ORIGINAL PAPER

Artificial Intelligence for road quality assessment in smart cities: a machine learning approach to acoustic data analysis

Senthil Kumar Jagatheesaperumal[1], Simon Elias Bibri[2*], Shrivarshni Ganesan[1] and Poongkalai Jeyaraman[1]

Abstract

In smart cities, ensuring road safety and optimizing transportation efficiency heavily relies on streamlined road condition monitoring. The application of Artificial Intelligence (AI) has notably enhanced the capability to detect road surfaces effectively. This study presents a novel approach to road condition monitoring in smart cities through the development of an acoustic data processing and analysis module. It focuses on four types of road conditions: smooth, slippery, grassy, and rough roads. To assess road conditions, a microphone integrated road surface detector unit is designed to collect audio signals, and an ultrasonic module is used to observe the road depth information. The whole hardware unit is installed in the wheel rim of the vehicles. The data collected from the road surfaces are then analyzed using machine learning algorithms, such as Multi-Layer Perceptron (MLP), Support Vector Machine (SVM), and Random Forest (RF). The results demonstrate the effectiveness of the proposed method in accurately identifying different road conditions. From these results, it was observed that the MLP provides better accuracy of 98.98% in assessing road conditions. The study provides valuable insights into the development of a more efficient and reliable road condition monitoring system for delivering secure transportation services in smart cities.

Keywords Road surface, Surface detector, Acoustic data processing, Artificial intelligence, Machine Learning

1 Introduction

Road condition monitoring plays a crucial role in ensuring road safety and enhancing transportation efficiency in the context of smart cities. Timely detection and accurate assessment of road conditions is essential for the provision of safe and sustainable transportation services, effective infrastructure management, and optimized resource allocation. In recent years, advancements in Artificial Intelligence (AI) and Machine Learning (ML) have provided new opportunities for enhancing smart

city systems, including transport infrastructure (Bibri et al. 2023a, b). This involves improving road condition monitoring systems, enabling more efficient maintenance, and enhancing overall road safety. The integration of AI and ML algorithms in these systems enables the analysis of large volumes of data collected from various sources, such as sensors, cameras, and mobile devices. These algorithms can effectively process, analyze, and interpret complex data patterns, allowing for real time identification of road defects, pavement deterioration, and other hazards. By leveraging AI and ML capabilities, road condition monitoring systems can provide valuable insights for decision making processes, enabling proactive maintenance strategies, optimized resource allocation, and timely interventions to enhance road safety and sustainability.

Road safety remains a crucial concern for developing countries, significantly influencing their development

*Correspondence:
Simon Elias Bibri
simon.bibri@epfl.ch
[1] Department of Electronics and Communication Engineering, Mepco Schlenk Engineering College, Sivakasi 626 005, India
[2] School of Architecture, Civil and Environmental Engineering, Civil Engineering Institute, Visual Intelligence for Transportation, Swiss Federal Institute of Technology in Lausanne (EPFL), 1015 Lausanne, Switzerland

and exerting a substantial impact on public health, consequently affecting global mortality and injury rates. According to the World Health Organization, low- and middle-income nations account for 93% of global road fatalities, as reported by Verma et al. (2022). Every year, more than 1.3 million people die prematurely due to vehicle collisions, and an additional 20 to 50 million people sustain non-fatal injuries, which may lead to long-term disabilities. The safety of road users is heavily influenced by the design and infrastructure of the roads. Therefore, when designing roads, the safety of all road users should be taken into consideration, including road design practices to ensure the safety of all road users.as highlighted by Zheng et al. (2021).

Adverse weather conditions and high traffic volume may often result in suboptimal road conditions. Such weather related hazards include rough surfaces, potholes, slippery roads caused by rain, and uneven pavement in construction zones, all of which are significant causes of vehicle crashes (Omar and Mahdjoubi 2022). Additionally, poorly maintained roads pose a significant risk to drivers and are a leading contributor to vehicle collisions (Oviedo-Trespalacios et al. 2019).

Given the prevalence of traffic collisions caused by damaged road surfaces, a reliable and robust road surface detector is essential in such a scenario. Implementation of such road surface detectors needs to be configured to collect data on the current state of the roads, such as whether they are dry or wet, and transmit the information to both the driver and relevant authorities, as noted by Alonso et al. (2014). One critical application of road surface detectors is in detecting and classifying road cracks based on a large volume of acoustic data collected from road surfaces, using ML algorithms. Based on the collected data, authorities can take appropriate action to repair the roads. In addition to detecting excellent or bad road conditions, road surface detectors can also classify road surfaces as rough, smooth, slippery, and so on, helping drivers prevent vehicle collisions and save lives (Basavaraju et al. 2019).

Numerous datasets, including those available on Kaggle and Ravdess, can be leveraged for road surface detection (Al-refai et al. 2022; Farooq et al. 2020). Various classifiers can also be used for classification purposes. Input for classification can come from a range of sources, such as images from a camera, ultrasonic sensor readings, or acoustic data from sound sensors. Compared to other forms of input, auditory data offers several advantages in terms of its inherent capability to contain more information in less audio duration, higher efficiency, and ease of collection. Moreover, as Li et al. (2015) pointed out, the sound sensors employed for audio signal collection are both efficient and inexpensive, making them easy to maintain. Table 1 provides an overview of different sensors used in road condition monitoring and the corresponding signal processing techniques applied to extract meaningful information from the acquired data.

There are various new technologies utilized for detecting road surfaces, such as image processing and acoustic signal processing. In image processing, a camera is used to capture images of the road, which are then analyzed using comparison algorithms like Convolutional Neural Network (CNN) (Fan et al. 2019), Fully Connected Neural Network (FCN) (Chun and Ryu 2019), and Support Vector Machine (SVM) (Cao et al. 2020), among others, to classify the road into different categories. On the other hand, in acoustic signal processing, audio signals are collected, preprocessed, and transformed into spectrograms and Mel spectrograms, and then analyzed using comparison algorithms to classify the road surface into various categories. Each algorithm has its own inherent level of accuracy characteristics and it can be

Table 1 Comprehensive analysis of various sensors and the signal processing techniques

Sensors	Signal Processing	Description	Strengths	Limitations
Camera	Image Processing	Utilizes visual data captured by cameras	High resolution images for analysis	Limited performance in low light conditions
Microphone	Acoustic Signal Processing	Analyzes audio signals from road surfaces	Real-time monitoring	Susceptible to background noise and interference
GPS	Position Tracking	Determines location and movement of the vehicle	Accurate geolocation data	May experience signal loss in urban canyons
Accelerometer	Motion Sensor	Measures vehicle Acceleration and movement	Provides dynamic motion data	Sensitiveto vibrations and may require calibration
Gyroscope	Motion Sensor	Measures angular velocity and orientation of the vehicle	Offers rotational motion data	Prone to drift and noise over time
Ultrasonic Sensor	Distance Measurement	Uses sound waves to measure distances from obstacles	Non-contact distance sensing	Limited range and accuracy at longer distances
Infrared Sensor	Object Detection	Detects objects based on infrared radiation	Works in low light conditions	Dependent on line of sight for accurate detection

improved by using ensemble learning techniques. Generally, a number of AI models and techniques have been applied to transport and traffic management in the field of smart cities (see Bibri et al. (2023); (Bibri 2023); (Nishant et al. 2020) for the synthesis of many such studies), especially ML and Deep Learning (DL) based on ANN, kNN, CNN, RF, Decision trees (DT), SVM, linear regression, time series models, and others. The same applies to road condition monitoring for different purposes (Al-refai et al. 2022; Almannaa et al. 2023; Ferjani and Alsaif 2022)).

Based on the aforementioned studies, the key observation is that none of the existing works have focused on providing a hardware-based solution for the challenges faced by drivers across various road conditions. Here we highlight the core contributions are summarized as follows:

- A novel approach for road condition monitoring in smart cities through the use of an acoustic data processing module.
- Integrated microphone and an ultrasonic module with the road surface detector unit for collecting audio signals and for observing road depth information.
- Identification and assessment of both the road condition (smooth, slippery, grassy, and rough) and the depth of the crack in the roads.
- Use of ML algorithms such as MLP, SVM, RF, and kNN for analyzing the collected data from various road surfaces.

The purpose of this segment is to explain the overview of the different sections used in this paper. Section 1, gives the outline of the need for a road surface detector due to the prevailing road conditions. In Sect. 3, related works have been studied in order to improvise the model and accuracy of the classified results. The audio collection, preprocessing, training of the dataset by converting to Mel spectrogram, and providing notification about the road conditions to the authorities concerned are explained in Sect. 4. The components of the model are explained under Sect. 4.1, Architecture. The observed results in Sect. 4.2 are the classified output of the model. The classification includes rough roads, smooth roads, slippery roads, and grassy roads. Finally, the advantages, disadvantages, and overview of the output are discussed in Sect. 4.2.1. Figure 1 illustrates the order of steps in the proposed methodology for road classification based on the audio signals obtained from the hardware module mounted on vehicles based on various road conditions.

A sequence of stages in the proposed data processing phase for classification of the audio Signals acquired from the Vehicle Mounted hardware module.

