Towards a Practical Crowdsensing System for Road Surface Conditions Monitoring

Amr S. El-Wakeel, *Student Member, IEEE*, Jin Li, *Student Member, IEEE*, Aboelmagd Noureldin, *Senior Member, IEEE*, Hossam S. Hassanein, *Fellow, IEEE*, and Nizar Zorba, *Member, IEEE*

***Abstract*—The Internet of Things (IoT) infrastructure, systems, and applications demonstrate potential in serving smart city development. Crowdsensing approaches for road surface conditions monitoring can benefit smart city road information services. Deteriorated roads induce vehicle damage, traffic congestion, and driver discomfort which influence traffic management. In this paper, we propose a framework for monitoring road surface anomalies. We analyze the common road surface types and irregularities as well as their impact on vehicle motion. In addition to the traditional use of sensors available in smart devices, we utilize the vehicle motion sensors (accelerometers and gyroscopes) presently available in most land vehicles. Various land vehicles were used in this paper, spanning different sizes, and year model for extensive road experiments. These trajectories were used to collect and build multiple labeled data sets that were used in the system structure. In order to enhance the performance of the sensor measurements, wavelet packet de-noising is used in this paper to enable efficient classification of road surface anomalies. We adopt statistical, time domain, and frequency domain features to distinguish different road anomalies. The descriptive data sets collected in this paper are used to build, train, and test a system classifier through machine learning techniques to detect and categorize multiple road anomalies with different severity levels. Furthermore, we analyze and assess the capabilities of the smart devices and the other vehicle motion sensors to accurately geo-reference the road surface anomalies. Several road test experiments examine the benefits and assess the performance of the proposed architecture.**

***Index Terms*—Crowdsensing, machine learning, road information services (RISs), signal processing, smart city applications.**

Manuscript received September 15, 2017; revised December 20, 2017 and January 29, 2018; accepted February 12, 2018. Date of publication February 19, 2018; date of current version January 16, 2019. This work was supported in part by the Natural Sciences and Engineering Research Council of Canada under Grant STPGP 479248 and in part by the NPRP through the Qatar National Research Fund (a member of the Qatar Foundation) under Grant NPRP 9-185-2-096. *(Corresponding author: Amr S. El-Wakeel.)*

A. S. El-Wakeel and J. Li are with the Department of Electrical and Computer Engineering, Queen's University, Kingston, ON K7L 3N6, Canada (e-mail: amr.elwakeel@queensu.ca; jin.li@queensu.ca).

A. Noureldin is with the Department of Electrical and Computer Engineering, Royal Military College of Canada, Kingston, ON K7K 7B4, Canada, and also with the School of Computing and the Department of Electrical and Computer Engineering, Queen's University, Kingston, ON K7L 3N6, Canada (e-mail: aboelmagd.noureldin@rmc.ca).

H. S. Hassanein is with the School of Computing, Queen's University, Kingston, ON K7L 2N8, Canada (e-mail: hossam@cs.queensu.ca).

N. Zorba is with the Department of Electrical Engineering, Qatar University, Doha, Qatar (e-mail: nizarz@qu.edu.qa).

Digital Object Identifier 10.1109/JIOT.2018.2807408

I. Introduction

A. Background and Motivation

THE Internet of Things (IoT) is a network of devices with embedded technology that can collect, analyze, interact, and communicate data within itself or its environment through the Internet [1], [2]. According to the analyst firm Gartner Inc. [3], it is estimated that a total of 8.4 billion devices are connected to the Internet in 2017, and 20.4 billion IoT devices will be deployed by 2020. Consequently, in this ever-changing technological era, the IoT introduces endless opportunities and connections to enhance the efficiency of existing solutions leading to the global expansion of technological networks. IoT applications are seen commonly in smart homes, wearables, smart cities, and smart vehicles [4], [5]. These applications create an environment where every device communicates with other related devices to automate the home and industry and provides data to interested users [6], [7].

The evolution of smart cities continues to enhance efficiency, safety, and living standards. Smart city services range from healthcare and waste management to traffic and pedestrian monitoring in hopes of creating a sustainable intelligent system [8], [9]. Specifically, intelligent transportation applies sensing, analysis, control, and communication to ground transportation through the IoT to increase safety and mobility [10], [11].

Accordingly, location-based services (LBSs) could be used to enhance the quality of intelligent transportation. Crowdsensing as an aspect of crowdsourcing could benefit LBS, as sensed data can be transmitted with user participation or in an opportunistic form where no specific involvement of users is required [12], [13]. Additionally, micro-electro-mechanical systems (MEMSs) sensors embedded in smart devices, vehicle motion sensors, and GPS receivers contribute heavily to the crowdsensing LBS [13], [14]. Consequently, various intelligent transportation applications such as road information services (RISs), traffic congestion monitoring, real-time route planning, driver behavioral pattern systems, and road surface monitoring could be efficiently implemented, utilized, and assessed [15]–[18].

Road surface condition information is important for reasons related to vehicle damage and repair costs, safety, and control. Road surface monitoring is critical in transport infrastructure management; however, manual reporting can incur long delays and challenges for obtaining up-to-date information. Governments spend great efforts annually to provide safer

Fig. 1. Multiple road anomalies.

road networks, for instance, the U.K. government announced that it will spend $1.2 billion on road repairs between 2017 and 2018 [19]. Similarly, estimates from 2017 determined that vehicle damage from road anomalies cost an average American driver $523 in extra vehicle and maintenance fees, and this price may range depending on the state [20]. In Toronto, it costs around $25 to repair a pothole, and 360 000 potholes of all sizes and shapes were repaired in 2014, costing the city $6 million every year [21]. Likewise, in Montreal, potholes cost around $7 million annually [22]. Though financially costly, serious accidents can arise due to road anomalies. For instance, between 2000 and 2011, over 2 million traffic accidents were reported in Canada, 33% of which were related to poor weather or road conditions [23]. In 2015, around 50 000 motorists reported damage caused by potholes in the U.K., with an estimate of potholes damaging cars every 11 min [24]. Therefore, an important part of RIS aims to efficiently monitor road surface conditions and anomalies.

B. Problem Statement

As previously discussed, the effects of road surface anomalies and the government attention in monitoring and repairing affected roads they both highlight the importance of addressing irregularities and their effect on multiple common driving styles. This requires performing several driving surveys in different locations to label the most common road anomalies, and define their effects as per the example shown in Fig. 1. The several driving surveys should enable the identification and categorization of road surface types and anomalies over several road test trajectories. The road distress appears in different forms such as transverse cracks, longitudinal cracks, crocodile cracks, and road dents. Also, this distress might cause the removal of pavement, introducing various levels of potholes. In addition, road-related infrastructure services such as manholes,

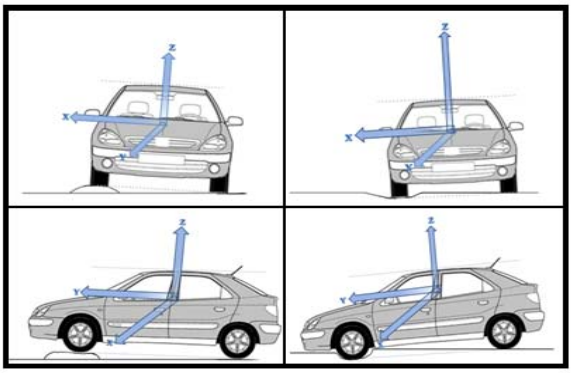


Fig. 2. Vehicle motion under single and double-sided road anomalies.

speedbumps, drain pits, road and bridge joints, deceleration strips, and railroad crossings may also damage vehicles when they are improperly maintained. Regarding the effects of the above-mentioned anomalies, they affect the vehicle motion in different perspectives. According to their nature, a vehicle can either attend the event with single-sided or double-sided wheels. Under certain conditions, the vehicle could attend the event with only one wheel due to swerving. Fig. 2 shows the effects of the event on the vehicle motion. Therefore, there is a growing demand toward crowd-based information services to monitor road health conditions and road hazards.

C. Objective

In this paper, we present a framework to collect crowd-based information to monitor road health conditions and road hazards. The proposed system utilizes the sensor technologies available in both vehicles and driver smart devices to offer more advanced, accurately localized road surface anomaly monitoring. While developments have been independently made in these areas of intelligent transportation systems, there is very limited work toward a framework that integrates the above to provide a real-time, personalized, driver service application. Therefore, this paper aims at the following.

1) Utilizing different types and grades of accelerometers and gyroscopes including low-cost consumer-grade MEMS-based sensors and fiber optics automotive grade sensors to enable rich collection of data for different road types and anomalies.
2) Exploring multiple common road irregularities while analyzing the effects of irregular roads over vehicle motions by adopting eight different car models that span multiple types, sizes and model year. This provides a comprehensive collection of a variety of driving scenarios.
3) Developing a wavelet packet de-noising method to enhance the quality of the data sensed by low-cost MEMS sensors, and to separate the vehicle motion dynamics and the effects of the road anomalies from other measurement noises and disturbances.
4) Processing the de-noised measurements by a feature extraction technique using statistical, time domain and

frequency domain approaches to extract rich and distinguishable features for each road anomaly type.

5) Adopting a multilevel support vector machine (SVM) classifier that is able to detect and classify eight road anomaly types and irregularities with different levels of severity.
6) Integrating positioning solution from GPS receiver with motion sensors, providing robust geo-referencing of road surface anomalies even in challenging/denied GPS environment.

II. Related Work

Currently, the authorities monitor road conditions and manual reporting is employed, but this form of tracking does not account for the constant changes in road conditions, where road damages can deteriorate quickly. On the other hand, some crowdsensing-based applications were developed for road anomaly detection. As in 2012, the city of Boston partnered with Connected Bits to develop an application known as "Street Bump," which requires the user to install the application on their mobile device to sense and report potholes [25], [26].

