



EPiC Series in Computing

Volume 58, 2019, Pages 454–463

Proceedings of 34th International Conference on Computers and Their Applications



Data Driven Analytics of Road Quality

Ali Rahim Taleqani, Raj Bridgelall, Jill Hough, Kendall E. Nygard
North Dakota State University, Fargo, North Dakota, USA
(ali.rahimtaleqani, raj.bridgelall, jill.hough,
kendall.nygard)@ndsu.edu

Abstract

There is a lack of research into the impact of road roughness on ride quality and route choice. The scarcity of ride roughness data for local and urban roads is likely one reason for the lack of such studies. Existing methods of obtaining ride roughness data are expensive and require expert practitioners and laborious data processing by trained personnel. Sensors in most current vehicles provide an alternative source for road roughness data. This study emulated the data needed from vehicle sensors by using the accelerometer and gyroscope of a smartphone. The authors used data collected from two different bus routes to classify segments of roads into objectively distinct roughness clusters. The output enables map service applications to suggest better routing options based on expected ride quality and also quantifies road roughness consistently to enable optimized maintenance planning and decision-making for roadway assets.

1 Introduction

Seasonal changes in weather and temperature cycles cause potholes to form when moisture in small holes and cracks expand and contract. According to a recent study by the American Automobile Association, potholes in the United States caused 15 billion dollars in vehicle repairs between 2011 and 2015 [1]. The damages range from tire punctures and bent wheels to more expensive suspension damage. The problem is not just limited to the United States. Studies in 2017 suggested that there has been an increase in pothole-related breakdowns in the U.K. as well [2].

Researchers developed several pothole detection systems during the past decade. The detection methods vary in either their data collection or computational approach. The primary method of detecting potholes relies on reports from road users. However, this method is inefficient because of the small portion of potholes reported, and the biased estimation of their size and severity [3]. Traditional approaches to pothole detection are time-consuming, inefficient with heavy traffic flow, and subjective [4], [5].

2 Related Works

Jaguar Land Rover developed a model to detect, predict and share data on potholes with other vehicles and road authorities. The technology will allow vehicles to gather data about the location and severity of potholes, broken drains, and manhole covers. Sharing the data will enable speed advisories for individual drivers and alert road authorities so that they can prioritize repairs [6]. Google patented a technology to assess road quality conditions using GPS and vehicle sensors, enabling the company to form a database of roads in need of repair [7]. Pavement condition assessment primarily has three tasks: 1) raw data collection; 2) distress detection; and, 3) distress assessment [8]. Li et al. proposed a method based on two-dimensional images and Ground-Penetrating Radar (GPR) data to automate efficient pothole detection. The model took advantage of data fusion by retrieving data from multiple sensors [9]. Several other authors applied a similar approach by using various sensors such as two-dimensional (2D) video and thermal imaging, three-dimensional (3D) laser scanning, global position system (GPS), and GPR [10]–[12]. Detection techniques fall into three main categories: 2D video/image-based method, 3D video/image-based methods, and hybrid approaches. Among two-dimensional (2D) methods, the fuzzy logic-based algorithm analyzes the brightness membership function of the gray image level to quantify pavement cracking [13]. Grid cell analysis divides the images into several cells and use several image processing algorithms to locate the cracks [4]. Other methods include crack geometrical or morphological features [14], multi-scale analysis functions such as wavelet transforms [15]–[17], beamlet transform-based approach [18], and artificial intelligence methods such as neural network [19] and genetic programming [20]. One of the prominent issues with 2D pavement crack detection is that it cannot distinguish dark areas from tire marks, oil spills, shadows or repairs from real pavement cracks [21]–[23].

