Road Anomaly Classification for Low-Cost Road Maintenance and Route Quality Maps

Naufil Hassan
Deptartment of Computer Science
Information Technology University
Lahore, Pakistan
naufil.hassan@itu.edu.pk

Ifrah Siddiqui
Deptartment of Computer Science
Information Technology University
Lahore, Pakistan
ifrah.siddiqui@itu.edu.pk

Suleman Mazhar
Deptartment of Computer Science
Information Technology University
Lahore, Pakistan
suleman.mazhar@itu.edu.pk

Hadia Hameed
Deptartment of Computer Science
Stevens Institute of Technology
New Jersey, US
hhameed@stevens.edu

***Abstract*—Traditional road maintenance methods are costly; requiring expensive equipment and manpower. Road quality categorization based on machine learning techniques, using real-time opportunistic data gathered from inexpensive open-source inertial systems, is a promising alternative. Existing open-source datasets for this problem are small and less representative of actual situation where data is imbalanced and skewed towards regular road surface instances. With the help of an inexpensive device and data collection platform developed by our lab, we have collected a large, heterogeneous dataset which is more realistic representative of the problem in real world settings. There are four kinds of Roadway Surface Disruptions (RSDs) considered in this work, namely, Cat eyes, Manholes, Potholes and Speed bumps. The feature set used consists of spectral features, time-series peaks, statistical features such as Kurtosis and Skewness and cepstral features such as Mel Frequency Cepstral Coefficients(MFCC). Feature selection was conducted using Sequential Forward Selection and Relief Algorithm. Support Vector Machine (SVM), Convolutional Neural Network (CNN), Random Forest (RF) and Nave Bayes (NB) were used for classification. The best results are reported by SVM with the True Positive Rate (TPR) of 95.2%. These anomaly classification results can be used as a low-cost road maintenance solution by road repairing authorities and the road quality maps thus generated can provide the passengers and drivers with the information of most comfortable route for their journey. Hence, the proposed unified classification framework provides a solution to both of the target audiences by considering relevant anomalies.**

***Index Terms*—Roadway surface disruption (RSD) , Convolution neural network (CNN), System modeling, Kurtosis, Skewness, Heterogeneous dataset, Mel Frequency Cepstral Coefficients (MFCC).**

I. Introduction

A large amount of budget is spent on road infrastructure maintenance annually [1]. The cost includes the money spent on municipality surveys for road quality assessment and repair charges. According to American Association of Automobiles, pothole damages cost around $15 billion to the drivers in the U.S only, in the past five years [2]. Keeping roadways bump-free is a difficult task and requires proper information of road networks all over the region. The unanticipated traffic loads, harsh weather conditions and usual deterioration, all contribute to the degradation of roadways over a short-term.

As a passenger or a driver, one would want a comfortable and smooth ride with least wear and tear of vehicle, but unfortunately current navigation systems are not capable of providing an information of such a route. In addition, the constrained budget of road maintenance authorities makes it difficult to maintain the quality of roads. This paper addresses both of these issues and presents a real-time system that can provide route quality maps to the commuters and information of worn out roads to the related authorities.

The system presented in this paper comprises of a low-cost embedded solution, designed and deployed by our lab, that can be mounted on vehicles and is capable of identifying the anomalies and categorizing them into sub-groups. We have opportunistically collected heterogeneous data using 12 different vehicles and shared the dataset publicly [3] for the research community to perform further analysis and make improvements. Previous public datasets lacked heterogeneity owing to the fact that they included equal instances of anomalies and normal events collected in a planned repeated-driving over the same anomaly events, which is highly unlikely in actual scenario. Our embedded solution is well suited for this cause due to the reasons that it is

- cheap and low cost
- small and robust enough to be easily placed on any location in the vehicle
- gathers real-time patterns of roadway surfaces along with the GPS coordinates

