UNIVERSITY OF OULU

FACULTY OF INFORMATION TECHNOLOGY AND ELECTRICAL ENGINEERING

Axel Wester

UTILIZING ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING FOR MONITORING AND MODELING ROAD CONDITIONS

Bachelor's Thesis
Degree Programme in Computer Science and Engineering
January 2023

Wester A. (2023) Utilizing Artificial Intelligence and Machine Learning for Monitoring and Modeling Road Conditions. University of Oulu, Degree Programme in Computer Science and Engineering, 27 p.

ABSTRACT

Road maintenance requires resources increasingly as climate change and high traffic volume in populous areas inflict a significant strain on the traffic infrastructure. In rural areas, the car is usually the only mode of transport and long driving distances with high average speeds are covered on a daily basis. The majority of maintenance resources are located in densely populated cities, making the maintenance of rural roads challenging and expensive.

New scalable methods to optimize the usage of road maintenance resources are demanded. This thesis reviews several artificial intelligence and machine learning based techniques and systems designed for monitoring, evaluating, and predicting road condition and deterioration. In the implementation part of the thesis, two classification models, based on logistic regression and support vector machines, are trained to classify five different types of normal or damaged road segments from vertical acceleration data measured with smartphone sensors. A classification accuracy of 70.9% was achieved with logistic regression and 73.9% with support vector machine. The results of the implementation provide more evidence that vibration-based road condition monitoring systems can identify road anomalies with good accuracy and could have practical utility in road maintenance related tasks.

Keywords: pavement distress detection, pavement performance modeling, supervised learning, classification

Wester A. (2023) Tekoälyn ja koneoppimisen hyödyntäminen tien kunnon tunnistamisessa ja mallintamisessa. Oulun yliopisto, Tietotekniikan tutkinto-ohjelma, 27 s.

TIIVISTELMÄ

Teiden huoltotoimenpiteet vaativat resursseja enenevissä määrin, sillä ilmastonmuutos ja vilkastuva liikenne tiheästi asutuilla alueilla kuormittavat liikenneinfrastruktuuria merkittävästi. Harvaan asutuilla alueilla auto on usein ainoa kulkuväline, ja asukkaat keskimäärin ajavat pidempiä matkoja suuremmalla keskinopeudella. Suurin osa teiden huoltoon vaadittavista resursseista keskittyy tiheään asutuille taajama-alueille, tehden harvaan asuttujen alueiden tiestön huollosta haastavaa.

Uusia skaalautuvia menetelmiä teiden huoltoon vaadittavien resurssien optimoimiseksi tarvitaan. Tässä tutkielmassa tarkastellaan erilaisia tekoälyyn ja koneoppimiseen pohjautuvia menetelmiä ja järjestelmiä teiden kunnon tarkastamista, arviointia ja mallintamista varten. Tutkielman suoritusosassa kaksi luokittelumallia, jotka pohjautuvat logistiseen regressioon ja tukivektorikoneeseen, koulutetaan erottamaan viisi erityyppistä normaalia tai vaurioitunutta tieosuutta älypuhelimen liikesensoreilla kerätyistä vertikaalisista kiihtyvyysanturimittauksista. Logistinen regressiomalli luokitteli testidataa keskimäärin 70.9% tarkkuudella, kun taas tukivektorikoneeseen perustuva malli saavutti vastaavasti 73.9% luokittelutarkkuuden. Suoritusosan tulokset antavat näyttöä siitä, että värähtelymittauksiin perustuvat tien kunnon tunnistamiseen suunnitellut järjestelmät voivat tunnistaa erinäisiä poikkeamia tien pinnassa hyvällä tarkkuudella, ja että näistä järjestelmistä voisi olla hyötyä teiden huoltoon liittyvissä toimenpiteissä.

Avainsanat: tievaurioiden tunnistaminen, tien kulumisen mallintaminen, ohjattu oppiminen, luokittelu

TABLE OF CONTENTS

FOREWORD

I want to thank Assistant Professor Jaakko Suutala for supervising this work and providing guidance and inspiration. Thanks also to my family and friends for all the support.

Oulu, January 30th, 2023

Axel Wester

LIST OF ABBREVIATIONS AND SYMBOLS

AI	artificial intelligence
CV	cross-validation
LR	logistic regression
LTPP	Long-Term Pavement Performance program
ML	machine learning
NN	neural network
PPM	pavement performance modeling
RCM	road condition monitoring
SVC	support vector classifier
SVM	support vector machine
y	target label
$\bar{x}$	sample mean
s_x	sample standard deviation
x_{min}	smallest observation
x_{max}	greatest observation
$Skew(\mathbf{x})$	sample skewness
$Kurt(\mathbf{x})$	sample Fisher's kurtosis
X	feature space
$\mathbf{x}$	data sample
$\mathbf{x}_f$	feature vector
$\mathbf{x}_{sf}$	standardized feature vector
M	margin width
β	model coefficient
γ	kernel scaling parameter
C	regularization parameter
l	likelihood function
p	probability function
K	kernel function

1. INTRODUCTION

Road networks play an essential role in modern societies. Today, road networks cover entire nations and connect them together. This infrastructure provides effortless mobility for people and commodities ensuring flexible lifestyles and a foundation for increased economic growth [1]. Well planned traffic infrastructure connects employees to workplaces, customers to services, and producers to markets efficiently.

Extensive road networks demand frequent maintenance. Increasing traffic load and extreme weather conditions amplified by climate change cast a multi-factor burden on the structures of roads. Weather and climate induced factors such as temperature, precipitation, ground frost and ice, as well as human induced factors including traffic, poor construction quality, and insufficient maintenance can make the road surface prone for damages [2, 3]. Roads in poor condition may reduce ride comfort, cause vehicular damages, and even increase risk for traffic accidents.

Prior maintenance, the condition of the road should be inspected and evaluated to determine the extent of the potentially required work. This evaluation has traditionally been performed manually, before automatized inspection systems and tools started to emerge. Manually performed inspection requires multiple human actions and decisions making the evaluation expensive, time-consuming, and imprecise. Automatizing some of the inspection related tasks can enhance the maintenance in terms of costs and time.

Data-analysis has a remarkable potential in pavement engineering including pavement structure analysis, condition evaluation, deterioration prediction, and identification of influencing factors and key features of pavement performance [4]. Artificial Intelligence (AI) and Machine Learning (ML) provides scalable and diverse methods for numerous prediction, evaluation, identification, optimization, automation, and decision-making related objectives. The scope of this thesis comprehends the utility of AI and ML for predicting, evaluating, and identifying road condition and road anomalies.

In the following chapters of this thesis, recently introduced approaches to utilize AI and ML techniques for road condition monitoring (RCM) and pavement performance modeling (PPM) are reviewed. Topics including identification and evaluation of the current state of the road from several types of measurements as well as prediction of the deterioration and service life of the road using AI and ML are covered. To conclude, an approach to utilize two ML techniques, Logistic Regression (LR) and Support Vector Machine (SVM), to classify road anomalies, such as bumps, potholes, and worn-out road, from vertical acceleration data, is proposed.