Many authors proposed three-dimensional (3D) techniques to address the disadvantages of 2D gray-level image analysis methods [24]. They use stereo vision and image registration to acquire 3D characteristics of the pavement [25], [26]. This method exhibits the lowest cost/performance ratio. However, such robustness comes with the challenge of achieving real-time processing capability. 3D laser scanning methods have been widely used because they are simple to implement and are able to discriminate real cracks from image interferences [27], [28]. Unfortunately, the 3D laser scanning approach could not reflect measurement details when the cracks have no noticeable depth changes, for example, when cracks are filled with sand. Despite their strong performance, the laser-based methods are costly, sensitive to natural or artificial lighting and weather conditions, the shape of the object, distance and other anomalies. The stereovision-based techniques are very computationally complex. Additionally, such sensors must have low power requirements and have limited buffer and storage capabilities. Fox et al. developed a crowd-sourcing platform to overcome these challenges. They aggregated sensor data from multiple vehicles to detect potholes. They designed and evaluated three crowd-sourcing schemes that involved vehicles and the Cloud to implement the detection system [29]–[31]. Cekmis used the fuzzy logic method to evaluate ride quality of bicycling given three factors: bicycle facility, slope, and vehicle flow [32]. Astarita et al. developed an application to monitor road surface quality and to detect road bumps and potholes using sensor signals. Their study focused on providing real-time information to road users similar to vision-based methods [33].

One of the challenges encountered with the methods described above is the absence of a reliable, cost-effective, and proactive method of detecting potholes. The contribution of this research is a proactive model for detecting the locations of roadway anomalies and their relative severity based on the roughness intensity they generate. The proposed solution is to gather sensor data from multiple vehicle traversals and to cluster the roughness of road segments into some number of classes. In general,

subjective opinions about the relative level of ride quality across individual segments of a route are not available. To fill this gap, the model developed provides an objective classification to rank the ride quality anticipated from traveling a given route. A goal of the modeling is to develop an objective method of ranking the relative roughness among road segments. The model quantifies the roughness of each cluster by using the road impact factor (RIF) transform [34]. The RIF transform produces RIF indices, which are spatial and statistical characterizations of road surface unevenness due to anomalies such as potholes, cracks, joints, and utility covers. Connected vehicles can upload inertial and GPS data to a cloud-based system that computes the RIF to identify and classify anomalies. The cloud service can subsequently deliver the output of the roughness classification model to onboard route planning systems that can help a driver choose routes with the best road quality. The benefit would be improved ride quality, lower vehicle operating costs, and the elimination of possible unsafe driver reactions to avoid anomalies such as potholes [35].

3 Methodology

Clustering algorithms identify similar objects in a spatial database. To support non-flat geometry clustering and uneven cluster size, this research used density-based spatial clustering of applications with noise (DBSCAN). DBSCAN address three primary issues in clustering methods: minimal requirements of domain knowledge to determine input parameters, the discovery of clusters with arbitrary shape, and good efficiency on large databases [36], [37]. Moreover, DBSCAN is a deterministic algorithm, meaning that it always generates the same clusters when given the same data in the same order. Unlike many other clustering methods, DBSCAN does not require a predetermined number of clusters. The algorithm needs three parameters: 1) the neighbor list size, 2) the radius that delimitates the neighborhood area of a point (eps), and 3) the minimum number of points that must exist in the eps-neighborhood.

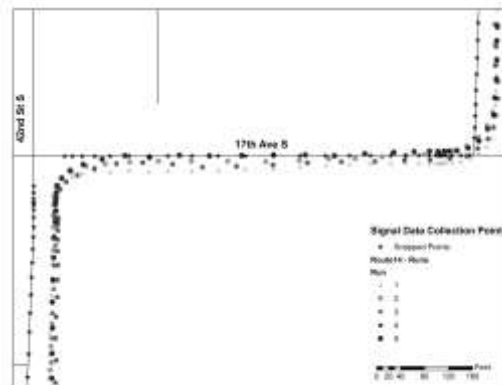


Figure 1: Sensor signal collection points along a segment of Route 14.