Lately, the smartphones, tablets, and other smart mobile devices were used for assessing and monitoring road conditions through inputs from mainly the accelerometers, gyroscopes, and GPS sensors. Recent technical research has led to various approaches in addressing this problem. Brisimi *et al.* [27] introduced a detection and decision support system called Street Bump which classifies roadway "bumps" into actionable (potholes and manhole covers) and nonactionable (train tracks and speedbump). Predetermined and trained machine learning algorithms identify the ones of immediate attention with up to 88% accuracy. Kalim *et al.* [28] presented a mobile app that utilizes accelerometers, gyros, and GPS on smart devices to collect data while driving. Using machine learning techniques the anomalies are only classified to potholes and speedbumps with an average detection accuracy of 92.5%. Embedded GPS receiver was used to localize the anomalies. The proposed crowdsourcing-based road surface monitoring system from [29] can detect potholes with up to 90% accuracy and road roughness levels using accelerometers and GPS devices, yet it only focuses on pothole detection. Pothole in the dark [30], proposed a system that detects and determines the dimensions of potholes. Leveraging smartphone sensors and a land vehicle, they trained a system that considers the vehicle parameters and detects potholes. The main limitation of this system is that it is trained to detect only potholes at a relatively low speed below 30 km/h that can be irrelevant for some scenarios.

In [31], a half car model was derived and implemented to describe the car motion and any encountered disturbances (in the form of potholes, speedbumps, and road joints). Front and rear left suspension reflection, vertical acceleration, and roll rate are the measurement inputs. Threshold-based detection techniques were used to classify the detected anomalies. The system achieved a labeling accuracy of $\sim$ 86% and false negative rate (FNR) of $\sim$ 2%. Additionally, Fox *et al.* [32] involved multilane pothole detections from crowdsourcing, using accelerometer data from embedded vehicle sensors with inclined roads. It also presents various mathematical models and algorithms to assess the simulated data. However, it does not consider other possible road anomalies and does not obtain training data through real-world driving.

Generally, a robust road surface condition monitoring system should address two main aspects: 1) the anomaly detection with its classification system structure and 2) the geo-referencing system for road anomalies. Most recent systems lack several aspects required for robust detection and classification. There are some systems highly focused on detecting a limited variety of anomalies, specifically potholes [29], [30]. Others lack appropriate experimental validation or consider unrealistic scenarios for their experiments, this spans low vehicle speed, restricted sensor placements, and using simulated data [30]–[32]. Moreover, smartphone-based road surface monitoring systems rely on GPS for geo-referencing the detected events, which is not adequate due to positioning errors in urban canyons. While considering GPS signal blockage and multipath, localization errors are significantly increased in downtown cores and urban areas [33], [34]. The lack of adequate event localization leads to insufficient system operation; hence, event detection without accurate localization is meaningless.

Considering all the aspects of monitoring road anomalies would enable robust and continuous reporting systems of the road surface conditions, to benefit present and future smart cities operation. Additionally, drivers' prior knowledge of the road quality will strongly contribute to their route planning. As the current route planning key players such as Google Maps and Apple Maps Connect suggest best routes based only on the shortest distance routes and the fewer traffic roads. Providing continuous and accurate road quality information will definitely open the door toward efficient and road maintenance, dynamic route planning, and traffic management. In addition, road surface condition information could be highly valuable for autonomous vehicles to ensure comfort and safe trips.

III. System Configuration and Data Collection

Fig. 3 shows the proposed system configuration. Our methods process both the vehicle and smartphone sensors to recognize the road-related anomalies, such as the existence of potholes, manholes, and other road surface objects. Sensor measurements from both vehicle and driver's smartphone are first synchronized and time tagged to the GPS time.

In this paper, we developed multiple trajectories of varying lengths to be used in identifying, training, and for testing purposes. For the trajectories, we utilized multiple inertial measurement units (IMUs), GPS receivers, and integrated positioning units. Three smart devices, two smartphones (Nexus 5 and Samsung Galaxy S4), and one tablet (Samsung GT-N8010) were used. These devices are embedded with six degrees of freedom IMUs and GPS receivers. In addition, we also used a low-cost MEMS grade six degrees of freedom IMU (Crossbow), and two integrated positioning units (VTI and Novatel). The VTI includes a MEMS grade

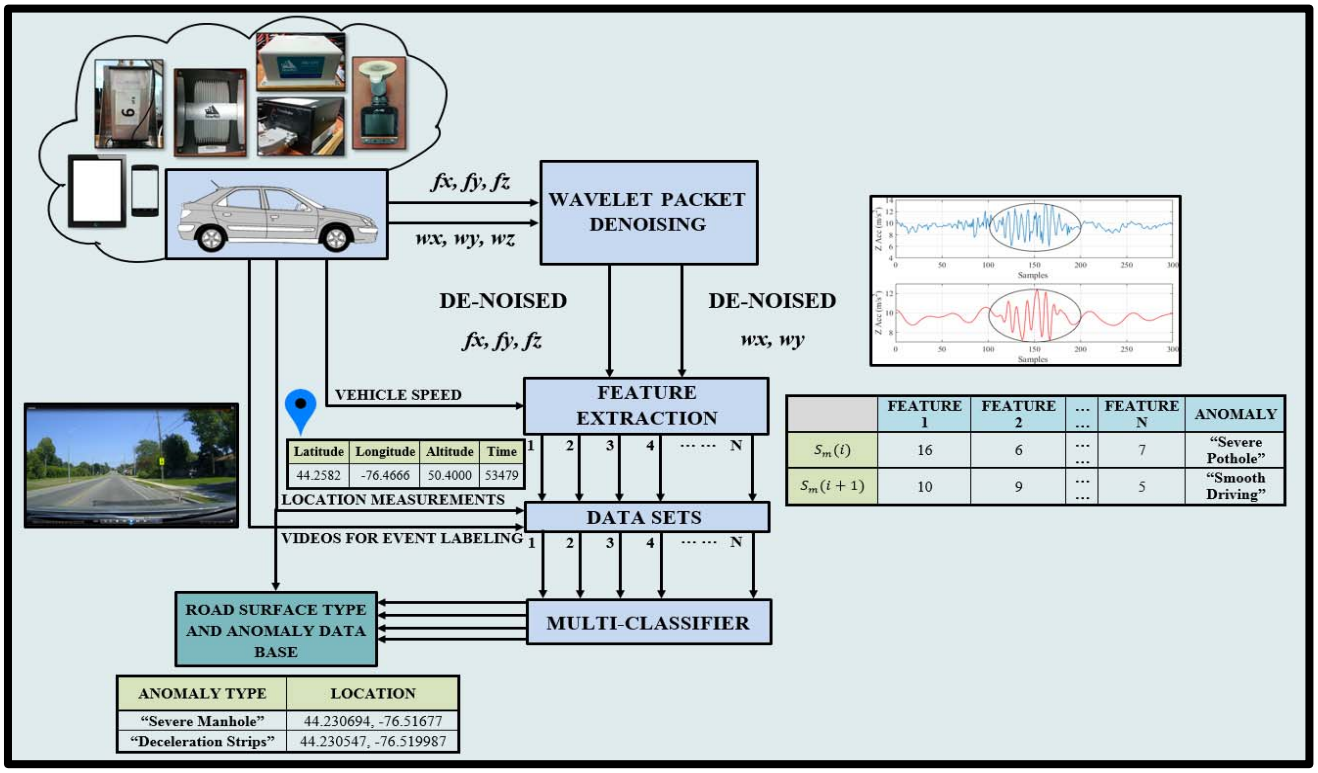


Fig. 3. System configuration for monitoring road surface types and anomalies.

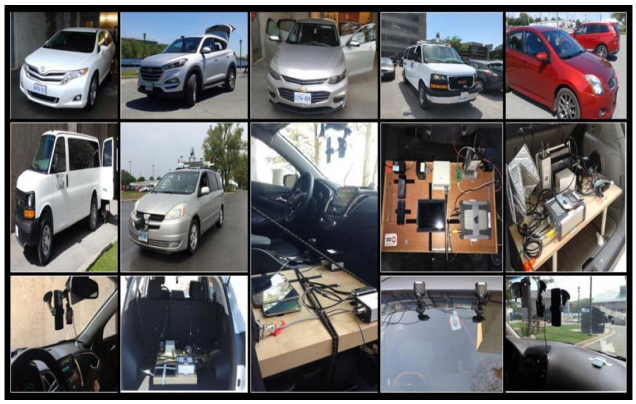


Fig. 4. Multiple vehicles, testbed, and sensors.

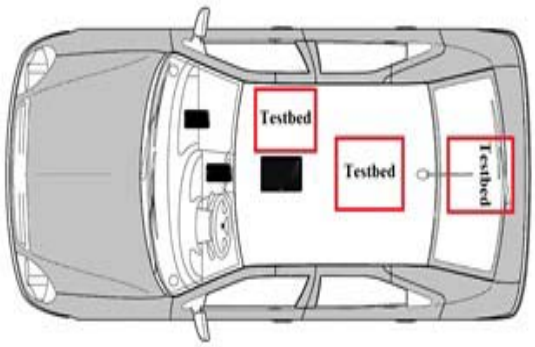


Fig. 5. Testbed and smart devices locations in the vehicle.

IMU and an OEM GNSS receiver. The NovAtel SPAN integrated solution includes a tactical grade span-CPT IMU with MEMS accelerometers, fiber-optic gyros, and the OEM GNSS receiver. We also used a MiVue 388 Dash Cam for recording the trajectories.