2. LITERATURE REVIEW

A great quantity of literature has been published about AI and ML as tools in pavement engineering and maintenance related tasks in recent years. Several studies about using AI for RCM [5] and PPM [6] prove its applicability in the area of research. This chapter gives an overview of RCM and PPM techniques proposed in the literature.

2.1. Data Acquisition

In order to inspect and analyze the road condition, data with information about the road pavement or its surroundings needs to be collected. In the literature, multiple types of data have been utilized in RCM and PPM and collected with several different platforms. All data acquisition systems have their strengths and weaknesses, but the results prove that many systems may be useful when sufficient ML algorithms are utilized. Information about the data acquisition approaches is summarized in Table 1 in the end of this chapter.

2.1.1. Data Types

Vibration data, including acceleration, rotation, and velocity measurements, is one of the most common data types for RCM. When a moving car hits an anomaly on the road surface, abnormal movement is caused to the vehicle especially in the vertical direction. By identifying these abnormalities in the data captured by movement sensors, the anomalies on the road surface can be detected. To collect vibration data, the measurement platform needs to be traversed on the surface of the road to make contact with the road anomalies. Thus, a vehicle is required for vibration data collection. Platforms already equipped with movement sensors, e.g., smartphones [7, 8, 9] and modern cars, have a great potential to be harnessed for vibration data collection.

It is a challenging task to collect information of the complete road surface with vibration-based systems. Cars are in contact with the road only through their wheels, making them unable to capture every anomaly in the road in one run [9]. Seasonal or arbitrary anomalies, e.g., ice, snow, and litter, may cause inaccuracies in the data [5]. Actual road damages and objects intended to be on the road, such as speed bumps, may be difficult to distinguish from vibration data. The bias caused by vehicle model or type, when performing measurements with several different vehicles, may also require consideration when designing vibration-based systems [7].

From digital images and videos, it is possible to examine the entire road surface. Also, other properties such as size, severity, or location of an anomaly may be evaluated [5, 10]. Visual data collection is not dependent of a physical contact with the road surface, allowing data collection from a greater distance, e.g. using unmanned aerial vehicles [10, 11]. Vision based systems are dependent of the visual setting, making them possibly limited by light and weather conditions.

When expanding from the two-dimensional visual data into three dimensions, by using e.g. LiDAR technology [12] or stereo cameras [13], even more detailed information about the surface can be acquired. The equipment used to collect three-dimensional data can be expensive, while, on the other hand, vibration sensors and cameras are relatively cheap. These measurements may also be limited by bad weather conditions, which can affect the accuracy and reliability of the data especially in the depth dimension when the pavement surface is flooded. As the number of dimensions of the data increases, also the computational requirements of the used equipment and technology grows.

For long term road condition prediction and modeling, weather information is one of the most prominent data types. Precipitation and temperature have been adopted in many PPM studies [6] and considered to have a significant role in long term pavement deterioration [2]. Weather data is inexpensive to collect but usually data over several years is needed. Another significant data group especially for analyzing long term pavement deterioration is traffic information [3]. Information about the quantity, velocity, and weight of the vehicles passing through a specific road section is relatively easily measured with static sensory equipment. Also, several other features for PPM may be considered such as pavement structural and material information, maintenance data, age, and more [3, 14].

A somewhat less experimented area is the use of acoustical measurements, that is, e.g., evaluating road condition based on tire noise recordings. Different acoustic signals have been applied in some studies with success for evaluating pavement condition and type [15, 16]. Audio tends to be vulnerable for noise from external agents in the surroundings, especially in congested traffic. Thus, audio may be more applicable for overall road condition evaluation of a rather lengthy road segment, than accurate anomaly detection.

2.1.2. Measurement Platforms

Smartphones and some other handheld devices are equipped with multiple sensors and cameras for data collection. Movement sensors that are suitable to collect vibration data, cameras for visual data, and global positioning systems for location records [9]. Smartphones are relatively economic data acquisition platforms and a great number of smartphones are already carried by people driving on a daily basis, enabling a possibility for broad data collection without the need to produce specific devices.

Ground vehicles, especially cars, are common data acquisition platforms being the most common vehicles moving on roads. Ground vehicles can be equipped with almost any types of data measurement devices or sensors. Even audio can be collected by mounting microphones on practical parts of the vehicle as seen in [16]. Harnessing the already existing vehicle traffic, inexpensive and broad measurements can be collected without making separate drives for data collection. A large quantity of automobiles providing measurements from the same measurement point on the road, can allow crowdsourcing techniques to improve the confidence of the predictions [9].

Unmanned aerial vehicles (UAV), such as drones, are applicable for quick visual data collection [10, 11]. UAVs can capture high-resolution imagery or even three-dimensional data of the road surface. They can cover relatively large road segments swiftly, although limited to their maximum flight-time. Bad weather conditions, especially heavy rain or strong winds, may limit their use.

Satellites have several similar advantages that UAVs have. A large field of view provides broad visual information from large road networks [17]. Analyzing satellite imagery can be less expensive than other methods that require new equipment for data acquisition. The accuracy of the satellite imagery-based analyses may depend on the resolution and quality of the imagery. The frequency of new images of the same area may be limited depending on the satellite, possibly making it challenging to track the changes in the condition of a specific road segment over time.

Stationary measurement devices and platforms are crucial for the development of PPM systems. They are mostly applied to provide information about local weather conditions and traffic metrics from the same location over a long time period. They may also be applicable for several other measurements, such as vibro-acoustical measurements [15].

2.2. Road Condition Monitoring

Road condition monitoring refers to the action of retrieving information about the current condition of the road. AI and ML can be applied to automatize and expand RCM with broad, accurate, and fast systems.

In [7, 8], smartphone-based RCM systems analyzing vibration data were proposed. [7] used SVMs, neural networks (NN), and decision trees to classify between potholes, cracks, and smooth road. They considered several features in time, frequency, and wavelet domains for analysis and reached over 88% accuracy with all trained models. [8] classified normal road versus potholes from vibration data with random forest, SVM, and LR models. Random forest performed best with a 97% accuracy.

A combination of vibration data and digital images collected by smartphones was seen in [9]. They designed a broad, distributed RCM system for road anomaly detection. They trained artificial NNs and object-detection techniques to identify road anomalies locally with smartphones. The identification results and location of the anomalies collected by several vehicles were automatically uploaded into a cloud-based fusion module for further analysis and processing. In the fusion module, outliers could be filtered out, and the confidence of the final classification was improved. The results were displayed in a map on a smartphone app for the users to see where road anomalies are located.

Various image processing techniques were applied in [18] to extract features from pavement images to detect potholes using least squares SVM and artificial NNs. From high resolution satellite imagery, [17] classified the quality of three types of road segments using several convolutional NN architectures reaching an accuracy up to 80%.

Three object-detection algorithms were considered for classifying six distress types from UAV captured images in [11]. The best performing algorithm achieved 57% mean average precision. System designed by [10] achieved 95% precision for detecting damages on transport routes from images captured by UAVs.