The RIF transform takes the product of any acceleration signal component or a composite of acceleration signals combined from any number of accelerometer directions $g(t)$, and the longitudinal velocity $v(t)$ from a speed sensor as follows:

$$R_v^L = \sqrt{\frac{1}{L} \int_0^L |g(t) \cdot v(t)|^2 dt}$$

The RIF-index, R_v^L , is interpreted as the average g-force magnitude experienced per unit of distance L traversed [34]. The transform then accumulates the square of that product across the segment length analyzed. Dividing the accumulated square-of-products by the segment length and then taking the square root yields the RIF-index. The RIF-transform computes a RIF-index for road segments of pre-determined fixed lengths by windowing the raw accelerometer signal. For every window size, the classifiers will check the number of clusters as well as a silhouette coefficient to evaluate the consistency within clusters.

Cluster Analysis. The silhouette is an interpretation and validation method for evaluating the consistency within clusters of data. This method compares the average distance to elements in the same cluster with the average distance to elements in other clusters. In other words, the silhouette measures how similar an element is to its cluster (cohesion) relative to other clusters (separation). The coefficient ranges from 1 to +1. A high silhouette value indicates that the object is appropriately matched to its assigned cluster, and not well matched to the neighboring clusters. Values near 0 indicate that the object is on the border of two clusters. Negative values indicate that an object may be an outlier for the assigned cluster, or that the number of clusters for the data is inappropriate.

Data. The case study used the dataset from previous research [38]. A data collection application called PAVVET installed on an iPhone® 4S collected several sensor signals for two transit routes of distinctly different levels of roughness in Fargo, ND. The signals logged were from the device accelerometer, GPS, and velocity sensors. The accelerometer produced the G_x , G_y , and G_z inertial signals, which are the acceleration levels sensed in the vertical, lateral, and longitudinal directions, respectively, all normalized to $9.81 \frac{m}{s^2}$ to produce the units in g-force. The GPS receiver produced the latitude and longitude in decimal format. The GPS receiver also produced the estimated ground speed in $\frac{m}{s}$. The sample rate for the inertial and gyroscopic signals were the temporal rate that each row of the data repeated. It varied between 64 Hz and 150 Hz, depending on the user setting on the application at the time that the data was collected.

Since the iOS® constrained the rate of the GPS output (latitude, longitude, speed) to approximately 1 Hz, it took some time for the GPS to lock and produce valid coordinates for some portions of the data sets. Therefore, data cleaning was necessary to remove those observations. The GPS positions registered for each traversal of a segment varied due to the asynchronous nature of the GPS updates, as well as the errors in position registration. To minimize errors, the authors used ArcGIS® 5.0 to snap GPS positions of each signal to the nearest corresponding road segment (1). To simplify the illustration, the Figure 1 shows data from five traversals, and the final snapped data points used in this study. For consistency, the method used the coordinates of snapped data points for clustering, and then later for ground truth validation. The traversal data were of different lengths, based on when the device triggered data logging. Hence, a GIS platform helped to approximately equalize the length of each traversal by truncating samples outside of the minimum distance traversal.

Feature Extraction. The feature extraction window of the RIF transform should ideally begin at the same spatial position of each signal. This is because the method involves a machine learning technique to extract roughness features that characterize spatial properties of the road independent of the temporal movements (speed changes) of the vehicle. Therefore, to find the approximate same position of each signal, the algorithm must find a common starting and ending position where all the signals overlap maximally. All data outside of these areas are truncated. A new column called distance is added and then the first sample of the truncated signals is marked as distance zero. The distance from the first sample to the last sample is interpolated based on the speed and sample period such that,

$$D_n = D_{n-1} + V_{n-1} \times (t_n - t_{n-1}) \times \frac{1}{1000}$$

where n is a data point in the dataset and t is time in milliseconds. The formula shows that the change in distance is the product of the speed and the change in time (sample period). The division by 1000 changes the time units to seconds. The attribute used to express roughness in three linear dimensions is:

$$G_T = \sqrt{G_x^2 + G_y^2 + G_z^2}$$

The offsets from each of the component signals G_x , G_y , and G_z is removed before applying the RIF transform. Finally, several features are extracted from GT including mean (Gt_Mean), median absolute deviation (Gt_MAD), first quantile (Gt_Quantile25), third quantile (Gt_Quantile75), energy (Gt_Energy), spectral entropy (Gt_SpectEntr), average RIF (RIF_Mean), and RIF variance (RIFSTD) across all windows. For example, the Gt_Mean is the mean of all G_T signals within a window of size 25 cm. After extracting the features for each traversal, the model computed their ensemble average across traversals, within the same window. Using a log transform to remove the data skewness and rescaling all features using the Min-Max method normalized the data before applying the clustering algorithm.