2 Conceptual definitions

Smart cities are urban environments that leverage advanced technologies, especially IoT, data analytics, and AI to enhance the quality of life for citizens, improve resource management, and promote sustainability. These cities use digital technologies and interconnected systems
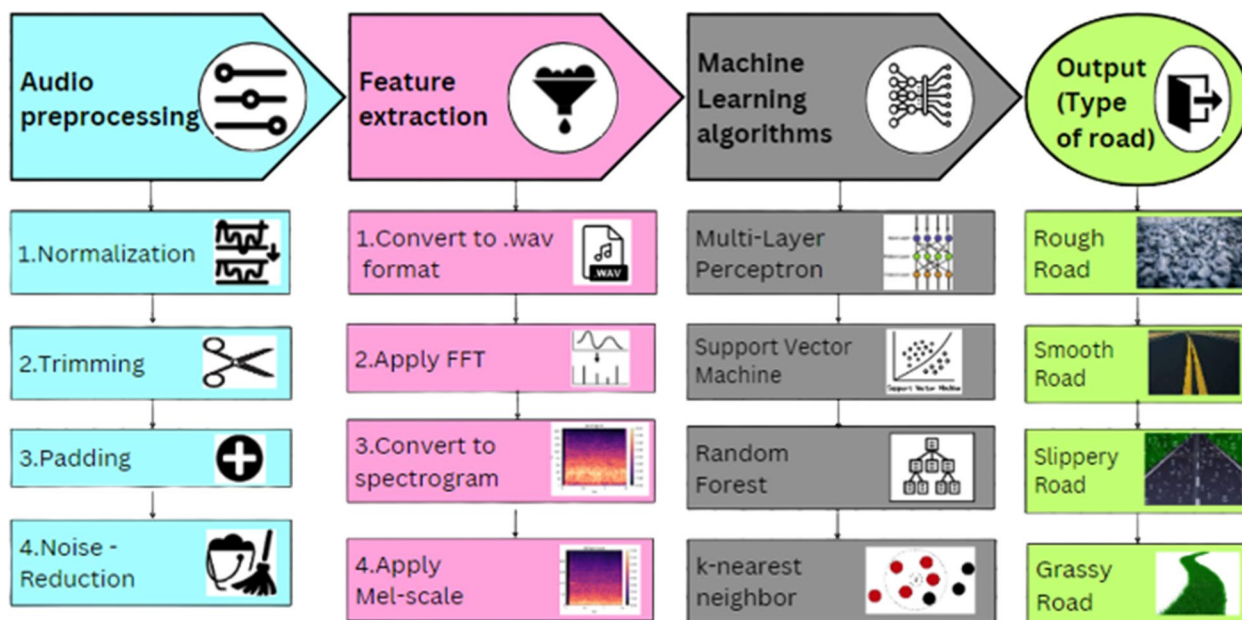

Fig. 1 The sequence of stages in the proposed data processing phase for classification of the audio signals acquired from the vehicle mounted hardware module

to collect and analyze data, enabling informed decision-making, efficient service delivery, and improved urban infrastructure, especially transportation, thanks to urban intelligence functions (Bibri 2020; Bibri and Krogstie 2020). For example, the application of AI and ML in road condition monitoring underscores the development of smart cities, where data-driven technologies contribute to safer, more efficient urban transportation networks and infrastructure.

AI is often referred to as the simulation of human intelligent behavior through the creation of computers or machines capable of such emulation. In this study, AI is defined as "any device/system that perceives its environment and takes actions for its goals" (Poole, Mackworth, and Goebel 1998). Essentially, artificially intelligent machines can learn by acquiring information from their environment (Russell and Norvig 2016), improve their performance through experiential knowledge, and tackle complex tasks akin to human problem solving. Accordingly, AI empowers systems in smart cities to learn from data and adapt their behavior to new information. ML is a subset of AI that focuses on enabling computers to learn and improve from experience without being explicitly programmed. Mitchell (1997) defines it as "a computer program learning from experience 'E' with respect to some class of tasks 'T' and performance measure 'P', if its performance at tasks in 'T' as measured by 'P', improves with experience 'E.'" ML involves the development of algorithms that allow machines to analyze and interpret data, recognize patterns, make predictions, and take actions based on the patterns identified. ML algorithms utilize statistical techniques to iteratively learn from data and improve their performance over time. ML encompasses various techniques, including supervised learning (where models learn from labeled examples), unsupervised learning (where models find patterns in unlabeled data), and reinforcement learning (where models learn through interaction with an environment and feedback mechanisms).

Concerning AI and ML techniques, they encompasses Artificial Neural Networks (ANN), Multilayer Perceptron (MLP), Support Vector Machines (SVM), Linear Regression (LR), Decision Trees (DT), Random Forests (RF), K-Nearest Neighbour (KNN), Adaptive Neuro-Fuzzy Inference System (ANFIS), Batch-Normalization (BN), Convolutional Neural Networks (CNN), Deep Neural Networks (DNN), and Genetic Algorithms (GA).

The integration of AI and ML in road condition monitoring has advanced road condition monitoring in smart cities in terms of assessing the state of road surfaces, identifying defects, and addressing maintenance needs promptly. Traditional methods of road condition monitoring often require manual inspections and periodic assessments, which can be time-consuming and less accurate. AI and ML technologies enable automated and real-time analysis of road conditions using data from various sources (Kafrawy et al. 2021). As exemplified and documented throughout this study, AI and ML algorithms can identify anomalies, predict deterioration patterns, and assess the need for repairs. AI and ML provide several advantages in road condition monitoring within smart cities, including real-time monitoring, accuracy in identifying road defects, predictive maintenance, efficiency of condition assessment, and cost savings on emergency repairs and road infrastructure management.

3 Related works and challenges

This section covers various studies that have explored the use of different machine-learning techniques for road condition classification based on acoustic data, as well as the different types of sensors used for data collection. Also, this section aims to highlight the gaps and limitations of previous research and provide a foundation for the proposed methodology. Table 2 shows the summary of key inferences and challenges observed from recent popular works. After reviewing the existing literature related to road condition monitoring using audio signals, the techniques used, their advantages, and their limitations were analyzed. One such study presented in (Gagliardi et al. 2022), collected raw audio signals from the wheel-road interaction, which were converted into Mel spectrograms. The authors utilized the CNN algorithm for the real-time classification of different road types. The information about the prevailing road conditions was transferred to road authorities via Bluetooth communication. Two models of CNN architecture, the original and quantized, were employed, achieving an accuracy of 93% and 90%, respectively. However, this model was found to be less effective in changing light conditions and environmental noise. Further, it was suggested to development of a smartphone application and a web platform that contains essential information such as road location.

In (Yuan et al. 2020), the authors propose a method for road damage detection using edge and cloud computing. They first collected videos of various road scenes, which were converted into images. The road damage was then detected using a road segmentation algorithm, combining edge and cloud computing. Here, Gray-level co-occurrence matrix features were used for classification in edge computing. Moreover, real-time road information is provided to drivers, resulting in highly accurate and fast damage detection with warnings to drivers using that road. It was observed that such methods require limited labor with reduced time and cost. However, it is not

Table 2 Summary of inferences and challenges from recent popular related works

Ref	Year	Application	Method	Inferences	Challenges
Yuan et al. (2020)	2021	Road surface detection	Edge and cloud computing	Image processing is used for classification	Not compatible for huge volume of road networks
Zhang et al. (2021)	2022	Road surface defect detection	RF, SVM, FCNN	Vibration data from IMUI sensor used for data used for data processing with FCNN	RF shows worst performance
Alhussan et al. (2022)	2022	Pothole and plain road classification	PSDTO, SMOTE, AMDTO	Potholes identified using the images collected and tasks performed by the algorithms	Expensive setup cost
Abdelrauf al (2022)	2022	Real-time weather- related route advisory to drivers	SVM, Naive Bayes, Decision Trees, RF	Rain and road surface detection using CCTV cameras with a pre – trained Vision Transformer model	Limited access to CCTV camera images
Zhao et al. (2022)	2022	Road surface anomaly detection	SVM, CNN	Hough Transform is used for the detection of flaws in the road surface using image processing	Impact on the vibration signals due to the temperature and the vehicle speed

suitable for huge road networks as it needs massive storage and high computational power. To address this issue, a lightweight model can be incorporated, and new learning techniques can be used in the future for better performance with limited data.

In accordance with (Zhao et al. 2022), the TLD framework is employed to detect post disaster road damage by acquiring aerial images and using the spoke wheel operator for generating an initial road template. The TLD framework is utilized to identify and rectify damaged roads based on color invariance, thereby providing significant potential for road damage detection in emergency response and rescue operations following disasters. However, obtaining pre-disaster remote sensing imagery in rural areas is a challenge due to data source mismatch. This approach can be adapted for assessing the level of road damage.

The authors in (Guo et al. 2021) proposed a method to detect road cracks using RGB images by refining the original image gradient. They employed FCN and CNN as deep learning techniques for road surface classification. The model achieved good generalization of various features of cracks and could be applied to smaller images than previous methods. However, it has default weightings and lacks automation. To address these issues, the authors proposed on developing an adaptive trainer mechanism to control the weighting of loss in each output and a validator for automation.

The study in (Zhang et al. 2021) presents the use of an IMU sensor for detecting defects on road surfaces. To do this, vibration data is collected using sensors, and a database is constructed for data storage. Different ML techniques such as RF, SVM, Light Gradient Boosting Machines (LightGBM), and FCNN are applied for damage detection. Additionally, the Extended Kalman Filtering (EKF) algorithm can be utilized to classify road types. The FCNN technique is used to process the vibration data from the IMU sensor, which yields highly precise detection of defects. Future enhancements include joint detection using both camera and IMU sensors and extending its use to different vehicles. The results indicate that the FCNN used in this study performs the best, while RF exhibits the poorest performance.

