



# International Journal of **A**dvanced **R**esearch in **E**ducation and **T**echnolog**Y** (IJARETY)

Volume 13, Issue 2, March-April 2026

Impact Factor: 8.152



INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA



# Road Quality Survey Bot using Arduino and Machine Learning

ShashiKiran S<sup>1</sup>, Priyanka S<sup>2</sup>, Sahana PS<sup>2</sup>, Swathi SG<sup>2</sup>, Sanjana U<sup>2</sup>

Associate Professor, Department of Electronics and Telecommunication Engineering, Jawaharlal Nehru New College of Engineering, Karnataka, Shivamogga, India<sup>1</sup>

Department of Electronics and Telecommunication Engineering, Jawaharlal Nehru New College of Engineering, Karnataka, Shivamogga, India<sup>2</sup>

**ABSTRACT:** The state of roads may sometimes be difficult to perceive due to intense climate conditions, absence of road signs, or simply human inattention, which may be harmful to both vehicles and drivers. The automatic monitoring of the road states represents a promising solution to warn drivers about the status of a road in order to protect them from injuries or accidents. In this paper, we present a novel application for data collection regarding road states. Our application entitled “Road Scanner” allows on-board users to tag four types of segments in roads: smooth, bumps, potholes, and others. For each tagged segment the application records multimodal data using the embedded sensors of a smartphone. The collected data concerns mainly vehicle accelerations, angular rotations, and geographical positions recorded by the accelerometer, the gyroscope, and the GPS sensor, respectively, of a user phone. Moreover, a mediumsize dataset was built and machine learning models were applied to detect the right label for the road segment.

**KEYWORDS:** smartphone sensors, road scanning, data collection, in situ labelling, machine learning.

## I. INTRODUCTION

A road is a vital part of people’s daily living. Sadly, in many situations, drivers can fail to detect road conditions due to serious circumstances or lack of attention. In many countries around the world, roads may suffer from low maintenance due to high costs and budget cuts. Traffic congestion has been increasing worldwide as a result of increased motorize urbanization, population growth and changes in population density. Congestion reduces utilization of the transportation and infrastructure and increases travel time, air pollution, fuel consumption and most significantly, traffic accidents. There is an exponential increase in the population of Mumbai. As people live in far places from their offices, they believe in commuting through road transport or trains. This has led to faster corroding of the roads. These roads are left attended because of huge traffic, delays and accidents which can cause mental strain and in some cases can be fatal. At present, there are various ways to detect the potholes either manually or automated. The main aim of our project is to detect the pothole automated and also repair it without us being present there. Along with that our system is quite cheaper than other proposed systems. To reduce the potholes, we have decided on a mechanism where it could simultaneously detect and store the data of the pothole in the database. For transmission of data, we use Lo Ra WAN module which can send data of the pothole in the database. Then it would send a robot without any manual assistance to that place to correct it. This reduction of manual labour will also lead less in time consumption for the correction of pothole. In our system, image processing is an integral part of the detection of potholes. Plus, for the transmission of the images and data, we are using Lo Ra instead of Zig bee, as Lo Ra can transmit data over longer ranges, i.e. ~10 km. Also the project deals with the filling of potholes which has not been attempted earlier on, here instead of asphalt we are using chip filling which can greatly reduce the recurring of the filling.

## II. THEORETICAL BACKGROUND

- Road condition monitoring aims to detect defects such as potholes, cracks and uneven surfaces so that maintenance can be planned efficiently.
- Pavement management systems treat road roughness and damage as measurable engineering quantities, for example through indices like the International Roughness Index (IRI), which relate surface irregularities to ride comfort and safety.
- Traditional monitoring relies on manual visual inspection or expensive survey vehicles, which are slow and not scalable for large networks, motivating low-cost automated solutions using sensors and intelligent algorithms.

- An Arduino board provides a simple embedded platform to read sensors, timestamp the data and control the movement of the robot.
- The microcontroller samples accelerometer and other sensor values at a fixed rate, applies basic noise reduction (for example, moving-average filtering) and transmits the measurements to a computer or stores them for later analysis.
- Using Arduino allows the system to be low cost, portable, and easily replicated for continuous road surveys, instead of relying on specialized commercial equipment.
- Supervised machine learning is used to map the extracted features from sensor signals to discrete road condition classes such as "good", "moderate", "pothole" or "severe damage".
- In this framework, a labelled dataset is first created by driving the bot over different road segments and manually marking the true condition; algorithms like Random Forest, Support Vector Machine or neural networks then learn the relationship between features and labels.
- After training, the model can automatically predict the condition of new segments in real time, providing more frequent and consistent assessments than manual inspection while maintaining good accuracy on unseen data.
- In your project these ideas are combined into one system: Arduino and sensors perform embedded data acquisition on the moving bot, Python libraries handle pre-processing and feature extraction, and machine learning models classify the road quality.
- The output can be stored with GPS coordinates to generate maps of road health, helping authorities prioritize maintenance and enabling students or researchers to experiment with different algorithms and sensor configurations on a practical, real-world problem.