The Scikit-learn [39] library package provided the algorithms to perform all of the clustering. The DBSCAN method applied to different combinations of features extracted from four different window sizes (25 cm, 50 cm, 100 cm, 150 cm) identified features that provided the highest silhouette coefficient. The rest of the paper is organized as follows: The next section gives an overview of the methodology for clustering the RIF indices. Section 3 presents the data and describes the data preprocessing. Section 4 presents the results and findings. The final section provides concluding comments.

4 Experimental Results

Given all features, Table 1 and summarizes the clusters and their proportion of data points for each window size. The noise is interpreted as algorithm error where it could not classify the data points into any clusters. They are very small and can be ignored in the analysis. Given the model with all-features, for almost all window size, there is only one dominant cluster for both routes. Hence, given all features, the model could not distinguish segment roughness into more than one cluster for any of the window sizes. In other words, by using all extracted features, the model could not sufficiently distinguish smooth roadway segments from those that contain anomalies. Furthermore, using all of the extracted features did not generate the highest silhouette values (Figure 2).

Window	Route 14				Route 15			
	Noise	Class 0	Class 1	Remaining	Noise	Class 0	Class 1	Remaining
25 cm	0.81%	6.68%	92.45%	0.06%	2.52%	17.67%	79.30%	0.51%
50 cm	1.71%	28.47%	69.76%	0.06%	6.16%	31.58%	47.44%	14.83%
75 cm	3.32%	36.67%	58.01%	0.00%	10.59%	8.32%	75.70%	5.39%
100 cm	4.71%	13.73%	81.18%	0.38%	12.77%	0.59%	86.43%	0.21%
125 cm	4.92%	0.52%	94.47%	0.10%	12.76%	0.42%	86.08%	0.74%
150 cm	5.06%	0.12%	94.58%	0.24%	14.00%	0.95%	84.68%	0.38%

Table 1: Share of each cluster of route 14 and 15 (all features).

Figure 2 reveals a critical fact about the model: the silhouette coefficient generally decreases as window size increases. In general, most window sizes result in only two clusters. Another overall pattern observed is that, as the window size increases, the method produces less balanced clusters.

Further investigation identified two features that produced the highest silhouette coefficient: 1) G_T Spectral Entropy; 2) Variance of RIF-indices, as shown in Figure 2. Hence, further analysis used only these two features. The entropy is the measure of the amount of disorder in a system. The spectral entropy shows differences in the energy distribution of a signal.

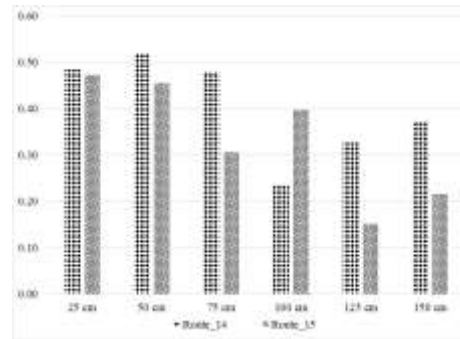


Figure 2: Silhouette coefficient for route 14 and 15.

In particular, a high entropy value indicates the existence of a constant similarity of energy (small variations) and a low entropy value indicates the presence of information such as road irregularities. The RIF is a statistical characterization of the quality of the ride. It is proportional to the average energy of the overall accelerations experienced per unit of longitudinal distance when traveling a road segment at a specified speed. Hence, the RIF-index of the G_T signal is proportional to the overall ride roughness experienced. In particular, relatively high values of RIF-indices, such as above 0.2, indicates that pavement unevenness is contributing to poor ride quality. Considering the window size and silhouette coefficient, as shown in Table 2, the window size of 50 cm contributes the highest average silhouette coefficient and the lowest standard deviation among other window sizes. The class membership is more balanced for a 50-cm window size. Table 3 summarizes the clusters and their proportion of data points for each window size for a model with two feature.