According to Alhussan et al. (2022), the classification of potholes and plain roads is explained in their study. Image data are collected, augmented, and features are extracted for further classification using various algorithms. This model uses Adaptive Mutation and Dipper-Throated Optimization (AMDTO) for feature selection and optimization of the RF classifier. The proposed AMDTO Algorithm uses ML techniques such as genetic algorithms, Binary PSDTO Algorithm, and Optimized SMOTE Algorithm. The potholes are identified from the images obtained and the algorithms are used to perform respective tasks based on AMDTO and RF. The proposed method achieves an accuracy of 99.795%, which outperforms other approaches such as WOA + RF with 97.5%, GWO + RF with 98.6%, and PSO + RF with 98.1%. However, this model has a major drawback of having an expensive setup cost. The authors in (Wei et al. 2020) proposed a method for automatic road extraction from aerial and satellite images, which involves boosting segmentation, multiple starting points tracing, and fusion. To perform this task, they employed ML techniques such as ANN, SVM, and maximum likelihood in conjunction with DL technology using FCNN and

CNN. The results showed that the FCNN outperformed the CNN by 7% and 40% for the connectivity and completeness indicators, respectively. However, the model's performance can be further improved by using a semi-supervised learning approach that requires fewer training samples for road surface and centerline detection and regularization of road networks. One major limitation of such models are the availability of high-quality labels.

According to the study by Nakashima et al. (2020), road surface detection can be performed using the reflection intensity of an ultrasonic sensor. The intensity variations of these reflections are used to analyze the road surface.

Pattern recognition, regression analysis, and classification are done using the SVM algorithm. They calculate the average of the reflection intensity from the horizontal axis and the standard deviation of the reflection intensity from the vertical axis for accuracy improvement by reducing the overlaps. However, even though the system has a longer measurement time it is useful for visually impaired people, elderly people, and vehicles.

The detection of potholes is accomplished through Road Surface Modeling and Disparity Transformation as explained in (Fan et al. 2019). This technique utilizes a 3D road surface dataset for detection. The accuracy of pothole detection is achieved by comparing actual and modeled disparity maps. The algorithm employed is a novel disparity transformation algorithm and a disparity map modeling algorithm. Otsu's thresholding method is used to extract the undamaged road surfaces from the transformed disparity map. The pothole detection system relies on robust stereo vision with morphological filters used to reduce image noise. The method produces an accuracy of approximately 98.7% and an overall pixel-level accuracy of 99.6%. However, the parameters used for pothole detection are not applicable to all cases. Future research could segment the reconstructed road surfaces into groups of localized planes through a segmentation algorithm.

In their paper (Abdelraouf et al. 2022), Abdelraouf et al. propose a real-time route advisory system that takes into account weather conditions. They use a self-created dataset from roadside CCTV cameras and employ various ML algorithms, including SVM, RF, kNN, Naive Bayes, and Decision Trees, based on both vision-based methodology and sequence-to-sequence technique. The model detects road surface conditions on freeways using traffic CCTV cameras and pretrained Vision Transformer models. The performance of the Vision Transformer is boosted by 5.61% and 5.97% for rain and road conditions, respectively, resulting in an overall F1-score of 96.71% and 98.07%. However, a limitation of the study is the limited number of sequential segments available due to restricted access to adjacent CCTV camera images. The

authors suggest that the model could be improved by testing it on a larger scale.

The Edge Sensing Module and Attention Module, using the RoadNet benchmark, have multiple applications including vehicle navigation, urban planning, intelligent transportation systems, and geographic information systems, as stated in (Liu et al. 2022). These modules improve the perception of road edges by reinforcing perception with the AM guide of the network. The ESM and AM are combined to create an encoder-decoder structure for road surface detection. This technique is supported by a cascaded automatic road detection network, called CasEANet, based on ESM and AM. This method employs a CNN. CasEANet helps solve the issue of unsmooth edges in detecting road conditions. The F1 score, overall accuracy, and balanced error rates of the CasEANet are 0.946, 98.6%, and 0.0219, respectively, which outperform other state-of-the-art models. The drawback is that the method relies on manual annotation, which is time-consuming and requires more labor. To mitigate this, the method can be further developed into automated and accurate road detection techniques.

The Cascaded Multi-Task Road Extraction Network was discussed in (Lu et al. 2022) for extracting road surface, centerline, and edges. The authors used the Deep-Globe road dataset and a large-scale road dataset, the LSCC dataset for this purpose. The authors formed the cascade multitask framework by connecting road surface segmentation (SS), centerline extraction (CE), and edge detection (ED). The LinkNet50 algorithm was used to perform these tasks. The authors employed the methodology of very high-resolution (VHR) remote sensing imagery, topology-aware learning, and hard example mining (HEM) technology. The proposed framework's superiority was ensured by the average path length similarity (APLS) road topology metric, which exhibited the best performance. The main drawback of this model is that the road extraction task is very challenging, and there are numerous road discontinuities. However, advanced methods can be used to enhance the actual application requirements.

The authors of the paper referenced as (Chen et al. 2022) developed an IoT-based system for road icing detection and prediction. The system uses time-series data and employs a model classification approach based on adversarial networks. The proposed deep neural network model, Trans-CGAN, comprises two main components, i.e., imbalanced data classification and time-series prediction. The model employs Long Range Radio (LoRa) and Long Short-Term Memory (LSTM) techniques. The Trans-CGAN model outperforms other existing models in road icing detection. The system can be installed in multiple locations to gather more extensive

Jagatheesaperumal *et al. Computational Urban Science*   (2023) 3:28

Page 7 of 17

road condition data, thereby enhancing the model's adaptability.

The detection of road potholes in real-time using "crowdsourcing" images and GNSS positioning from citizens can be accomplished through vibration assessment, machine vision analysis, and laser scanning techniques. The combination of fast reflectometry and vibration-based methods, along with spatiotemporal trajectory fusion, enables real-time sensing of road potholes. In the article by Chen et al. (2022), a model is proposed that utilizes this approach to achieve real-time ground correction with a high spatial and temporal resolution, providing a useful strategy for road pothole detection. Additionally, the Beidou Grid Code can be implemented to optimize the cost, power consumption, and computational pressure of the geospatial observation system.

The work by Wei et al. (2021). proposes a Scribble-Based Weakly Supervised Deep Learning method for road surface extraction from remote sensing images using centerline detection. The method is applied to the Cheng dataset, Wuhan dataset, and DeepGlobe dataset using weakly supervised learning, road label propagation algorithm, and scribble annotations. The geophysical image processing methodology reduces the requirement for training data. The proposed ScRoadExtractor outperforms classic scribble-supervised segmentation methods by 20%, but there is noise due to the limited capacity of the graph cut method. In another similar work by Doring et al. (2021), a Capacitive Sensor System is used for the wetness quantification of road surfaces using capacitive sensor data. Learning algorithms and optimization criteria are used to classify road surfaces into eight classes. A feature selection algorithm based on a 2×4 planar capacitive transducer array using nearest neighbor methods is employed. The feature average wheel speed improves classifier performance significantly, with the classifier achieving a BAC of 0.93 and the binary version yielding a BAC of 0.998. However, the method is not suitable for high-wheel-speed vehicles. In the future, the feature selection algorithm can be modified to make it compatible with high-speed vehicles.

In Zhao et al.'s (2022) study, Distributed Fiber Optic Sensing (DFOS) was utilized to detect road surface anomalies through image capture. The process involved data collection, Hough transforms, image preparation for training, and classification using both SVM and CNN methods. The model relied on image processing, Hough transforms, and ML. DFOS signals were processed to estimate vehicle speed and detect road surface anomalies, and the study showed that using DFOS signals is feasible for monitoring road surface anomalies. The CNN method was found to outperform LBP and SVM methods in terms of accuracy. However, one major drawback

is that vibration signals may be affected by temperature and vehicle speed. To improve the performance of this technology, it can be combined with sliding window technology.

Liu et al. (2018) discuss the analysis of road networks in complex urban scenes using very high-resolution (VHR) remotely sensed images. The authors utilize a multitask CNN-RoadNet to predict road surfaces, edges, and centerlines from VHR remote sensing images. The paper covers various techniques such as the supervision method, loss function, user interaction, bilinear blending, and training configuration. To handle large VHR images that cannot be holistically trained or tested with finite-GPU resources, the techniques of cropping and bilinear blending approach are employed. Additionally, the proposed user interaction operation effectively eliminates shadows and occlusions along the road regions. This technology can be used in real-world map applications. However, complexity issues in the dataset can be addressed by extracting road topology information from predicted maps.

Pan et al. (2018) discuss the use of unmanned aerial vehicle (UAV) multispectral imagery to detect potholes and cracks in asphalt pavement. The process involves image acquisition and segmentation, preparation of sample data sets and feature selection, and the application of SVM, ANN, and RF algorithms with multi resolution segmentation methodology. The condition of asphalt pavement is monitored using a flexible UAV platform equipped with multispectral remote sensors. The study reports an overall accuracy of 98.3% for the classification of potholes, cracks, and non-distressed pavements using UAV MSI. However, due to spatial resolution limitations, the UAV pavement images used in the study cannot capture cracks with a width of 13.54 mm. To further evaluate the performance of these models and parameters in detecting potholes and cracks, more UAV pavement images of various types of roads are needed.

Daraghmi et al. (2020). propose a road surface evaluation and indexing technique based on crowdsourcing. The authors use vertical acceleration power spectral density to detect road surface roughness and employ blind source separation by the least mean square method to collect and transmit signals, which are then processed to extract useful information. The majority of voting algorithms such as energy-based, weight-based, and score-based methods are used to rank roads, and roughness is indexed by the vibration index. This technology provides accurate, efficient, and cost-effective methods for detecting road surface roughness, even on well-paved roads. The proposed model can be expanded to other countries and the vibration index can be examined under different road conditions in the future.