Higher dimensional data about the road surface were analyzed by [12] and [13]. A good performance in crack detection was achieved in [12] with a semi-supervised three dimensional data analyzing algorithm. [13] implemented a pothole detection algorithm reaching an accuracy of over 98% by analyzing data from a stereo vision-based system.

From the acoustical data approaches, [15] focused to examine the quality of the road structure with an approach that captures also concealed damages that are not visible on the road surface. This was possible by detecting vibro-acoustic signals from pavement structures initiated by passing vehicles and captured by stationary measurement devices. On the other hand, [16] focused to detect the road surface quality from the tire-road noise measured from a driving car. Their unsupervised learning technique was able to recognize two types of road surfaces on the inspected routes.

2.3. Pavement Performance Modeling

Pavement performance modeling is the art of predicting the future state of the road pavement based on historical information or current state. Most studies try to predict the value of a metric resembling the overall performance of road segments. Such metrics are, for example, the pavement condition index, international roughness index, and rutting depth.

Pavement deterioration is a slow process. Therefore, consistent data from a long period of time is needed to train a useful and accurate deterioration model. Not that many datasets exist that provide comprehensive information about pavement quality and affecting properties for a long time period. Many studies rely on data gathered by national research agencies, such as the Long-Term Pavement Performance Program (LTPP), which is one of the most comprehensive data sets providing data about the quality of pavements in the United States and Canada.

Data provided by the Korean National Highway Pavement Management System was utilized in [19] to predict road quality properties for the next one year from weather, pavement maintenance, and traffic data from the last ten years. Recurrent NNs were chosen to predict international roughness index, number of cracks, and the rutting depth for the next one year with success.

An empirical approach to predict the remaining service life of a pavement was demonstrated by [14], where pavement temperature and structural characteristics including pavement thickness were analyzed. Their proposed method, utilizing support vector regression and a particle filter, achieved a precision of 95% and only 2% mean squared error in estimating the remaining service life in the test data.

Traffic, weather, time, structural, and initial condition variables in the LTPP database were used to predict international roughness index and rut depth in [20]. A gradient boosting decision tree-based algorithm implemented reached a coefficient of determination of 90%.

[21] applied also the LTPP database to predict road condition within two to six year intervals. Same authors extended their work to focus even more to climate and climate change related factors in [22]. They examined several algorithms including k-nearest neighbors, naive Bayesian algorithms, decision trees to predict the pavement condition index or corresponding class achieving over 90% accuracy with some of the models.

Table 1. Overview of the data types in the literature

Data Type	Acquisition	Application	Pros	Cons
vibration [5 pp. 8-9], [7], [8], [9]	movement sensors, ground vehicles	RCM	economic	platform dependent
image [5 pp. 9-21], [9], [10], [11], [17], [18]	cameras, mobile platforms	RCM	good coverage, flexible	may be limited by light and weather
three dimensional [5 pp. 9-21], [12], [13]	laser, ground penetrating radar, stereo camera, LiDAR, mobile platforms	RCM	detailed information	requires sophisticated equipment and computational resources
acoustic [15], [16]	static or mobile sensors depending on use case	RCM	potential for versatile use cases	vulnerable for noise
weather information [19], [20], [21], [22]	static sensors	PPM	economic	data from long time periods required
traffic information [19], [20], [21], [22]	static sensors	PPM	relatively economic	data from long time periods required
maintenance and structural information [19], [20], [21], [22]	historical databases, construction and maintenance agencies	PPM	no acquisition equipment needed necessarily depending on feature	data from long time periods required, limited sources

3. THE DATA

The data used in this work was collected by [23]. It is a collection of vertical acceleration measurements measured using smartphone sensors while driving on five different types of road segments in the city of Chihuahua, Mexico. The five types of roads that were recorded, include asphalt bumps, metal bumps, potholes, regular road, and worn-out road. There are 100 measurement samples of each type, summing up to 500 samples. The sample with the least amount of measurement points has 59 points, meanwhile the sample with greatest amount has 378. The measurements were recorded using a 50 Hz sampling rate, making the duration of the shortest measurement 1.18 seconds and the duration of the longest measurement 7.56 seconds. The properties of the complete dataset are summarized in Table 2. In this work, the varying size of the samples is not a concern, because only statistical features that are comparable despite the sample size are used. Nonetheless, it is worth mentioning that in some other situations it may be necessary to have samples of even length, or at least considering an optimal sample length, e.g. when using ML models that are fed a raw time series signal, or when deploying an advanced automatized vibration-based RCM system in practice.

The data is split into two parts, one for training and one for testing the ML models. In this work, a 30% test size is used, which means that 150 data samples are left out for testing, and 350 samples are used for training, from the available 500 samples. It is crucial to separate the testing data from the training data in order to properly examine the performance of the trained models with data that they are not familiar with. Despite the models would fit well into the training data, thus have low bias, that would not guarantee that they are good at making predictions. Conversely, low bias tends to lead to high variance, meaning that the model can not make accurate predictions from new data samples. That would be what is called overfitting.

Table 2. Properties of the dataset

Property	**Value**
Number of classes	5
Samples per class	100
Sum of datapoints	89847
Sampling rate	50 Hz
Datapoints per sample	179.7
Average duration per sample	3.594 s
Smallest observation	1.098 m/s²
Greatest observation	28.162 m/s²

4. METHODOLOGY

Machine learning comprehends a collection of techniques to build mathematical and statistical models for making predictions and decisions based on a sample of data. This chapter demonstrates the essentials of the ML techniques applied in this work. The project was implemented using the programming language Python and the libraries NumPy [24] and Pandas [25, 26] for data manipulation, Scikit-learn [27] for machine learning, and Matplotlib [28] for visualization.

4.1. Feature Extraction

The time series acceleration data signals, visualized in Figure 1, are not directly fed to the ML models. Instead, features summarizing information over the complete sample are extracted from the signals. Statistical features are commonly applied with ML models and used in this work. From an acceleration measurement sample $\mathbf{x}$, the extracted features are the sample mean $\bar{x}$, standard deviation s_x, smallest observation x_{min}, greatest observation x_{max}, skewness $Skew(\mathbf{x})$ and Fisher's kurtosis $Kurt(\mathbf{x})$. Especially the standard deviations, minima, and maxima provide information about the variation range of the measurement values in the samples, possibly indicating an existing road anomaly captured in the data. The skewness and kurtosis are measures that describe the shape of the data signals. These features form the initial feature vector $\mathbf{x}_f$ in Equation (1).

