ANDROID APPLICATION FOR POTHOLE DETECTION ON ROAD USING EMBEDDED SENSORS IN SMARTPHONES

G.Vidya Sagar[1], V. Annapurna[2]

[1]M.tech in Embedded Systems, JNTUA College Of Engineering, Ananthapuramu
[2] Lecture, Department of ECE, JNTUA College Of Engineering, Ananthapuramu.

***Abstract:**Monitoring of the road condition has acquired a critical significance during recent years. Different reasonsbehind broadening research on this field are: to start with, it will guarantee safety and comfort to different road users; second, smooth streets will cause less damage to the car. Here motivate to create a real-time Android Application Road Sense it willshow the road location trace on a geographic map using GPS and save all recorded workout entries, automatically predicts the quality of the road based on gyroscope and tri-axial accelerometer. Decision tree C4.5 classifier is applying on training data to classify road segments and to build proposed model. That show experimental results consistent accuracy of 97.6%. By using of this approach, expect to a quality road mapsof a selected regions visualize. Hence, it can provide feedback to local authorities and drivers. Besides, benefit toRoad Manager from proposed system to evaluating the state of particular road networks and make a checkup on road construction projects, whether they meet or not the required quality.*

***Index terms:**Smartphone, Road monitoring, Accelerometer, Gyroscope, Pothole, Real time, Machine learning, Android*

1. INTRODUCTION

According to WHO (World Health Organization)statistics provided as, road accidents have become a one of top 10 leading cause of deaths in the world. Specifically, for every year road accidents claimed nearly 1.25 million lives (2015). Studies show that most of the road accidents causes by poor condition of roads. Bad roads are a big problem for vehicles and drivers; this because of the deterioration of roads leads to more expensive maintenance, not only for the roads itself but also for vehicles. Accordingly, road monitoring systems of surface conditions are very most possible solutions to improve traffic safety, reduce accidents and protect vehicles fromdamageduetobadroads.Both road managers anddrivers are interested in having sufficient information concerning road infrastructure quality (safe or dangerous road). Strengthenapproaches for road monitoring systems of surface conditions involve the adoption costly and sophisticated hardware equipments such as ultrasonic or specific accelerometers withsystems of data acquisition. These approaches incur a high installation and maintenance cost and requires large manual effort, which can induce error while deploying or collecting the data. Another alternative to use sensing technologies, gain this information forsolving the problem road monitoring systems of surface conditions. Now, smartphones are popularly utilized. The meager part of system is equipped with various sorts of sensors like camera, accelerometer, GPS, gyroscope, microphones, etc. Thus, one of such helpful applications to monitor street conditions is smartphone based road monitoring systems of surface conditions.

This paper introduces a framework of road condition monitoring which based on embedded sensors (accelerometer, gyroscope and GPS) built with in smartphones to give us the quality, conditions of different road sections using different machine learning techniques. The contributions part of present paper is manifold and can be concluded as follows:As a first contribution, design a machine-learning algorithm(C4.5Decisiontree)toclassifyroadsegmentas compared to past works that use simple thresholds, SVM and fuzzy logic. This experiment shows the system able to detect and classify events related to road conditions with an accuracy of 97.6%.

Proposed system, unlike existing solutions that require external hardware, is an inexpensive simple yet efficient solution that able to monitor road quality conditions. It is accomplished on android smartphones and that highly portable and maintain easily. User applications provide usefulto drivers and local authoritiesfeedback by plotting the evaluated road location on a Map and saving all recorded entries.

Creating an application of Android that allows real-time and automatic collection and analysis of gyroscope withaccelerometer data order to get reliable road surface labels in differentiate to previous works that mostly use offline methods (videos, images for data labeling).

While most of previous works employ unique modal accelerometer data, right now using gyroscope sensor in association with accelerometer sensor to derive more accurate road quality prediction. The rest of part in this organized as, presents a background on the three different machine learning algorithms used in this system, introduces some recent research works related to theroad monitoring systems of surface conditions and describes the general idea and the proposed architecture.