Table 3 provides a comparative analysis of the proposed scheme with existing schemes along with the literature gaps present in the state-of-the-art schemes.

4 Acoustic processing and analysis

This section provides an overview of the proposed approach and the techniques used for preprocessing, feature extraction, and classification of the audio signals. It also discusses the ML algorithms used for classification and the evaluation metrics used to assess the performance of the proposed approach. The first step in processing acoustic data is converting it into a usable format. Next, the necessary features are extracted from the converted data. These features are then fed into an appropriate machine-learning algorithm, and the resulting classification output is obtained.

4.1 Data set

The input data for the proposed model is derived from audio signals. The audio signals from various sources were collected in real-time and different types of road surfaces were recorded. The collected audio signals were then converted into the Waveform Audio File (WAV) format. Subsequently, these audio signals were classified into four different classes: smooth road, rough road, slippery road, and grassy road. The classified audio signals were then stored in separate folders for future use, and each folder contains a significant amount of data. These datasets serve as input data for further processing.

4.2 Audio preprocessing

Audio signal processing is an essential component in the capture, enhancement, storage, and transfer of audio content. This process involves converting audio signals between analog and digital formats, adjusting frequency ranges, reducing unwanted noise, adding effects, and accomplishing other objectives. Figure 2 refers to the audio preprocessing techniques, such as Normalization, Trimming, Padding, and Noise Reduction, which are utilized in this study. Subsequently, they were applied over the chosen dataset as illustrated in the Fig. 3 observed for sample audio acquired from a road surface.

4.2.1 Normalization

Audio normalization refers to the process of applying a constant amount of gain to an audio recording to adjust the amplitude to a desired level. This technique is commonly used in audio production to ensure that the overall loudness of a track is consistent and optimized for different listening environments (Alonso et al. 2014). Normalization does not affect the signal-to-noise ratio or relative dynamics of the recording because the same amount of boost is applied uniformly across the entire track.

The signal-to-noise ratio or relative dynamics of the audio recording are unaffected by audio normalization. This is due to the consistent distribution of gain applied during normalizing throughout the whole track. As a result, the original dynamics are preserved and uncompromised sound quality is guaranteed. Both the audio signal and the background noise maintain their

Table 3 Comparative analysis of proposed scheme with existing schemes

| References	Core application towards road condition	Description of Algorithm	Performance optimization indicators	Literature gaps
(Astrand et al. 2020)	Underground road condition	Rao Blackwellized extended particle filter and Kalman filter	Large neighborhood search	Only underground mine dataset is used
Rajput et al. 2022)	Infers rough and damaged road segments	Unsupervised learning	Bus trajectory processing	Only trip records four routes are used for analysis
Cui et al. 2019)	Condition monitoring authentication	Certificateless aggregate sign cryption scheme	Mutual authentication, untraceability, and anonymity	Fog computing and VANETs are not widely deployed
Cabral et al. 2019)	Paved and unpaved roads classification	Smartphone inertial sensors	Perceptual linear prediction coefficients	Feature extraction from speech signals not directly reflects road conditions
Jahan et al. 2022)	Paved and unpaved roads classification	Blackbox module	Road data logger	No clear indication of road classifications
Kortmann et al. 2019)	Surface profile assessment	On-board vehicle sensors	Assess the frequency and slope of the signal due to inertia	Lacks acoustic processing modules
Boucetta et al. 2021)	Road deterioration notification	Deep learning framework	Accuracy, edge weights, severity index	Classifies crack types alone not the road conditions
Proposed scheme	Acoustic data analysis to classify rough, smooth, grassy, and slippery roads	MLP, SVM, RF, and KNN	Accuracy, crack depth, and web alert	
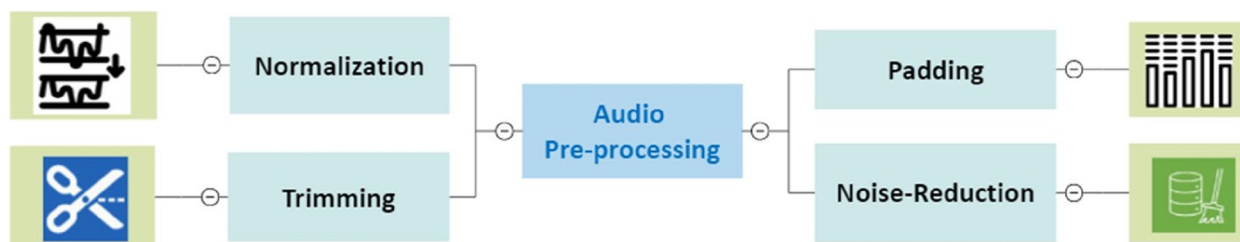
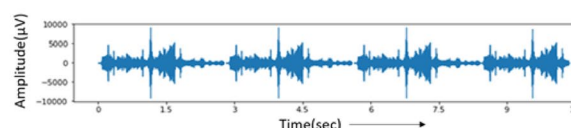**Fig. 2** Audio preprocessing methods for the audio signals collected from the wheel-road interaction
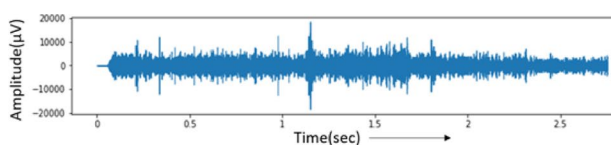
(a) Normalised input audio
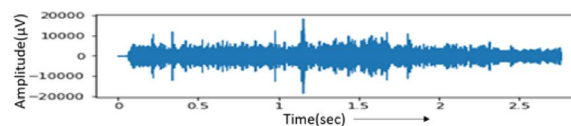
(b) Trimmed audio
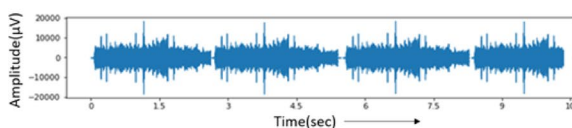
(c) Padded audio

(d) Noise reduced audio

Fig. 3 Audio preprocessing steps for the collected audio signals using the hardware setup

proportional amplitudes. Being non-destructive, it enables modifications and reversibility as necessary and does not change the actual audio data. Audio normalization is essential for maintaining constant volume levels for diverse audio signals acquired from multiple road segments during road condition monitoring. The accuracy and dependability of the road condition categorization are ultimately improved by this method, which prevents abrupt changes in loudness and enables a seamless transition between the audio parts. Additionally, audio normalization maintains audio fidelity by avoiding any clipping or distortion brought on by excessive volume levels, which enhances the experience of monitoring the state of the roads in general.

4.2.2 Trimming
Trimming is a common audio editing technique that involves removing a portion of the audio at the beginning or end of a file. This process is important for improving the flow and pacing of the audio content and ensuring that it starts and ends smoothly. Additionally, trimming can be used to remove unwanted noise or silence from the audio signal, which can enhance its clarity and overall quality (Das et al. 2022).

In the context of road surface analysis, audio trimming can be used to remove unwanted noise or silence from the recorded vehicle sounds before processing the signal. This can help to improve the accuracy and reliability of the analysis by reducing the impact of external factors that may affect the signal. Furthermore, trimming can be combined with other audio processing techniques such as filtering and normalization to further enhance the quality of the audio signal and improve the accuracy of the road surface analysis.

4.2.3 Padding
The process of extending trimmed audio is known as audio padding. Audio padding is used to ensure that all audio signals are of equal length, which is an essential requirement for many audio analysis applications. In this process, the audio sample is repeated multiple times until the desired duration is achieved (Ahsan et al. 2019). Padding helps in avoiding any loss of information due to different lengths of audio signals and ensures that the analysis is consistent across all signals. This technique is widely used in various audio applications, including road surface analysis, to achieve accurate and reliable results.

4.2.4 Noise reduction

The process of noise reduction is vital in improving the quality of the audio signals obtained from road surfaces. It helps to eliminate unwanted sounds such as background noise, wind noise, and other sources of noise that may interfere with the analysis of the road surface signals. By removing these noises, it becomes easier to detect and analyze the road surface anomalies such as potholes and cracks. The noise reduction process involves various techniques such as filtering and equalization, which help to suppress or eliminate unwanted sounds from the recorded audio signals (Guo et al. 2020). This ultimately leads to a more accurate and reliable analysis of the road surface conditions, which is crucial for ensuring safe and efficient driving.

4.3 Feature extraction

To begin with, the dataset is filtered to contain only WAV format data to ensure consistency. The transformation of the dataset is done using Fast Fourier Transform, which is faster and more efficient, converting the time domain representation to the frequency domain representation. The next step is to create a spectrogram of the frequency domain signal to analyze the features present. The spectrogram provides a visual representation of the signal strength. The spectrogram is then converted to a Mel Spectrogram by utilizing the Mel scale. The Mel scale filters the audio signals, which are time-windowed, to highlight the relevant features. By using this technique, a clearer representation of the road surface conditions can be obtained from the audio signals. Additionally, the Mel Spectrogram can be further analyzed using deep learning techniques to classify and identify specific types of road surface conditions.

4.4 Machine learning algorithms

The classification of road types is performed using four different machine-learning algorithms. These algorithms take the feature extracted from the audio signal as input and classify the road surfaces into four types. The algorithms employed include MLP, SVM, RF, and kNN. The MLP is a widely used algorithm for parallel distributed processing, computational neuroscience, and supervised learning. SVM is a supervised ML technique that is used for classification and regression. RF is another supervised ML algorithm that is commonly employed in classification and regression problems. KNN, on the other hand, is a supervised learning classifier that uses proximity to make predictions about the grouping of a single data point.