This system can be installed on vehicles that are meant to patrol the areas regularly, for example garbage collection vehicles, taxis and freight-carrying vehicles etc. The data gathered from all these sources would incur no extra cost and it will be updated on daily basis as these vehicles would move around and would have to perform their duties as per schedule. For this work we have collected data by using our embedded system on 12 different vehicles which were driven by different drivers. The system contains accelerometer and GPS. Accelerometer is responsible for recording 3-axis acceleration patterns while GPS gathers longitude and latitude coordinates of the vehicle at run-time. The dataset collected is divided into four different types of RSDs namely, Manholes, Potholes, Speed bumps and Cat eyes. The usual road without any anomaly is tagged as 'Normal' in the data. Annotation was done with a separate GUI at the time of data collection and the anomalies were marked by a person present in the vehicle. After the data is successfully collected, it is passed through annotation correction for careful inspection of any mistakes might be committed at the time of collection. Different features are extracted from the time-series signal of 3-axis patterns and are fed to the machine learning classifiers for training. The dataset required for the two target audience is different as per their needs. So the dataset prepared for route quality maps contains all the classes but it is made less complex by combining all the RSDs into a single class named as Anomaly and the Normal events are kept in the other class, so in this case the problem is treated as a binary-class problem. Whereas, for road repairing authorities, speed bumps and cat eyes are filtered from the data and the rest of the classes are kept. In this paper we have treated the problem in both cases, i.e. as a binary-class problem and as a multi-class problem.

Considering normal events as negative class, the performance metrics being reported in this paper include True Positive Rate (TPR), False Negative Rate (FNR) and False Positive Rate (FPR).

- TPR is for comparing results with other papers
- FNR is for commuters because they will be requiring the classifier that predicts the normal events better
- FPR is for the road repairing authorities.

II. Related Work

Classifying RSDs has been through many stages ranging from the use of camera based techniques to static sensors. But the inertial embedded sensors are being used more recently, as they are very cost effective and ensure constant road monitoring. In literature, the datasets available are somewhat synthetic because they lack heterogeneity of classes. In the real-world scenario anomalies occur intermittently during a drive and much of the journey contains a good amount of normal events. Hence, the data must be skewed towards normal events rather than having equal proportion of all events. Much of the work is done using smart phone's accelerometer data and then signal processing techniques are applied. There is a significant gap in the literature of application based generalized framework.

Hadia et al. proposed a system that classifies roadway surface anomalies using different machine learning algorithms. The proposed system was not generalized for all vehicles, rather it had a separate feature set and a separate classifier for each vehicle [4]. Gonzalez et al. used a novel approach of representing the patterns of acceleration of RSDs by Bag of Words [5]. They have used Artificial Neural Networks (ANN), Decision Trees (DT), Support Vector Machine (SVM), Random Forest (RF), K-Nearest Neighbor (KNN), Nave Bayes (NB) and Kernel Ridge (KR). However, ANN has given the best results for multi-classification of five distinct classes, reporting TPR of 93.8%. This work has reported these results on the dataset they have collected themselves and shared publicly but the dataset lack heterogeneity. Also, only the z-axis acceleration patterns have been used in the dataset which restricts the exploration of acceleration patterns in other two axes, i.e. x-axis and y-axis. Mohan et al. have used some thresholding based techniques to classify the RSDs [6]. They identified that the major issue being faced is the lack of proper method for annotation of the dataset, on the basis of which the classifiers are being trained. The paper attempts to perform rich sensing which exploits the accelerometer, microphone and GPS sensor of mobile phone to gather data that can then be localized with good precision. Eriksson et al. have also used smartphones to collect data and then annotated it manually [7]. In their work, they have tried to detect potholes, railway crossings, manholes, and extended joints. But the same issue of non-heterogeneous dataset exists in their work. Perttunen et al. have explored Fast Fourier Transform (FFT) and Mel Frequency Cepstral Coefficient (MFCC) based features along all of the three axes i.e. x, y and z [8]. The paper also reported the dependency of these features on vehicle's speed and had attempted to remove that dependency. They have reported FPR and FNR of 3 percent and 18 percent respectively. The limitation of this work is the consideration of abrupt instances only. As they have very distinguishable effect on the speed of vehicle, they can be classified very easily while comparing with normal events. Active machine learning is now being used in online systems for quickly adopting the learning models with a changed environment [11] i.e. removing dependency of the vehicle. Transfer learning techniques are also being investigated in labelling activity recognition data to reduce the effort and cost [12].