### III. RELATED WORK

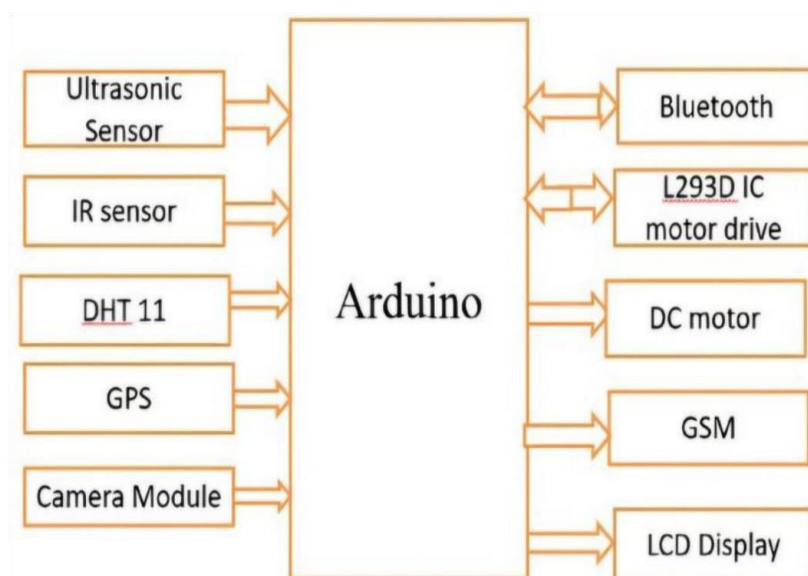
- The concept of road quality assessment using embedded systems and machine learning was first explored by Fan et al., who developed a smartphone-based pothole detection system leveraging accelerometer data and random forest classifiers to achieve high precision in identifying road anomalies through vibration patterns.
- This approach inspired subsequent robotic implementations by integrating Arduino microcontrollers with inertial sensors for real-time monitoring.
- In, researchers proposed an Arduino-based sensing module with G-sensors and Bluetooth connectivity to detect road slopes and accelerations, mapping defective segments via GPS while suggesting ML enhancements for automated classification.
- The exemplar-based pothole detection robot in employed ultrasonic sensors driven by Arduino for obstacle avoidance and autonomous navigation, combining detection with potential filling mechanisms.
- Morphological and filtering techniques for anomaly isolation appear in, where time-frequency features from accelerometers feed into ML models insensitive to vehicle speed variations.
- In, real-time depth sensing via Arduino processes vibration signals with filtering before ML verification of potholes.
- The work in this paper is divided into two stages: 1) Road Anomaly Detection using Arduino accelerometer and morphology-based processing, 2) Quality Classification via Machine Learning models.
- Anomaly detection applies morphological open-close filters on vibration data, followed by gradient edge detection, thresholding, dilation, erosion, and connected component labelling to isolate potholes and cracks.
- Thereafter, extracted features train an ML classifier for road quality grading and mapping.
- In many countries across the world, authorities in charge of maintaining roads surface primarily rely on statistical analyses of the acquired data, visual field inspections, or vehicles outfitted with specialized measurement equipment to keep an eye on roads condition.
- These monitoring techniques are labour-intensive, ineffectual, and time-consuming.
- Furthermore, they usually lack adequate data coverage to provide a complete overview of a road surface status, and they are susceptible to human errors in many situations.

### IV. METHODOLOGY

- Our methodology in this work relies mainly on a user's smartphone.
- Modern mobile phones host a variety of sensors capable of capturing rich data and inferring significant amounts of information.
- This approach is effortless, and simple to develop, without any pre-deployed infrastructure dependencies or heavy investment in specific sensors and hardware.
- Our research plan will be carried over three steps.
- First is the stage of the phone app implementation.
- It consists of developing a smartphone application for data recording.