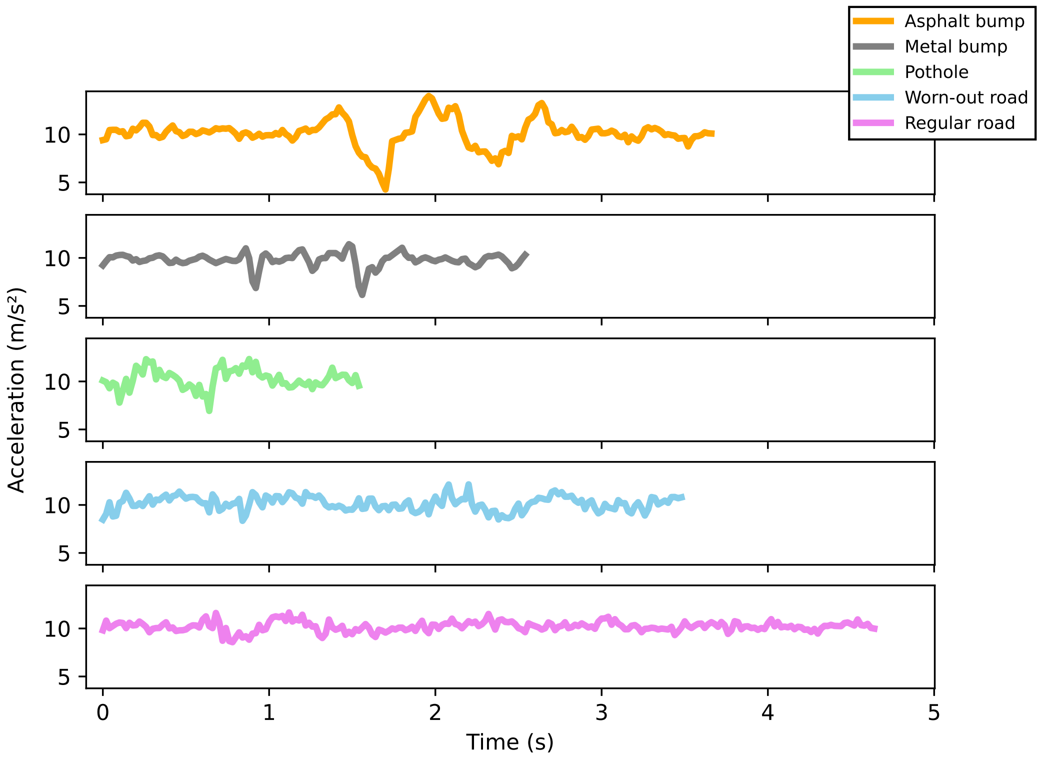

Figure 1. Example time series signals from each class.

The distributions of the computed features for each class in the dataset are visualized as box plots in Figure 2. From the distributions it can be seen that some features may have a greater significance in the analysis than others. Specifically, the sample mean seems to have quite similar distribution in each class, while the standard deviation seems to vary more on average between the classes. The feature values of the data samples labeled as metal bumps seem to differ significantly from the rest, especially in terms of kurtosis and skewness.

$$\mathbf{x}_f = \begin{bmatrix} \bar{x} \\ s_x \\ x_{min} \\ x_{max} \\ Skew(\mathbf{x}) \\ Kurt(\mathbf{x}) \end{bmatrix} \tag{1}$$

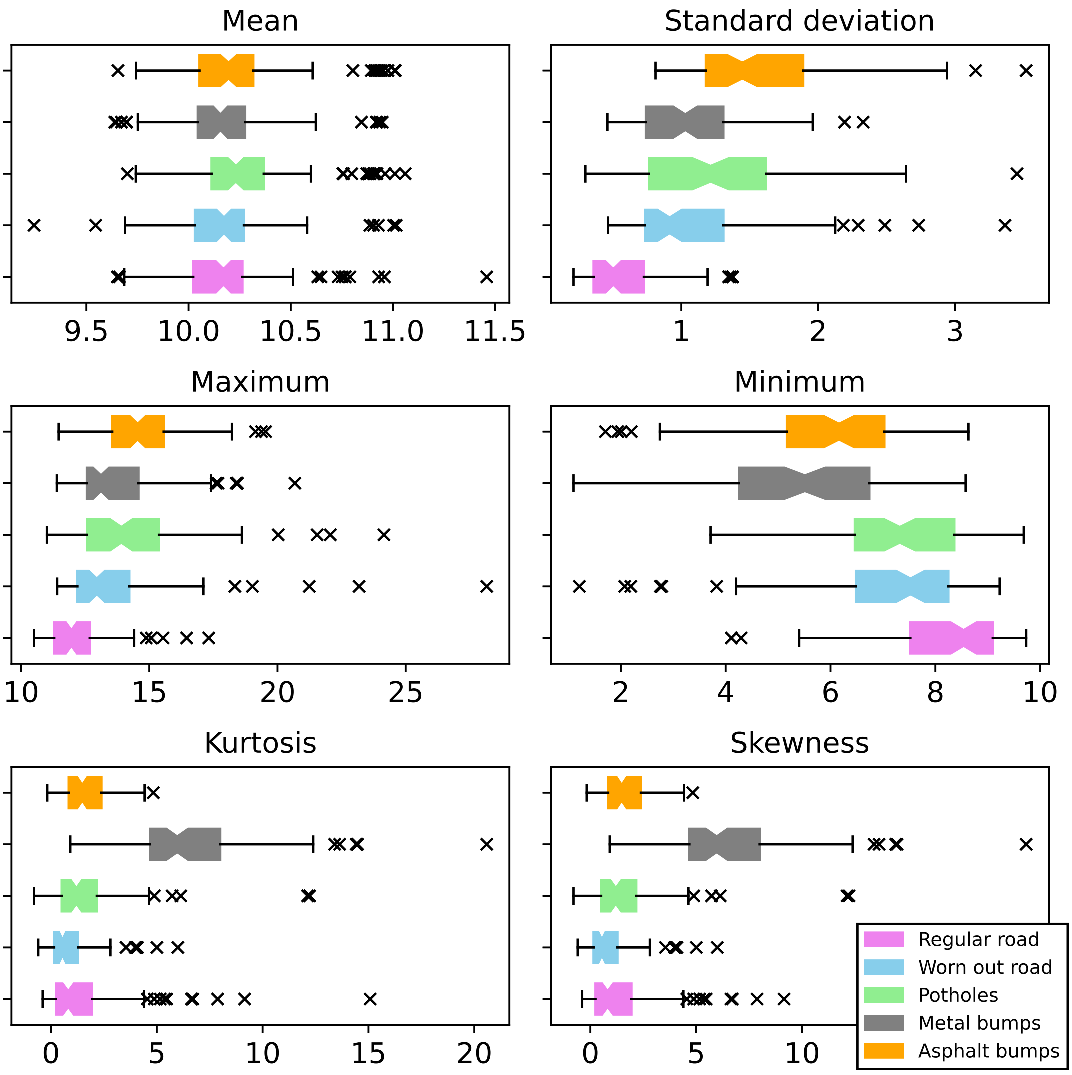

Figure 2. Distributions of the feature values in each class.

After computing the features of a data sample, the feature values are standardized to zero mean and unit variance according to the distributions in the training data. All the data, including test data, is standardized according to the feature values of the training data, ensuring that the models remain unfamiliar with the test data. A visualization of an example data signal and the derived and standardized features are presented in Figure 3. Standardization scales all the features into the same range, which is desirable for some ML models that are, for example, based on the spatial distances between the input data points in the feature space. The standardized features of the time series signals in Figure 1 are presented in Figure 4.

