Toward Practical Crowdsourcing-Based Road Anomaly Detection With Scale-Invariant Feature

YUANYI CHEN1,2, MINGXUAN ZHOU1, ZENGWEI ZHENG1, AND MEIMEI HUO1

1Department of Computer Science and Computing, Zhejiang University City College, Hangzhou 310015, China
2Department of Electrical Engineering & Computer Science, University of California at Irvine, Irvine, CA 92697, USA

Corresponding author: Yuanyi Chen (yuanyic2@uci.edu)

This work was supported in part by the Young Scientists Fund of the National Natural Science Foundation of China under Grant 61802343, in part by the Zhejiang Provincial Natural Science Foundation of China under Grant LGF19F020019, and in part by the Hangzhou Key Laboratory for IoT Technology and Application.

ABSTRACT Road anomaly detection with crowdsourced sensor data has become an increasingly important field of research over the last few years. Traditional ways for road anomaly detection are either threshold-based detection techniques or feature-based detection techniques. However, road anomaly patterns from crowdsourcing data are often shifted in time and exhibit local distortions/noise, thus existing methods rely on the original sensor data greatly limit the accuracy of road anomaly detection. In this paper, we present a road anomaly detection model by learning scale-invariant features from the differences between small local segments of road anomaly samples. Specifically, the proposed model consists of two phases: 1) Road anomaly segmentation. The phase is designed to roughly extract road anomaly subsequence using piecewise aggregate approximation representation of sensor series data, and 2) Road anomaly detection. In this phase, we observe the differences among road anomaly classes are attributed to small local segments, then we learn scale-invariant features from these small local segments for road anomaly detection. To demonstrate the utility of our proposed model, we have performed a comprehensive experimental evaluation on two real-world datasets and one large-scale simulation dataset. The experimental results show our proposed model outperforms all baselines significantly in terms of road anomaly detection.

INDEX TERMS Road anomaly detection, mobile crowdsourcing, accelerometer readings, sliding window, scale-invariant feature.

I. INTRODUCTION

Road anomalies such as potholes or metal bumps can cause vehicle damage and potentially injure drivers and passengers. Billions of dollars are spent on maintenance and repair of road anomalies every year. For example, AAA motor club estimated that pothole damage costs U.S. drivers 3 billion dollars per year [1]. Similarly, the UK government will spend 1.2 billion euros on road repairs between 2017 and 2018 [2]. Worse still, almost of third of all motor vehicle crashes in US are related to poor road conditions, resulting in more than two million injuries and 22,000 fatalities [3]. Some technological solutions have been proposed to automatically detect and report road anomalies to government agencies and drivers in order to road maintenance and knowledge of anomalies locations. For instance, vision-based solutions [4]–[6] utilized shape, edges or texture differences with regular road surface to identify road anomalies, [7]–[9] utilized stereo, 3D vision and depth sensors to obtain more detailed road surface information. However, these solutions need expensive hardware equipment such as laser profilers [10] or specific video systems to collect road surface information, which is time consuming and labor intensive for anomaly detection of large-scale road networks.

Recently, user-generated sensor data make various smart city services such as indoor localization [11], indoor POI recommendation [12]–[14] and road anomaly detection [15]–[18] could be efficiently implemented, utilized, and assessed. Specifically, crowdsourcing-based sensor data from passenger’s mobile devices or vehicles provide rich information to monitor and detect road anomalies. The underlying principle is the collected sensor data imply inherent patterns when the vehicle encounters anomalies, for example, hitting a metal bump, as displayed in Figure 1. Even though a number
of studies on road anomalies detection with crowdsourcing-based sensor data [15], [16], [19]–[21] have been proposed recently (for a review see Section 2), these approaches suffer from a number of limitations. For example, the studies in [15], [16], [19]–[21] utilized threshold-based techniques to detect road anomalies, which can only detect whether the road is damaged without the specific types of damage (such as a pothole or metal bump), which can be seen as an identification problem. In addition, the detection performance of these methods is poor due to data noise or acquisition anomalies. It is important to note that this identification problem aims to indicate if an anomaly is present in a (continuous) sequence of accelerometer readings, thus being a binary classification problem. In this study, we not only need to detect and segment an road anomaly from a continuous signal, but also need to assign it a correct label (e.g. pothole, metal bump and speed bump, etc.), thus being by nature a multi-class classification problem. Detecting the specific anomaly type has more importance value: 1) for road maintenance department, knowing the specific type of the anomaly is pothole or normal road equipment (speed bump or metal bump) is helpful for making their countermeasures; 2) for drivers, knowing the specific type of the anomaly helps to take different measures, such as decrease their speed before the speed bump, or choose to avoid pothole and metal bump, which can improve the driving experience. 3) Analyzing the specific types of road anomalies can comprehensively evaluate the road conditions. Usually, road facilities such as speed bumps and metal bumps have less impact on a section of road conditions, but abnormal damages such as potholes have more impact on evaluation road conditions.