These sensors are gathered on a testbed or in an open setup and then mounted on multiple vehicles as shown in Fig. 4. We utilized eight land vehicles: two vans, two sedans, two crossovers, one mini-van, and one hatchback. We intentionally chose them to span multiple sizes, vendors, wheel size, quality of suspensions, model, and make, which provides a broad range of real-driving scenarios. The testbeds and open setup are positioned and oriented differently in the vehicles as shown in Figs. 4 and 5.

The positions of the testbed range from the front passenger seat, middle backseats (minivan and van), and trunk (crossover and hatchback). In an open setup, smartphones and tablets are secured on the arm rest, windshield, dashboard, front passenger seats, and cup holder. Using the mentioned forms of the experimental setup, we held multiple trajectories as shown in Fig. 6 that spanned different road types and anomalies in the downtown core, residential neighborhoods, and urban streets.

Using the experimental setup, we approached the different events using various vehicles, sensor grades, driving behaviors, and speeds. This leads to a rich collection of different data sets, and describe every single event by attending them in different ways. In order to obtain highly descriptive data, we logged all sensors at their maximum and relevant data rates. For smart devices, we logged the accelerometers at approximately 100 Hz, the gyroscopes at approximately 200 Hz,

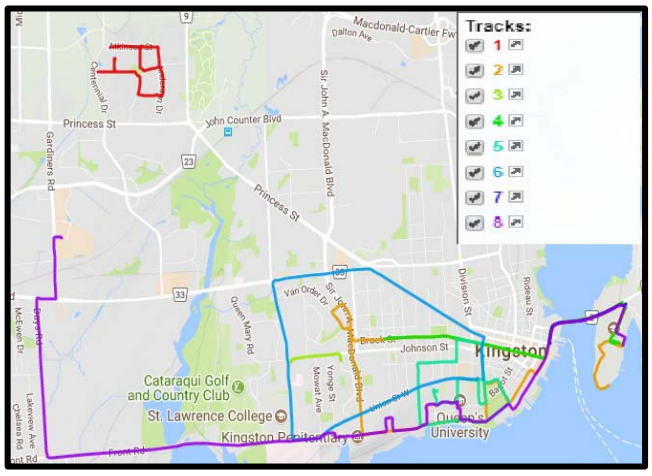


Fig. 6. Multiple trajectories used for system building, testing, and training.

and the GPS measurements at 1 Hz. In regards to the external IMUs, Crossbow (Xbow), and IMU CPT were logged at 100 Hz for both accelerometers and gyroscopes while the VTI full IMU was logged at 20 Hz. Additionally, the VTI and NovAtel OEM GPS receivers were logged at 1 Hz. Similarly, the integrated positioning solution provided by both units (VTI and NovAtel SPAN integrated solution) were obtained at 1 Hz.

During the trajectories, the attended anomalies were: potholes, manholes, transverse cracks, longitudinal cracks, railroad tracks, speedbumps, deceleration strips, paved roads, and road dents. For every single trajectory, we built a data set for each sensor that was used and labeled each attended event using the recorded videos for guidance. Later the data sets from each trajectory were combined, these data sets were used independently or combined with other data sets from different trajectories. The data sets of varying sizes were arranged for building, training, and testing of classifiers.

IV. Experimental Results and Discussion

A. Signal De-Noising

Principally, the sensed and collected data reflects vehicle motion regular dynamics, road anomalies effects, and noise. For IMU's, specifically, there are multifarious sources of noise. Basically, IMU's noise can be branched to long-term errors (ones with low-frequency components) and short-term errors (ones with high-frequency components).

Consequently, long-term errors could be presented as a result of different causes. For instance, the noise from external or internal heat distribution variations or temperature is represented by exponentially correlated noise. This type of noise appears in low frequencies and is identified as a time-varying additive noise. In addition, random walk, another form of the low noise frequency component, could appear as result of shot or thermal noise in a photodetector of optical gyros [33]–[35]. On the other hand for short-term errors, they could be presented as a result of different causes. For instance, white sensor noise is distributed in all the frequencies of a sensor according to its bandwidth, this type of noise is a result of electronic instruments such as power supplies, quantization

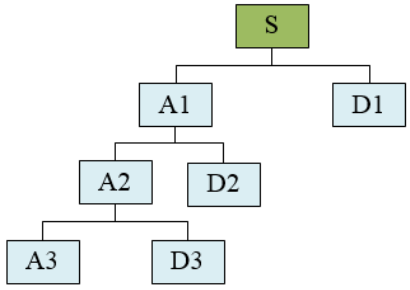


Fig. 7. Wavelet decomposition tree.

errors during digitization or even through the semiconductor devices intrinsic noise [36], [37]. Clearly, suppressing the noise effects on the sensed data is a challenging process as the long-term errors are mixed with vehicle dynamics low frequencies [38], [39]. Also, removing the effects of the high-frequency noise could eliminate the road anomalies effects leading to high detection errors.

In order to achieve robust signal de-nosing, the best case scenario is that a full prior information about the signal is available. In such case, signal and noise joint probabilities are known and the average estimation error could be minimized [40]. One of the trivial de-noising techniques is the Fourier transform (FT); however, the main disadvantage of the FT is that once a frequency domain transformation is performed, any relevant time domain information is lost. Accordingly, windowed discrete time FT (WDFT) was proposed to overcome the lack of time localization in ordinary FT. In WDFT, a time window is predefined assuming local signal periodicity. For each window, the WDFT produces a sequence of complex values whose magnitudes are those of the discrete frequencies of the input. A major drawback of WDFT is that a window with fixed width always rise the time-frequency tradeoff, as narrow window width leads to appropriate time localization, but on the other hand insufficient frequency resolution is obtained. Furthermore, WDFT can experience phenomena named spectral leakage, as to model a frequency that does not have an integral number of periods in the record length, energy is spread into all the other frequencies [41]. In order to overcome the challenges of WDFT, wavelets are proposed and adopted. In wavelets, the analysis grants the usage of relatively long time windows where low-frequency components are required. Meanwhile, where high-frequency content is needed, shorter intervals are adopted. Consequently, in the wavelet multiresolution analysis (WMRA) process, the signal is divided into multiple resolution levels as shown in Fig. 7.

In WMRA, a wavelet function, Ψ_t, is used to seek the details in a signal in an operation equivalent to high pass filtering while a scaling function, ϕ_t, is designed to smooth the input signal to seek its approximation in a process equivalent to low pass filtering [38]–[40]. The wavelet and scaling functions are usually orthonormal functions. On the other hand, even with adopting thresholds, some motion dynamics or road anomalies influences could be deteriorated [42].

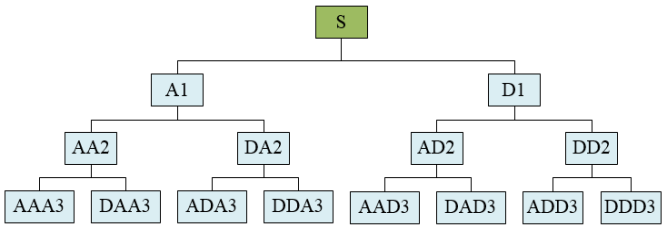


Fig. 8. Wavelet packet decomposition tree.

To bypass the challenges of wavelet de-noising, a wavelet packet de-noising is adopted in our analysis. Wavelet packets, as shown in Fig. 8, apply an initial decomposition step that separates the signal into approximation (A) and details (D).

Furthermore, both of A and D are decomposed for several levels in order to break the signal into fine resolution components [40], [42]. Regarding our de-noising approach, we utilized bi-orthogonal wavelet packet de-noising analysis. Basically, the analysis utilizes wavelet packet bases that are able to break down the frequency axis into apportioned slots with multiple sizes and these slots are translated in time to cover the entire time-frequency plane. As any space, $\boldsymbol{V_j}$, could be split to subspaces of approximation, $\boldsymbol{V_{j-1}}$, and details, $\boldsymbol{W_{j-1}}$, where

$$\boldsymbol{V_j} = \boldsymbol{V_{j-1}} + \boldsymbol{W_{j-1}}. \tag{1}$$

In the case of orthogonal bases, this is done by breaking the orthogonal basis of $\boldsymbol{V_j}$ to orthogonal bases of approximation and details to $\{\boldsymbol{\phi_{j-1}}(\boldsymbol{t} - \boldsymbol{2}^{\boldsymbol{j-1}}\boldsymbol{n})\}_{\boldsymbol{n}\in\mathbb{Z}}$ and $\{\boldsymbol{\Psi_{j-1}}(\boldsymbol{t} - \boldsymbol{2}^{\boldsymbol{j-1}}\boldsymbol{n})\}_{\boldsymbol{n}\in\mathbb{Z}}$, respectively, this split is specified by a pair of conjugate mirror filters $\boldsymbol{h}[\boldsymbol{n}]$ and $\boldsymbol{g}[\boldsymbol{n}] = (-\boldsymbol{1})^{\boldsymbol{1-n}}\boldsymbol{h}[\boldsymbol{n} - \boldsymbol{1}]$. Considering bi-orthogonal wavelet packets, they exploit nonorthogonal wavelet bases which are built with two pairs of perfect reconstruction filters (h, g) and $(\tilde{h},\ \tilde{g})$ [40], [43]. For thresholds, we used soft threshold of Stein unbiased risk estimator (SURE) introduced in [44]. As mentioned in [40], reducing the thresholding risk could occur by choosing a threshold less than $\sqrt{2\log_e N}$, where σ is the standard deviation and N is the signal length. For the SURE thresholding, T

$$T = \sqrt{2\ \log_e(N\log_2(N)}. \tag{2}$$

In order to assess the performance of the discussed wavelet packet de-noising, we applied the technique on the system inputs listed in Table I. These inputs are the linear accelerations and the rotational velocities gathered by various adopted IMUs. Figs. 9 and 10 show the linear accelerations and angular rotations for the Crossbow IMU, wavelet packet de-noised Crossbow IMU and IMU-CPT during a road anomaly.