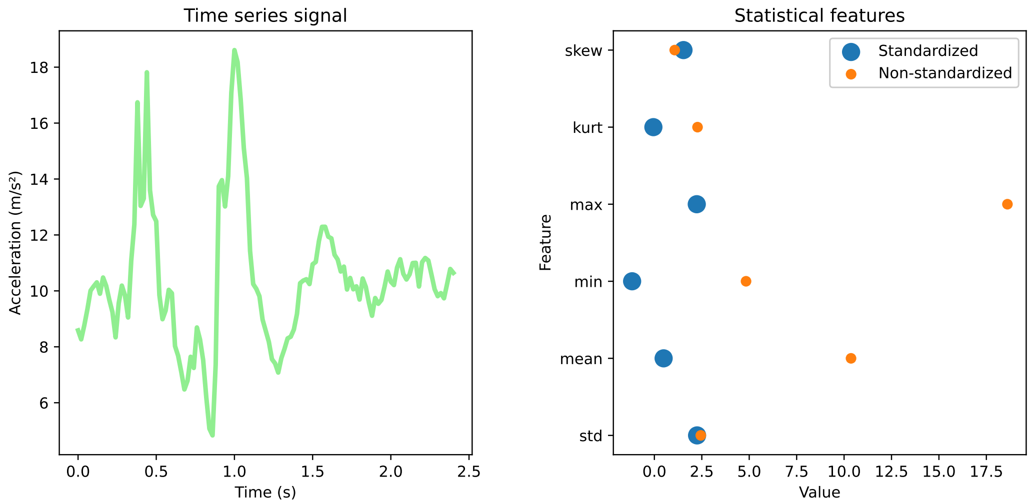

Figure 3. Time series acceleration data sample and corresponding statistical features.

The standardized values of these features form the feature space X, presented in Equation (2),

$$X = \begin{bmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \\ X_5 \\ X_6 \end{bmatrix}, \tag{2}$$

where X_i represent the standardized features' distributions as random variables. This feature space is further analyzed with the ML models described in the following section. A group of feature vectors, the training data, in the feature space is used to train the models, and another group, the testing data, is used to evaluate the models. In the optimal outcome, the final models can make accurate predictions from any input point in the feature space.

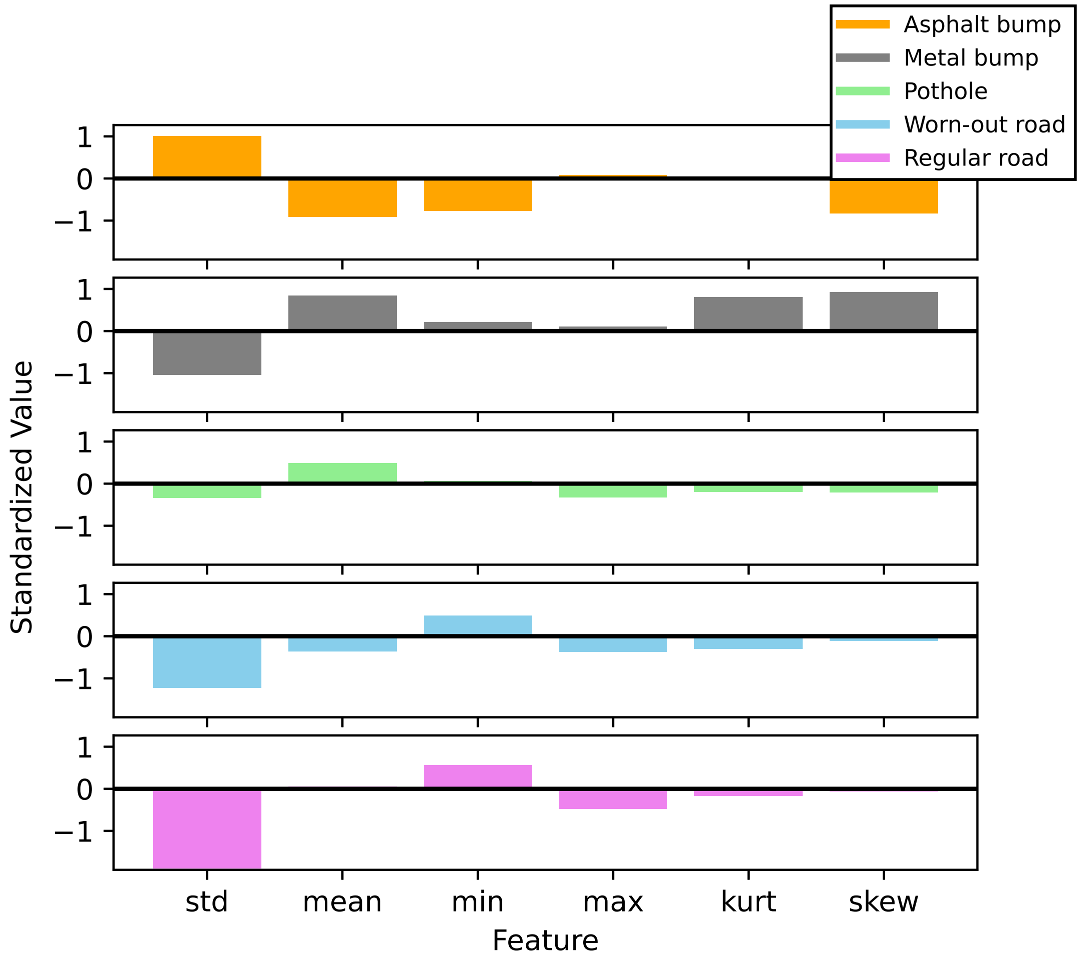

Figure 4. Example standardized features derived from each class.

4.2. The Models

The two models selected for this work are logistic regression and support vector machine. LR and SVM belong to the section of supervised learning models in ML. This implies, that they are trained with a set of feature vectors and their corresponding target values. One could imagine the process being like demonstrating someone explicitly with several examples, that some specific inputs should lead to some specific output. From these examples, the ML model is expected to learn to derive correct outputs even from new data inputs it is unfamiliar with.

Classification is the action of identifying to which category of outputs the input belongs to. Thus, the output space in classification consists of discrete values. In this work, LR and SVM are used to classify which type of road or road anomaly is recorded in the acceleration data sample.

4.2.1. Logistic Regression

Although the term logistic regression refers to regression rather than classification, LR is a common binary classification model. The logistic model computes the probability of an input belonging to a class [29 pp. 133-135]. If the probability is calculated to be at least 0.5, the input is considered to belong to the class. Equation (3) presents the logistic model for the input feature space,

$$p(X) = \frac{e^{\beta_0+\beta_1 X_1+\cdots+\beta_6 X_6}}{1+e^{\beta_0+\beta_1 X_1+\cdots+\beta_6 X_6}}, \tag{3}$$

in which p is the probability of the predicted event and β_i are the model coefficients that are estimated during model training.

To estimate the model coefficients β_i a method called maximum likelihood estimation [29 pp. 135-136] is used. In maximum likelihood estimation, the objective is to maximize the likelihood of the events in the training data by searching optimal values for the model coefficients. The mathematical definition of the likelihood function to be maximized is shown in Equation (4),

$$l(\beta_0,\ldots,\beta_6) = \prod_{i:y_i=1} p(\mathbf{x}_{sf,i}) \prod_{i':y_{i'}=0} (1-p(\mathbf{x}_{sf,i'})), \tag{4}$$

where $\mathbf{x}_{sf,i}$ are the standardized feature vectors belonging to the training data, and y_i are their corresponding target labels.