Recently, a few efforts [17], [18], [22]–[26] have been made to detect the specific type of road anomalies, by firstly extracting the time domain or frequency domain features from collected sensor data, then feeding these features into a classification model to detect road anomalies. The main bottleneck of these methods is the differences among different type of road anomaly are attributed to small local segments, rather than the global feature due to the following two factors: 1) the crowdsourced-sensor data usually corrupted with noise and outliers; 2) the road anomaly patterns are often shifted in time and have different scales.

For a practical road anomaly detection system, several requirements are necessary: reasonable detection accuracy; no additional hardware components on user’s side; scalable to large-scale deployment. On this basis, we propose a practical road anomaly detection model using crowdsourcing-based acceleration sensor data. Note that our proposed road anomaly detection model includes two phases: 1) offline building the anomaly detection model. This phase aims to train detection model with acceleration data and the corresponding road anomaly type. Currently, we generate training dataset by collecting both acceleration data and the video, and label the anomaly type and acceleration sequences by timestamp alignment. We only detect three anomalies that are significantly different, thus it’s easy to distinguish the type of anomalies. The judgement of start time or duration of a road anomaly will be a bit difficult but can still be handled by locating the rapid change in acceleration data. For the future work, we consider to utilize image processing technique to automatically label the training set. To check the effect of anomalies labeling, people only need to check whether an apparent anomaly in the video is labeled without the need to judge the duration or start time; 2) online detecting road anomaly with crowdsourcing data. In this phase, the only work that non-professional people need to do is to download the application and open it no matter they are driving or sitting in a vehicle. Our application will collect the acceleration data, then detect the exact type of the anomaly to help them make driving decisions such as slow down. Therefore, non-professional users can also utilize the proposed anomaly detection model without professional knowledge. The contributions of our research are three-fold:

- We propose a sliding window-based method to segment road anomaly subsequences, and remove false road anomaly samples using piecewise aggregate approximation (PAA) representation of acceleration sensor data.
- We observe the differences among road anomaly classes are attributed to small local segments, then propose a road anomaly detection model by learning scale-invariant features with high discriminative power from the differences of small local segments.
- We evaluate our approach based on three datasets including two real-world datasets and one simulation dataset, and the results show the advantages of our approach beyond multiple baseline algorithms.

The remainder of the paper is organized as follows: Section II surveys related work on road anomaly detection with acceleration sensor data. Section III describes the proposed road anomaly detection model in detail. Section IV reports and discusses the experimental results. Finally, we present our conclusion and future work in Section V.

II. RELATED WORK

In this section, we first survey related work on road anomaly detection with crowdsourcing-based acceleration sensor data,
then discuss how these works differ from our work. In general, existing studies on this topic can be divided into two categories:

**A. THRESHOLD-BASED DETECTION TECHNIQUES**

The accelerometer sensor measures corresponds to longitudinal (Yaxis-Acc), vertical (Zaxis-Acc) and transversal (Xaxis-Acc), which are affected by the inertia of vehicle and can be utilized to detect road anomalies with threshold-based solutions [15], [16], [19]–[21]. The study [15] proposed “z-peak” to detect potholes by determining whether the collected accelerometer data surpass a threshold value. [16] further improved “z-peak” by increasing bumps and potholes identification reliability, and introduced four threshold-based approaches: 1) Z-thresh, which measures the acceleration amplitude at Z-axis; 2) Z-diff to measure the difference between the two amplitude values; 3) STDDEV (Z) to find the standard deviation of vertical axis acceleration and 4) G-Zero are used to identify potholes. The work in [19] detected road surface conditions by jointly considering predetermined thresholds of the transverse and vertical acceleration. In [20], road bumps are detected by examining the peak vertical acceleration and the duration for which the acceleration dips below a defined threshold. Reference [21] firstly proposed a half car model to describe the car motion any encountered disturbances (e.g., potholes and speed bumps), then used threshold-based techniques to classify road anomalies. However, threshold-based solutions suffer a few limitations include the need of accelerator sensor with high sample frequency (such as 380 Hz in [15] and 310 Hz [20]) that not available in modern smartphones, poor detection performance due to data noise and can only detect single type of road anomaly.