We have intentionally chosen to show the results of the packet de-nosing of the Crossbow IMU, as it has the lowest performance among all the IMUs in our experiments. In addition, the IMU-CPT were used as a reference IMU due to its higher performance among the others. The standard deviation results displayed in Table I show that wavelet packet de-noising of the Crossbow IMU have been enhanced on average by approximately 48%.

TABLE I
STANDARD DEVIATIONS FOR XBOW IMU, WAVELET PACKET DE-NOISED XBOW IMU, AND IMU-CPT

INERTIAL SENSOR MEASUREMENT	STD OF XBOW	STD OF WP DE-NOISED XBOW	STD OF IMU-CPT
f_z	3.108	2.528	2.290
f_y	1.006	0.376	0.248
f_x	1.079	0.608	0.596
ω_z	1.239	0.281	0.200
ω_y	3.551	2.110	1.607
ω_z	4.999	2.730	1.274

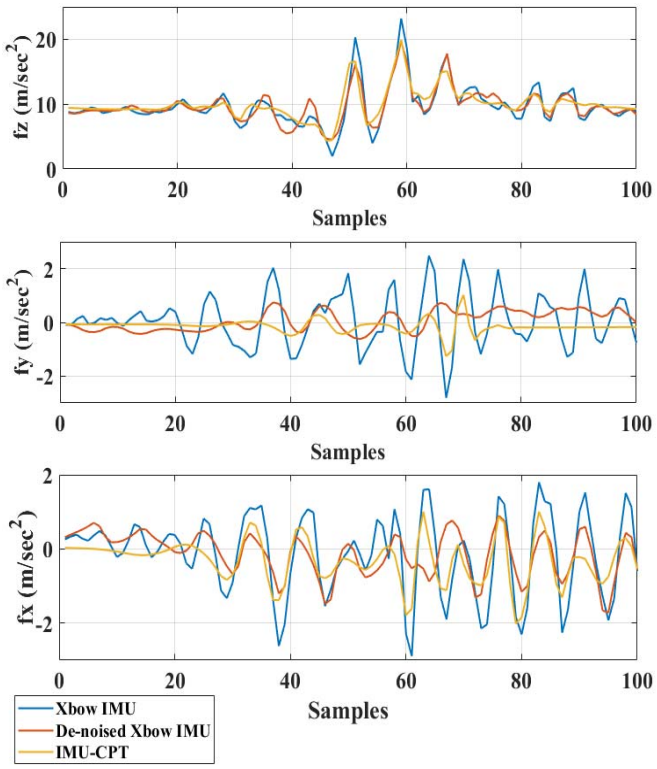


Fig. 9. Linear accelerations for Xbow IMU, wavelet packet de-noised Xbow IMU, and IMU-CPT during road anomaly.

Furthermore, all the sensed linear accelerations and angular rotations gathered by the whole utilized IMUs are de-noised using the same technique, before being applied to the feature extraction process as shown in Fig. 3.

In addition to the Crossbow IMU results, and in order to examine the impact of the wavelet packet de-noising on the inertial sensors embedded in the smart devices, we provide the results related to the Samsung tablet in Figs. 11 and 12. They plot the linear accelerations and angular rotations for the accelerometers and gyroscopes sensors of the tablet device, wavelet packet de-noised tablet sensors and IMU-CPT during the same road anomaly shown in Figs. 9 and 10.

Consequently, as displayed in Figs. 11 and 12, wavelet packet de-noising enhanced the quality of the tablet accelerometers and gyroscopes. Table II shows the standard deviations of the linear accelerations and the angular rotations of the tablet

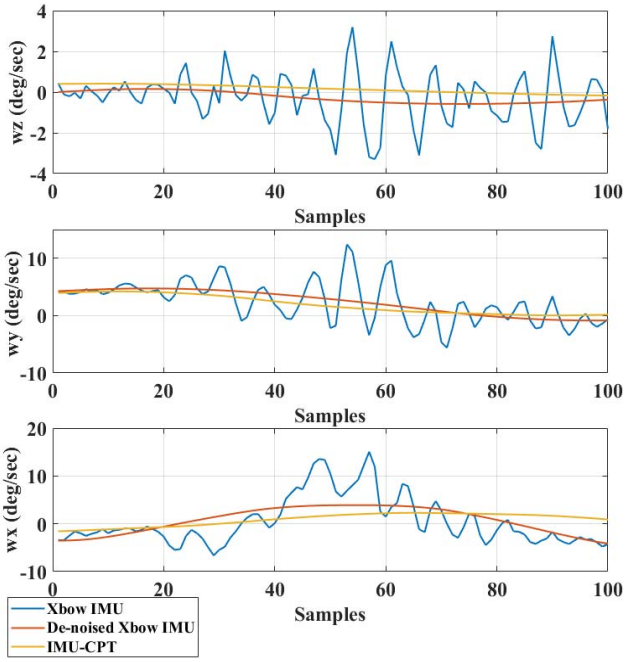


Fig. 10. Angular rotations for Xbow IMU, wavelet packet de-noised Xbow IMU, and IMU-CPT during road anomaly.

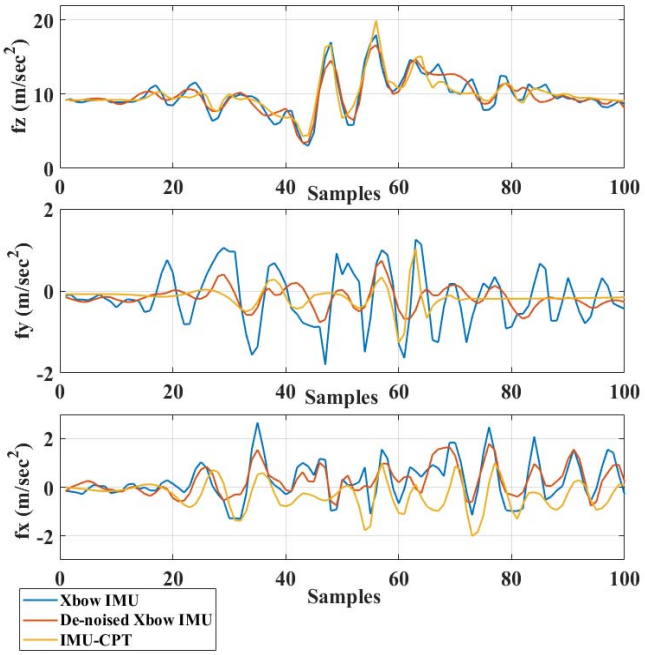


Fig. 11. Linear accelerations for tablet accelerometers, wavelet packet de-noised tablet accelerometers, and IMU-CPT during road anomaly.

sensors before and after wavelet packet de-noising compared to IMU-CPT as a reference. The wavelet packet de-noising enhanced the tablet sensors measurements by approximately 40% on average.

B. Feature Extraction

Highlighting and analyzing the effects of road surface anomalies on the vehicle motion as shown in Fig. 2, we noticed that the disturbance of the motion occurs at the sensed

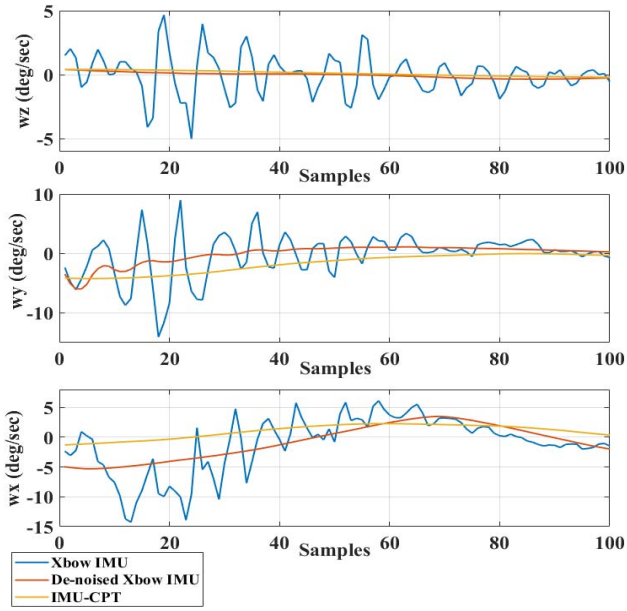


Fig. 12. Angular rotations for Samsung tablet gyroscopes, wavelet packet de-noised tablet gyroscopes, and IMU-CPT during road anomaly.

TABLE II
STANDARD DEVIATIONS FOR TABLET SENSORS, WAVELET PACKET DE-NOISED TABLET SENSORS, AND IMU-CPT

INERTIAL SENSOR MEASUREMENT	STD OF TABLET	STD OF WP DE-NOISED TABLET	STD OF IMU-CPT
f_z	2.568	2.300	2.290
f_y	0.667	0.262	0.248
f_x	0.836	0.614	0.596
ω_z	1.575	0.212	0.200
ω_y	3.822	2.308	1.607
ω_z	4.764	2.848	1.274

transversal and angular rotation of the IMUs. However, considering the data of the angular rotation along the Z-direction could be misleading as it is mainly used in identifying the vehicle heading angle. Accordingly, we combine the other five system inputs as listed in Tables I and II.

Based on the nature of the events, the effects of the attended anomaly with single-sided wheels (ex. manhole) appears on the three-axes of the transversal acceleration and on the angular rotation along the x-axis. Meanwhile, the double-sided wheel anomalies (ex. transverse cracks) mainly affect the $Z-Y$ transversal accelerations and the angular rotation around the y-axis. Additionally, vehicle speed, V, is considered in order to avoid the misleading effects of getting in or out of a parking event.