In the literature we have observed expensive and inaccurate solutions to the problem, unrealistically balanced datasets, absence of geo-tags in some works, no discussion about vehicles dynamics dependency on the accelerometer patterns and end-to-end distinct solutions for any targeted audience. This paper seeks to fill up this gap of literature and provide generalized, precise, low-cost and useful application oriented results.

III. Experimental Setup

A. Hardware Details

A low-cost, dedicated hardware is designed which logs the acceleration patterns and longitude/latitude coordinates using ADXL362 and VK2828U7G5LF sensors respectively. All this data is stored in an on-board storage (SD card). The brain behind all the instructions given to the sensors is PIC18F26K22 micro-controller. The data-logger is powered by 3.1Ah Lithium-ion battery. Data is collected at 100Hz frequency but due to the inaccuracy of internal clock, there

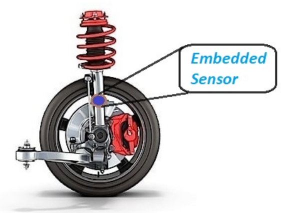

Fig. 1. *Embedded sensor attached near the tyre of the car, for data collection*

is a margin of 10% error. An average sampling rate of 93Hz is used for maintaining consistency of the data.

B. Data Collection Details

This process consists of 12 cars, a driver, a co-pilot and two data-loggers. The first data-logger is mounted outside the car, on its shock absorber, as shown in Fig. (1). The second data-logger is placed inside the car, on the dashboard. The data is being annotated on-the-go by co-pilot, using our own developed Graphical User Interface (GUI) in MATLAB. By the end of the journey, a file is generated by the data-logger that contains vehicle's speed, GPS coordinates, time stamps and acceleration readings of 3-axes, per second. The GUI is designed in such a way that the anomalies are associated with certain numbers, i.e. if a speed bump is identified by a button labeled as "5", so now if the car traverses through the speed bump, the co-pilot presses button "5" as soon as the car hits the speed bump. The button is kept pressed until the car has covered the full event and is released when car is back on the normal road. In this way, the time is logged between the instances when the button was pressed and then released. Based on this information we can compare the time stamps from the file generated by GUI and data-logger to annotate the data.
It was ensured that each car gathers at least 50 instances of each RSD. For this purpose, the average journey was 25-30km long and average time duration was 45 minutes. If the co-pilot has wrongly tagged an anomaly, he can press the button for "Mistake" right after it and then we can correct the label with manual inspection. Those readings of the sensor were labeled as "Normal" for which no button was pressed. The 3-axes of accelerometer x, y and z are aligned with transversal, longitudinal and vertical axes of the car. Fig. (2) shows the experimental setup for data collection.

C. Data Cleaning and Pre-Processing

The data collected by data-logger is at 100Hz and has an error rate of 10%. So, in order to maintain the consistency of data, we have linearly interpolated the number of samples to 93, in every second. The time series data is then divided into non-overlapping chunks of 1 second and these chunks are then labeled by comparing the time stamps logged by GUI. Similarly, the mistakes encountered by human error were covered up by careful visual inspection.

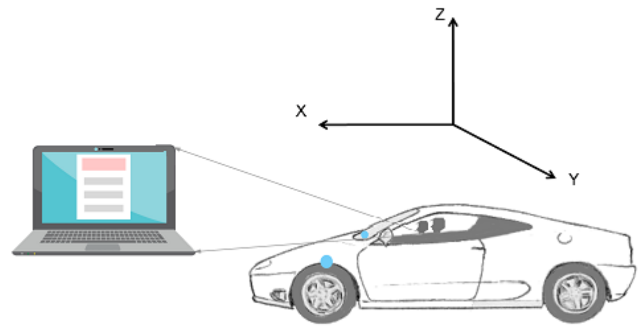

Fig. 2. *Data Collection*: Data loggers are mounted on dash board and near tyre of the car. Driver and co-pilot are sitting in the car where co-pilot uses MATLAB based GUI for annotation. The 3 axes of accelerometer x, y and z are aligned with transversal, longitudinal and vertical axes of the car.