**B. FEATURE-BASED DETECTION TECHNIQUES**

A few literatures [18], [22]–[26] formulated road anomaly detection as a multi-class classification problem and a large number of classification features have been considered. Roughly speaking, these features can be divided into two categories: 1) Statistical feature, which are mainly describing typical values of the collected acceleration dataset and measuring the data distribution behavior. For instance, the work in [22] performed detection using Support Vector Machines (SVM) by extracting a list of 120 candidate features (e.g., maximum, mean and standard deviation of the raw data). In [27], 12 features such as mean of the accelerometer values, standard deviation and variance are utilized to train road anomaly classification model. Reference [18] firstly utilized a sliding window to group the data to calculate features include weighted means, standard deviations, and maximum values of the raw inertial data, then trained anomaly classification model with SVM. In [23], gaussian model-based mining algorithm is proposed for the abnormal detection, x-z ratio filtering is applied for anomaly classification as a pothole or bump; 2) Frequency feature. Some studies believed that the feature extracted from the frequency of the accelerometer signals could be more discriminative in comparison with the statistical feature extracted from time-domain. For example, the study [24] utilized a wavelet-based feature extraction together with statistical approach for road anomaly detection. Reference [28] further extracted features using stationary wavelet transformation for both accelerometer and gyroscope signals, and used Self Organized Maps to detect road anomaly. In [25], power spectra density is utilized to calculate the roughness of potholes based on GPS sensor and three-axis accelerometer. Still other studies [17], [26] performed road anomaly detection by jointly considering both statistical feature and frequency feature. For example, the work [17] performed pothole detection by extracting feature such as mean, peak-to-peak ratio, root mean square, standard deviation, variance, power spectrum density and wavelet packet decomposition. Similarly, [26] adopted a DENCLUE-like algorithm to detect road anomaly with 15 features (e.g., mean, median, mode, range, maximum, minimum and root mean square).

Recently, some studies [29]–[31] detected road anomaly from a new perspective. The work in [29] proposed a new clustering algorithm for processing road anomaly detection using Mahalanobis distance to quantify the similarity between a new report and the existing clusters. P3 [30] utilized k-means algorithm to cluster all perceptions of the same pothole according to the depth field. [31] represented sensor data with a novel technique inspired by the bag of words (BOW) representation and the results suggest that the BOW representation boosting anomaly classification performance. On the other hand, road anomaly detection using crowdsourcing sensor data can be regarded as a special instance of time-series classification problem. Most existing time-series classification literatures focus on designing distance measures of two time series with better classification performance, such as Euclidean distance, dynamic time warping (DTW) [32], Weighted DTW [33], Time Warp Edit [34] and Move-Split-Merge [35]. Still other studies proposed distance measures by jointly considering the time domain and the first order differences of time series domain, such as Complexity Invariant distance [36], Derivative DTW [37] and Derivative Transform Distance [38]. Recently, some dictionary-based methods (such as bag of patterns [39] and bag of SFA symbols [40]) are proposed to classify time series, which transform the original time series into representative words for approximating and reducing the dimensionality of time series.

Unfortunately, road anomaly detection using existing time series classification techniques and feature-based methods from the original sensor data will suffer the following three challenges: 1) Existing time series classification techniques assume all time series data in a problem are equal length, real valued and have no missing values. However, the length of abnormal road sensing series depend on a few factors (e.g., the size of road anomaly, vehicle speed and collecting frequency) thus cannot be the same; 2) Existing methods assume that there already have segmented anomaly subsequence for feature extraction, but in fact the collected sensor
TABLE 1. Notations used in the paper.

<table>
<thead>
<tr>
<th>SYMBOL</th>
<th>DESCRIPTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>N, M, K</td>
<td>The number of $D_r, D_c$, and $C$</td>
</tr>
<tr>
<td>$D_r$</td>
<td>Train dataset: a few subsequence labeled by road anomalies</td>
</tr>
<tr>
<td>$D_c$</td>
<td>Test dataset: a few acceleration series records</td>
</tr>
<tr>
<td>$C$</td>
<td>The set of road anomalies</td>
</tr>
<tr>
<td>$A$</td>
<td>An acceleration series record</td>
</tr>
<tr>
<td>$\tau$</td>
<td>The size of sliding window</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>The threshold for selecting candidate anomaly split points</td>
</tr>
<tr>
<td>$</td>
<td>A(i)</td>
</tr>
<tr>
<td>$a_i$</td>
<td>Represent the Z axis acceleration reading at time $t_i$</td>
</tr>
<tr>
<td>$z_i$</td>
<td>The Z axis acceleration reading</td>
</tr>
<tr>
<td>$f_d(t_i, \tau)$</td>
<td>Vertical acceleration fluctuation distance in sliding window</td>
</tr>
<tr>
<td>$S^g_{1:2}$</td>
<td>Acceleration subsequence that starts at $p$ and ends at $q$.</td>
</tr>
<tr>
<td>$A_i$</td>
<td>An acceleration series records need to detect road anomalies.</td>
</tr>
</tbody>
</table>

data are only continuous time series data. Therefore, we need to design a segmentation method for automatically generating road anomaly samples for anomaly detection; 3) Existing methods focus on extracting statistical or frequency feature from raw sensor series data to detect road anomaly. Unfortunately, the anomaly patterns embedded in raw sensor data might be shifted in time, in addition to potentially being locally distorted or scaled. Thus, the anomaly patterns for detection might be present only on short local segments, rather than on global structure. Therefore, extracting feature from raw sensor data for road anomaly detection cannot achieve reasonable detection accuracy.

