Automation in Construction 139 (2022) 104309

Contents lists available at ScienceDirect

Automation in Construction

journal homepage: www.elsevier.com/locate/autcon

Review

Machine learning algorithms for monitoring pavement performance

Saúl Cano-Ortiz, Pablo Pascual-Muñoz *, Daniel Castro-Fresno

GITECO Research Group, Universidad de Cantabria, 39005 Santander, Spain

A R T I C L E I N F O

Keywords:
Machine learning
Road performance
Data collection and road maintenance

A B S T R A C T

This work introduces the need to develop competitive, low-cost and applicable technologies to real roads to detect the asphalt condition by means of Machine Learning (ML) algorithms. Specifically, the most recent studies are described according to the data collection methods: images, ground penetrating radar (GPR), laser and optic fiber. The main models that are presented for such state-of-the-art studies are Support Vector Machine, Random Forest, Naïve Bayes, Artificial neural networks or Convolutional Neural Networks. For these analyses, the methodology, type of problem, data source, computational resources, discussion and future research are highlighted. Open data sources, programming frameworks, model comparisons and data collection technologies are illustrated to allow the research community to initiate future investigation. There is indeed research on ML-based pavement evaluation but there is not a widely used applicability by pavement management entities yet, so it is mandatory to work on the refinement of models and data collection methods.

1. Introduction

The satisfactory level of road serviceability is decisive to ensure economic growth, safety of passengers and sustainability. Periodic surveys of asphalt pavement are usually managed by human inspectors, but this long-established approach of road inspection is sluggish and subjected to variations in assessment outcomes. Accordingly, automatic pavement condition inspection via pavement performance prediction models are the future solution as an important element of road management systems.

Highway maintenance and rehabilitation intentions are maximizing the condition of a pavement network in order to minimize costs, increase functional level of service, reduce greenhouse gas emissions and improve road users' safety.

Manual visual inspection is the fundamental pattern of assessing the physical and functional conditions of civil infrastructures. However, accidents still occur due to deficient inspections and condition assessment [1]. Additionally, as reported by the American Society of Civil Engineers (ASCE), pavement defects cost US motorists $ 67 billion a year of repairs [2]. Conclusively, road surface evaluation to detect defects is genuinely important to ensure traffic safety.

On the application of ML techniques to pavement performance prediction, different characteristics can be distinguished: the utilization of International Roughness Index (IRI) [3] as a pavement performance indicator, the artificial neural network predominance on predicting the IRI as the ML technique most commonly used and the constant creation of Long-Term Pavement Performance databases (LTPPs) [4].

This review does not focus exclusively on the development of the most widely used techniques, but gives a brief introduction to the different algorithms developed, where artificial neural networks for prediction and convolutional neural networks for image processing and detection will be shown in detail. Subsequently, the different indicators or magnitudes to be detected are explained. Finally, the different techniques for data extraction techniques are explained together with a review of the state-of-the-art of the associated ML algorithms.

2. Pavement performance prediction

2.1. Description

The main purpose in pavement management is to assess the pavement state in order to predict future conditions. The mathematical functions of performing such tasks are pavement performance prediction models (PPPMs) as a fundamental axis for the pavement management (J. [5]). For pavement performance prediction [6], there are two categories considered, static (absolute models) and dynamic (relative models). A static approach, $P_t = f(X_t, t)$, where P_t is the pavement performance for a given time t, and X_t are auxiliary or explanatory variables (e.g. structural characteristics, climatic conditions, traffic, etc.) at time t [4]. The lagged values of the output are not considered as inputs in static

* Corresponding author.
E-mail address: pascualmp@unican.es (P. Pascual-Muñoz).

https://doi.org/10.1016/j.autcon.2022.104309
Received 24 September 2021; Received in revised form 9 April 2022; Accepted 29 April 2022
Available online 10 May 2022

models. A possible example of a static approximation is the regression model [7]. Consequently, this approach limits pavement performance prediction since pavements deteriorate in an incremental form. A dynamic approach, $P_t = f(P_{t-1}\ldots P_{t-n}, X_t, X_{t-1}\ldots X_{t-n})$, where P_t and X_t have identical meanings to the static approximation (at time t) and n is the number of past observations. Dynamic models forecast pavement deterioration using historical pavement performance data. Therefore, dynamic models should provide a more accurate prediction of pavement future conditions than static models.