Window Size	Mean	Standard Deviation
25 cm	0.478	0.18
50 cm	0.487	0.16
75 cm	0.393	0.17
100 cm	0.317	0.29
125 cm	0.238	0.25
150 cm	0.294	0.31

Table 2: Average and standard deviation of silhouette coefficient for route 14 and 15 across window

Given the difference between two classes, as shown in Table 4 the mean of RIF-indices for Class 1 are higher for Route 14. Moreover, the variation between two classes for Route 15 is less than Route 14. The fact that the mean RIF index of Class 1 segments is consistently higher than that of Class 0, and that the Class 1 membership proportion is much higher for Route 14 than for Route 15 suggests that Route 14 produces lower ride quality than Route 15. The deduction that Route 14 is rougher than Route 15 is consistent with the finding that Class 1 has lower spectral entropy.

Window Size	Route 14			Route 15		
	Noise	Class 0	Class 1	Noise	Class 0	Class 1
25 cm	n/a	0.0183	0.3733	0.00	0.0527	0.3331
50 cm	0.0001	0.1163	0.3752	n/a	0.2078	0.2890

Table 4: The weighted average for RIF-Mean for each cluster (two features).

Window Size	Route 14			Route 15		
	Noise	Class 0	Class 1	Noise	Class 0	Class 1
25 cm	0.00%	6.96%	93.04%	0.01%	19.24%	80.75%
50 cm	0.01%	29.34%	70.65%	n/a	50.57%	49.43%

Table 3: Number of clusters and share of each cluster of route 14 and 15 (two features).

Figure 4 and Figure 3 illustrate the 25 cm and 50 cm window size cluster for Routes 14 and 15, respectively. As shown in these figures, the two output clusters are easily distinguishable by the entropy of the signal and the variance of the RIF-indices. The figures also graphically illustrate that the RIF indices for Class 1 segments distribute with both higher mean and higher variance than those of Class 0 segments. This explanation is consistent with the findings that Class 1 has lower spectral entropy.

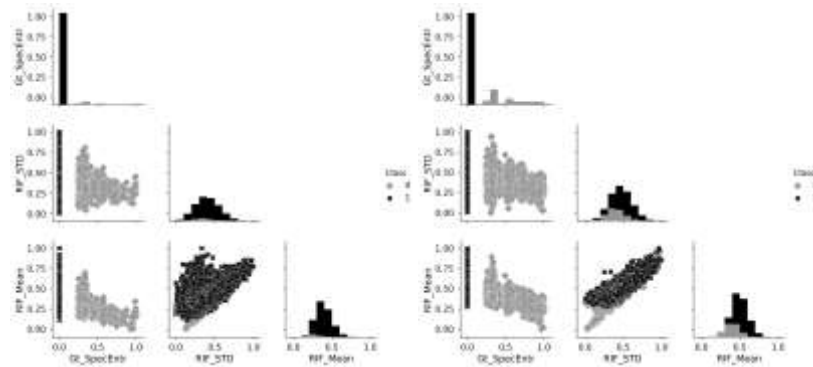


Figure 4: Pairwise comparison of feature of route 14 win size = 25 cm (left) 50 cm (right).