Our proposed approach differs from the above-mentioned works in the following two aspects: 1) We automatically generate road anomalies samples by utilizing sliding window to segment continuous sensor series data and PAA-based classification model; 2) Instead of extracting feature from raw sensor data, we mine a set of scale-invariant features with high classification accuracy from labeled anomaly subsequences, which can boost road anomaly detection accuracy.

III. ROAD ANOMALY DETECTION MODEL

In this section, we first introduce the key data structures and notations for the problem definition of road anomaly detection then present our solution.

A. PROBLEM DEFINITION

For ease of the following presentation, we define the key data structures and notations used in the proposed method. Table 1 lists the relevant notations used in this paper.

Definition 1 (Acceleration Series Record): An acceleration series record consists of a set of acceleration readings measured successively at uniform time intervals, which are collected by mobile devices of drivers or passengers in crowdsourcing manner and denote by $A = \{a_1, \ldots, a_i, \ldots\}$. Our method only utilizes Z axis acceleration, thus $a_i$ is a tuple $< z_i, t_i >$ which means the Z axis acceleration reading $z_i$ is collected at time $t_i$.

Definition 2 (Acceleration subsequence): An acceleration subsequence $S^g_{1:2}$ is a set of continuous Z axis acceleration readings from an acceleration series record, that starts at position $p$ and ends at $q$.

Definition 3 (Train dataset): A train dataset $D_r$ is a collection of $N$ acceleration subsequences $\{S(1), S(2), \ldots, S(N)\}$ labeled by $K$ road anomalies $C = \{c_1, \ldots, c_i, \ldots, c_k\}$. Thus, $D_r$ is a set of pairs of acceleration subsequence and their labeled road anomalies (e.g., pothole, metal bump and speed bump), $D_r = \{(S(i), c_k) : 1 \leq i \leq N, 1 \leq k \leq K\}$.

Based on the above definitions, we formulate the problem of road anomaly detection as: Given: 1) Train dataset $D_r = \{(S(i), c_k) : 1 \leq i \leq N, 1 \leq k \leq K\}$ generated by manual annotation; 2) A few acceleration series records $D_c = \{A_1, \ldots, A_i, \ldots, A_M\}$ collected by crowdsourcing; Objective: find the correspond anomaly acceleration subsequences in $D_c$ and detect the anomaly type of these anomaly subsequences.

Basically, one can firstly extract all subsequences of $D_c$, then determine the anomaly type of each subsequence by building classification model or similarity match. However, for an acceleration series record of length $L$, the number of all possible subsequences are $L(L+1)/2$. Thus, for $D_c$, the number of all possible subsequences of are $M \times L(L+1)/2$, where $M$ is the number of acceleration series records. Clearly, running the algorithm mentioned above on a smartphone platform involves high power consumption and requires high processing capabilities. On the contrary, our solution for this problem consists of two phases: (1) Road anomaly segmentation. This phase extracts candidate anomaly acceleration subsequences by sliding window-based method instead of evaluating all the possible subsequences; (2) Road anomaly detection. This phase determines the anomaly type of acceleration subsequences by learning scale-invariant feature.

B. ROAD ANOMALY SEGMENTATION USING SLIDING WINDOW

A few factors can influence the vertical acceleration readings during driving, such as mobile phone pose, slight bumps and noise in acceleration measurement. By investigating the spatial–temporal characteristics of vertical acceleration readings when passing normal and anomaly roads, we observe a valuable characteristics can be exploited to extract road anomaly subsequence from the continuous sensor data. As shown in Figure 2, the vertical acceleration readings usually fluctuate gently near gravity acceleration (9.8 m/s²) when passing normal roads, but the vertical acceleration readings will fluctuate dramatically when passing a road anomaly, such as pothole and metal bump. Therefore, this characteristic can reflect road anomalies to a certain degree and can be used to distinguish normal and anomaly subsequence, which is also demonstrated in [22], [41].

Based on the spatial–temporal characteristics of vertical acceleration readings, we segment anomaly subsequence by...
two stages, as shown in Figure 3. Specifically, we first identify split points of anomaly subsequence based on the vertical acceleration readings “fluctuate” characteristic when passing a road anomaly, and merge some adjacent points according to the length of subsequences. Then, we extract candidate anomaly subsequences according to the merged split points and remove false anomaly subsequences using PAA-based feature.