IV. System Modelling

To obtain good results from a system, it is very important to understand its dynamics. The signal we get from the accelerometer is actually a convoluted signal with additive noise in it.

$$a(t) = [a_x(t)\ a_y(t)\ a_z(t)] \tag{1}$$

$$a(t) = a^*(t) + n(t) \tag{2}$$

where, $t = 1, 2, 3, ...$ $\quad a^*(t) \in R^3 \quad n(t) \in R^3$

$$n(t) \sim \mathcal{N}\left(0, \sigma_n^2\right) \tag{3}$$

$$a^*(t) = a_{event}(t) * h(t) \tag{4}$$

Here *a(t)* represents the data we are getting from the system and it is a three-dimensional data. *a*(t)* is the actual signal and *n(t)* is the additive Gaussian noise with zero mean and standard deviation of σ^2, hence it can be represented by the Eq. (3). This noise can be due to temperature drifts and electromagnetic fields.
Further decomposing *a*(t)* we get Eq. (4) where the $a_{event}(t)$ is the signal of road surface and *h(t)* is the effect of car hydraulics/dynamics on this signal. These two signals get convoluted with each other and hence there is a difference of signatures of same anomalies by different cars. This problem is catered in this paper by processing the signal in its frequency domain and using its cepstral features, as explained in the next section.

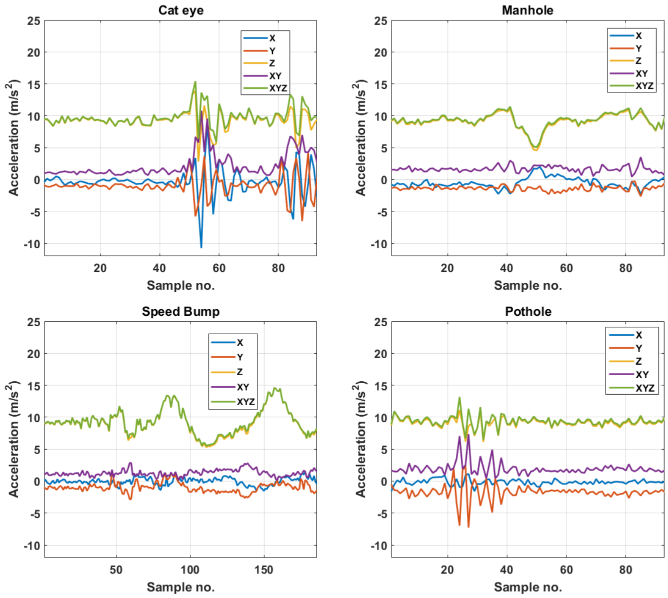

Fig. 3. Accelerometer patterns of different anomalies.

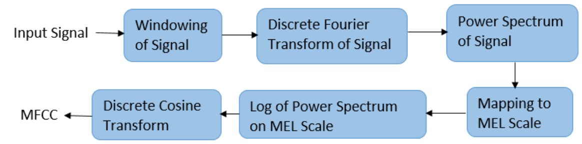

Fig. 4. Block Diagram illustrating the steps involved in computing MFCC.

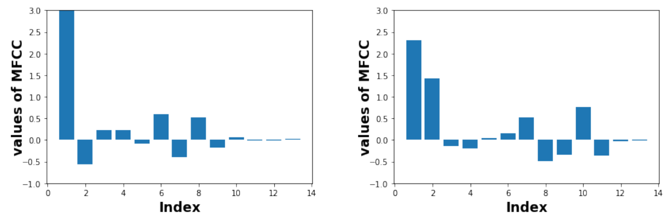

Fig. 5. MFCC co-efficients of Normal and Anomaly Class.

V. Feature Extraction and Selection

Several features were extracted from the time series data that have been used in the literature. As per the experimental setup, it is very clear that z-axis acceleration patterns will help in classifying the anomalies. However, well established feature selection algorithms have been used for ranking the potential features from all axes.

The data has been dealt with in terms of windows and each window is considered as an instance. Hence different features have been extracted from each of the instances from five dimensions i.e. X, Y, Z, XY and XYZ. The features are divided into several groups where Group-A contains FFT based features, Group-B contains time series peaks, Group-C contains Kurtosis, Group-D contains Skewness and Group-E contains MFCC.