As mentioned in the beginning of this subsection, LR is a binary classification model, but the goal in this work is to identify the correct class from five different alternatives. One approach to overcome this limitation is to train five separate models to predict the probability of each class. This technique is called one-versus-the-rest, as it predicts the probability of only one class versus the rest of the classes. When making predictions with LR models using this approach, a probability with each model is calculated and the prediction is chosen to be the class with the highest probability of the five predictions.

4.2.2. Support Vector Machine

The goal in classification with support vector machines is to generate a hyperplane that separates the data points of the target classes in the input feature space [29 pp. 367-386]. When the hyperplane is determined optimally, the new data points are expected to fall on the correct side of the hyperplane among other data points with the same class label. Thus, the predictions are made by determining which side of the hyperplane the input data point falls. The essential elements of the SVM are the support vector classifier (SVC) and a method called "the kernel trick" [29 pp. 380-383].

A SVC is a binary classifier, that utilizes so-called support vectors to fit a soft margin into the training data [29 pp. 373-379]. The support vectors are a group of three data points that are used to fit the hyperplane. Unlike maximum margin classifiers, that attempt to strictly divide the data by fitting a hyperplane between the data points of the classes without any tolerance for bias, and often resulting into overfitting, SVC has a soft margin, which accepts some of the training data points within the margin, or even on the wrong side of the hyperplane. By accepting some bias by using a soft margin in the model while training, the risk to overfit the model decreases, and lower variance in the final model is usually achieved. SVC is defined by the optimization problem in Equations (5) - (8), as follows,

$$\max_{\beta_0,\beta_1,\ldots,\beta_6,\epsilon_1,\ldots,\epsilon_n,M} M \tag{5}$$

$$\text{subject to} \sum_{j=1}^{6} \beta_j^2 = 1, \tag{6}$$

$$y_i(\beta_0 + \beta_1 X_{i,1} + \beta_2 X_{i,2} + \cdots + \beta_6 X_{i,6}) \geq M(1-\epsilon_i), \tag{7}$$

$$\epsilon_i \geq 0, \sum_{i=1}^{n} \epsilon_i \geq C, \tag{8}$$

where C is a fine-tuning parameter adjusting the amount of data points allowed within the margin or wrong side of the hyperplane, and M in Equation (5) is the width of the margin.

The SVC is optimal only when the data is suitably separated with a hyperplane, but with some data that is not possible. As a solution, SVC is extended with the aforementioned "kernel trick", to form the eventual SVM. In the kernel trick, the input feature vector is "artificially" transformed into a higher-dimensional feature space with a kernel function to manipulate the data to be easier to separate with a hyperplane. Here, a radial basis function [29 p. 382] was used as the kernel function for the SVM model, defined in Equation (9),

$$K(\mathbf{x}_{sf}, \mathbf{x}'_{sf}) = e^{-\gamma||\mathbf{X}_{sf}-\mathbf{X}'_{sf}||^2}, \tag{9}$$

where γ is another fine-tuning parameter, the kernel scaling parameter.

The SVM is likewise designed for binary classification, thus an applied solution for multiclass classification is needed. Similar to LR, here the one-versus-the-rest technique, referred as "one-versus-all" in [29 pp. 385-386], is used to extend the SVM for multi-class classification. A separate SVM is trained for each class to detect it from the rest. The classification with the resulting model is made according to the results of the individual models.

4.3. Cross-Validation

The regularization parameter C for both models and the kernel scaling parameter γ for SVM were fine-tuned using k-fold cross-validation (CV) [29 pp. 203-205]. Specifically, 7-fold CV was used in this work, keeping the training size relatively large with 300 samples, and having 50 samples for validation in each round. In the k-fold CV, the training data is split randomly into k parts. Then, $k-1$ of the parts are used to train the model, and one is left out to validate the performance of the trained model. The procedure is repeated switching the validation part every round. For each round, different values for the tuning parameters, (C, γ), are tested to determine which provide best performance, and selected into the final model.

The classes *LogisticRegression*[1] and *SVC*[2] in the Scikit-learn library were used to implement the ML models. In Table 3 and Table 4 are summarized the given parameters for the constructors of the class-objects.

Table 3. Parameters for *LogisticRegression*

Parameter	**Value**
penalty	'l2'
tol	10^{-4}
C	50, *determined with CV*
intercept_scaling	1
solver	'liblinear'
multi_class	'ovr'
verbose	None

Table 4. Parameters for *SVC*

Parameter	**Value**
C	100, *determined with CV*
kernel	'rbf'
gamma	0.1, *determined with CV*
decision_function_shape	'ovr'
tol	10^{-3}
shrinking	True

[1]https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LogisticRegression.html
[2]https://scikit-learn.org/stable/modules/generated/sklearn.svm.SVC.html

5. RESULTS

After the model hyperparameters were fine-tuned with CV and the final models trained with the training data, the performances of the models were evaluated by testing how well the models were able to predict the correct classes of the 150 samples of testing data. Due to the relatively small dataset, the accuracy of the models varied slightly each time training and testing the models with different random train-test splits. To achieve a robust assessment of the model performances, 10 iterations of separately dividing the data, training the models, and testing the models was performed. The following results are averages of those 10 iterations.

The confusion matrices in Figure 5 and Figure 6 visualize the distributions of the predictions on the test data. In both cases, a pleasing dark diagonal was achieved, resembling a relatively high accuracy of the models. Both models seemed to struggle to distinguish potholes from asphalt bumps and worn-out road from regular road. Both models predicted the asphalt bumps especially well, which may be explained by the class standing out in the values of skewness and kurtosis seen in figure Figure 2 in the preceding chapter.

There was a small difference in performance between the two models, the SVM performing slightly better. Average accuracy for LR over 10 iterations was 70.9%, and 73.9% for SVM. In addition to accuracy, three other classification performance metrics were calculated, separately for each class. Precision depicts the rate of true positives of all positive predictions. Recall is the rate of correctly classified positives from all the positives in the test data. The F1-score unifies information about both preceding metrics. The metrics for LR and SVM are summarized in Table 5 and Table 6, respectively.

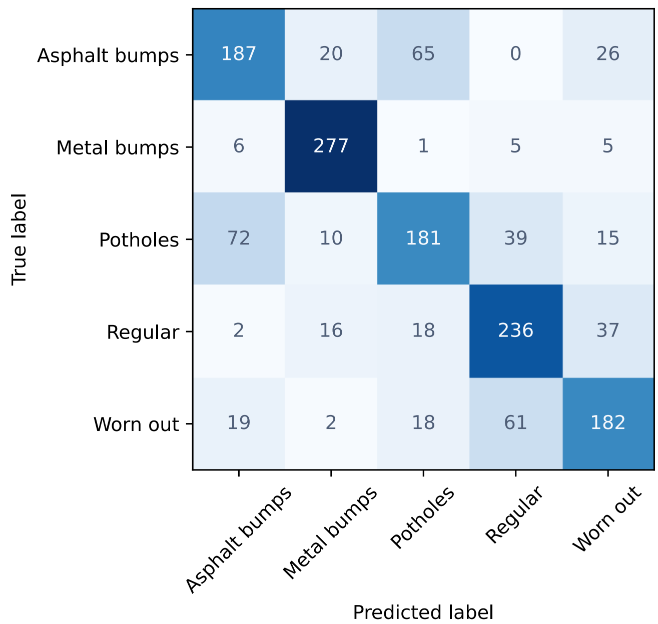

Figure 5. Confusion matrix for LR over 10 iterations.