As for the representative variables X_t, the most relevant ones for the predictive tasks have been highlighted in Table 1 [4].

In order to introduce the explanatory variables, now what each one quantifies will be analyzed. The Structural Number [19] is used as an indicator or index to settle the strength of a complete pavement structure; i.e., the sum of the strengths of all the layers. Pavement thickness is traditionally measured by GPR [20]; i.e., a short duration electromagnetic (EM) pulse that penetrates pavement materials and it is reflected at interfaces where its thickness is determined from velocity of signal in a specific pavement layer and two-way travel time at layer interfaces. Annual Average Precipitation, temperature and humidity accounts for amount of precipitation, temperature and humidity expected per year, respectively. Freezing Index is used for evaluating frost penetration for pavement design [21]. AADT is calculated by dividing the total volume of vehicle traffic by 365 days. Only certain explanatory variables have been shown previously, but in reality, there are many more.

2.2. Data life cycle

To structure the evolution of the data throughout the pavement performance evaluation process, several stages can be distinguished as shown in Fig. 1.

The initial step consists of collecting the data of interest using different sensors based on different physical methodologies, such as those shown in Section 4 (image-based, laser systems, etc.). Predominantly, pavement performance data are extracted from transportation agencies or open-source datasets. Once the data has been extracted, the expert agent is aware that the data received is not tabular and ideal. For this reason, a data pre-processing procedure is necessary. Data pre-processing is the method that involves the transformation of raw data into an understandable format. There are several techniques depending on the application. For example, sampling (model generalization insurance), missing data and data cleaning (error detection and removal).

After extracting the data, previously processed, the different machine learning algorithms capable of performing the different predictive tasks are studied (modelling phase). To choose the model whose predictive capacity is most optimal, evaluation metrics are used to compare the values predicted by the model and the observed results. Based on performance metrics, a process of model evaluation starts. In order to improve the performance, there are different approaches like collecting more data, extracting different data, tuning model hyperparameters or making a study with another algorithm. Once the previous steps are completed, a final evaluation using the test dataset is performed. Afterwards, a review of the different machine learning algorithms (data modelling) from methodologies or procedures (techniques for data collection via different sensors) will be interpreted.

Table 1
Main important explanatory variables (X_t).

Category	Description
Structure	Structural number and pavement thickness.
Climate	Annual average precipitation, annual average temperature, annual average freeze index, minimum/ maximum annual average humidity.
Traffic	Annual average daily traffic (AADT), average traffic speed and degree of noise.

2.3. Machine learning algorithms: brief approach

To introduce in an intuitive way the Machine Learning fundamental concepts, the following is a brief description of the most used algorithms and an in-depth study of the ML algorithms most used for the development of pavement management systems. ML models can be divided into three group categories [7]: Supervised learning, Unsupervised Learning and Reinforcement Learning.

2.3.1. Supervised learning (SL)

SL uses input data and output data, building a model to predict when applied to new data. If the target variable to predict (predictand) is categorical, one is dealing with classification problems and the most common algorithms are: (1) Decision Trees, (2) Support Vector Machine, (3) K-nearest neighbors, (4) Naïve Bayes, (5) Random Forest, (6) Logistic regression and Neural Networks - extensively explained in 2.3.4. However, if the target variable is continuous, this is about Regression Pavement Performance Models (RPPMs) and the most popular are Linear Regression, Neural Networks, Decision trees and Random Forests. These models are widely used for project-level.