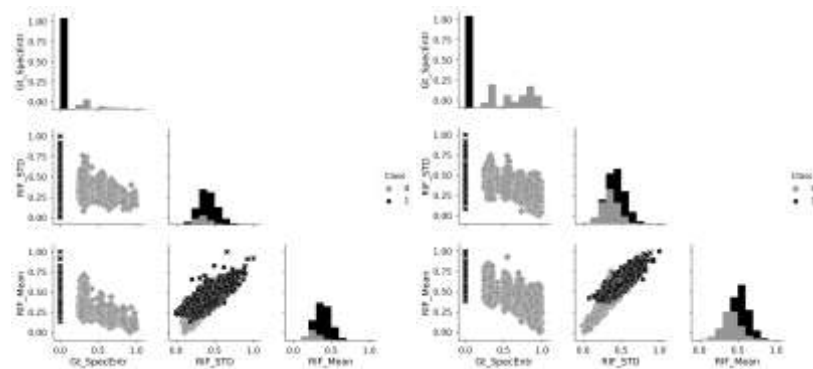


Figure 3: Pairwise comparison of feature of route 15 win size = 25 cm (left) 50 cm (right).

Google Street Views from 2015 provided ground-truth validation. Figure 6 and Figure 5 show top and side views of Class 1 segments on Routes 14 and 15, respectively. These segments exhibit crocodile

cracking which leads to higher values of RIF-indices, and hence low ride quality. In general, the specific contributions of the model are: 1) it could objectively cluster the ride quality if different routes and 2) given a specific route, it could differentiate the quality of road segments (lanes) to find the best possible ride quality.

5 Conclusion

This work contributed a new paradigm for ranking the relative ride quality of road segments based only on objective measures. The findings provide new inputs for route selection based on the expected ride quality during trip planning. Mapping service companies such as Google LLC, HERE Technologies, and TomTom NV could incorporate this model to include ride quality as another factor in trip advisories, and to guide future autonomous and connected vehicles. The benefit would be lower

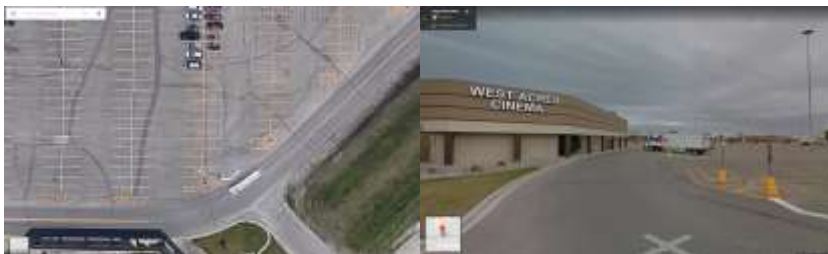


Figure 5: Google street view route 15.

maintenance costs because of improved ride quality by consistently selecting routes with the greatest proportion of Class 0 segments. The roughness classification problem identified the spectral entropy of the composite accelerometer signal and the variance of RIF-indices as features to objectively classify road segments into distinct levels of roughness, obviating the need to rely on subjective roughness opinion surveys. The model identified the best window size for objective roughness classification. The general result was a classification of road segments into two roughness classes 1) among different routes, and 2) within a specific route. Agencies can use the model with any mapping application to reduce road maintenance costs by optimizing the timing and location of remediation measures based on the relative percentages of anomalies found in the roughest class. Future research will examine the potential impacts of expected ride quality on optimization models for route selection that also involve travel distance and travel time.



Figure 6: Google street view route 14.