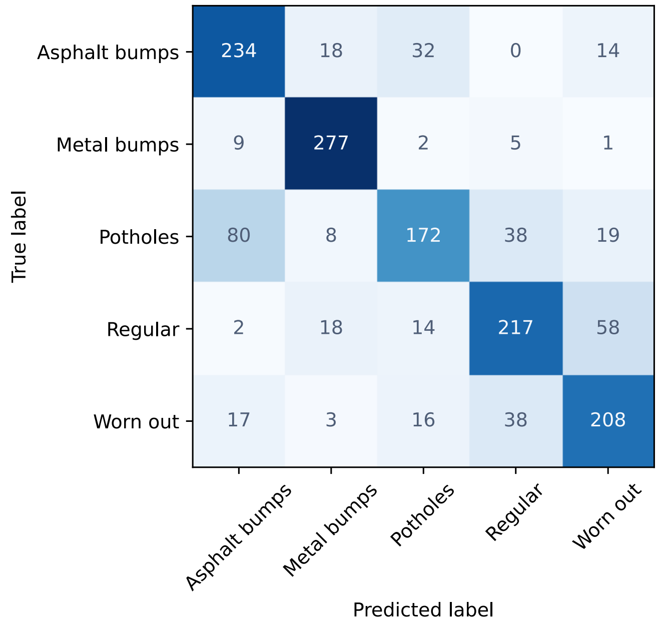

Figure 6. Confusion matrix for SVM over 10 iterations.

Table 5. Average metrics for Logistic Regression

Label	**Metric**		
	Avg. Precision	Avg. Recall	Avg. F1-Score
Asphalt bumps	0.66	0.63	0.64
Metal bumps	0.85	0.94	0.89
Potholes	0.65	0.57	0.60
Regular road	0.69	0.77	0.73
Worn out road	0.69	0.65	0.66

Table 6. Average metrics for Support Vector Machine

Label	**Metric**		
	Avg. Precision	Avg. Recall	Avg. F1-Score
Asphalt bumps	0.69	0.79	0.73
Metal bumps	0.85	0.94	0.89
Potholes	0.73	0.54	0.62
Regular road	0.73	0.70	0.72
Worn out road	0.69	0.75	0.71

6. DISCUSSION

Although in several studies higher classification accuracy for both LR and SVM was achieved [7, 8], the results of this thesis were in line with the related work in general terms. The decent classification results provide more evidence that vibration-based RCM systems could have the potential to be applied in practice. Despite the relatively good accuracy of the implementation, several limitations exist that potentially affected the results. In future work, any of the following improvements could be considered.

By using a more comprehensive and larger dataset, the accuracy of the models could improve. As observed in [7], acceleration data across several directions, instead of measurements only in the vertical direction, could improve the classification accuracy. Especially when classifying potholes versus transverse cracks, a pothole would impose a rotational movement sideways, because potholes are usually met with tires only on one side of the car. Meanwhile a transverse crack would cause a somewhat symmetrical impact hitting the tires on both sides of the car.

The focus in the classification process was to distinguish all five pre-defined categories of road segments. The data could have been considered to be categorized differently, e.g. dividing the samples in the initial classes into only two groups, where the other group would represent normal road including regular road and speed bumps, and the other damaged road types, compressing the initial classification task into a binary problem. When classifying between five different classes, the differences between the classes may be more delicate, making the classification task more challenging. In some of the related literature the classification was done between only two or three classes.

In this work, only time domain statistical features were analyzed, whereas frequency and wavelet domain features could have provided more dimensions to the analysis. Also, a proper feature analysis and selection could have been conducted to identify the most significant features for the models optimizing computational efficiency.

Moreover, more types of classification techniques could have been tested and compared with the data. Good results in the literature were achieved, for example, with decision trees [8], neural networks [7], and deep learning techniques [9].

As the area of research is getting more attention and new research about RCM emerge, more detailed and applied research is needed to encourage the deployment of these systems in practice. Regarding the vibration-based RCM systems, in-depth research about ways to compensate the inability to cover the complete road surface could be conducted.

7. SUMMARY

This thesis reviewed multiple approaches to utilize machine learning and artificial intelligence to inspect the current and predict the future condition of a road surface. The most conveniently deployable inspection systems were based usually on vibration or visual data. More complex systems requiring advanced equipment processed even three-dimensional data measured using laser, stereo camera, or ground-penetrating radar-based devices. The predictive models generally relied on long term information about weather conditions, traffic volume, maintenance, or pavement structure.

To contribute to the area of study, classification of road anomalies from vertical acceleration data with machine learning algorithms, logistic regression and support vector machine, was demonstrated. The logistic model achieved an average accuracy of 70.9% and the support vector machine 73.9%, when classifying between five categories of road types. In the big picture, the results of this thesis were in line with the related literature, although in several studies even better results with LR and SVM models were achieved in similar objectives. Nevertheless, the results of this thesis approve the concept of vibration-based road condition monitoring systems and the potential of machine learning and artificial intelligence in enhancing road maintenance related activities.

8. REFERENCES

[1] Ng C.P., Law T.H., Jakarni F.M. & Kulanthayan S. (2019) Road infrastructure development and economic growth. IOP Conference Series: Materials Science and Engineering 512(1), 012045. DOI: https://doi.org/10.1088/1757-899X/512/1/012045.

[2] Qiao Y., Dawson A.R., Parry T., Flintsch G. & Wang W. (2020) Flexible pavements and climate change: A comprehensive review and implications. Sustainability 12(3), 1057. DOI: https://doi.org/10.3390/su12031057.

[3] Llopis-Castelló D., García-Segura T., Montalbán-Domingo L., Sanz-Benlloch A. & Pellicer E. (2020) Influence of pavement structure, traffic, and weather on urban flexible pavement deterioration. Sustainability 12(22), 9717. DOI: https://doi.org/10.3390/su12229717.

[4] Dong Q., Chen X., Dong S. & Ni F. (2022) Data analysis in pavement engineering: An overview. IEEE Transactions on Intelligent Transportation Systems 23(11), pp. 22020–22039. DOI: https://doi.org/10.1109/TITS.2021.3115792.

[5] Ranyal E., Sadhu A. & Jain K. (2022) Road condition monitoring using smart sensing and artificial intelligence: A review. Sensors 22(8), 3044. DOI: https://doi.org/10.3390/s22083044.

[6] Karimzadeh A. & Shoghli O. (2020) Predictive analytics for roadway maintenance: A review of current models, challenges, and opportunities. Civil Engineering Journal 6(3), pp. 602–625. DOI: https://doi.org/10.28991/cej-2020-03091495.