These sensed linear accelerations and angular rotations are present in all the data sets. For extracting descriptive features of road anomalies, the data sets are "time windowed" at 1 s each. During a time window, a vector of more than 75 features is being constructed. In order to construct efficient data sets, we mainly considered the events that have their signature on the sensed data to occur in the same time window. However,

for some events, where the signature of the anomaly is divided between two successive time windows, we made an overlap between these time windows to assure the best collection of the anomaly signature over the sensed data.

There are multiple feature extraction techniques adopted for road surface condition monitoring [27], [30] and other applications such as motion mode recognition [45]. The main feature extraction techniques can be categorized into statistical, time-domain, frequency-domain, and time-frequency domain features. In this paper, we adopted statistical, time domain, frequency domain, and other features.

Statistical features are mainly describing typical or central values of a data set and give measures on the data distribution behavior [45], [46]. Regarding the statistical features used, at each time window, the main statistical features computed are: mean, median, mode, range, maximum, minimum, root mean square (RMS), peak to RMS range, root sum of the squares, standard deviation, variance, interquartile range, percentile range, cumulative maximum, and minimum. We also developed some statistical features that combine the effects of more than one acceleration component at the same time epoch. These combined features are helpful in distinguishing similar behavior events as potholes and manholes. Considering time domain features which describe the signal variation with respect to the time domain, we mainly used multiple zero and threshold crossing rates and also the local maxima and minima points within each time window.

Frequency domain features describe the signals' behavior in the frequency domain. We computed frequency domain features such as mean and median frequencies, and also computed the estimated power spectral density within a time window. The other features used include multiple cross-correlations and cross-covariance between transversal and longitudinal accelerations. The main types of used features and their definitions are summarized in Table III. All these computed features were used in constructing a unique feature vector at each time window and for each type of data set to provide a high description level of each labeled road surface anomaly.

TABLE III
VARIOUS FEATURES AND THEIR DEFINITIONS

Feature	Domain	Definition
Mean	Statistical	$\text{Mean (s)} = \frac{1}{K}\sum_{m=1}^{K} s[m]$
Mode	Statistical	The most frequent value in a time window
Median	Statistical	Threshold value that separates the higher and lower 50% of samples in a time window
x^{th} Percentile	Statistical	Threshold value that separates the (100% - x %) higher samples from the lower x% of samples in a time window
Interquartile Range	Statistical	The range between (100%-x%) and x% of the samples in time widow
Variance	Statistical	$var(s) = \sigma_s{}^2 = \overline{(s-\bar{s})^2}$
Standard Deviation	Statistical	$std(s) = \sqrt{var(s)} = \sigma_s$
Root Mean Square	Statistical	$RMS(s) = \sqrt{\frac{1}{K}\sum_{m=1}^{k} s[m]^2}$
Local Maxima	Time	The number of maxima points in a time window
Local Minima	Time	The number of minima points in a time window
Threshold Crossing Rate	Time	$tcr(s) = \frac{1}{K-1}\sum_{m=1}^{K} \mathbb{C}\{s[m]s[m-1] < T\}$
Mean Frequency	Frequency	Estimates the mean normalized frequency of the power spectrum of a time window
Median Frequency	Frequency	Estimates the median normalized frequency of the power spectrum of a time window
Cross-Correlation of Linear Accelerations	other	$\frac{\sum_{m=1}^{K}(f_y[m]-\bar{f}_y)(f_x[m]-\bar{f}_x)}{\sqrt{\sum_{m=1}^{K}(f_y[m]-\bar{f}_y)^2\sum_{m=1}^{K}(f_x[m]-\bar{f}_x)^2}}$
Cross-Correlation of angular rotations	other	$\frac{\sum_{m=1}^{K}(w_y[m]-\overline{w_y})(w_x[m]-\overline{w_x})}{\sqrt{\sum_{m=1}^{K}(w_y[m]-\overline{w_y})^2\sum_{m=1}^{K}(w_x[m]-\overline{w_x})^2}}$

C. Classifiers Building, Training, and Testing

The held road experiments led to various data sets of more than 1000 road anomalies, plus data for mild road driving. Through analysis, it was observed that the roads located downtown and receive high traffic and less maintenance have a road anomaly or irregularity on average at every 1 s. Concurrently, the average anomalies rise to reach an average of 1 anomaly per 8 s in relatively new and well-maintained residential neighborhoods. These numbers show the demand and the challenges of robust anomaly detection systems.

The labeled events in the data sets spanned various anomalies with multiple severity levels. Mainly, and based on their frequent presence and their significant effects on vehicle motion, we chose to build the system to monitor potholes, manholes, transverse cracks, longitudinal cracks, road dents, railway tracks crossings, speedbumps, deceleration strips, and paved stone road. Some anomalies such as railway track crossings, speedbumps, and deceleration strips have well-known locations to authorities and road operators. However, these could be harmful to the vehicles if they are improperly installed or not well-maintained, hence we decided to detect and classify their level of severity as well. On the other hand, there are some road types such as paved stone roads which can be wrongly detected as longitudinal cracks, therefore they are also being detected and classified.

As the first step in building the classifier, we adopted various classification techniques such as decision trees [47], [48], SVM [27], [49], k-nearest neighbor [50], and ensemble

meta classifiers [51]. Additionally, we used the data sets built for each trajectory to train each type of classifier with the labeled feature vector constructed from each event. The results showed a higher average performance within multiple classification techniques achieved by the support vector machine and bagged trees.

Respectively, the main concept of the SVM is in trying to locate most of the data or features, $\boldsymbol{S_m}$, that fit a class in a sphere. The sphere comes with radius, r, where it can be minimized as follows:

$$\begin{aligned} &\text{minimum} \;\; r^2 + C \sum_{m=1}^{k} \Gamma_m \\ &\text{Such that} \;\; (\|\boldsymbol{C} - \boldsymbol{S_m}\|)^2 \leq r^r + \Gamma_m \\ &\qquad \Gamma_m \geq 0 \end{aligned} \tag{3}$$

where C is a parameter for the number of errors and sphere volume trade off control and Γ_m is a factor that allows some data samples to endure outside a sphere. Consequently, for two data or feature vectors $\boldsymbol{S_m}$ and $\boldsymbol{S_n}$ it can be proven that

$$\begin{aligned} &\max \;\; \sum_{m=1}^{k} \gamma_m \langle \boldsymbol{S_m}.\boldsymbol{S_n} \rangle - \sum_{m,n=1}^{k} \gamma_m \gamma_n \langle \boldsymbol{S_m}.\boldsymbol{S_n} \rangle \\ &\text{s.t.} \;\; 0 \leq \gamma_m \leq C \text{ and } \sum_{m=1}^{k} \gamma_m = 1 \end{aligned} \tag{4}$$

as $\langle \boldsymbol{S_m}.\boldsymbol{S_n} \rangle$ is the feature vectors' inner product, and γ_m and γ_n are Lagrange multipliers.

Training sets could be present at any distribution rather than spherical and the feature vectors' inner product could be substituted by a kernel function. Kernel functions [49] span linear, polynomial, and radial basis function kernels.

On the other hand, ensemble meta classifiers [51] adopt classification techniques where multiple classifiers of a different or similar type are being trained over the same or subsets of a training set. Specifically, in bagging, a classifier is being trained in leveraging subsets of training sets. The construction of the subsets is a result of random selection from the main training set. Accordingly, we combined the whole constructed data sets gathered by different sensors and different cars utilized in all trajectories. Afterwards, we used almost 60% of the data sets in building and training the SVM module as well as bagging classifiers. The remaining 40 % was used for the performance assessment.

For the SVM, a cubic kernel function was used to build a multilevel classifier with an automatic kernel scale. In the training process, all 75 features were used in order to achieve better results. For the ensemble meta classifier, we built and trained a bagged trees classifier. Hence, in the early stages of training, both classifiers were observed to have a better performance of the multilevel SVM over the bagged trees. As a result, we assessed the performance of the SVM classifier with the remaining data sets. Table IV presents various types of road types or anomalies with their corresponding number of occurrences. The data sets for the anomalies presented in Table IV are used for testing the classifier. Furthermore, we adopted a confusion matrix that presents the results for eight various road types and anomalies with different levels of severity. Additionally, in Table V, the SVM multilevel classifier showed an average performance of approximately 90% true positive rate (TPR) while achieving approximately 10% of average FNR. The highest TPR of approximately 95% was present in predicting two different kinds of road types and anomalies (mild road driving and longitudinal cracks). On the other hand, the highest FNR of approximately 20% was achieved while detecting deceleration strips. In many occasions when attending the deceleration strips, they were being detected as transverse cracks (20%) because of improper construction or less maintenance.

TABLE IV
VARIOUS ROAD TYPES OR ANOMALIES AND THEIR NUMBER OF OCCURRENCES USED FOR CLASSIFIER TESTING

ROAD TYPE OR ANOMALY	NUMBER OF OCCURRENCES
Mild Road	60
Mild Pothole	30
Severe Pothole	20
Mild Manhole	40
Severe Manhole	25
Transverse Crack	50
Longitudinal Cracks	40
Mild Dent	20
Severe Dent	10
Deceleration Strips	10
Speedbumps	20
Mild Railway Crossing	20
Severe Railway Crossing	10
Paved Stone Road	50

Moreover, to assess the analysis of the extracted features in building the classifier, we calculated the average FNR of single-sided events (ones attended with one side of the wheels) and double-sided events to check how often a miss-detected event is being predicted within the same general class. Consequently, the single-sided events (mild potholes, severe potholes, mild manholes, and severe manholes) achieved an average FNR of approximately 11%. We found that approximately 91% of the miss-detected single-sided events were detected as other single-sided events at different types and multiple severity levels. On the other hand, the double-sided events (ones attended with both sides of the wheels) of the same general class (transverse cracks, deceleration strips, mild railway crossing, and severe railway crossing) achieved an average FNR of 13%. While 83% of the miss-detected double-sided events were being predicted as single-sided events.