The statistical features (i.e Kurtosis and Skewness) have proved to be very helpful in classification and this can be easily understood from Fig. (3) where it is very clear that Speed bump pattern has different tail as compared to other patterns and Manholes have a skewed pattern. The equation used to compute these statistical features are given by Eq. (5 and 6).

$$Kurtosis = \frac{m^4}{(m^2)^2} = n\frac{\sum_{i=1}^{n}(X_i - X_{avg})^4}{\sum_{i=1}^{n}((X_i - X_{avg})^2)^4} \quad (5)$$

$$Skewness = \frac{m^3}{(m^2)^{3/2}} = \sqrt{n}\frac{\sum_{i=1}^{n}(X_i - X_{avg})^2}{\sum_{i=1}^{n}((X_i - X_{avg})^2)^{3/2}} \quad (6)$$

The cepstral features(i.e MFCC) also turned out to be helpful in classifying the RSDs. MFCCs are usually used as features in speech recognition applications where they are used to represent the vibrations of vocal tract. In the context of the work presented in this paper, our approach is to use these MFCCs to get an accurate representation of the vibrations produced when the vehicle passes through a certain RSD. Different anomalies will produce different types of vibrations, so a true representation of these vibrations, can help in the classification of these anomalies.

The steps involved in the computation of MFCCs are shown in Fig. (4). Computing MFCCs is governed by Eq. (7).

$$Coef(t,k) = \sqrt{\frac{2}{N}}\sum_{i=1}^{N} log[E_{mel}(t,i)]cos[k(i-0.5)\frac{\pi}{N}] \quad (7)$$

Here N is the number of filters, $E_{mel}(t,i)$ is the i^{th} filters energy at time t and the order of filter is represented by k in the equation where k=1,2,3,.....,p.

We have used MATLAB's MIR tool box to compute MFCC features which gives us fixed length vector of 13 coefficients for each instance.

MFCC's also proved to be useful for resolving the problem of convolution of the two signals i.e signal of car response and actual anomaly response. Since MFCCs represent the signals in frequency domain, so the time convoluted signals are transformed to additive signals in the frequency domain. MFCC coefficients are actually the amplitudes of Mel-frequencies, calculated by DCT and these amplitudes are treated as a feature to be fed to the classifier. There is a significant difference among the MFCC co-efficients of the signals representing normal and anomaly class as shown in Fig. (5).

The feature vector of one window contains approximately 284 features where FFT based features are extracted from the combination of axes information [4] i.e. FFT-XY and FFT-XYZ. The feature vector is then passed through two feature selection algorithms namely, Sequential Forward Selection (SFS) and Relief Algorithm [9]. For the later, we have

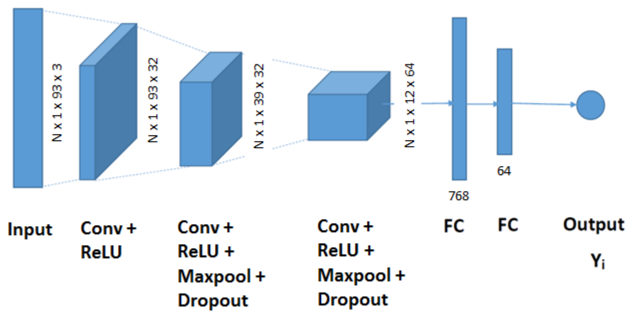

Fig. 6. Architecture of Convolutional Neural Network

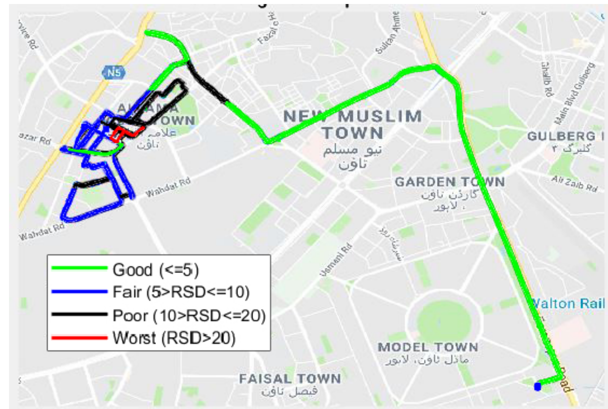

Fig. 7. Route quality maps generated for drivers and passengers as per the criteria given in table 1

used implementation of *Relieff* from MATLAB Statistics and Machine Learning Toolbox.