[7] Basavaraju A., Du J., Zhou F. & Ji J. (2020) A Machine Learning Approach to Road Surface Anomaly Assessment Using Smartphone Sensors. IEEE Sensors Journal 20(5), pp. 2635–2647. DOI: https://doi.org/10.1109/JSEN.2019.2952857.

[8] Wu C., Wang Z., Hu S., Lepine J., Na X., Ainalis D. & Stettler M. (2020) An Automated Machine-Learning Approach for Road Pothole Detection Using Smartphone Sensor Data. Sensors 20(19), 5564. DOI: https://doi.org/10.3390/s20195564.

[9] Ramesh, A. and Nikam, D. and Balachandran, V. N. and Guo, L. and Wang, R. and Hu, L. and Comert, G. and Jia, Y. (2022) Cloud-based collaborative road-damage monitoring with deep learning and smartphones. Sustainability 14(14), 8682. DOI: https://doi.org/10.3390/su14148682.

[10] Silva L.A., Sanchez San Blas H., Peral García D., Sales Mendes A. & Villarubia González G. (2020) An Architectural Multi-Agent System for a Pavement Monitoring System with Pothole Recognition in UAV Images. Sensors 20(21), 6205. DOI: https://doi.org/10.3390/s20216205.

[11] Zhu J., Zhong J., Ma T., Huang X., Zhang W. & Zhou Y. (2022) Pavement distress detection using convolutional neural networks with images captured via UAV. Automation in Construction 133, 103991. DOI: https://doi.org/10.1016/j.autcon.2021.103991.

[12] Feng H., Li W., Luo Z., Chen Y., Fatholahi S.N., Cheng M., Wang C., Junior J.M. & Li J. (2022) GCN-Based Pavement Crack Detection Using Mobile LiDAR Point Clouds. IEEE Transactions on Intelligent Transportation Systems 23(8), pp. 11052–11061. DOI: https://doi.org/10.1109/TITS.2021.3099023.

[13] Fan R., Ozgunalp U., Hosking B., Liu M. & Pitas I. (2020) Pothole detection based on disparity transformation and road surface modeling. IEEE Transactions on Image Processing 29, pp. 897–908. DOI: https://doi.org/10.1109/TIP.2019.2933750.

[14] Karballaeezadeh N., S D.M., Shamshirband S., Hajikhodaverdikhan P., Mosavi A. & wing Chau K. (2019) Prediction of remaining service life of pavement using an optimized support vector machine (case study of Semnan–Firuzkuh road). Engineering Applications of Computational Fluid Mechanics 13(1), pp. 188–198. DOI: https://doi.org/10.1080/19942060.2018.1563829.

[15] Praticò F.G., Fedele R., Naumov V. & Sauer T. (2020) Detection and monitoring of bottom-up cracks in road pavement using a machine-learning approach. Algorithms 13(4), 81. DOI: https://doi.org/10.3390/a13040081.

[16] Ramos-Romero C., Asensio C., Moreno R. & de Arcas G. (2022) Urban road surface discrimination by tire-road noise analysis and data clustering. Sensors 22(24), 9686. DOI: https://doi.org/10.3390/s22249686.

[17] Brewer E., Lin J., Kemper P., Hennin J. & Runfola D. (2021) Predicting road quality using high resolution satellite imagery: A transfer learning approach. PLOS ONE 16(7), e0253370. DOI: https://doi.org/10.1371/journal.pone.0253370.

[18] Hoang N.D. (2018) An Artificial Intelligence Method for Asphalt Pavement Pothole Detection Using Least Squares Support Vector Machine and Neural Network with Steerable Filter-Based Feature Extraction. Advances in Civil Engineering 2018, 7419058. DOI: https://doi.org/10.1155/2018/7419058.

[19] Choi S. & Do M. (2020) Development of the Road Pavement Deterioration Model Based on the Deep Learning Method. Electronics 9(1), 3. DOI: https://doi.org/10.3390/electronics9010003.

[20] Guo R., Fu D. & Sollazzo G. (2022) An ensemble learning model for asphalt pavement performance prediction based on gradient boosting decision tree. International Journal of Pavement Engineering 23(10), pp. 3633–3646. DOI: https://doi.org/10.1080/10298436.2021.1910825.

[21] Piryonesi S.M. & El-Diraby T.E. (2020) Data Analytics in Asset Management: Cost-Effective Prediction of the Pavement Condition Index. Journal of Infrastructure Systems 26(1), 04019036. DOI: https://doi.org/10.1061/(ASCE)IS.1943-555X.0000512.

[22] Piryonesi S.M. & El-Diraby T. (2021) Climate change impact on infrastructure: A machine learning solution for predicting pavement condition index. Construction and Building Materials 306, 124905. DOI: https://doi.org/10.1016/j.conbuildmat.2021.124905.

[23] González L.C., Moreno R., Escalante H.J., Martínez F. & Carlos M.R. (2017) Learning Roadway Surface Disruption Patterns Using the Bag of Words Representation. IEEE Transactions on Intelligent Transportation Systems 18(11), pp. 2916–2928. DOI: https://doi.org/10.1109/TITS.2017.2662483.

[24] Harris C.R., Millman K.J., van der Walt S.J., Gommers R., Virtanen P., Cournapeau D., Wieser E., Taylor J., Berg S., Smith N.J., Kern R., Picus M., Hoyer S., van Kerkwijk M.H., Brett M., Haldane A., del Río J.F., Wiebe M., Peterson P., Gérard-Marchant P., Sheppard K., Reddy T., Weckesser W., Abbasi H., Gohlke C. & Oliphant T.E. (2020) Array programming with NumPy. Nature 585(7825), pp. 357–362. DOI: https://doi.org/10.1038/s41586-020-2649-2.

[25] Wes McKinney (2010) Data Structures for Statistical Computing in Python. In: Stéfan van der Walt & Jarrod Millman (eds.) Proceedings of the 9th Python in Science Conference, pp. 56 – 61. DOI: https://doi.org/10.25080/Majora-92bf1922-00a.

[26] The pandas development team (2022), pandas-dev/pandas: Pandas. URL: https://doi.org/10.5281/zenodo.7344967.

[27] Pedregosa F., Varoquaux G., Gramfort A., Michel V., Thirion B., Grisel O., Blondel M., Prettenhofer P., Weiss R., Dubourg V., Vanderplas J., Passos A., Cournapeau D., Brucher M., Perrot M. & Duchesnay E. (2011) Scikit-learn: Machine learning in Python. Journal of Machine Learning Research 12(85), pp. 2825–2830.

[28] Hunter J.D. (2007) Matplotlib: A 2d graphics environment. Computing in Science & Engineering 9(3), pp. 90–95. DOI: https://doi.org/10.1109/MCSE.2007.55.

[29] James G., Witten D., Hastie T. & Tibshirani R. (2021) An Introduction to Statistical Learning: with Applications in R. Springer US, New York, NY. DOI: https://doi.org/10.1007/978-1-0716-1418-1_4.