2. APPLICATION OVERVIEW

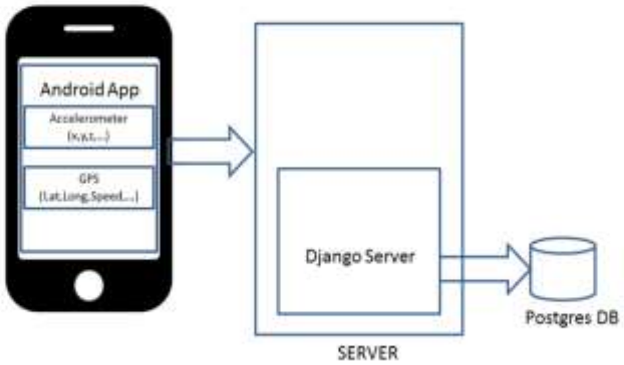

Fig1: overview of proposed system

3. SYSTEM DESIGN

A. Training phase

In training phase, train the classifier model using machine-learning techniques based up on the collected data. During a preprocessing stage, a LPF(low pass filter) is applying to remove high frequency components, and then compute magnitude of gyroscope accelerometer values. In Feature Extraction, effective features extracted from specific types of road conditions patterns on acceleration and rotation around gravity. Afterwards, the features are selected in training phase and a classifier model would be generated which can realize fine-grained identification. Finally, the classifier model is generated and saved.

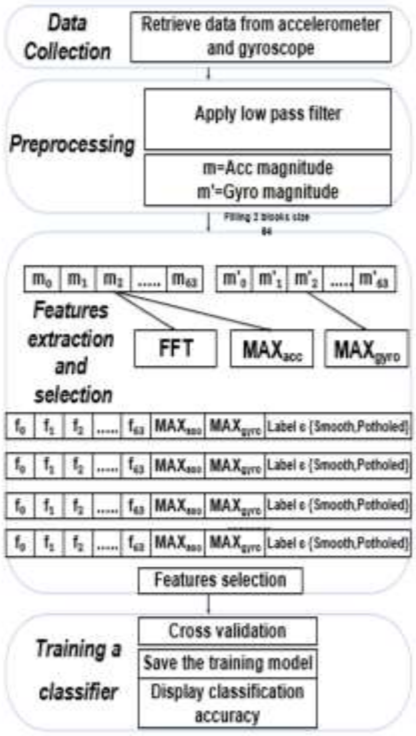

Fig2:Training phase

1) Collecting Data from Smartphone Sensors:

The Data collection phase is the more important one; it is responsible for collecting road information. To develop an Android Applicationcollect readings from the gyroscope and triaxial-accelerometer sensors. The data of sensors road surface quality were collected using gyroscope and accelerometer both sensors built in the Galaxy mobile phone, mounted on the car dashboard shown in Fig.2, along with vehicle path. The frequency of the sensorssampling is 50 Hz. Several data collection drives were performed with a varied speed, the condition of roads label is pre-set before the collection starts. Once the user stops the data acquisition, the application stores the learning data-set as acomma separate values (csv) file. For the work reported in this paper, a drive of about 40 minutes in length (15 km) was selected among the drives, as it seemed to be a smooth segment. It is also difficult to classify the anomalies into different categories at same time. So, the classification of this study was not pointed to identify different road anomalies, but to differentiate potholes from the smooth road. In total, we obtain 2000 samples of data.

2) Preprocessing:

Accelerometer data readings usually contained irrelevant data (noises). Therefore, a pre-processing phase should be applied for reduce noise and improve the road quality recognition. Due to several factors such as jerks or vibrations, turning, veering, braking, and well also subtle changes in sensor orientation, a considerable amount of noise added to these signals. A LPF (Low-Pass Filter) can be helpful to remove high frequencies signals (noise) present in the input signal by applying a suitable threshold to the filter output reading. Once the noise reduction is completed, the filtered accelerometer samples (x, y, z) and the filtered gyroscope samples (x', y', z') are then preprocessed where each sampled smoothed vector was combined as single magnitude where

$$m = \sqrt{x^2 + y^2 + z^2}$$

$$m^1 = \sqrt{x^{12} + y^{12} + z^{12}}$$

3) Feature Extraction:

When machine-learning algorithms are processed, representative tuple of features than raw data is more effective input. Thus, it is necessary to extract effective features from road conditions patterns. To reduce input data, raw data which is obtained from the sensors is initially segmented into several windows, and features as window of samples frequencies are extracted. This Feature Extractionsteps, serves for recognizing of roads quality as inputs used in the classification algorithms. As mentioned, the window size is cardinal parameter that which influences computation and also power consumption of sensing algorithms. A slide window of window length 64 applied for feature extraction stage, Are proves to helpsin extracting frequency-domain features. In comparison to heuristic features, time and frequency-domain features can describe the information within the time-varying signal. Unlike time domain, the frequency-domain features need another phase of transforming the received data (in time domain) from previous phases of the pipeline. This stage generates features of frequency-domain using very fast and efficient versions of the FFT (Fast Fourier Transform). After computing magnitude m and m', the work-flow of this training phase buffers up 64 successive magnitudes (m_0:::m_{63}) , (m^0_0:::m^0_{63}) before the FFT computing,resulting a feature vector (f_0:::f_{63}) or a vector of Fourier coefficients. FFT is time series of amplitude, over totime magnitude across frequency. Since, the x, y, z are accelerometer and x', y', z' are gyroscope readings and the magnitude are time domain variables; it is necessary to convert these time-domain data into frequency domain as that represent incompact manner of distribution in that the classifier will use to build a model in further phase of this pipeline. While computing the Fourier coefficients, training phase also stores the maximum (MAX_{acc}) magnitude of the (m_0:::m_{63}), the maximum (MAX_{gyro}) magnitude of the (m^0_0:::m^0_{63}) and roads supplied labels (Smooth, Potholed) using a native android application called Road Data Collector.

B. Prediction phase

The prediction phase is installed on smartphone, which senses real-time vehicular dynamics to detect and identify road conditions. This system first senses the collection of readings from gyroscope and accelerometer embedded sensors on smartphone. After getting real-time readings from sensors, the preprocessing is performed on sensors readings. Afterwards,system extracts features from patterns of road conditions, and then predicts the road quality based classifier model trained in training stage and shows tracing of road locations on geographic map. Finally, a history of all reported road conditions is saved.

4. ROAD SURFACE QUALITY DETECTION ALGORITHM

C4.5 Decision tree:

C4.5 is an algorithm use to generate a decision tree developed by Ross Quinlan.Tree based learning algorithms are best and mostly use for learning methods of supervised. Tree based methods are allowwith high accuracy of predictive models, stability and easy of analysis. Unlike linear models and it maps non-linear relationships well. They are alterable at solving any type of problems at hand (classification or regression).

Decision Tree:

Decision tree is one type of supervised learning algorithm whichhaspre-defined target variable that is used mostly in classification problems. It worked for continuous and categorical both input and output variables. Under this technique, splitting of sample or population into two or more homogeneous sets or sub-populations dependsupon most significant splitter in input variables.

C4.5decision tree is used here to classify the road segments. A binary decision tree created to model the classification process. Once treebuilds, each tuple in dataset is passed through decision tree to assign a class to the tuple. Decision tree builds by considering the parameters as speed of the vehicle, z axis of the accelerometer data, etc.

5. ABOUT APPLICATION

Android Application:

Android applications have built to collect the data the information of the potholes and bumps. Application makes uses of built in mobile accelerometer, GPS sensors data to analyze information to identify the potholes and the location where exactly the pothole is. Once the ride is started the application gathers the sensors data and is dumped into a two text files in CSV format, one text file contains the accelerometer data and the other contains the GPS data. Data is collected lot per second. Accelerometer sensor gives 60 samples per each second and GPS sensor gives a sample per each second. So, there would be huge data and redundant data and transferring the data directly to server can take huge of network bandwidth and time. To overcome this, once the ride is stopped the application compresses the text files. After compressing, file sizes are reduced by 60-70% which saves user's internet cost.

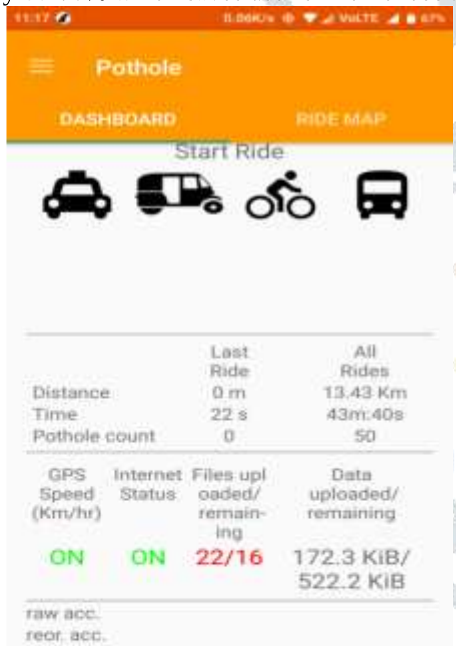

Fig3:Application dash board

Server Side:

To develop server for web development here used Django python framework. REST API used for communicating with the android application. Server receives information from application of JSON object form, process it and store into 'PostgreSQL' database. Ride details comprises of the information follow

- Start time of the ride
- Stop time of the ride
- Distance covered on during ride
- Number of Potholes encountered during ride

Docker:

Docker is software that performsvirtualization operating-system-level is also called as containerization. Docker is tool that an application of package and its dependenciesin virtual container which run on other Linux servers. This used to enable flexibility and to portability on area of application can run.As docker provides a linux environment in a consistently have setup the server and database in docker.

6. RESULTS AND ANALYSIS

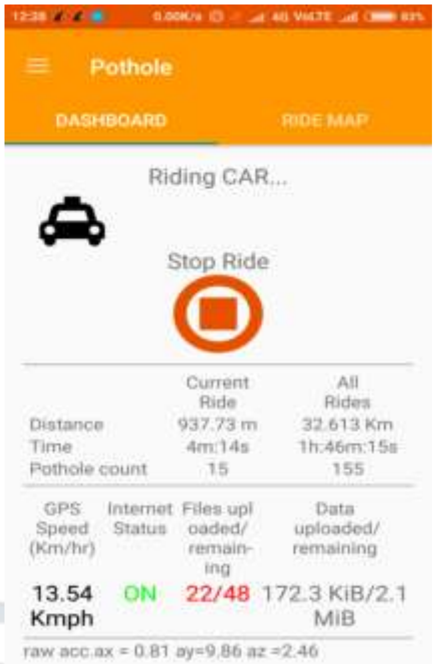

Fig4: Dash board when ride started

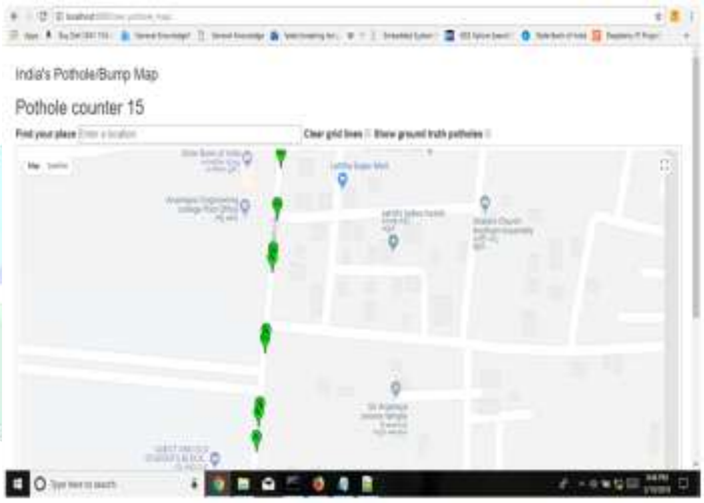

Fig5:pothole plotted in map

7. PERFORMANCE EVALUATION

The overall goal of this evaluation, to determine the accuracy and effectiveness of thissystem able to detect/predict road conditions. This section, evaluate system in two steps: 1.Analytical Validation:Evaluate performance of different classifiers based on several metrics. 2. Experimental Validation: Test the utility of the totalsystem in real driving environments.

Analytical Validation The learning data-set with the selection of features is passed input to various classification algorithms in order of selecting the most appropriate model. To evaluate effectiveness of different classifiers, cross validation use as the evaluation method. More specifically, the smartphone sensor dataset is divided into test and training sets. The goals of cross-validation to define a subset of data or a validation set for the test of model for training phase, order to scrutinize the problems such as over-fitting. To represent whether the classifier better than others, a 10 fold cross-validation were performed, where 1/10 of both data were used only for testing purpose. Have to use several measures to evaluate performance of classifiers used.

a) Classification accuracy: The accuracy can explain the overall classification performance for all the activity classes as the follows

$$\text{Accuracy}= \frac{(TP+TN)}{(TP+TN+FP+FN)} (1)$$

This shows the classification accuracies of various machine learning algorithms with and without feature selection. C4.5 is the most accurate classifier compared to Naive Bayes and SVM with the average accuracy is 97.6%, CFS algorithm increases the average accuracy by 0.1%. By observing that C4.5 perform well when using CFS algorithm; this high accuracy was achieved by fusing data fromgyroscope and accelerometer both sensors used to reduce the number of false positives. SVM (Support Vector Machine) performs poorly in accuracy (95.25%) and Naive Bayesian Classifier is a close competitor in terms of accuracy (96.90%).