(1) **Classification and Regression Trees (CART)**. CART [22] have the aim of predicting discrete and continuous variables, respectively. With regard to its intuitive structure, each node corresponds to a test attribute, each branch corresponds to an attribute value, each leaf (terminal node) represents a final class, and each path is a conjunction of attributed values. Tree construction is established with several algorithms but the fundamental idea of all of them is to evaluate the attribute power of separation (maximum impurity, minimal entropy). To avoid overfitting, the post-pruning technique (pruned overfitted fully grown tree, removing less useful nodes) is commonly used. CART advantages are having an easy approach to explain and understand representation and being applicable to classification and regression problems. Its disadvantages are poor prediction accuracy comparing with other approaches and instability when changing the data (need of cross-validation).

(2) **Support Vector Machine (SVM)**. SVM [23] is the most popular kernel machine for classification. SVM finds an optimal hyperplane that best separates the features into different domains. In other words, the hyperplane is a function used to differentiate between features (e.g., hyperplane is a line for 2D problem and a plane for 3D problem). In Eq. (1), $f(x)$ refers to the hyperplane, $x = \{x_i\}_{i=1\ldots N}$ are the features, $w = \{w_i\}_{i=1\ldots N}$ vectors and $b = w_0$, the bias. Basically, the goal is to maximize the distance (a.k.a margin) between the points closest to a given hyperplane (also called support vectors) and the hyperplane. Then, the hyperplane for which the margin is maximum is the optimal hyperplane or maximum margin hyperplane. Mathematically, it is a minimization problem with the objective of finding the optimal values w_i which results in solving a quadratic equation and at the computational level, it is a quadratic programming (QP) problem -i.e., mathematical optimization problem involving quadratic equations.

$$f(x) = w^T x + b \tag{1}$$

(3) **K-nearest neighbors (k−NN)**. K-NN is a non-parametric and instance-based algorithm applied to regression and classification problems. k-NN has a very simple functioning where distances are computed, ordered, and dependent on the number of neighbors selected (cross-validation is typically used to select an optimal value). To compute distances, different metrics are used (Euclidean, Minkowski, Manhattan, etc.). It is a “pure” lazy learning algorithm; i.e., it retrieves the k least distant instances

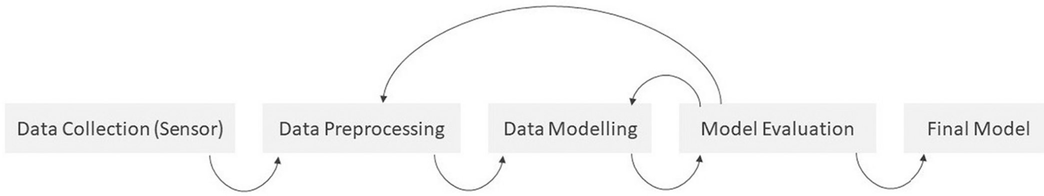

Fig. 1. Proposed general scheme of data life cycle.

and it infers the predicted value or class according to the metric but it does not learn a discriminative function (model fitting) and there is not training phase [10]. In terms of benefits, k-NN is straightforward to understand, versatile (classification and regression) and produces high accuracy. Regarding the drawbacks, k-NN implies high memory requirements (computationally expensive), sensitive to scale of data and performance severely degraded in high dimensional problems.

(4) **Naïve Bayes**. Significant changes in the probabilities reflect the dependence between the predictand and the predictors (predictability). Then, Naïve Bayes is a classification technique that uses this independence resource via the Bayes Theorem, thus producing a reduction of parameters. Briefly, it is a probability-based model that seeks to maximize the function linking the different events. Some studies have trained Naïve Bayes algorithms for predicting the condition of the highway pavements based on previous pavement condition ratings [11] or the classification of surface conditions to predict the Condition Survey Rating Scale (CSRS) category based on explanatory variables such as average rut depth or drainage condition [24].

(5) **Random Forest** (**RF**). By aggregating many trees, the instability of the trees can be reduced, and their performance improved. This idea is one of the fundamental concepts of bagging that consists of the random selection of M subsamples (bootstrapping) given N observations. Subsequently, M trees are fully grown in parallel (overfitted) and finally, a prediction for new input data is given based on the prediction from M individual trees (mean value generally). RF is an improvement over bagged trees.