Regarding the false positive rate (FPR), the average FPR for the whole anomalies and types is 9.7%. As the highest FPR is achieved in the case of deceleration strips with 40%, while the lowest is obtained in the case of speedbumps with 0%. For single-sided events, the case of severe manhole experienced the highest FPR with 16.6%. On the other hand, for double-sided events, deceleration strips experienced the highest FPR with 40%.

TABLE V
CONFUSION MATRIX FOR ROAD SURFACE ANOMALIES DETECTION

	Predicted Road Anomaly																
Actual Road Anomaly	*MR*	*MP*	*SP*	*MM*	*SM*	*TC*	*LC*	*MRD*	*SRD*	*DS*	*SB*	*MRC*	*SRC*	*PSR*	**TP**	**FN**	**FP**
MR	95%									5%					95%	5%	6.6%
MP		90%	3.3%	3.3%				3.3%							90%	10%	16.6 %
SP			85%	5%	5%		5%								85%	15%	2%
MM		5%		90%				5%							90%	10%	10%
SM			4%		92%				4%						92%	8%	8%
TC	4%					94%				2%					94%	6%	10%
LC	2.5%			2.5%			95%								95%	5%	1.25 %
MRD		10%						90%							90%	10%	15%
SRD					10%				90%						90%	10%	10%
DS						20%				80%					80%	20%	40%
SB											90 %		5%	5%	90%	10%	0%
MRC		5%				10%						85%			85%	15%	5%
SRC												10%	90%		90%	10%	10%
PSR	2%			2%		2%								94%	94%	6%	2%
Average															90%	10%	9.7%

Mild Road: **MR**, Mild Pothole: **MP**, Severe Pothole: **SP**, Mild Manhole: **MM**, Severe Manhole: **SM**, Transverse Cracks: **TC**, Longitudinal Cracks: **LC**, Mild Road Dent: **MRD**, Severe Road Dent: **SRD**, Deceleration Strips: **DS**, Speedbumps: **SB**, Mild Railway Crossing: **MRC**, Severe Railway Crossing: **SRC**, Paved Stone Road: **PSR**

In our opinion, the main reasons for the miss-predicted events occur as different levels of severity might have the same features when sensed with different car models. For example, the mild pothole sensed by a sedan or hatchback car with a relatively old year make and low suspension performance might have the same features as a severe pothole being sensed by a van with large tire sizes and good suspension performance. Additionally, when sensing an event with equivalent car models, the features of the same anomaly may vary based on the driver behavior, way of attendance, and the velocity during the attendance. Even with high qualities of detection, a dynamic road monitoring system is essential to avoid the miss-detection that could occur because of different car models and the further deterioration of roads due to high-traffic, less frequent maintenance, and weather conditions.

D. Geo-Referencing of Road Surface Anomalies

As mentioned earlier, robust monitoring of road surface anomalies cannot be achieved without adequate and continuous localization and positioning. Relying on commercial GPS receivers embedded in smart devices could lead to improper geo-referencing for monitored events. Generally speaking, GPS receivers are vulnerable to outages and multipath specifically in urban canyons and downtown scenarios [33], [34].

To assist with such GPS problems, integrated navigation systems shows high potential to enhance the overall performance of geo-referencing [52], [53]. There are multiple techniques of integrated navigation with GPS such as in inertial navigation systems, visual odometry, radar sensors, LIDAR's, or other approaches such as map-matching [33], [34], [54], [55]. For geo-referencing, we use GPS receivers embedded in smart devices and external positioning units that include GPS receivers, inertial navigation systems, and their integration. To assess the capabilities of GPS receivers by comparing them to integrated navigation systems, Figs. 13–15 show the 2-D position in different driving locations. In these figures, the positioning solutions are obtained by GPS receivers embedded in smart devices, OEM GPS receiver, and the integrated positioning solution obtained by integrating OEM GPS receiver and the inertial navigation systems of the VTI unit and NovAtel SPAN unit. In

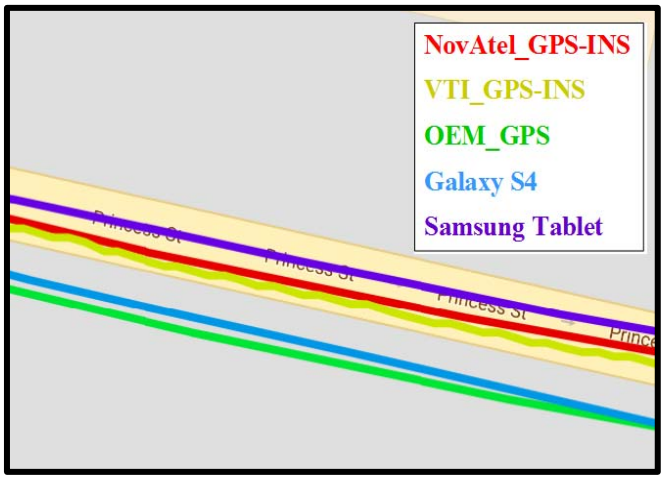


Fig. 13. 2-D position utilizing smart devices GPS receivers, OEM GPS receiver, integrated positioning solutions of VTI unit, and NovAtel SPAN unit, downtown portion 1.

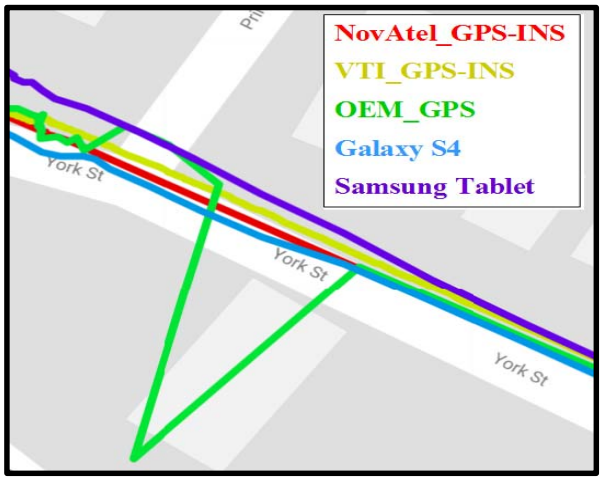


Fig. 15. 2-D position utilizing smart devices GPS receivers, OEM GPS receiver, integrated positioning solutions of VTI unit, and NovAtel SPAN unit, residential area.

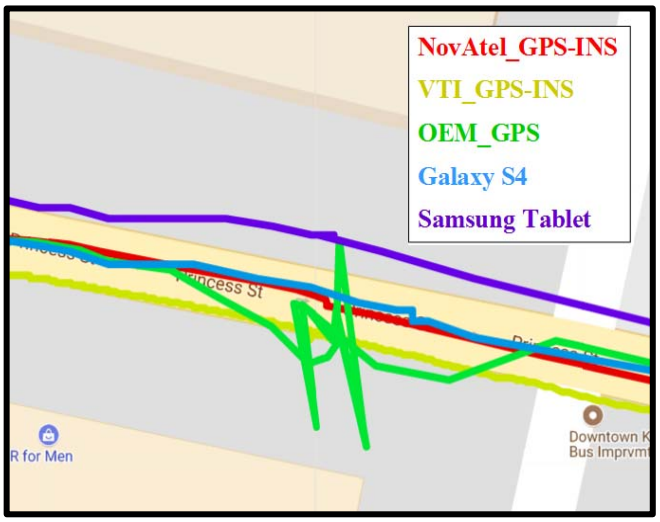


Fig. 14. 2-D position utilizing smart devices GPS receivers, OEM GPS receiver, integrated positioning solutions of VTI unit, and NovAtel SPAN unit, downtown portion 2.

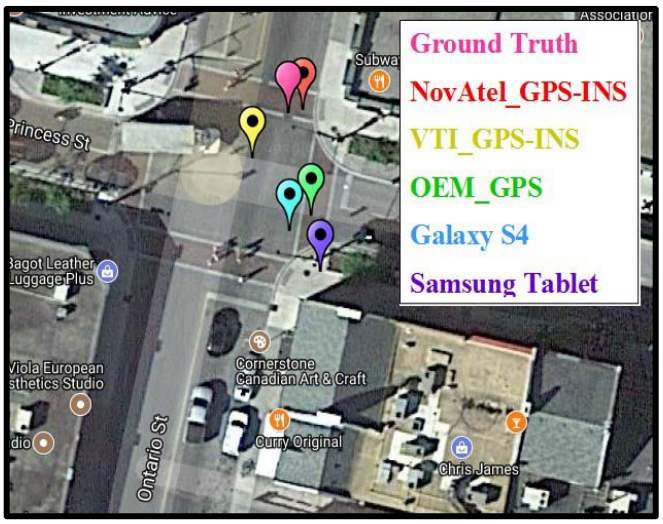


Fig. 16. 2-D road anomaly geo-referencing utilizing smart devices GPS receivers, OEM GPS receiver, integrated positioning solutions of VTI unit, and NovAtel SPAN unit, downtown area.

Fig. 13, through downtown driving scenario, the GPS position obtained by Samsung Galaxy S4 smartphone and OEM GPS receiver showed a continuous drift of the road, while the GPS positioning obtained by the Samsung tablet was on the road, as the positioning solutions obtained by integrated positioning of the VTI unit and NovAtel SPAN unit. On the other hand, in Fig. 14, the GPS position obtained by the smartphone was within the road as the integrated positioning solutions, while the OEM GPS receiver showed high positioning errors and the tablet showed off-road drifted position, these results are within the second portion of downtown driving scenario.