The choice of MLP, SVM, RF, and k-NN algorithms have proven successful in managing challenging classification problems and they are frequently employed.

Their appropriateness for assessing road conditions is supported by prior investigations and benchmarking trials (Kafrawy et al. 2021). Their adaptability in collecting various patterns and features in acoustic data acquired from the road surfaces is in line with the dataset's nature. Initial tests verified their effectiveness on the chosen dataset, hence validating their choice. It follows that the selection of MLP, SVM, RF, and KNN creates a solid foundation for the road condition monitoring framework and contributes to accurate and trustworthy classification findings.

4.5 Classification output

The classifier gets the extracted features of the audio signals as the input by applying the aforementioned algorithms. The road surfaces are classified into four types: rough road, smooth road, slippery road, and grassy road according to the conditions applied to the audio signals. This classification output further leads to measuring the depth of the crack if the road is detected as rough.

Optimizing the performance of ML models requires parameter adjustment. Each of the models utilized in our study—MLP, SVM, RF, and KNN—has had its parameters carefully tuned. We were able to pinpoint the ideal settings that optimize the prediction power of the models by methodically examining various combinations of hyperparameters. The methods section of the details on the precise parameter values utilized for each model. By ensuring transparency and reproducibility, we make it possible to carry out similar tests and confirm the efficacy of our strategy.

5 Experimental setup

The hardware components necessary for data collection and transmission to the cloud are crucial, and their systematic workflow is outlined here. This section discusses the essential hardware requirements for data collection and transmission to the cloud. Additionally, the systematic workflow of the surface detection system for identifying the type of road surface on which the vehicle is traveling is described.

As illustrated in Fig. 4, the surface detection system identifies the type of road surface on which the vehicle is moving. To accomplish this, the open-source microcontroller Arduino UNO is employed because it is low-cost, flexible, and simple to program. Ultrasonic sensors, sound sensors, and GPS modules are connected to the microcontroller to gather data for subsequent processing. The ESP8266 Wi-Fi module is used to transmit the collected data to the cloud for further processing of the audio signal. In the cloud, the gathered data is collected and saved in a database as a reference for real-time classification. The outcome is then presented to road authorities for inspection. To collect
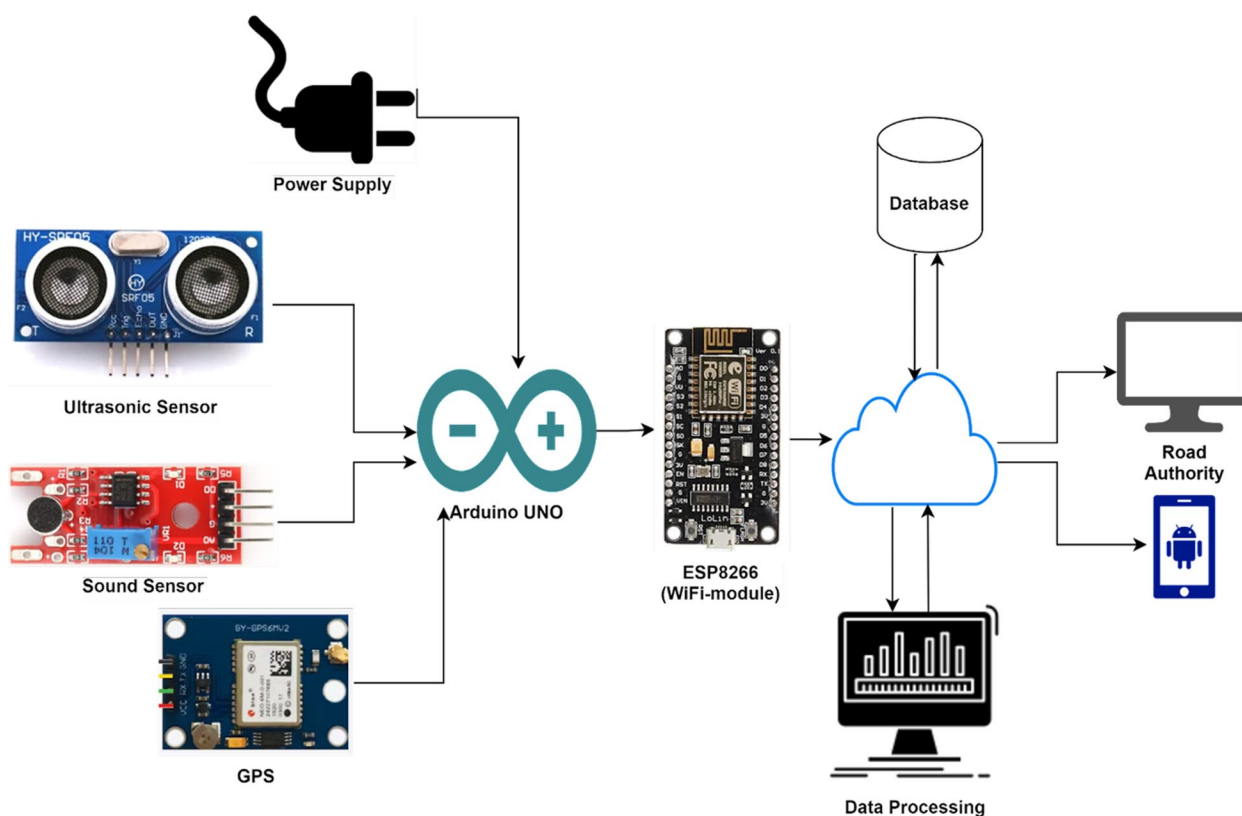
Fig. 4 Functional blocks in the road surface identification system, showcasing the integration of core hardware modules

data for further processing, the open-source microcontroller, Arduino UNO, is used to interface with ultrasonic sensors, sound sensors, and GPS modules. The collected data is then transmitted to the cloud using the ESP8266 Wi-Fi module for further processing of the audio signal. The corresponding hardware setup of the deployed system is shown in Fig. 5.

In the cloud, the collected data is stored in a database for future reference, which can be used as a reference in real-time for classification. The result is then displayed to the road authorities. The proposed model enables the road authorities to be notified of the road condition, based on the type of road surface observed, and take necessary actions accordingly. The road surface classification is performed using the specified algorithms, and the output is sent to the authorities through the cloud. With the help of this model, the road authorities can obtain updates on the road surface and take timely actions to resolve the problem, without the need for human intervention. This can lead to efficient and effective road maintenance, resulting in safer and smoother transportation for the public. Additionally, the implementation of such a system raises concerns about hardware requirements for real-time road surface analysis. This necessitates the selection of the appropriate microcontroller, sensors, and modules to

ensure reliable data collection and transmission. Furthermore, the system's efficiency can be improved by optimizing the hardware's power consumption to extend its lifespan. Finally, the hardware's durability and stability in harsh weather conditions must be considered to ensure the system's continuous operation.

6 Evaluation results and discussion

This section provides a detailed analysis of the experimental results and a discussion of the findings. The results are compared with the state-of-the-art methods to show the effectiveness of the proposed approach.

The results of each stage of the model are presented and discussed in this section. The collected audio signal is in WAV format and is normalized as shown in Fig. 3a. To enhance the signal features, the normalized audio is multiplied by a gain. The trimmed audio signal is obtained by removing the silence in the signal, as shown in Fig. 3b, which sharpens the audio signal. In order to maintain signals in the same time limit, padding is performed, as depicted in Fig. 3c. The audio signal is repeated during padding instead of adding zeroes to make it more effective and efficient. The noise-reduced signal is illustrated in Fig. 3d, where the external noise in the audio signal is eliminated using a noise reduction method to prevent the
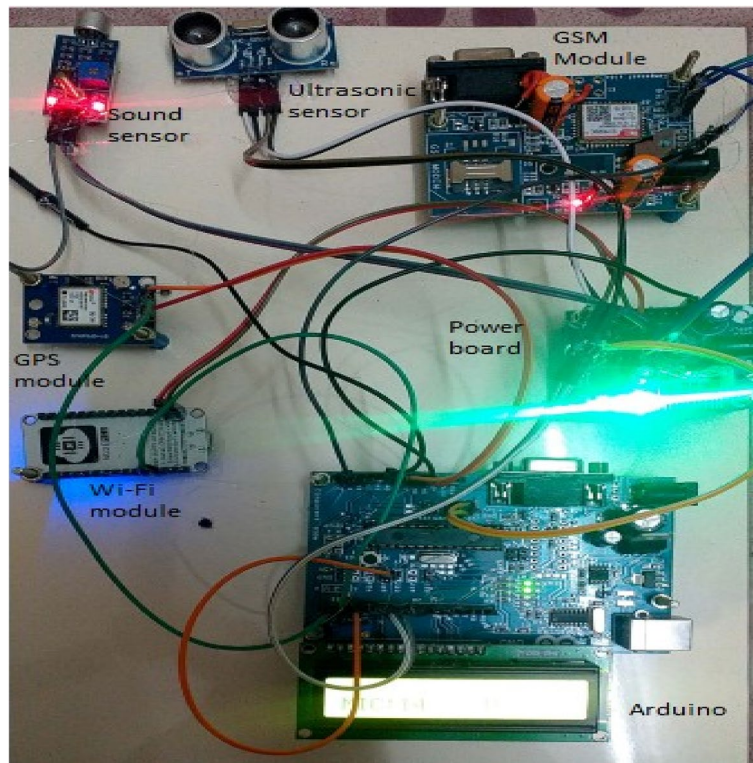
Fig. 5 Hardware setup deployed in the vehicles for assessing various road conditions

information in the audio signal from being corrupted by unwanted noise.