(6) **Logistic Regression**. Logistic regression performs the classificatory duty by maximizing a quantity known as likelihood, where this amount is no more than the product of the probability of density distributions. In Eq. (2), L is the likelihood which depends on the parameters α_i and the input-output pair of samples (x_i, y_i) − x_i is the vector of features and y_i the observed class-, and pdf is the probability density function. Logistic regression is widely used for binary classification.

$$L(\alpha_i; x_i, y_i) = \Pi pdf(y_i|x_i) \tag{2}$$

2.3.2. Unsupervised learning (UL)

UL is utilized to find patterns in data and draw inferences from data sets that only have input data. It is often used for exploratory data analysis and clustering. The most common algorithms are (1) Hierarchical clustering, (2) K-means and (3) K-medoids.

(1) **Hierarchical clustering.** The main objective of the cluster analysis is to group or segment a collection of objects, understood as a set of measurements, into subsets or "clusters," so that those within each cluster are more closely related to one another than objects assigned to different clusters [25]. To determine whether two instances are similar, there are two main types of estimations, distance (e.g., Minkowski) and similarity measures (e.g., Extended Jaccard coefficient).

(2) **K-means.** This is a partitioning method, i.e., it relocates instances by moving them from one cluster to another, starting from an initial partitioning. Such methods typically require that the number of clusters will be pre-set by the user (K clusters). The basic idea is to find a clustering structure that minimizes a certain error criterion that measures the "distance" of each instance to its representative value. The algorithm starts with an initial set of cluster centers. In each iteration, each instance is assigned to its nearest cluster center according to the Euclidean distance between the two. Then the cluster centers are recalculated via Eq. (3) where $N = \{N_1, N_2 \ldots N_K\}$ are the k-clusters and $(x_1, x_2 \ldots x_n)$, vectors.

$$\mu_i^{t+1} = \frac{1}{N_i^{(t)}} \sum_{x_j \epsilon N_i^{(t)}} x_j \tag{3}$$

(3) **K-medoids.** Very similar to K-means, each cluster is this case is represented by the most centric object in the cluster, rather than by the implicit mean that may not belong to the cluster. Moreover, it is more robust than k-means in the presence of noise or outliers, but its processing is computationally costly.

2.3.3. Reinforcement learning

Unlike SL and UL, RL works with data from a dynamic environment, with the objective of looking for the optimal sequence of actions that will produce the most reward in the long run. RL are divided into two groups: model-based and model-free. Reinforcement learning has three most important distinguishing features. The learning system's actions influence its later inputs (closed-loop). The learner does not have direct instructions, the learner must discover which actions yield the most reward, and where the consequences of actions play out over extended time periods [26]. The most commonly used algorithms for predictive tasks (artificial neural networks) and for image processing (convolutional neural networks) are shown in detail below.

2.3.4. Artificial neural networks (ANNs)

Neural networks were first proposed in 1944 by Warren McCullough and Walter Pitts. Thereafter, the extraction of knowledge from them has been a multidisciplinary breakthrough [27].

A neural network is an interconnected set of single processing units (neurons, nodes or cells), communicating with each other, where the processing capacity of the network or intensity of the connection is defined by weights [28]. To introduce the terminology, the example proposed in [29] (Fig. 2, Fig. 3) will be used.

This study starts from a dataset that relates the amount of bitumen ("dosage") and the efficacy of a given pavement. Thus, the objective is to learn a function that indicates what efficacy is given a specific bitumen quantity. Therefore, a neural network is constructed as in Fig. 3, where the input neuron receives information about the amount of dosage and the output neuron must be able to predict the efficacy of the bitumen mixture.