In residential area driving scenario, the positioning solution of the smartphone showed initial drifted solution then started to align with the integrated positioning solutions, while the positioning solution of the tablet showed off-road drift and the OEM GPS receiver displayed high positioning errors during most of the presented portion before aligning with the integrated positioning solutions. In residential area driving scenario, the positioning solution of the smartphone showed initial drifted solution then started to align with the integrated positioning solutions, while the positioning solution of the tablet showed off-road drift and the OEM GPS receiver displayed high positioning errors during most of the presented portion before aligning with the integrated positioning solutions. As a conclusion, Figs. 13–15 showed that relying only on the GPS receivers either the ones embedded on smart devices or the external ones lead to deteriorated positioning accuracy and therefore misleading geo-referencing for the road surface anomalies. Integrated positioning systems exhibited accurate performance even when the GPS solution was drifted or experienced high positioning uncertainties and errors. Such performance of integrated positioning units indicated the importance of their utilization in road surface anomalies monitoring systems. In order to highlight the capabilities of GPS receivers and integrated positioning systems,

Fig. 16 provides the 2-D geo-referencing of a detected road anomaly compared to the ground truth. The integrated geo-referencing obtained by NovAtel and VTI presented the least errors compared to the ground truth. On the other hand, the standalone GPS receivers of the smart devices and the external ones achieved lower geo-referencing accuracy which is not sufficient for robust monitoring of road surface anomalies.

V. Conclusion

The rapid development of the IoT is prompting further growth in crowdsensing applications to provide smart and robust solutions for challenges arising in smart cities. An important aspect of RISs as a smart city application is the monitoring of road surface conditions. We proposed a robust framework for road anomaly detection based on extensive experimental activities held in Kingston, ON, Canada. In such activities, we adopted eight different car models mounted with multiple smart devices and integrated positioning units. The sensed and collected data from these experiments were used to build data sets with various and multiple types of events. All the sensed data were applied to a wavelet packet denoising technique, which enhanced the performance of the feature extraction by decreasing the noise effects. Also, we adopted feature extraction techniques to effectively describe the effects of the road irregularities on the sensed data. These data sets were used to build training and assessment multilevel SVM classifier, and it was able to efficiently detect and classify multiple anomalies with varying levels of severity, achieving an average TPR performance of 90%. We analyzed the capabilities of GPS receiver and integrated positioning systems in the localization of the monitored events, and we highlight the challenges and needs for accurate and acceptable geo-referencing. In order to achieve a dynamic robust system, we suggest continuous monitoring of the road anomalies, building data sets that could be used for further training. For geo-referencing, we discussed various GPS receivers on smart devices along with integrated positioning system units. Results showed the need for integrated positioning units for achieving an appropriate level of accuracy for the geo-referenced events. However, for maintaining sustainable and a high positioning accuracy level, we suggest using integrated positioning units at high data rates for getting better resolution specifically at high vehicle speeds.

References

[1] D. Minoli, K. Sohraby, and B. Occhiogrosso, "IoT considerations, requirements, and architectures for smart buildings—Energy optimization and next-generation building management systems," *IEEE Internet Things J.*, vol. 4, no. 1, pp. 269–283, Feb. 2017.

[2] H. Ma, L. Liu, A. Zhou, and D. Zhao, "On networking of Internet of Things: Explorations and challenges," *IEEE Internet Things J.*, vol. 3, no. 4, pp. 441–452, Aug. 2016.

[3] *Gartner Says 8.4 Billion Connected 'Things' Will Be in Use in 2017, Up 31 Percent From 2016*, Gartner Press Release, Stamford, CT, USA, Feb. 7, 2017. [Online]. Available: http://www.gartner.com/newsroom/id/3598917

[4] K. Sood, S. Yu, and Y. Xiang, "Software-defined wireless networking opportunities and challenges for Internet-of-Things: A review," *IEEE Internet Things J.*, vol. 3, no. 4, pp. 453–463, Aug. 2016.

[5] A. Kamilaris and A. Pitsillides, "Mobile phone computing and the Internet of Things: A survey," *IEEE Internet Things J.*, vol. 3, no. 6, pp. 885–898, Dec. 2016.

[6] R. Girau, S. Martis, and L. Atzori, "Lysis: A platform for IoT distributed applications over socially connected objects," *IEEE Internet Things J.*, vol. 4, no. 1, pp. 40–51, Feb. 2017.

[7] S. M. A. Oteafy and H. S. Hassanein, "Resilient IoT architectures over dynamic sensor networks with adaptive components," *IEEE Internet Things J.*, vol. 4, no. 2, pp. 474–483, Apr. 2017.

[8] A. Gharaibeh *et al.*, "Smart cities: A survey on data management, security, and enabling technologies," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 4, pp. 2456–2501, 4th Quart., 2017.

[9] T. Anagnostopoulos *et al.*, "Challenges and opportunities of waste management in IoT-enabled smart cities: A survey," *IEEE Trans. Sustain. Comput.*, vol. 2, no. 3, pp. 275–289, Jul./Sep. 2017.

[10] M. Handte, S. Foell, S. Wagner, G. Kortuem, and P. J. Marrón, "An Internet-of-Things enabled connected navigation system for urban bus riders," *IEEE Internet Things J.*, vol. 3, no. 5, pp. 735–744, Oct. 2016.

[11] H. Menouar *et al.*, "UAV-enabled intelligent transportation systems for the smart city: Applications and challenges," *IEEE Commun. Mag.*, vol. 55, no. 3, pp. 22–28, Mar. 2017.

[12] M. A. Rahman and M. S. Hossain, "A location-based mobile crowdsensing framework supporting a massive ad Hoc social network environment," *IEEE Commun. Mag.*, vol. 55, no. 3, pp. 76–85, Mar. 2017.

[13] E. Wang, Y. Yang, J. Wu, W. Liu, and X. Wang, "An efficient prediction-based user recruitment for mobile crowdsensing," *IEEE Trans. Mobile Comput.*, vol. 17, no. 1, pp. 16–28, Jan. 2018.

[14] X. Wang, X. Zheng, Q. Zhang, T. Wang, and D. Shen, "Crowdsourcing in ITS: The state of the work and the networking," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 6, pp. 1596–1605, Jun. 2016.

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

[16] C. Miyajima and K. Takeda, "Driver-behavior modeling using on-road driving data: A new application for behavior signal processing," *IEEE Signal Process. Mag.*, vol. 33, no. 6, pp. 14–21, Nov. 2016.

[17] Z. Liu, S. Jiang, P. Zhou, and M. Li, "A participatory urban traffic monitoring system: The power of bus riders," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 10, pp. 2851–2864, Oct. 2017.

[18] X. Fan, J. Liu, Z. Wang, Y. Jiang, and X. Liu, "Crowdsourced road navigation: Concept, design, and implementation," *IEEE Commun. Mag.*, vol. 55, no. 6, pp. 126–128, Jun. 2017.

[19] *Potholes 'Talk of the Canteen*,' BBC News, London, U.K., Mar. 28, 2017. [Online]. Available: http://www.bbc.com/news/business-39407387

[20] Business Insider. (Jan. 11, 2017). *Here's What the 'Pothole Tax' Costs the Average American Driver Every Year.* [Online]. Available: http://www.businessinsider.com/road-maintenance-us-costs-trump-2017-1

[21] City of Toronto. (2017). *Potholes.* [Online]. Available: https://www1.toronto.ca/wps/portal/contentonly?vgnextoid=db67a84c9f6e1410VgnVCM10000071d60f89RCRD&vgnextchannel=9afb4074781e1410VgnVCM10000071d60f89RCRD

[22] *Pothole Damage Leads to Booming Business for Auto Repair Shops This Spring*, CBC News Bus., Toronto, ON, Canada, Mar. 28, 2016. [Online]. Available: http://www.cbc.ca/news/business/pothole-damage-leads-to-booming-business-for-auto-repair-shops-this-spring-1.3509375

[23] Transport Canada-Road Safety. (2011). *NCDB*. [Online]. Available: http://www.tc.gc.ca/eng/motorvehiclesafety/tp-tp15145-1201.html

[24] The Telegraph. (Jan. 26, 2015). *Potholes Cause Damage to Cars Every 11 Minutes*, *Figures Show.* [Online]. Available: http://www.telegraph.co.uk/news/uknews/road-and-rail-transport/11368326/Potholes-cause-damage-to-cars-every-11-mins-figures-show.html

[25] Street Bump. (2017). *About Street Bump*. [Online]. Available: http://www.streetbump.org/about

[26] CNN. (Feb. 16, 2012). *Street Bump App Detects Potholes*, *Tells City Officials*. [Online]. Available: http://www.cnn.com/2012/02/16/tech/street-bump-app-detects-potholes-tells-city-officials/index.html

[27] T. S. Brisimi, C. G. Cassandras, C. Osgood, I. C. Paschalidis, and Y. Zhang, "Sensing and classifying roadway obstacles in smart cities: The street bump system," *IEEE Access*, vol. 4, pp. 1301–1312, 2016.

[28] F. Kalim, J. P. Jeong, and M. U. Ilyas, "CRATER: A crowd sensing application to estimate road conditions," *IEEE Access*, vol. 4, pp. 8317–8326, 2016.

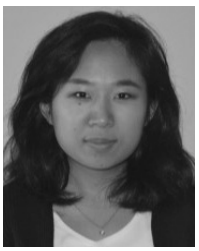
[29] K. Chen, G. Tan, M. Lu, and J. Wu, "CRSM: A practical crowdsourcing-based road surface monitoring system," *Wireless Netw.*, vol. 22, no. 3, pp. 765–779, 2016.