**Step 1: Identify Anomaly Split Points:** Based on the observation that the vertical acceleration readings will fluctuate drastically when passing through a road anomaly, we utilize the fluctuation distance of vertical acceleration readings in a small time window to identify anomaly split points. Formally, given an acceleration series record \( A = \{a_1, \ldots, a_i, \ldots\} \), we define \( fd(t_i, \tau) \) to represent the vertical acceleration fluctuation distance in sliding window \((t_i - \tau/2, t_i + \tau/2)\), as shown in Equation 1.

\[
fd(t_i, \tau) = \| S_{t_i - \tau/2}^{t_i + \tau/2} - G(\tau) \|
\]  

(1)

where \( \tau \) is the window size, \( G(\tau) \) is a subsequence that has \( \tau \) elements and all the elements are equal to gravity acceleration (9.8 m/s²).

After calculating the fluctuation distance of vertical acceleration readings in each sliding window, we select candidate anomaly split points using threshold-based approach. For example, given the raw vertical acceleration readings as shown in Figure 3a, if we set the sliding window size as 10 and slide 5 elements forward each time, there are 38 sliding windows in total and the slide distance of each sliding window is shown in 3b. If the variation coefficient \( \lambda \) in sliding window \((t_j - \tau/2, t_j + \tau/2)\) is significantly higher than average (as shown in Equation 2), we can infer the vehicle is passing through a road anomaly at time \( t_j \). Thus we select candidate anomaly split points based on variation coefficient \( \lambda \) with a threshold-based approach, and merge the candidate anomaly split points according to the interval between adjacent split points. As shown in Figure 3c, we generate four anomaly split points after adjacent points merging. Finally, we segment...
candidate anomaly subsequences according to the merged
anomaly split points (as shown in Figure 3d).

\[\lambda = \frac{fd(t_j, \tau) \times N_{A(i)}}{\sum_{j=1}^{1/\tau/2+1} fd(t_j, \tau)}\]  \hspace{1cm} (2)

where \(N_{A(i)}\) is the sliding window number of \(A(i)\).

**Step 2: Remove false segmentation:** As mentioned above, we identify road anomaly split points according to the fluctuation distance of vertical acceleration readings. This method may bring some false positives, since other factors (e.g., passenger’s hand shaking and slight bumps on the ground) may cause the acceleration reading fluctuate dramatically for one or two points, which make the fluctuation distance larger than threshold thus is recognized as a split point. Fortunately, the overall trend during a road anomaly subsequence is stable and robust. On the basis, we remove false positives by the following two steps: 1) Extract feature for candidate anomaly subsequences using piecewise aggregate approximation (PAA); 2) Classify normal and anomaly subsequences using the PAA-based feature with random forest.

**C. ROAD ANOMALY DETECTION BY LEARNING SCALE-INvariant FEATURE**

After segmenting candidate anomaly subsequences, the road anomaly detection problem is formulated as predicting the type of future anomalies subsequences using a few expert-labeled anomalies samples. The intuition of the proposed method is the acceleration sequence data imply inherent patterns when the vehicle encounters road anomalies (e.g., hitting a pothole or speed bump as displayed in Figure 1), which is also reported in previous studies [15] and [18]. In this way, we perform road anomaly detection by learning scale-invariant features for respeting the inherent patterns of each anomaly class. However, anomaly patterns might be shifted in time and the difference might be present on short local segments rather than on global structure. For example, three samples for three kinds of anomalies are shown in Figure 4, the discriminative segment is highlighted for each class, we observe most samples have the same global structure but differ only in the highlighted segments.

To learn the scale-invariant feature for each anomaly class with high discriminative power, we utilize shapelet discovery [42] by evaluating the prediction quality of numerous candidates extracted from anomaly subsequences. Here, scale-invariant means the learnt scale-invariant features are invariant to shifts and scales of acceleration sequence patterns. Generally, the intra-class variations occur in local segments of acceleration sequence, which are either shifted in time or distorted. But the proposed scale-invariant feature extracts local segments from acceleration sequence, thus can capture patterns that are invariant to shifts and scales.

Formally, shapelets are a set of \(u \in U\) patterns and each having length \(H\), and denoted as \(U \in \mathbb{R}^{U \times H}\). We consider all possible subsequences of a kind of acceleration training sequence as potential candidate scale-invariant features. The minimum distances between scale-invariant features \(U_{u,h}\) and all the subsequences of an acceleration sequence \(S\) were used as a metric for ranking the information gain accuracy of that candidate on the target acceleration sequence, the minimum distance of a set of scale-invariant features to acceleration sequences can be perceived as a kind of data transformation, namely the shapelet-transformation representation (more details about the shapelet-transformation representation can be found in [42] and [43]), as denoted in Equation 3.

\[\Gamma_h(i) = \min_{h=1}^{H} \sum_{j=1}^{H} (\Gamma_j^{j+H-1}(i) - U_{u,h})^2\]  \hspace{1cm} (3)

where \(1 \leq j \leq |A(i)| - H + 1, 1 \leq i \leq N\) and \(1 \leq h \leq H\).
The challenge of this representation is to find the shapelets \( U \), for which the resulting representation \( \Gamma \) helps achieve the highest classification accuracy. More exactly, we learn the scale-invariant feature for training dataset \( D_{tr} \) by the following four steps: 1) Iterating over all the subsequences of the train dataset \( D_{tr} \) and considers each subsequence as a candidate shapelet; 2) Evaluating the distance between each candidate and all training samples as the distance to \( \phi \). 3) Utilizing the candidate shapelets based on F-stats score and selecting the top-\( k \) shapelets with the highest classification accuracy as scale-invariant feature.