References

- [1] “Pothole Damage Costs Drivers \$3 Billion Annually Nationwide | AAA SoCal,” 2016. [Online]. Available: <http://news.aaa-calif.com/news/pothole-damage-costs-drivers-3-billion-annually-nationwide>. [Accessed: 06-Feb-2018].
- [2] “UK ‘going backwards’ on pothole crisis - The Transport Network,” 2018. [Online]. Available: <https://www.transport-network.co.uk/UK-%27going-backwards%27-on-pothole-crisis/14775>. [Accessed: 06-Feb-2018].
- [3] “Reporting Potholes and Other Roadside Issues - Potholes - MnDOT.” [Online]. Available: <https://www.dot.state.mn.us/information/potholes/reporting.html>. [Accessed: 06-Feb-2018].
- [4] Y. Huang, B. Xu, Yaxiong Huang, and Bugao Xu, “Automatic inspection of pavement cracking distress,” *J. Electron. Imaging*, vol. 15, no. 1, p. 13017, Jan. 2006.
- [5] K. C. P. Wang and O. Smadi, “Automated imaging technologies for pavement distress surveys,” *Transp. Res. E-Circular*, no. E-C156, 2011.
- [6] “Pothole Detection and Warning System - Jaguar Land Rover.” [Online]. Available: <https://www.landrover.com/experiences/news/pothole-detection.html>. [Accessed: 11-Feb-2018].
- [7] “Google Patents New Pothole Detection System - Motor Trend WOT.” [Online]. Available: <http://www.motortrend.com/news/google-patents-new-pothole-detection-system/>. [Accessed: 11-Feb-2018].
- [8] C. Koch and I. Brilakis, “Pothole detection in asphalt pavement images,” *Adv. Eng. Informatics*, vol. 25, no. 3, pp. 507–515, 2011.
- [9] S. Li, C. Yuan, D. Liu, and H. Cai, “Integrated Processing of Image and GPR Data for Automated Pothole Detection,” *J. Comput. Civ. Eng.*, vol. 30, no. 6, p. 4016015, 2016.
- [10] R. Birken, G. Schirner, and M. Wang, “VOTERS: design of a mobile multi-modal multi-sensor system,” *Proc. Sixth Int. Work. Knowl. Discov. from Sens. Data - SensorKDD '12*, pp. 8–15, 2012.
- [11] J. Huang, W. Liu, and X. Sun, “A pavement crack detection method combining 2D with 3D information based on dempster-shafer theory,” *Comput. Civ. Infrastruct. Eng.*, vol. 29, no. 4, pp. 299–313, 2014.
- [12] M. Solla, S. Lagüela, H. González-Jorge, and P. Arias, “Approach to identify cracking in asphalt pavement using GPR and infrared thermographic methods: Preliminary findings,” *NDT E Int.*, vol. 62, pp. 55–65, 2014.
- [13] H. D. Cheng, J.-R. Chen, C. Glazier, and Y. G. Hu, “Novel approach to pavement cracking detection based on fuzzy set theory,” *J. Comput. Civ. Eng.*, vol. 13, no. 4, pp. 270–280, Oct. 1999.
- [14] N. Tanaka and K. Uematsu, “A crack detection method in road surface images using morphology,” *Proc. IAPR Work. Mach. Vis. Appl.*, 1998.
- [15] X. Jiang, Z. J. Ma, and W. Ren, “Crack detection from the slope of the mode shape using complex continuous wavelet transform,” *Comput. Civ. Infrastruct. Eng.*, vol. 27, no. 3, pp. 187–201, 2012.
- [16] H. Tao, J. M. Zain, M. M. Ahmed, A. N. Abdalla, and W. Jing, “A wavelet-based particle swarm optimization algorithm for digital image watermarking,” *Integr. Comput. Aided. Eng.*, vol. 19, no. 1, pp. 81–91, 2012.
- [17] J. Xiang and M. Liang, “Wavelet- Based Detection of Beam Cracks Using Modal Shape and Frequency Measurements,” *Comput. Civ. Infrastruct. Eng.*, vol. 27, no. 6, pp. 439–454, 2012.
- [18] L. Ying and E. Salari, “Beamlet Transform- Based Technique for Pavement Crack Detection and Classification,” *Comput. Civ. Infrastruct. Eng.*, vol. 25, no. 8, pp. 572–580, 2010.
- [19] B. J. Lee and H. Lee, “Position- Invariant Neural Network for Digital Pavement Crack Analysis,” *Comput. Civ. Infrastruct. Eng.*, vol. 19, no. 2, pp. 105–118, 2004.