b) TP Rate and FP Rate: TP (True positive) rate is positive case percentagecorrectly classified as belongs to positive class.

$$\text{TP Rate} = \frac{TP}{(TP+FN)} (2)$$

FP (False positive) rate isnegative case percentage misclassified as belongs to positive class.

$$\text{FP Rate} = \frac{FP}{(FP+TN)} (3)$$

c) Precision, Recall and F-measure: The precision is another metric which can explain as ratio to the correctly classified positive instances to the wholeNo.of instances classified as positive and is mathematically calculated as follows:

$$\text{Precision} = \frac{TP}{(TP+FP)} (4)$$

d) The recall again called as the TP rate or sensitivity, is measure of how good is the classifier to correctly predict actual positives samples

$$\text{Recall} = \frac{TP}{(TP+FN)} (5)$$

The F-measure is measures test's accuracy. The F measure can be explained as weighted average of recall andprecision

$$\text{F}-\text{measure} = \frac{2\times\text{precision}\times\text{recall}}{(\text{Precision}+\text{Recall})} (6)$$

Both recall andprecision are therefore based onmeasuring and understanding of relevance. Recall is measure of quantity orcompleteness, Precision can be measure of quality orexactness.

Table: Results comparison for C4.5 and SVM

Classifier	C4.5	SVM
Accuracy	0.97	0.95
TP Rate	0.98	0.95
FP Rate	0.23	0.265
Precision	0.984	0.951
Recall	0.981	0.953
F-measure	0.988	0.950

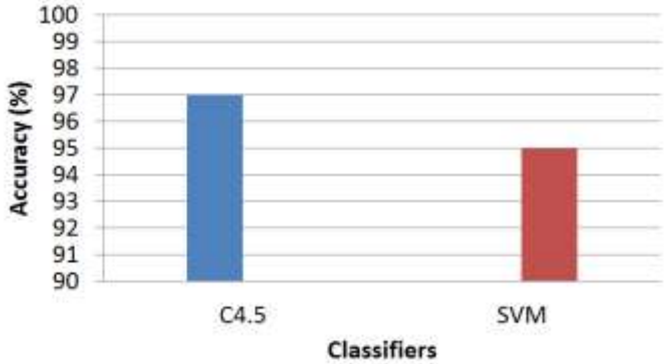

Fig6: Graphical representation for C4.5 and SVM

8. REFERENCES

[1] AzzaAllouch, AnisKoubˆaa, Tarek Abbes, and Adel Ammar," RoadSense: Smartphone Application to Estimate Road Conditions using Accelerometer and Gyroscope,"2017 IEEE

[2] W. H. Organization. (2015) Global status report on road safety. [Online]. Available: http://www.who.int/violence injury prevention/road safety status/2015/en.pdf/

[3] R. Madli, P. Pattar , S. Hebbar and V. Golla, "Automatic detection and notification of potholes and humps on roads to aid drivers," IEEE Sensors Journal, vol. 15, no. 8, pp. 4313–4318, 2015.

[4] R. Ramjee, V. N. Padmanabhan, and P. Mohan, "Nericell: rich monitor-ing of road and traffic conditions using mobile smartphones," in Proc. 6th ACM Conf. Embedded Netw. Sensor Syst, Raleigh, NC, USA, 2011,323–336.

[5] J. Eriksson, L. Girod, B. Hull, R. Newton, S. Madden, and H. Balakr-ishnan, "The pothole patrol: using a mobile sensor network for road surface monitoring," in Proc. 6th Int. Conf. Mobile Syst.Appl. Services, USA New York, NY, 2008, pp. 29–39.

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

[7] S. A. HoseiniTabatabaei, A. Gluhak, and R. Tafazolli, "A survey on smartphone-based systems for opportunistic user context recognition," ACM Comput. Surv(CSUR), vol. 45, no. 3, pp. 1–51, jul 2013

[8] A.T.Campbell.(2013)Smart phone programming. [Online].Available: http://www.cs.dartmouth.edu/campbell/cs65/myruns/ myrunsmanual.html