After learning scale-invariant feature of training dataset \( D_{tr} \), we detect road anomalies in test dataset \( D_{te} \) as shown in Algorithm 1. First, as shown in Line 2–3, we extract anomaly subsequences of test dataset based on sliding window method. Then, as depicted in Line 4–5, we generate the final anomaly subsequences using random forest with PAA-based feature. In Line 7–14, we transform the raw train dataset and test dataset into a new representation by calculating the minimum distances to the learned shapelets of \( D_{tr} \). Finally, we detect the anomaly type of anomaly subsequence based on a random forest model as shown in Line 15.

### IV. EXPERIMENT EVALUATION

In this section, we report on the results of a series of experiments conducted to evaluate the performance of the proposed anomaly detection model. We first describe the settings of experiments including data sets, comparative methods and evaluation metric. Then, we report and discuss the experimental results.

#### A. EXPERIMENTAL SETTINGS

1) DATA SETS

Three datasets about road anomaly detection using acceleration series records: one dataset collected by ourselves (Dataset 1), one public dataset from [27] (Dataset 2) and one simulation dataset using Carsim ¹ (Dataset 3), are used for experimental evaluation. More details of these datasets are reported in Table 2.

- **Dataset 1.** We develop an android application to collect GPS, acceleration readings and video, a dataset about 500m on 10 roads are collected and the speed of collected vehicle is about 30km/h. For labeling road anomaly subsequences, we manually synchronize the recording timestamp of acceleration readings and video. Finally, three types of road anomalies are marked according to the dataset attached to [27]: potholes, metal bumps and speed bumps.

- **Dataset 2.** This dataset is generated with Pothole lab [27], a platform can be used to generate virtual roads with a configurable number and nature of road anomalies. For practical situations, we insert the road anomalies data of Pothole lab into normal data at a ratio of 1:50 that suggested by the studies [24], [31].

- **Dataset 3.** We follow this work [22] to generate a large-scale road anomaly dataset using Carsim simulation. In total, we generated a dataset includes 2000 potholes, 1200 metal bumps and 1200 speed bumps.

2) COMPARATIVE METHODS

- **SVM-WW:** the method [27] detected different anomalies such as potholes, metal bumps and speed bumps based on z-axis acceleration readings. Firstly, the accelerometer series records are prepossessed by applying a sliding window, then 12 features (such as

1 ¹https://www.carsim.com/products/carsim/index.php
Mean of the accelerometer values, Standard Deviation, Variance and Coefficient of Variation) are used to train a SVM model for road anomaly detection.

- **CPD-SFS**: the method [18] firstly grouped acceleration readings with a sliding window scheme, then a feature list includes 85 candidate features are generated that contains functions of the acceleration and speed measurements that may have pothole identifying signatures. Finally, the road anomaly is detected with a SVM classifier by training the feature list.

- **SVM-MDDP**: this method [44] firstly denoised acceleration readings by applying wavelet packet technique, then extract more than 70 features include statistical, time domain and frequency domain features for training a SVM-based anomaly detection model.

- **MLM-WB**: Inspired in the bag of words representation, the method [31] represented raw accelerometer data with some simplified representations called word bag and learned a codebook formed by prototypical patterns. Then, accelerometer series records are first segmented and compared to every one of the codewords in the codebook, and the segmentation subsequence will be represented with the code word that had shown the minimum distance. Finally, a histogram will be created when all segmentation subsequence are compared and this histogram is regarded as feature to train a road anomaly detection classifier.

3) **EVALUATION METRIC**

We utilize F-measure to evaluate the performance, which is a weighted harmonic average between precision and recall, which has been widely utilized in many classification problems including binary classification, multi-label classification and structured output prediction. Let \( v_i \) and \( \overline{v_i} \) denote the ground truth and prediction anomaly type for a sample \( s_i \in \phi \), the F-measure is defined as:

\[
F - \text{measure} = \frac{1}{|\phi|} \sum_{s_i \in \phi} \frac{2 \times |v_i \cap \overline{v_i}|}{v_i + \overline{v_i}}
\]  

(4)

Following the work in [27], we identify a true positive if there is an overlap between the prediction and the real samples. For example, the highlight of Figure 5a shows the ground truth of a anomaly sample with start index and end index, while the highlight of Figure 5b and Figure 5c are the predicted samples by two anomaly detection model. We can observe the predicted sample in Figure 5b has an overlap with ground truth from index 10 to 15, while there is no overlap between the predicted sample in Figure 5c and ground truth. Thus, the predicted sample in Figure 5b is detected as a true positive and the predicted sample in Figure 5c as a false positive.