Structurally, between the two neurons is a layer of neurons, or rather a hidden layer consisting of two neurons. The image represents a simple system, consisting of an input layer and an output layer; however, there

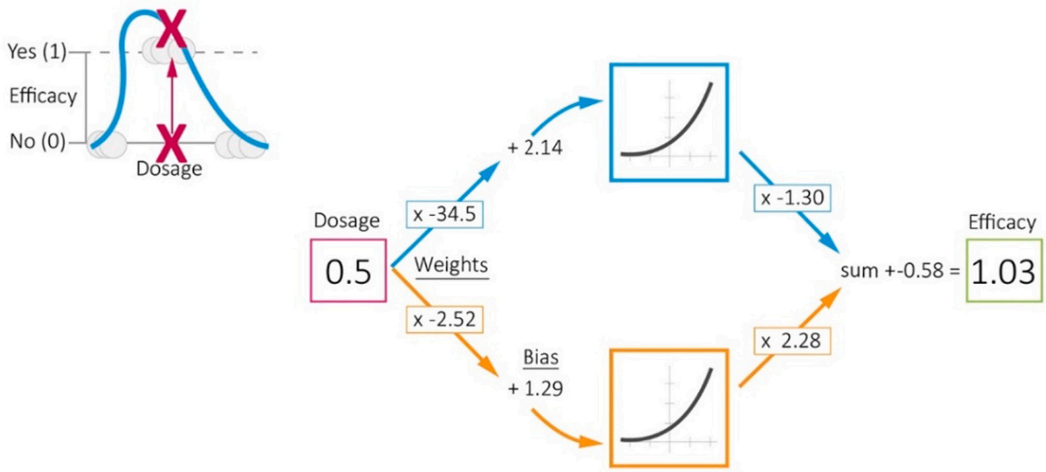

Fig. 2. Didactic representation for understanding the terminology of neural networks.

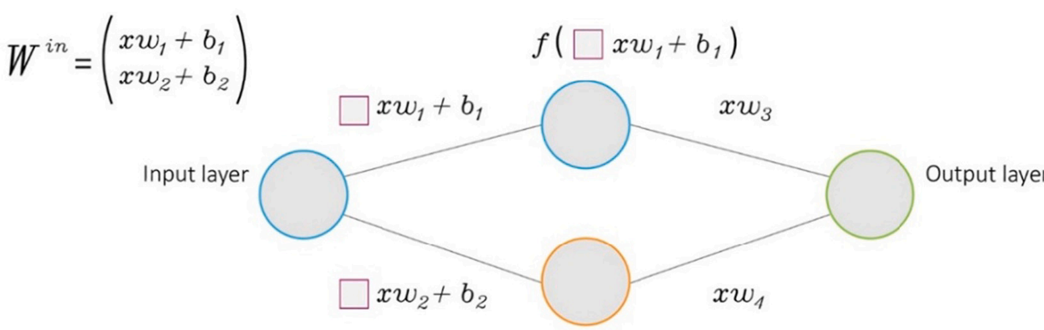

Fig. 3. Didactic representation of Fig. 2 with mathematical notation.

can be multiple hidden layers between the two, where each layer can have different numbers of neurons. The greater the number of layers, the greater the complexity of the neural network. In this context, the neurons of the hidden layers and the output layer receive as input the corresponding weighted or scaled values by means of weights; i.e., the connection between neuron i of one layer and neuron j of another is defined mathematically by a weight, w_{ij}. In addition, a quantity called bias θ_k is added. On the other hand, the information that comes out of each neuron is the input value that it will receive from one or more consecutive neurons after applying a function to it, called activation function (some examples of activation functions are: Rectified Linear Unit - ReLU, Leaky ReLU, etc.). Most importantly, they must be chosen according to their application. In short, each neuron receives an input from neighboring neurons and uses it to compute an output, propagating the information. At the mathematical level [30]:

$$s_k(t) = \Sigma\, w_{jk}(t) y_j(t) + \theta_k(t) \rightarrow F_k(s_k(t)) \tag{4}$$

where $s_k(t)$ is the weighted information and F_k is the non-linear activation function. In our didactic example, the activation function is called soft plus. Consequently, if the weights and biases are known, an architecture like the one proposed in the previous example can be drawn, which will allow to obtain a predictor function (blue graph in Fig. 2). Another very important question is about the determination of the optimal weights and biases that lead to the construction of the function, and the answer lies in the method of backpropagation.