Figures 6a, 7a, 8a, and 9a show the audio signals collected from rough, smooth, slippery, and grassy roads, respectively. For each audio signal, its corresponding spectrogram is shown in Figs. 6b, 7b, 8b, and 9b, which is then converted into Mel spectrogram by applying the Mel scale and shown in Figs. 6c, 7c, 8c, and 9c, respectively.

The detector model is fixed on the vehicle's rim to monitor and observe the condition of road surfaces. An ultrasonic sensor is used to measure the depth of any detected cracks in centimeters. Placed in the rim of the

vehicle integrated with the microcontroller module, the sensor measures both the depth of the crack and the rim's radius. If the depth is significant, a message is sent to the road authorities. The detector's results are outlined below: Fig. 10a shows the surface where the detector was tested, and it indicates the result identified as a smooth surface. Figure 10b displays the readings of the ultrasonic and sound sensors on the LCD screen. This result can also be viewed on the cloud platform as shown in Fig. 10c through the transmission of information using a Wi-Fi module. Similarly, Fig. 11a shows the surface where the detector detected as a slippery surface, and the

(a) Audio Signal　　　　　(b) Spectrogram　　　　　(c) Mel Spectrogram
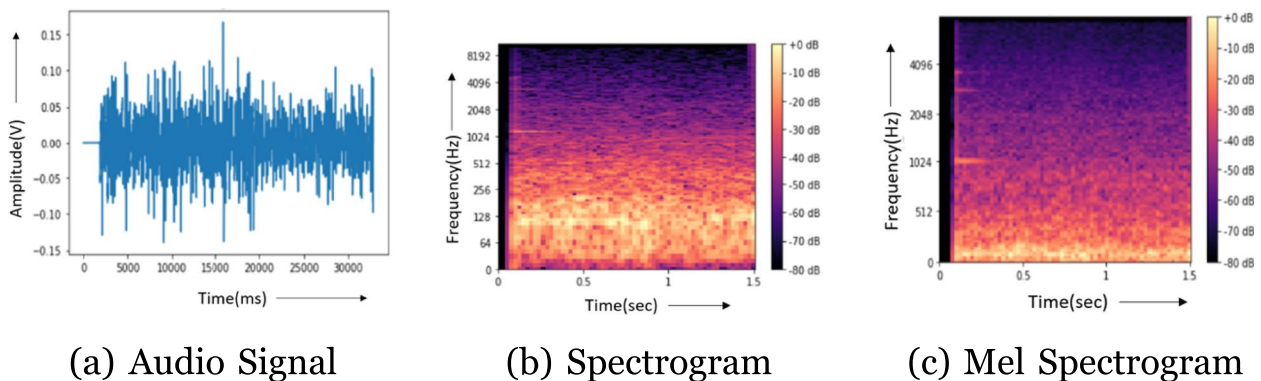
Fig. 6 Rough road surface analysis extracted from the collected audio signal
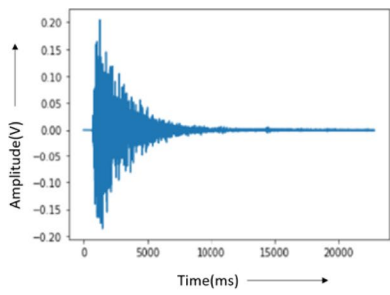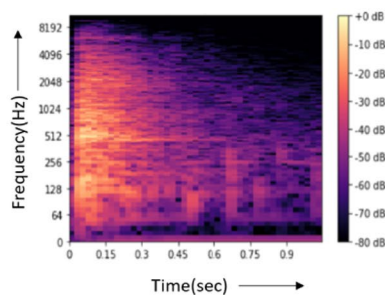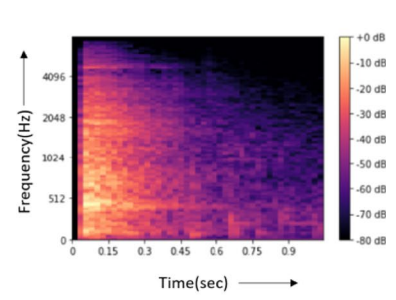
(a) Audio Signal   (b) Spectrogram   (c) Mel Spectrogram
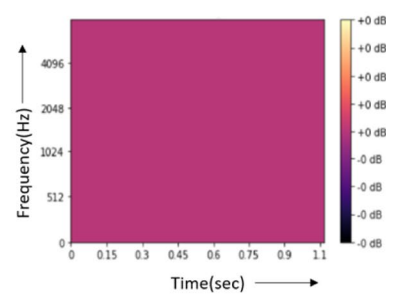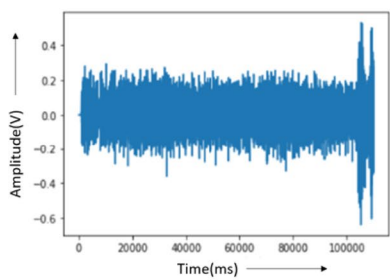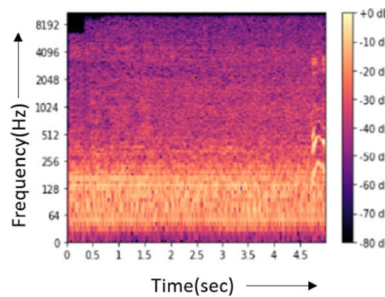
Fig. 7 Smooth road surface analysis extracted from the collected audio signal
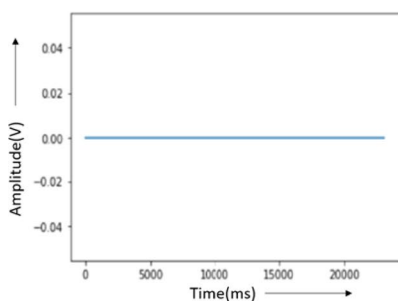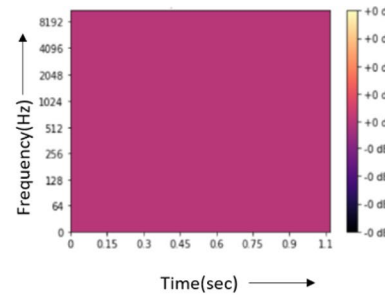
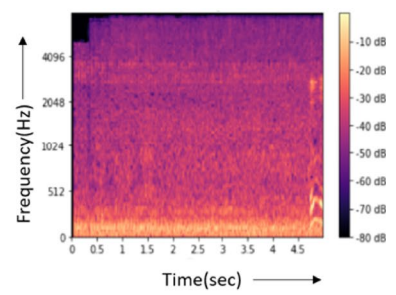(a) Audio Signal   (b) Spectrogram   (c) Mel Spectrogram

Fig. 8 Slippery road surface analysis extracted from the collected audio signal
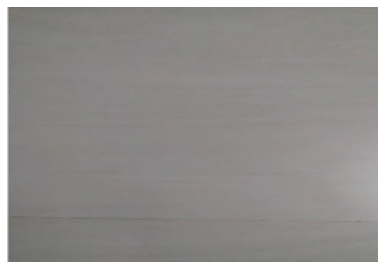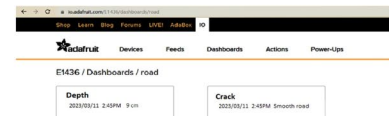

(a) Audio Signal   (b) Spectrogram   (c) Mel Spectrogram

Fig. 9 Grassy road surface analysis extracted from the collected audio signal
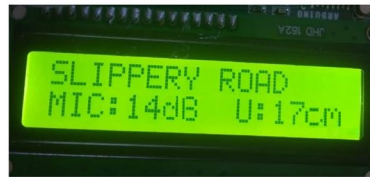

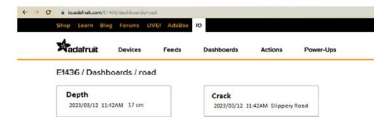(a) Smooth Surface   (b) Output on the hardware   (c) Web page output

Fig. 10 Smooth road output obtained from the classification result

(a) Slippery Surface

(b) Output on the hardware

(c) Web page output

Fig. 11 Slippery road output obtained from the classification result

corresponding readings of the ultrasonic and sound sensors, along with the type of road condition, are displayed on the LCD screen (Fig. 11b). This result is also transmitted to the cloud platform as shown in Fig. 11c. Figure 12a displays the surface that was detected as a grassy road, and the readings of the ultrasonic and sound sensors are shown on the LCD screen (Fig. 12b). The results are also transmitted to the cloud platform as shown in Fig. 12c. Figure 13a shows the surface where the detector detected a rough surface. The output on the LCD screen (Fig. 13b) displays the readings of the ultrasonic and sound sensors along with the depth of the crack. This result is also transmitted to the cloud platform (Fig. 13c) using a Wi-Fi module. If the detected surface is a rough road and the crack depth is maximum, an alert message is sent to the road authorities to take timely action.