- The application will be valid for both ios and Android phones.
- The application will allow users to label manually and accurately road segments corresponding to four surface states, namely smooth, bump, potholes, and others.
- For each labelled segment, the application must record multimodal data describing vehicle accelerations (on three-axis as shown in Figure, angular rotations, and geographical positions using the accelerometer, the gyroscope, and the GPS sensors, respectively, of a user's smartphone. ⊥ The application will be available for public download shortly.
- This application will constitute a great tool for researchers to collect in situ labelled data and develop a related machine learning approach.
- Appl. Sci. 2023, 13, 683 5 of 18 Figure 1. Phone orientation inside the vehicle where X, Y, and Z represent the space dimensions.
- Using our developed application, the second step of our research is to collect reliable data that will be available and shared on-demand.
- To automate the road surface inspection, we will explore the best state-of-the-art machine learning models.
- Validated models on large data will be deployed for real-time scanning and online user alerting.
- That's the third step.
- In this step, a new version of the application will be developed to make it more collaborative.
- Users will be able to see other user tags and share their labels.
- This version will be further empowered by trained models deployed for online detection.
- Please note that this paper covers only the first and the second steps of our research plan.
- The architecture of the proposed application is described in detail in the next section.
- Application Architecture and Usage.
- Architecture Overview In this section, we give an overview of the general architecture of the proposed application.
- The application is called "Road Scanner".
- The overall solution architecture is described in Figure.
- To interact correctly with our proposed solution, two interfaces are needed.
- First the smartphone application and second a dedicated website to visualize and export data.
- The process of data storage is as follows: once the data are recorded from the mobile sensors (i.e., accelerometer, gyroscope, and GPS), it is initially stored locally on the smartphone.
- This offline process is useful, especially in the case of a loss of connectivity with the application's server.
- Afterward, the application will be continuously synchronizing records with the server to keep the consistency of the data and to ignore any data replication.
- Indeed, once the connection is established, the stored data are sent to a remote cloud server called "Road Scanner Server" via specified API calls.
- All the received data are then saved to a PostgreSQL database.

### V. BLOCK DIAGRAM



## VI. EXPERIMENTAL RESULTS

Figures shows the results .This section will address the results for every research question addressed in this study. The conclusions of primary research that served as the basis for developing research questions are discussed. There are tables added with the total results for each research topic as well as thorough justifications and analysis. It's important to keep in mind that.

(a)



(b)



It is complex, time-consuming, expensive, and erroneous to effectively use and analyse unstructured data. Unstructured data consists of different types of printed and hand written forms, financial documents, bills, patient records, clinical notes, academic records, blogs, etc. This unstructured data is relevant for further analysis since the information extraction techniques assist in retrieving insightful and important information from this data in various domains, extraction methods are applied. To effectively perform automatic information extraction from business documents, the information extraction methodologies are classified in the current study. Selecting a common information extraction technique might be difficult for many application domains and unstructured documents with complicated or diverse formats. Nowadays in every domain, unstructured documents are generated on a daily basis. It becomes very essential to process this large amount of data and generate a profit out of it. Various domains where unstructured document processing is required with existing research work are discussed in detail in Sect. 3.

## VII. SOFTWARE COMPONENTS

The Arduino IDE is used to write, compile, and upload the embedded C/C++ code that runs on the Arduino board.

- Reads sensors (accelerometer, ultrasonic, IR, etc.) at fixed sampling intervals.
- Performs basic filtering (e.g., moving average) and sends data to Python via serial or stores it on SD card.
- Controls motors and other actuators to move the bot along the road and perform basic navigation.
- Can trigger events such as logging GPS coordinates when a large vibration spike (potential pothole) is detected.
- Python is used on the PC or companion board (e.g., laptop/Raspberry Pi) to handle data processing and machine learning.

## VIII. HARDWARE COMPONENTS

### 8.1 ARDUINO UNO

Arduino UNO is a microcontroller board based on the ATmega328P. It has 14 digital input/output pins. It has 6 analog

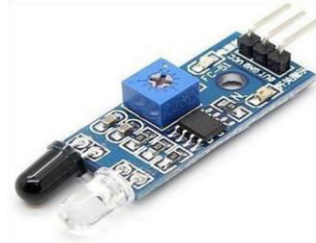


inputs, a 16 MHz ceramic resonator, a USB connection, a power jack, an ICSP header and a reset button. It contains everything needed to support the microcontroller; simply connect it to a computer with a USB cable. It is a Replaceable chip.

### 8.2 IR SENSOR

It has an IR LED (transmitter) and a photodiode (receiver).When an object comes in front of the sensor, the IR light reflects back, and the receiver detects it. The sensor outputs a digital signal LOW (0): Object detected

HIGH (1): No object detected



### 8.3 LCD

The term LCD full form is **Liquid Crystal Display**. The display is named 16×2 LCD because it has 16 Columns and 2 Rows. It can be displayed (16×2=32) 32 characters in total. Each character will be made of 5×8 Pixel Dots. These displays are mainly based on multi-segment light-emitting diodes.



### 8.4 Ultrasonic sensor

The ultrasonic sensor is fixed at the front of the bot, facing the direction of movement and It bounce continuously sends ultrasonic pulses forward. The transmitter emits sound waves they travel in front of the bot if there is an obstacle the waves .back the receiver detects the returning echo.