VI. Classification

Random Forest (RF), Support Vector Machine (SVM), Convolutional Neural Network (CNN) and Nave Bayes (NB) are used for classification of data. Three of these algorithms are fed with hand-crafted features, whereas CNN is a feature agnostic algorithm and has not been explored in the literature. We have used RBF kernel for SVM with cost $2^{3.16}$ and gamma of $10^{0.985}$ while treating the problem as a binary classification. For multi-class classification, we have used one-versus-all technique. Random Forest was set to 10 number of trees and Gini index criterion. No prior probabilities were set for NB.
The CNN architecture we have implemented is a 5-layer architecture consisting of 3 convolutional layers and 2 fully connected layers. The architecture is shown in Fig. (6). Eq. 9, 10 and 11 represent the vector of an instance in 3-axes. Each vector has length of 93 samples. By using these vectors we form a matrix of dimensions 93x3 and feed it to architecture. Y_i is the output of neural network which can be one of the labels as given in Eq. (11). We have the problem of imbalanced dataset and to tackle it we have used Focal Loss [10] which handles this scenario very well. In Eq. (13) , p_t represents the probability of the class to be predicted. The γ factor is a hyper parameter that decides how much penalty i.e. *(1- p_t)*, is to be given to the network for a certain misclassified example.

$$a_x(t) = [a_x(1)\ a_x(2)\ a_x(3)\ ....\ a_x(93)] \tag{8}$$

$$a_y(t) = [a_y(1)\ a_y(2)\ a_y(3)\ ....\ a_y(93)] \tag{9}$$

$$a_z(t) = [a_z(1)\ a_z(2)\ a_z(3)\ ....\ a_z(93)] \tag{10}$$

$$Y_i \in \{Manhole\ \ Pothole\ \ SpeedBump\ \ Cateye\ \ Normal\} \tag{11}$$

$$FocalLoss = F_L(p_t) = -(1-p_t)^r log(p_t) \tag{12}$$

TABLE I
Criteria for the Evaluation of Roads

Good	≤ 5 anomalies
Fair	$5\leq 10$ anomalies
Poor	$10\leq 20$
Worse	≥ 20

VII. Experiments and Results

We have combined all the data from different cars, hence the total of 59916 instances are present. The distribution of different classes is as follows; 93.2% of Normal events, 2.5% Speed bumps, 1.8% Potholes, 1.3% Manhole and 1.2% Cat eyes.
The results obtained from different classifiers are tabulated in Table (II). The best results are reported by SVM with TPR 95.2, FPR 2.551 and FNR 4.797. The confusion matrices of binary and multi-class classification are shown in Table III and IV respectively.

The main emphasis of this work is generation of route quality maps for drivers/passengers, and road-condition maps for road repairing authorities. The criteria for generating road-condition maps is given in Table (I). The maps for route quality and road-repairing authorities are shown in Fig. 7 and Fig. 8 respectively. The patches are color coded for the drivers in former and the anomalies are pointed out in the for authorities in the later.

TABLE II
Results of Machine Learning Algorithms

Classifier	FPR	FNR	TPR
SVM	2.6	4.8	95.2
NB	10.1	7.6	92.4
RF	9.6	21.4	78.6
CNN	14.2	11.9	88.1

Fig. 8. Road condition maps for road repairing authorities

TABLE III
CONFUSION MATRIX OF BINARY CLASSIFICATION USING SVM

	Anomaly	Normal	Total Predicted
Anomaly	1548	78	1626
Normal	114	4354	4468
Total Actuals	1662	4432	6094

VIII. DISCUSSION

From system modeling, it is very clear that the data is dependent on the car dynamics. Hence for better classification results, it is necessary to remove this dependency from the data if we have to increase our data and combine the dataset for all the different cars.We have tried to remove this dependency by using MFCC features.
This approach of processing the signal in frequency domain significantly improved our TPR from 94.08% to 95.2%. However this problem can also be solved by using the data from both sensors, i.e. the one mounted on dashboard and the other mounted near tyre. Similarly, from visual inspection of anomaly patterns, it is interpreted that cat eyes pattern resembles to that of potholes and speed bumps' resembles to manholes. This is also very much evident from the confusion matrices. Kurtosis and Skewness were computed for this cause, i.e. to differentiate between these classes.