- [20] T. Nishikawa, J. Yoshida, T. Sugiyama, and Y. Fujino, "Concrete crack detection by multiple sequential image filtering," *Comput. Civ. Infrastruct. Eng.*, vol. 27, no. 1, pp. 29–47, 2012.
- [21] L. Bursanescu and F. Blais, "Automated pavement distress data collection and analysis: a 3-D approach," in *3-D Digital Imaging and Modeling, 1997. Proceedings., International Conference on Recent Advances in*, 1997, pp. 311–317.
- [22] J. Laurent, D. Lefebvre, and E. Samson, "Development of a new 3D transverse laser profiling system for the automatic measurement of road cracks," in *Symposium on Pavement Surface Characteristics, 6th, 2008, Portoroz, Slovenia*, 2008.
- [23] S.-J. Yu, S. R. Sukumar, A. F. Koschan, D. L. Page, and M. A. Abidi, "3D reconstruction of road surfaces using an integrated multi-sensory approach," *Opt. Lasers Eng.*, vol. 45, no. 7, pp. 808–818, 2007.
- [24] S. Mathavan, K. Kamal, and M. Rahman, "A Review of Three-Dimensional Imaging Technologies for Pavement Distress Detection and Measurements," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 5, pp. 2353–2362, 2015.
- [25] J. Lee and K. Yoon, "Temporally Consistent Road Surface Profile Estimation Using Stereo Vision," *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 5, pp. 1618–1628, 2018.
- [26] E. Salari and G. Bao, "Automated pavement distress inspection based on 2D and 3D information," in *Electro/Information Technology (EIT), 2011 IEEE International Conference on*, 2011, pp. 1–4.
- [27] H. S. Park, H. M. Lee, H. Adeli, and I. Lee, "A new approach for health monitoring of structures: terrestrial laser scanning," *Comput. Civ. Infrastruct. Eng.*, vol. 22, no. 1, pp. 19–30, 2007.
- [28] G. Sansoni, M. Trebeschi, and F. Docchio, "State-of-the-art and applications of 3D imaging sensors in industry, cultural heritage, medicine, and criminal investigation," *Sensors*, vol. 9, no. 1, pp. 568–601, 2009.
- [29] A. Fox, B. V. K. V. Kumar, J. Chen, and F. Bai, "Crowdsourcing undersampled vehicular sensor data for pothole detection," *2015 12th Annu. IEEE Int. Conf. Sensing, Commun. Networking, SECON 2015*, pp. 515–523, 2015.
- [30] A. Fox, B. V. K. V. Kumar, J. Chen, and F. Bai, "Multi-Lane Pothole Detection from Crowdsourced Undersampled Vehicle Sensor Data," *IEEE Trans. Mob. Comput.*, vol. 16, no. 12, pp. 3417–3430, 2017.
- [31] C. W. Yi, Y. T. Chuang, and C. S. Nian, "Toward Crowdsourcing-Based Road Pavement Monitoring by Mobile Sensing Technologies," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 4, pp. 1905–1917, 2015.
- [32] A. Cekmis, "Fuzzy modeling for route characteristic," *Procedia Comput. Sci.*, vol. 120, pp. 112–119, 2017.
- [33] V. Astarita et al., "A Mobile Application for Road Surface Quality Control: UNIquALroad," *Procedia - Soc. Behav. Sci.*, vol. 54, pp. 1135–1144, 2012.
- [34] R. Bridgelall, "Connected Vehicle Approach for Pavement Roughness Evaluation," *J. Infrastruct. Syst.*, vol. 20, no. 1, p. 4013001, Mar. 2014.
- [35] P. Kumar and V. Singh, "Advanced traveler information system for Hyderabad City," *Transp. Syst. IEEE*, vol. 6, no. 1, pp. 26–37, 2005.
- [36] M. Daszykowski and B. Walczak, "Density-Based Clustering Methods," *Compr. Chemom.*, vol. 2, pp. 635–654, 2010.
- [37] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise."
- [38] R. Bridgelall, J. Hough, and D. Tolliver, "Characterising pavement roughness at non-uniform speeds using connected vehicles," *Int. J. Pavement Eng.*, vol. 8436, pp. 1–7, 2017.
- [39] F. Pedregosa et al., "Scikit-learn: Machine learning in Python," *J. Mach. Learn. Res.*, vol. 12, no. Oct, pp. 2825–2830, 2011.