4) **PARAMETER SETTING**

For all the three datasets, we extract anomaly subsequences from the raw accelerometer series records with fixed size sliding window (the window size is set as 10 and sliding by 5 each time). We set the threshold \( \lambda = 1.3 \) for inferring whether the vehicle is passing through a road anomaly based on fluctuation distance.

B. EXPERIMENTAL RESULTS

We conduct two groups of experiments and report their results. The first group is to perform parameter turning for anomaly subsequence segmentation using PAA-based features. The second group is to compare the effectiveness of the proposed detection model with four state-of-the-art competitors.

1) **IMPACT OF MODEL PARAMETERS**

For generating PAA-based feature, we first divide the original acceleration series data into equally sized frames and secondly compute the mean values for each frame as feature. The PAA scaling factor means the length of these equally sized frames. Tuning algorithm parameters, such as the PAA scaling factor for constructing PAA-based feature, are critical to the performance of removing false segmentation.

For each dataset, we perform five-fold cross validation and report the corresponding F-measure for both normal and anomaly class. We present the results of road anomaly detection in Figure 6 by varying the PAA scaling factor from 1 to 100. We observe the F-measure of normal and anomaly samples are basically stable at around 90% from Figure 6a, but the performance deteriorates significantly when the PAA scaling factor is between 36 and 42. The is because that dataset 1 is relatively small that only contains 131 anomaly samples, the partitioned test sets may contain few anomaly samples thus resulting in performance exception for some PAA scaling factor. For dataset 2, we can observe the F-measure slightly increase with the increasing of PAA scaling factor from 1 to 50, and decrease when PAA scaling factor increase continually from 50 to 69. As shown in Figure 6c, the trend of F-measure slightly decreases when the PAA scaling factor increase from 1 to 100, but the best performance is achieved when the PAA scaling factor is equal to 3.
For the three dataset, the best performance for road anomaly detection is achieved when the PAA scaling factor is 72, 74 and 3 (the F-measure of anomaly class are 91.2%, 95.6% and 98.4%, respectively).

2) PERFORMANCE COMPARISON OF ROAD ANOMALY DETECTION

For a comprehensive comparison, we investigate the performance of anomaly detection model by increasing the ratio of labeled samples from 50% to 90%. Each detection procedure with different ratio of labeled samples is repeated 10 times, we report the mean of the performance produced with 10 times as the ultimate experiment results.

For an experiment on a certain ratio ($x\%$) of labeled samples, we use $x\%$ dataset as training dataset and validation dataset, and the other $(1-x)\%$ dataset as test dataset. We have adopted two technologies to prevent over-fitting: 1) We use k-fold cross-validation ($k=5$) to select the tuning parameter that minimizes the prediction error of validation samples, which can avoid model over-fitting to some extent since k-fold cross-validation optimizes a tuning parameter that governs the number of features that are randomly chosen to grow each tree from the bootstrapped data; 2) We limit the hyperparameters range of random forest to control overfitting, the hyperparameters include: the number of trees. Basically, the more trees the less likely the model is to overfit; the depth of trees. Reducing the tree depth can add regularization, which will decrease the model overfitting; the feature of trees. Smaller features of each tree is randomly assigned, the less likely the model to overfit.

We compare the performance of 5 methods (SVM-WW, CPD-SFS, SVM-MDDP, MLM-WB and our method) on dataset 1, as shown in Figure 7. From this figure, we observe that our approach achieves the best performance in most cases, showing the advantage of learning scale-invariant feature for anomaly detection. For example, the F-measure of our method for metal bump is 57.1% when the ratio of labeled samples is 70%, while the F-measure of other baseline methods for metal bump are 20%(SVM-WW), 33.3%(CPD-SFS), 44.7%(SVM-MDDP), 13.6%(MLM-WB) with the same ratio of labeled samples. We also observe the performance of our method increases to some extent as the ratio of labeled samples increases. For example, the F-measure of our method for pothole is 72.1% when the ratio of labeled samples is 50%, while 77.8% when the ratio of labeled samples is 60%. This clearly demonstrates that our method can make full use of labeled sample to mine scale-invariant feature with high discriminative power for anomaly detection. Due to dataset

![FIGURE 6. Road anomaly detection performance with different number of PAA scaling factor. (a) Dataset 1. (b) Dataset 2. (c) Dataset 3.](image)

![FIGURE 7. Road anomaly detection performance on Dataset 1. (a) Pothole. (b) Speed bump. (c) Metal bump.](image)
1 is small in size, we find that some methods cannot detect anomaly in certain situations, for example, CPD-SFS and MLM-WB cannot find speed bump anomaly when the ratio of labeled samples is 50%. But our method can detect anomaly in any cases, showing our method is more stable and robust.