The performance of four ML algorithms in determining crack depth from acoustic signals are compared in Table 4. Accuracy, precision, recall, F1-score, area under the ROC curve (AUC), mean absolute error (MAE), and root mean square error (RMSE) are some of the validation criteria. The maximum accuracy, precision, recall, F1-score, and AUC are displayed by the RF-based technique, suggesting superior performance in crack depth detection. The RF-based technique also exhibits the lowest MAE and RMSE, suggesting that it can make predictions that are more accurate. The MLP-based technique, in contrast, performs admirably, with high accuracy and AUC, whereas SVM and KNN exhibit marginally worse performance. The RF-based strategy seems to be the most promising for crack depth identification, based on the observed results.
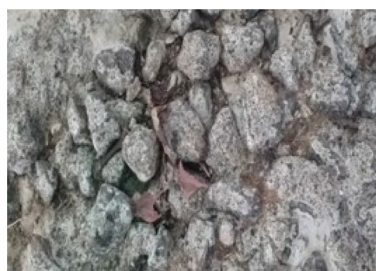

(a) Grassy Surface

(b) Output on the hardware
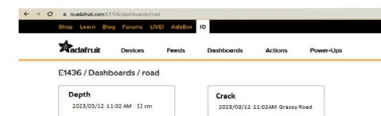
(c) Web page output

Fig. 12 Grassy road output obtained from the classification result

(a) Rough Surface

(b) Output on the hardware

(c) Web page output

Fig. 13 Rough road output obtained from the classification result

Table 4 Comparison of crack depth detection using Acoustic signals with different ML algorithms

Validation Parameters	MLP	SVM	RF	KNN
Accuracy	0.85	0.81	0.88	0.79
Precision	0.87	0.83	0.89	0.81
Recall	0.83	0.79	0.87	0.77
F1-score	0.85	0.80	0.88	0.78
Area Under ROC Curve (AUC)	0.92	0.87	0.94	0.85
Mean Absolute Error (MAE)	0.045	0.052	0.040	0.056
Root Mean Square Error (RMSE)	0.067	0.075	0.062	0.078

The confusion matrices of four algorithms and their corresponding accuracies are presented in this section. Figure 14a shows the confusion matrix of the MLP algorithm, which achieves an accuracy of 98.98%. The SVM algorithm has an accuracy of 89.80%, and its confusion matrix is shown in Fig. 14b. The confusion matrix for the RF algorithm, which has an accuracy of 97.96%, is presented in Fig. 14c. Lastly, Fig. 14d displays the confusion matrix of the KNN algorithm, which achieves an accuracy of 96.94%. Based on the comparison of these four algorithms, MLP outperforms the others with the highest accuracy. Table 5 describes the accuracy variation in testing and training data size for all four algorithm types: MLP, SVM, RF, and KNN. Four combinations of size variation have been tried and among all, it is proved that MLP produces better results in all the combinations.
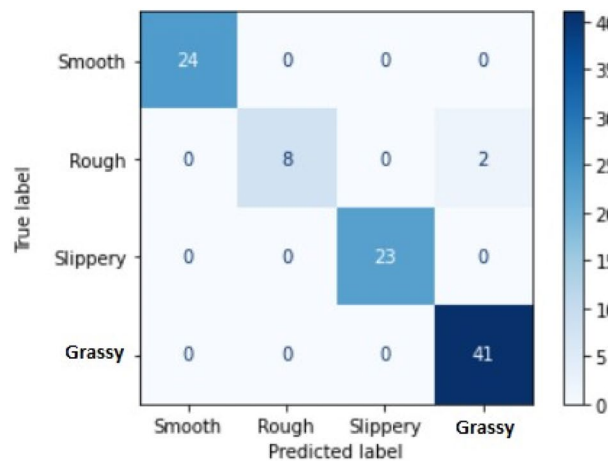
7 Conclusion

This paper presents a novel approach for road condition monitoring in smart cities using an acoustic data processing module integrated with the vehicle wheel rim. By collecting audio signals and road depth information using
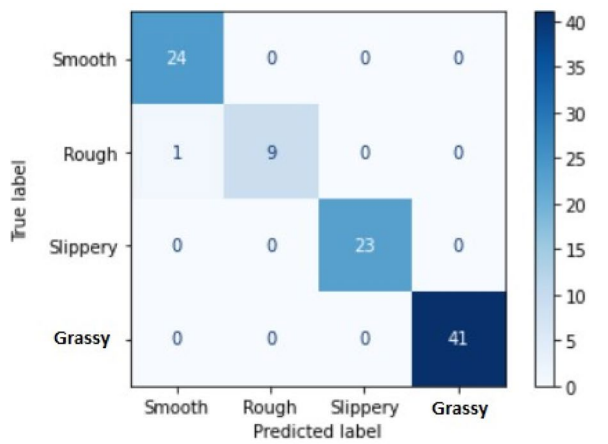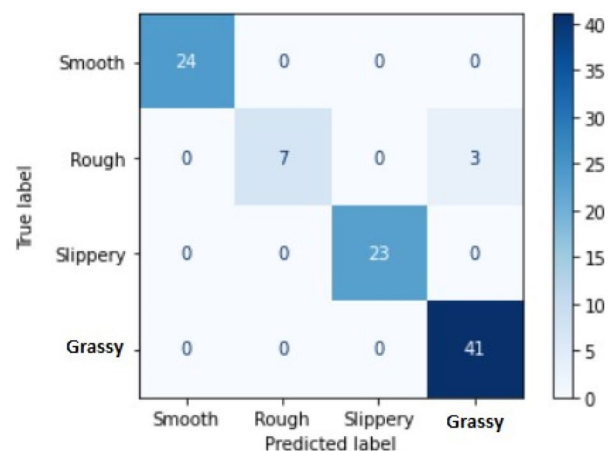

(a) MLP
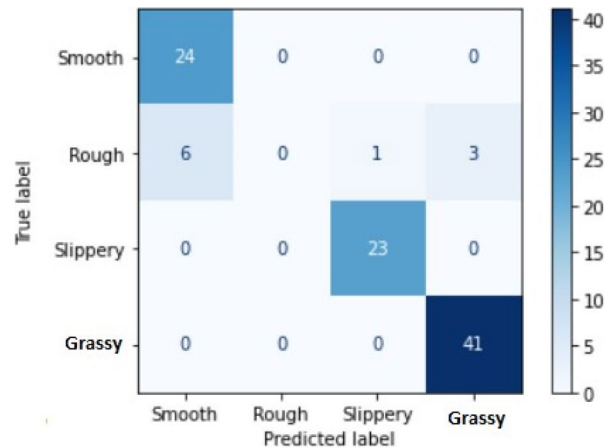
(b) SVM

(c) RF

(d) KNN

Fig. 14 Confusion matrix of various algorithms for assessing the performance

Table 5 Accuracy variation in training and testing data size for various algorithms

Training size (%)	Testing size (%)	MLP (%)	SVM (%)	RF (%)	KNN (%)
80	20	99.90	93.59	99.90	98.72
75	25	98.98	89.80	97.96	96.94
70	30	97.44	88.89	96.58	95.73
60	40	96.79	90.38	97.44	95.51

an ultrasonic module, the road surface is classified into four types: smooth, slippery, grassy, and rough roads. ML algorithms such as MLP, SVM, RF, and KNN were used to classify the road surfaces, with MLP providing the best accuracy of 98.98%. Compared to conventional methods, this model is more cost-effective, accurate, and less labor-intensive. If the road is classified as rough and has a maximum depth of the crack, an alert message is sent to the concerned road authorities, along with the location of the damaged road. This helps authorities take timely actions to resolve the problem, potentially saving many lives from vehicle crashes caused by damaged roads. In the future, this model could enable smart city plans using vehicle-to-vehicle communication, allowing approaching vehicles to be informed about damaged roads and to take alternative routes if necessary.

Declarations

References

Abdelraouf, A., Abdel-Aty, M., & Wu, Y. (2022). Using vision transformers for spatial-context-aware rain and road surface condition detection on freeways. *IEEE Transactions on Intelligent Transportation Systems, 23*(10), 18546–18556.

Ahsan, M.I., Kertesz, C., Mesaros, A., Heittola, T., Knight, A., Virtanen, T.: Audio-based epileptic seizure detection. In: 2019 27th European Signal Processing Conference (EUSIPCO), pp. 1–5 (2019). IEEE

Alhussan, A. A., Khafaga, D. S., El-Kenawy, E.-S.M., Ibrahim, A., Eid, M. M., & Abdelhamid, A. A. (2022). Pothole and plain road classification using adaptive mutation dipper throated optimization and transfer learning for self driving cars. *IEEE Access, 10*, 84188–84211.

Almannaa, M., Zawad, M.N., Moshawah, M., Alabduljabbar, H.: Investigating the effect of road condition and vacation on crash severity using machine learning algorithms. International journal of injury control and safety promotion, 1–11 (2023)

Alonso, J., L´opez, J., Pav´on, I., Asensio, R. M., Arcas, C., & Bravo, A. (2014). On-board wet road surface identification using tyre/road noise and support vector machines. *Applied acoustics, 76*, 407–415.

Al-refai, G., Elmoaqet, H., & Ryalat, M. (2022). In-vehicle data for predicting road conditions and driving style using machine learning. *Applied Sciences, 12*(18), 8928.

Astrand, M., Jakobsson, E., Lindfors, M., & Svensson, J. (2020). A system for underground road condition monitoring. *International Journal of Mining Science and Technology, 30*(3), 405–411.

Basavaraju, A., Du, J., Zhou, F., & Ji, J. (2019). A machine learning approach to road surface anomaly assessment using smartphone sensors. *IEEE Sensors Journal, 20*(5), 2635–2647.