### 8.5 GPS

Li-Fi Receiver is a custom-design front-end for visible light communications (VLC). It has a wide bandwidth (20 MHz) to support even the most demanding video streaming applications. The receiver features a photo detector with a field-of-view of 170° enabling a robust performance in non-line-of-sight conditions.



### 8.6 ESP32 Camera module

Captures real-time road images while the bot moves. Sends images wirelessly via Wi-Fi or Bluetooth for analysis. Tags each image with GPS location for accurate mapping.



## IX. CONCLUSION

To find opportunities for improvements in this field, the SLR explores current information extraction methods for unstructured document processing. To conduct the PRISMA approach of literature search for this SLR, the guidelines suggested by Kitchen ham and Charters were followed. A total of 68 research papers were ultimately chosen to address the study topics based on inclusion, exclusion, and quality evaluation criteria. The scholars' contributions to publications have increased significantly during the past nine years, it may be said. It illustrates the significance and developments in this field of research study. In this study, we described the "Road Scanner" route application, which enables on-board users to distinguish between four different types of road segments: smooth, bumpy, potholed, and others. The application uses the accelerometer, gyroscope, and GPS integrated into smartphones to record multimodal data for each tagged segment. A data set was gathered and combined with other online datasets to serve as a comparison for the classification methods in order to evaluate the application. In fact, a number of machine learning models were used to map the sensors-data that were gathered to the appropriate route state. Python and well-known data science tools were used to implement all models and evaluation criteria.

## REFERENCES

- [1] Kumar, T.; Acharya, D.; Lohani, D. Modeling IoT Enabled Classification System for Road Surface Monitoring. In Proceedings of the 2022 14th International Conference on Communication Systems & Networks(COMSNETS); January 2022; pp. 836–841.
- [2] Mihoub, A.; Lefebvre, G. Social Intelligence Modeling Using Wearable Devices. In Proceedings of the Proceedings of the 22nd International Conference on Intelligent User Interfaces; Association for Computing Machinery, New York, NY, USA, 7 March 2017; pp. 331–341. <https://doi.org/10.1145/3025171.3025195>.
- [3] Mihoub, A.; Lefebvre, G. Wearables and Social Signal Processing for Smarter Public Presentations. ACM Trans. Interact. Intell. Syst. 2019, 9, 1–24. <https://doi.org/10.1145/3234507>.
- [4] Chugh, G.; Bansal, D.; Sofat, S. Road Condition Detection Using Smartphone Sensors: A Survey. J. Electron. Electr. Eng. 2014, 7, 595–602.
- [5] Siriborvornratanakul, T. An Automatic Road Distress Visual Inspection System Using an Onboard In-Car Camera. Adv. Mul timed. 2018, 2018, e2561953. <https://doi.org/10.1155/2018/2561953>.
- [6] Shaghlil, N.; Khalafallah, A. Automating Highway Infrastructure Maintenance Using Unmanned Aerial Vehicles. In Construction Research Congress 2018; ASCE: Reston, VA, USA, 2018; pp. 486–495. <https://doi.org/10.1061/9780784481295.049>.
- [7] Krichen, M. Anomalies Detection Through Smartphone Sensors: A Review. IEEE Sens. J. 2021, 21, 7207–7217. <https://doi.org/10.1109/jsen.2021.3051931>.
- [8] Sattar, S.; Li, S.; Chapman, M. Road Surface Monitoring Using Smartphone Sensors: A Review. Sensors 2018, 18, 3845. <https://doi.org/10.3390/s18113845>.
- [9] Li, X.; Goldberg, D.W. Toward a mobile crowdsensing system for road surface assessment. Comput. Environ. Urban Syst. 2018, 69, 51–62. <https://doi.org/10.1016/j.compenvurbsys.2017.12.005>.
- [10] Aljaafreh, A.; Alawasa, K.; Aljaafreh, S.; Abadleh, A. Fuzzy Inference System for Speed Bumps Detection Using Smart Phone Accelerometer Sensor. J. Telecommun. Electron. Comput. Eng. 2017, 9, 133–136
- [11] Sattar, S.; Li, S.; Chapman, M. Developing a near real-time road surface anomaly detection approach for road surface monitoring. Measurement 2021, 185, 109990. <https://doi.org/10.1016/j.measurement.2021.109990>.
- [12] Souza, V.M.A.; Giusti, R.; Batista, A.J.L. Asfalt: A low-cost system to evaluate pavement conditions in real-time using smartphones and machine learning Pervasive Mob. Comput. 2018, 51, 121–137. <https://doi.org/10.1016/j.pmcj.2018.10.008>.



## International Journal of Advanced Research in Education and Technology

ISSN: 2394-2975

Impact Factor: 8.152