Figure 8 reports the performance of all anomaly detection model on dataset 2. Nevertheless, our method shows better performance consistently than other baseline methods with different ratio of labeled samples as we detect anomaly by learning scale-invariant feature. For example, the F-measure of our method for speed bump is 80% when the ratio of labeled samples is 60%, while the F-measure of other baseline methods for metal bump are 58.93%(SVM-WW), 33%(CPD-SFS), 72%(SVM-MDDP), 60.1%(MLM-WB) with the same ratio of labeled samples. In addition, we observe SVM-MDDP that detect anomaly by fusing statistical and frequency domain features, always outperforms other baseline methods (SVM-WW, CPD-SFS and MLM-WB) that merely utilizes one type of statistical, frequency domain features or raw data representation. For instance, set the ratio of labeled samples as 70%, SVM-MDDP outperforms SVM-WW by 20.5%, CPD-SFS by 6.1% and MLM-WB by 17.8% on pothole detection in terms of F-measure, respectively. This result suggests that, fusing statistical, frequency and time domain feature is beneficial for improving the performance of anomaly detection.

3) PERFORMANCE WITH DIFFERENT SPEED AND FREQUENCY

Two factors are important for road anomaly detection, which are vehicle speed and data collecting frequency. To investigate the performance with different speed and frequency, we utilize CarSim simulation to generate various datasets with different speed and frequency, and each dataset contains acceleration series records of 300 road anomalies (100 for potholes, 100 for speed bumps and the rest for metal bumps). For the dataset with different speed, setting the collecting frequency as 100 Hz, we generate acceleration series records with different vehicle speed (from 10 km/h to 100 km/h). For the dataset with different collecting frequency, setting the vehicle speed as 60 km/h, we generate acceleration series records with different data collecting frequency (from 10 Hz to 200 Hz). For each experiment, We randomly select 70% acceleration series records as training dataset to build

![Figure 8. Road anomaly detection performance on Dataset 2. (a) Pothole. (b) Speed bump. (c) Metal bump.](image-url)
anomaly detection model, and the rest 30% as testing dataset for evaluation detection performance.

Figure 10 reports the performance of road anomaly detection with different vehicle speed. We can see that the performance for the three kinds of anomaly detection is gradually increasing when improving the vehicle speed, and dropping when the vehicle speed is more than a certain speed. But the certain speed for different road anomaly is not the same, i.e., 70 km/h for pothole (the F-measure is 92.1%), 45 km/h for speed bump (the F-measure is 85.3%) and 75 km/h for metal bump (the F-measure is 84.5%). The reason is that the pattern embedded in Z-axis acceleration data series records when passing a road anomaly is more apparent with increasing the vehicle speed, thus resulting in better detection performance. However, the time spent on the same size of anomaly will gradually become shorter if increasing the vehicle speed, which will affect the quality of anomaly detection feature and further decreasing the detection performance. From Figure 10, we observe the proposed method can achieve relatively good anomaly detection performance when the vehicle speed is between 30 km/h and 80 km/h, which can be applied to most driving situations.
We further report the anomaly detection performance with different data collecting frequency in Figure 11. Clearly, the detection performance for all the three anomalies drops significantly with a low collecting frequency, which is no surprising since the extracted scale-invariant feature from acceleration series records in a slide window has poor discriminative power with a low collecting frequency. Similarly, the detection performance for all anomalies increase rapidly when the data collecting frequency is larger than 50 Hz. But the detection performance has a slight decrease when the data collecting frequency is larger than 120 Hz. The reason is that our method will align anomaly slide window with different length into the same length, which causes information loss during the stage. With increasing the sample frequency, the difference of the shortest length and the average length of anomaly slide window anomaly has widened, thus the anomaly detection performance decreases slowly when the sample frequency is too high.

V. CONCLUSION

This paper proposed a road anomaly detection model by learning scale-invariant features from crowdsourcing-based acceleration data. The proposed method aims to overcome two challenges of existing methods: the assumption that there already have segmented anomaly subsequences for extracting detection feature and the anomaly patterns embedded in acceleration sensor data might be locally distorted or scaled. Firstly, we extract candidate anomaly subsequences by sliding window scheme and remove false anomaly subsequences using PAA-based feature. Then, we learn the scale-invariant feature with high discriminative power for each anomaly class by evaluating the prediction quality of numerous candidate shapelets extracted from the anomaly subsequences. Experimental results on three datasets show that the proposed method achieves much better performance than the state-of-the-art baseline methods for road anomaly detection, showing the superiority of our approach and also supporting the assumption that the learnt scale-invariant feature among anomalies can boost anomaly detection.

As future work, we plan to enable real-time road anomaly detection by implementing our method using fog computing framework.

REFERENCES