Bibri, S. (2020). Data-driven smart sustainable cities: a conceptual framework for urban intelligence functions and related processes, systems, and sciences. In: *Advances in the Leading Paradigms of Urbanism and their Amalgamation. Advances in Science, Technology & Innovation*. Springer, Cham. https://doi.org/10.1007/978-3-030-41746-8_6

Bibri, S. E., & Krogstie, J. (2020). Environmentally data-driven smart sustainable cities: applied innovative solutions for energy efficiency, pollution reduction, and urban metabolism. *Energy Inform, 3*, 29. https://doi.org/10.1186/s42162-020-00130-8

Bibri, S. E., Alexandre, A., Sharifi, A., & Krogstie, J. (2023a). Environmentally sustainable smart cities and their converging ai, iot, and big data technologies and solutions: An integrated approach to an extensive literature review. *Energy Informatics*, 6(1), 9.

Bibri, S. E., Krogstie, J., Kaboli, A., & Alahi, A. (2023b). Smarter eco-cities and their leading-edge artificial intelligence of things solutions for environmental sustainability: a comprehensive systemic review. *Environmental Science and Ecotechnology,* in press.

Boucetta, Z., Fazziki, A.E., Adnani, M.E.: A deep-learning-based road deterioration notification and road condition monitoring framework. International Journal of Intelligent Engineering & Systems 14(3) (2021)

Cabral, F. S., Fukai, H., & Tamura, S. (2019). Feature extraction methods proposed for speech recognition are effective on road condition monitoring using smartphone inertial sensors. *Sensors, 19*(16), 3481.

Cao, W., Liu, Q., & He, Z. (2020). Review of pavement defect detection methods. *Ieee Access, 8*, 14531–14544.

Chen, D., Chen, N., Zhang, X., & Guan, Y. (2022). Real-time road pothole mapping based on vibration analysis in smart city. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 15*, 6972–6984.

Chen, Z., Xiong, G., Sun, Y., Li, Y., & Li, Y. (2022). An internet-of-things-enabled system for road icing detection and prediction. *IEEE Internet of Things Journal, 9*(20), 20257–20269.

Chun, C., & Ryu, S.-K. (2019). Road surface damage detection using fully convolutional neural networks and semi-supervised learning. *Sensors, 19*(24), 5501.

Cui, M., Han, D., & Wang, J. (2019). An efficient and safe road condition monitoring authentication scheme based on fog computing. *IEEE Internet of Things Journal, 6*(5), 9076–9084.

D¨oring, J., Beering, A., Scholtyssek, J., Krieger, K.-L.: Road surface wetness quantification using a capacitive sensor system. IEEE Access 9, 145498–145512 (2021)

Daraghmi, Y. A., Wu, T. H., & I˙k, T. U. (2020). Crowdsourcing-based road surface evaluation and indexing. *IEEE Transactions on Intelligent Transportation Systems, 23*(5), 4164–4175.

Das, S.B., Swain, D., Das, D.: A graphical user interface (gui) based speech recognition system using deep learning models. In: Advances in Distributed Computing and Machine Learning: Proceedings of ICAD- CML 2022, pp. 259–270. Springer, ??? (2022)

El Kafrawy, P., Fathi, H., Qaraad, M., Kelany, A. K., & Chen, X. (2021). An efficient svm-based feature selection model for cancer classification using high-dimensional microarray data. *IEEE Access, 9*, 155353–155369.

Fan, R., Bocus, M.J., Zhu, Y., Jiao, J., Wang, L., Ma, F., Cheng, S., Liu, M.: Road crack detection using deep convolutional neural network and adaptive thresholding. In: 2019 IEEE Intelligent Vehicles Symposium (IV), 2019;474–479. IEEE

Fan, R., Ozgunalp, U., Hosking, B., Liu, M., & Pitas, I. (2019). Pothole detection based on disparity transformation and road surface modeling. *IEEE Transactions on Image Processing, 29*, 897–908.

Farooq, M., Hussain, F., Baloch, N. K., Raja, F. R., Yu, H., & Zikria, Y. B. (2020). Impact of feature selection algorithm on speech emotion recognition using deep convolutional neural network. *Sensors, 20*(21), 6008.

Ferjani, I., & Alsaif, S. A. (2022). How to get best predictions for road monitoring using machine learning techniques. *PeerJ Computer Science, 8*, 941.

Gagliardi, A., Staderini, V., & Saponara, S. (2022). An embedded system for acoustic data processing and ai-based real-time classification for road surface analysis. *IEEE Access, 10*, 63073–63084.

Guo, J.-M., Markoni, H., & Lee, J.-D. (2021). Barnet: Boundary aware refinement network for crack detection. *IEEE Transactions on Intelligent Transportation Systems, 23*(7), 7343–7358.

Guo, Q., Ye, J., Chen, Y., Hu, Y., Lan, Y., Zhang, G., & Li, X. (2020). Inor—an intelligent noise reduction method to defend against adversarial audio examples. *Neurocomputing, 401*, 160–172.

Jahan, I.A., Jamil, I.A., Fahim, M.S.H., Huq, A.S., Faisal, F., Nishat, M.M.: Accident detection and road condition monitoring using blackbox module. In: 2022 Thirteenth International Conference on Ubiquitous and Future Networks (ICUFN), pp. 247–252 (2022). IEEE

Kortmann, F., Peitzmeier, H., Meier, N., Heger, J., Drews, P.: Enabling road condition monitoring with an on-board vehicle sensor setup. In: 2019 IEEE SENSORS, pp. 1–4 (2019). IEEE

Li, M., van Keulen, W., Tijs, E., van de Ven, M., & Molenaar, A. (2015). Sound absorption measurement of road surface with in situ technology. *Applied Acoustics, 88*, 12–21.

Liu, D., Zhang, J., Liu, K., & Zhang, Y. (2022). Aerial remote sensing image cascaded road detection network based on edge sensing module and attention module. *IEEE Geoscience and Remote Sensing Letters, 19*, 1–5.

Liu, Y., Yao, J., Lu, X., Xia, M., Wang, X., & Liu, Y. (2018). Roadnet: Learning to comprehensively analyze road networks in complex urban scenes from high-resolution remotely sensed images. *IEEE Transactions on Geoscience and Remote Sensing, 57*(4), 2043–2056.

Lu, X., Zhong, Y., Zheng, Z., Chen, D., Su, Y., Ma, A., & Zhang, L. (2022). Cascaded multi-task road extraction network for road surface, centerline, and edge extraction. *IEEE Transactions on Geoscience and Remote Sensing, 60*, 1–14.

Mitchell, T. M. (1997). *Machine learning*. McGraw-Hill Science/Engineering/Math.

Nakashima, S., Arimura, H., Yamamoto, M., Mu, S., & Lu, H. (2020). Improving the accuracy of road surface distinction based on reflection intensity variations using ultrasonic sensor. *IEEE Sensors Journal, 22*(18), 17399–17405.

Nishant, R., Kennedy, M., & Corbett, J. (2020). Artificial intelligence for sustainability: Challenges, opportunities, and a research agenda. *International Journal of Information Management, 53*, 102104.

Omar, H., Mahdjoubi, L.: Practical solutions for improving the sub-optimal performance of construction projects using dubai construction projects as an example. Engineering, Construction and Architectural Management (2022)

Oviedo-Trespalacios, O., Truelove, V., Watson, B., & Hinton, J. A. (2019). The impact of road advertising signs on driver behaviour and implications for road safety: A critical systematic review. *Transportation research part A: Policy and practice, 122*, 85–98.

Pan, Y., Zhang, X., Cervone, G., & Yang, L. (2018). Detection of asphalt pavement potholes and cracks based on the unmanned aerial vehicle multispectral imagery. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 11*(10), 3701–3712.

Poole, D., Macworth, A., & Goebel, R. (1998). *Computational Intelligence: a logical approach*. New York: Oxford University Press.

Rajput, P., Chaturvedi, M., & Patel, V. (2022). Road condition monitoring using unsupervised learning based bus trajectory processing. *Multimodal Transportation, 1*(4), 100041.

Russell, S. J., & Norvig, P. (2016). *Artificial intelligence: a modern approach*. Pearson Education Limited.

Verma, A., Sasidharan, S., Bhalla, K., & Allirani, H. (2022). Fatality risk analysis of vulnerable road users from an indian city. *Case studies on transport policy, 10*(1), 269–277.

Wei, Y., & Ji, S. (2021). Scribble-based weakly supervised deep learning for road surface extraction from remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing, 60*, 1–12.

Wei, Y., Zhang, K., & Ji, S. (2020). Simultaneous road surface and centerline extraction from large-scale remote sensing images using cnn-based segmentation and tracing. *IEEE Transactions on Geoscience and Remote Sensing, 58*(12), 8919–8931.

Yuan, Y., Islam, M. S., Yuan, Y., Wang, S., Baker, T., & Kolbe, L. M. (2020). Ecrd: Edge-cloud computing framework for smart road damage detection and warning. *IEEE Internet of Things Journal, 8*(16), 12734–12747.

Zhang, Y., Ma, Z., Song, X., Wu, J., Liu, S., Chen, X., & Guo, X. (2021). Road surface defects detection based on imu sensor. *IEEE Sensors Journal, 22*(3), 2711–2721.

Zhao, J., Wang, H., Chen, Y., & Huang, M.-F. (2022). Detection of road surface anomaly using distributed fiber optic sensing. *IEEE Transactions on Intelligent Transportation Systems, 23*(11), 22127–22134.

Zhao, K., Liu, J., Wang, Q., Wu, X., & Tu, J. (2022). Road damage detection from post-disaster high-resolution remote sensing images based on tld framework. *IEEE Access, 10*, 43552–43561.

Zheng, L., Sayed, T., & Mannering, F. (2021). Modeling traffic conflicts for use in road safety analysis: A review of analytic methods and future directions. *Analytic methods in accident research, 29*, 100142

Publisher's Note