017 he received the Latin America Google Research Award for his work on Intelligent Transportation Systems. He is a member of the IEEE.

Johan Wahlström received the BSc, MSc, and PhD degrees from the KTH Royal Institute of Technology, Stockholm, Sweden in 2013, 2014, and 2017, respectively. He is a postdoc researcher with Oxford University, since January 2018. His main PhD research topic was smartphone-based vehicle telematics. In 2015, he spent one month at University of Porto and six months at Washington University in St. Louis as a visiting PhD student. In 2016, he spent two months at the MIT-startup Cambridge Mobile Telematics. He was accepted into the program of excellence in electrical engineering at KTH in 2014, and in 2015 was the youngest recipient of the Sweden-America foundations research scholarship.

Raymundo Cornejo received the master's and DSc degrees in computer science from the Center for Scientific Research and Higher Education of Ensenada (CICESE). He is a professor with Autonomous University of Chihuahua, Chihuahua, Mexico, and member of the CONACYT Thematic Network on Ageing, Health and Social Welfare. His areas of interest include human computer interaction, pervasive computing, and aging.

Fernando Martínez received the master's degree in industrial electronic engineering from the Instituto Tecnológico de Chihuahua - Chihuahua, México, in 1997, and the PhD degree in computer science from The University of Nottingham, Nottingham, United Kingdom, in 2009. His current research interests include mobile computing, sensor networks and human-computer interactions.