IX. CONCLUSION

The main contributions of this work are:

- Heterogonous dataset is gathered; representing the real-world scenario and shared publicly for benchmarking
- This work is targeted at two audiences, i.e drivers/passengers and road-repairing authorities, and generate intelligent maps.
- Statistical features have not been explored in this domain, and they have significantly improved the results.
- CNN have been implemented which doesn't require hand-crafted features. Focal loss is implemented for catering imbalanced dataset.

TABLE IV
CONFUSION MATRIX OF MULTI CLASS CLASSIFICATION USING SVM

	Cateye	Manhole	Normal	Pothole	S.Bump
Cateye	192	14	8	68	8
Manhole	42	98	20	90	68
Normal	7	3	558	5	3
Pothole	107	39	18	233	45
S.Bump	17	26	9	24	500

- MFCC features' use and their clear reasoning has been provided that has helped us in removing car dynamics dependancy from the data.
- System modelling is discussed and the major pointers are identified that can be very helpful in improving the results.

REFERENCES

[1] DW.COM, "https://www.dw.com/en/trump-reveals-plan-to-repairamericas-creaky-infrastructure/a-42551209."
[2] Insurance Journal, "https://www.insurancejournal.com/magazines/magfeatures/2016/03/21/401900.html."
[3] Dataset, "https://bit.ly/2Fbfkaj."
[4] H. Hameed, S. Mazhar, and N. Hassan, "Real-Time Road Anomaly Detection, Using an On-Board Data Logger," IEEE 87th Vehicular Technology Conference (VTC Spring), 2018.
[5] L. C. Gonzalez, R. Moreno, H. J. Escalante, F. Martinez, and M. R. Carlos, "Learning Roadway Surface Disruption Patterns Using the Bag of Words Representation," IEEE Transactions on Intelligent Transportation Systems, vol. 18, no. 11, pp. 2916-2928, 2017.
[6] P. Mohan, V. N. Padmanabhan, and R. Ramjee, "Nericell: Rich Monitoring of Road and Traffic Conditions using Mobile Smartphones," Proceedings of the 6th ACM conference on Embedded network sensor systems - SenSys 08, pp. 323-336, 2008.
[7] J. Eriksson, L. Girod, B. Hull, R. Newton, S. Madden, and H. Balakrishnan, "The Pothole Patrol: Using a Mobile Sensor Network for Road Surface Monitoring," Proceeding of the 6th international conference on Mobile systems, applications, and services - MobiSys 08, pp. 29-39, 2008.
[8] M. Perttunen, O. Mazhelis, F. Cong, M. Kauppila, T. Leppnen, J. Kantola, J. Collin, S. Pirttikangas, J. Haverinen, T. Ristaniemi, and J. Riekki, "Distributed Road Surface Condition Monitoring Using Mobile Phones," Ubiquitous Intelligence and Computing Lecture Notes in Computer Science, pp. 6478, 2011.
[9] K. Kira and L.A. Rendell, "The feature selection problem: Traditional methods and a new algorithm," AAAI, Vol. 2, pp. 129-134, July 1992.
[10] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollar, "Focal loss for dense object detection," IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 2980-2988, 2018.
[11] T. Sztyler and H. Stuckenschmidt, "Online personalization of cross-subjects based activity recognition models on wearable devices," IEEE International Conference on Pervasive Computing and Communications (PerCom)," pp. 180-189, 2017.
[12] J. Wang, Y. Chen, L. Hu, X. Peng, and S. Y. Philip, "Stratified transfer learning for cross-domain activity recognition," IEEE International Conference on Pervasive Computing and Communications (PerCom), pp. 1-10, 2018.