[30] G. Xue *et al.*, "Pothole in the dark: Perceiving pothole profiles with participatory urban vehicles," *IEEE Trans. Mobile Comput.*, vol. 16, no. 5, pp. 1408–1419, May 2017.
[31] Z. Li *et al.*, "Optimal state estimation for systems driven by jump–diffusion process with application to road anomaly detection," *IEEE Trans. Control Syst. Technol.*, vol. 25, no. 5, pp. 1634–1643, Sep. 2017.
[32] A. Fox, B. V. K. V. Kumar, J. Chen, and F. Bai, "Multi-lane pothole detection from crowdsourced undersampled vehicle sensor data," *IEEE Trans. Mobile Comput.*, vol. 16, no. 12, pp. 3417–3430, Dec. 2017.
[33] A. Noureldin, T. B. Karamat, and J. Georgy, *Fundamentals of Inertial Navigation*, *Satellite-Based Positioning and Their Integration*. Heidelberg, Germany: Springer-Verlag, 2013.
[34] A. Noureldin, T. B. Karamat, M. D. Eberts, and A. El-Shafie, "Performance enhancement of MEMS-based INS/GPS integration for low-cost navigation applications," *IEEE Trans. Veh. Technol.*, vol. 58, no. 3, pp. 1077–1096, Mar. 2009.
[35] N. El-Sheimy, S. Nassar, and A. Noureldin, "Wavelet de-noising for IMU alignment," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 19, no. 10, pp. 32–39, Oct. 2004.
[36] J. Skaloud, A. M. Bruton, and K. P. Schwarz, "Detection and filtering of short-term (1/f) noise in inertial sensors," *Navigation*, vol. 46, no. 2, pp. 97–107, 1999.
[37] M. Jafari, T. A. Najafabadi, B. Moshiri, S. S. Tabatabaei, and M. Sahebjameyan, "PEM stochastic modeling for MEMS inertial sensors in conventional and redundant IMUs," *IEEE Sensors J.*, vol. 14, no. 6, pp. 2019–2027, Jun. 2014.
[38] W. Abd-Elhamid, A. Osman, A. Noureldin, and N. El-Sheimy, "Wavelet multi-resolution analysis for enhancing the performance of integrated GPS and MEMS-based navigation system," *Geomatica*, vol. 59, no. 1, pp. 61–72, Mar. 2005.
[39] E. I. Laftchiev, C. M. Lagoa, and S. N. Brennan, "Vehicle localization using in-vehicle pitch data and dynamical models," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 1, pp. 206–220, Feb. 2015.
[40] S. Mallat, *A Wavelet Tour of Signal Processing: The Sparse Way*, 3rd ed. New York, NY, USA: Academic Press, 2008.
[41] S. Mitra, *Digital Signal Processing—A Computer Based Approach*, 4th ed. New York, NY, USA: McGraw-Hill, 2011.
[42] L. Lau, "Wavelet packets based denoising method for measurement domain repeat-time multipath filtering in GPS static high-precision positioning," *GPS Solutions*, vol. 21, no. 2, pp. 461–474, Apr. 2017.
[43] C. K. Chui, *An Introduction to Wavelets: Wavelet Analysis and Its Applications*. Boston, MA, USA: Academic Press, 1992.
[44] C. M. Stein, "Estimation of the mean of a multivariate normal distribution," *Ann. Stat.*, vol. 9, no. 6, pp. 1135–1151, 1981.
[45] M. Elhoushi, J. Georgy, A. Noureldin, and M. J. Korenberg, "Motion mode recognition for indoor pedestrian navigation using portable devices," *IEEE Trans. Instrum. Meas.*, vol. 65, no. 1, pp. 208–221, Jan. 2016.
[46] M. Elhoushi, J. Georgy, A. Noureldin, and M. J. Korenberg, "A survey on approaches of motion mode recognition using sensors," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 7, pp. 1662–1686, Jul. 2017.
[47] M. Jaworski, P. Duda, and L. Rutkowski, "New splitting criteria for decision trees in stationary data streams," *IEEE Trans. Neural Netw. Learn. Syst.*, to be published.
[48] A. S. El-Wakeel, J. Li, M. T. Rahman, A. Noureldin, and H. S. Hassanein, "Monitoring road surface anomalies towards dynamic road mapping for future smart cities," in *Proc. IEEE Glob. Conf. Signal Inf. Process. (GlobalSIP)*, Montreal, QC, Canada, Nov. 2017, pp. 828–832.
[49] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, Sep. 1995.
[50] Y. Ying and P. Li, "Distance metric learning with eigenvalue optimization," *J. Mach. Learn. Res.*, vol. 13, no. 1, pp. 1–26, 2012.
[51] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining*, *Inference*, *and Prediction*, 2nd ed. New York, NY, USA: Springer, 2009.
[52] G. Xu, S. Gao, M. Daneshmand, C. Wang, and Y. Liu, "A survey for mobility big data analytics for geolocation prediction," *IEEE Wireless Commun.*, vol. 24, no. 1, pp. 111–119, Feb. 2017.
[53] A. S. El-Wakeel, A. Osman, N. Zorba, A. Noureldin, and H. S. Hassanein, "Robust geo-referencing for road information services in urban environments," *IEEE Trans. Intell. Transp. Syst.*, submitted for publication.
[54] D. Scaramuzza and F. Fraundorfer, "Visual odometry [tutorial]," *IEEE Robot. Autom. Mag.*, vol. 18, no. 4, pp. 80–92, Dec. 2011.
[55] A. S. El-Wakeel, A. Osman, A. Noureldin, and H. S. Hassanein, "Road test experiments and statistical analysis for real-time monitoring of road surface conditions," in *Proc. IEEE Glob. Commun. Conf. (GLOBECOM)*, Singapore, Dec. 2017, pp. 1–6.

Amr S. El-Wakeel (S'09) received the B.Sc. (Hons.) and M.Sc. degrees in electronics and communications engineering (ECE) from the Arab Academy for Science, Technology and Maritime Transport (AASTMT), Cairo, Egypt, in 2008 and 2013, respectively. He is currently pursuing the Ph.D. degree in electrical and computer engineering at Queen's University, Kingston, ON, Canada.

He was an Assistant Lecturer with the Department of ECE, AASTMT, from 2008 to 2014. He is currently a Research Assistant with the NavINST and TRL Research Laboratories, Queen's University. His research work appeared in IEEE flagship conferences and top tier journals and OSA journals. His current research interests include signal processing, mobile sensing, and wireless communication systems.

Mr. El-Wakeel serves as a TPC member for IEEE flagship conferences and a Technical Reviewer for IEEE, OSA, Elsevier, and Wiley journals.



Jin Li (S'17) is currently pursuing the electrical engineering degree at Queen's University, Kingston, ON, Canada.

She was with the NavINST Laboratory, Electrical Engineering Department, Royal Military College of Canada, Kingston, ON. Her summer project focused on road surface condition monitoring and software applications development. She has co-authored papers in the IEEE Globalsip Conference. Her current research interests include signal processing, machine learning, and communications.

Ms. Li was a recipient of the NSERC Undergraduate Research Assistant Award in 2017.

Aboelmagd Noureldin (S'98–M'02–SM'08) received the B.Sc. degree in electrical engineering and M.Sc. degree in engineering physics from Cairo University, Cairo, Egypt, in 1993 and 1997, respectively, and the Ph.D. degree in electrical and computer engineering from the University of Calgary, Calgary, AB, Canada, in 2002.

He is a Professor with the Department of Electrical and Computer Engineering, Royal Military College of Canada (RMCC), Kingston, ON, Canada, with a cross-appointment with the School of Computing and the Department of Electrical and Computer Engineering, Queen's University. He is also the Founder and the Director of the Navigation and Instrumentation Research Group, RMCC. He has authored or co-authored over 230 papers in journals and conference proceedings. His research work led to ten patents in the area of position, location, and navigation systems. His current research interests include GPS, wireless location and navigation, indoor positioning, and multisensor fusion.

Hossam S. Hassanein (S'86–M'90–SM'05–F'17) received the Ph.D. degree in computing science from the University of Alberta, Edmonton, AB, Canada, in 1990.

He is the Founder and the Director of the Telecommunications Research Laboratory, School of Computing, Queen's University, Kingston, ON, Canada, with extensive international academic and industrial collaborations. He is a leading authority in the areas of broadband, wireless and mobile networks architecture, protocols, control, and performance evaluation. His record spans over 500 publications in journals, conferences, and book chapters, in addition to numerous keynotes and plenary talks at flagship venues.

Dr. Hassanein was a recipient of several recognitions and Best Papers Awards at top international conferences and has led a number of symposia in ComSoc conferences. He is an IEEE Communications Society Distinguished Speaker (Distinguished Lecturer from 2008 to 2010). He is the Past Chair of the IEEE Communication Society Technical Committee on Ad Hoc and Sensor Networks.

Nizar Zorba (S'02–M'08) received the B.Sc. degree in electrical engineering from JUST University, Irbid, Jordan, in 2002, the M.Sc. degree in data communications and the M.B.A. degree from the University of Zaragoza, Zaragoza, Spain, in 2004 and 2005, respectively, and the Ph.D. degree in signal processing for communications from UPC Barcelona, Barcelona, Spain, in 2007.

He is an Associate Professor with the Electrical Engineering Department, Qatar University, Doha, Qatar. He has authored or co-authored over 100 publications in journals and conference proceedings and holds 5 patents. His current research interests include 5G networks optimization, demand-response in smart grids, and crowd management.

Dr. Zorba is currently the Secretary of the IEEE ComSoc Communication Systems Integration and Modeling Technical Committee.