First, before addressing the concept of backpropagation, it is important to stress the concept of the cost, error or loss function [31]. The network will be trained on a training data set, where both the amount of bitumen and the efficacy are known (observed data), with the objective of testing the trained neural network on a data set independent of those used in training (validation set), where only the amount of bitumen is known, and efficacy must be predicted. The goal is therefore to find the neural network configuration that best fits the data, maximizing its generalizability; i.e., to find the neural network configuration that minimizes the error function. Again, there are multiple cost functions such as Mean Squared Error (MSE), L_1 or L_2. The more optimal the parameters are, the smaller the value of the error function will be.

That is, to find the value of the parameters, it is necessary to minimize the cost function (optimization problem). To find the minimum it is really useful to use gradient descent. Consequently, the workflow of neural networks consists of random parameters initialization. Subsequently, the chain rule is applied, calculating the derivatives of the cost function with the aim of reaching its minimum, thus optimizing the parameters (backpropagation). Finally, the weights are updated based on the derivative calculation (gradient descent).

In the example proposed above, assuming that the unknown parameter is b_3 (Fig. 3), then for a cost function of the type MSE, one would have that:

$$Loss = \Sigma\, (observed_i - predicted_i)^2 \tag{5}$$

$$\Delta\omega = \frac{dLoss}{db_3} = \frac{dL}{dpredicted}\,\frac{dpredicted}{db_3} \tag{6}$$

$$\omega = \omega + \gamma\, \Delta\omega \tag{7}$$

where γ is called learning rate and is a hyperparameter. In order to minimize the cost function (Eq. (5)), the method of backpropagation is applied which consists on propagating the error, through the backward chain rule (gradient descent calculation), Eq. (6), and updating the weights (Eq. (7)).

In conclusion, from the descriptive example, it can be concluded that a neural network is a set of layers (input, hidden and output) formed by neurons, which are connected to each other by means of parameters (weights and biases). These parameters are unknown (initially generated randomly) and the way to find them consists of minimizing the cost function by means of the backpropagation method. Therefore, it is interesting to modify or adapt the hyperparameters, with the aim of maximizing the functionality of the neural network, varying the number of layers, the number of neurons per layer, the learning rate, the activation functions, etc. In addition, to try to improve the generalizability of the cost function, a regularization term can be added. There are different types of regularization such as the L_2 (Eq. (8)) or L_1 or dropout norm.

$$Loss = \Sigma\, Loss_i + \beta \|w\|^2{}_2 \tag{8}$$

where the first term has already been introduced and its main task is to fit the model as closely as possible to the data. The regularization term, on the other hand, tries to make the weights as small as possible, thus reducing the complexity of the network. In other words, the regularization term arises as a solution to a possible overfitting in the training, giving rise to a high variance in the validation dataset (loss of generalization). In addition, there are several modifications or improvements to gradient descent. Such improved algorithms are called optimizers. Next, convolutional neural networks, a subtype of ANNs, are analyzed.

2.3.5. *Convolutional neural networks (CNNs)*

The state-of-the-art related to feature extraction dealing with images for pavement evaluation using machine learning techniques based on image processing focuses on the same algorithm but with different architectures [32], convolutional neural networks. The main differences between traditional ANNs and CNNs are listed as follows:

- ANNs learn global patterns from the input feature space, contrarily CNNs learn local patterns. In the case of images, ANNs use all pixels and CNNs use small 2D windows or filters.
- The patterns learned by CNNs are invariant to translations. That is, if CNNs recognize a given pattern in the lower left corner, they will find it anywhere in the image. ANNs, however, would have to learn the pattern again if it appears in a new location.
- The most beneficial aspect of CNNs is the reduction of parameters, becoming a very efficient approach at processing images with high generalization power.

To begin dealing with CNNs it is necessary to introduce the concept of convolution and convolutional layers. Convolution is a linear operation that allows the extraction of image features. In fact, when considering Fig. 4, it can be observed the concept of convolution where the input image (receptive field) is a crack and its representation in pixels is convolved with another matrix called kernel, filter, or mask. The filter is shifted in such a way that it maps the input image. The convolution is the sum of the product of element-by-element matrices. As a result, the filter maps the entire receptive field into the so-called feature map.

The masks represent the connectivity between successive layers,

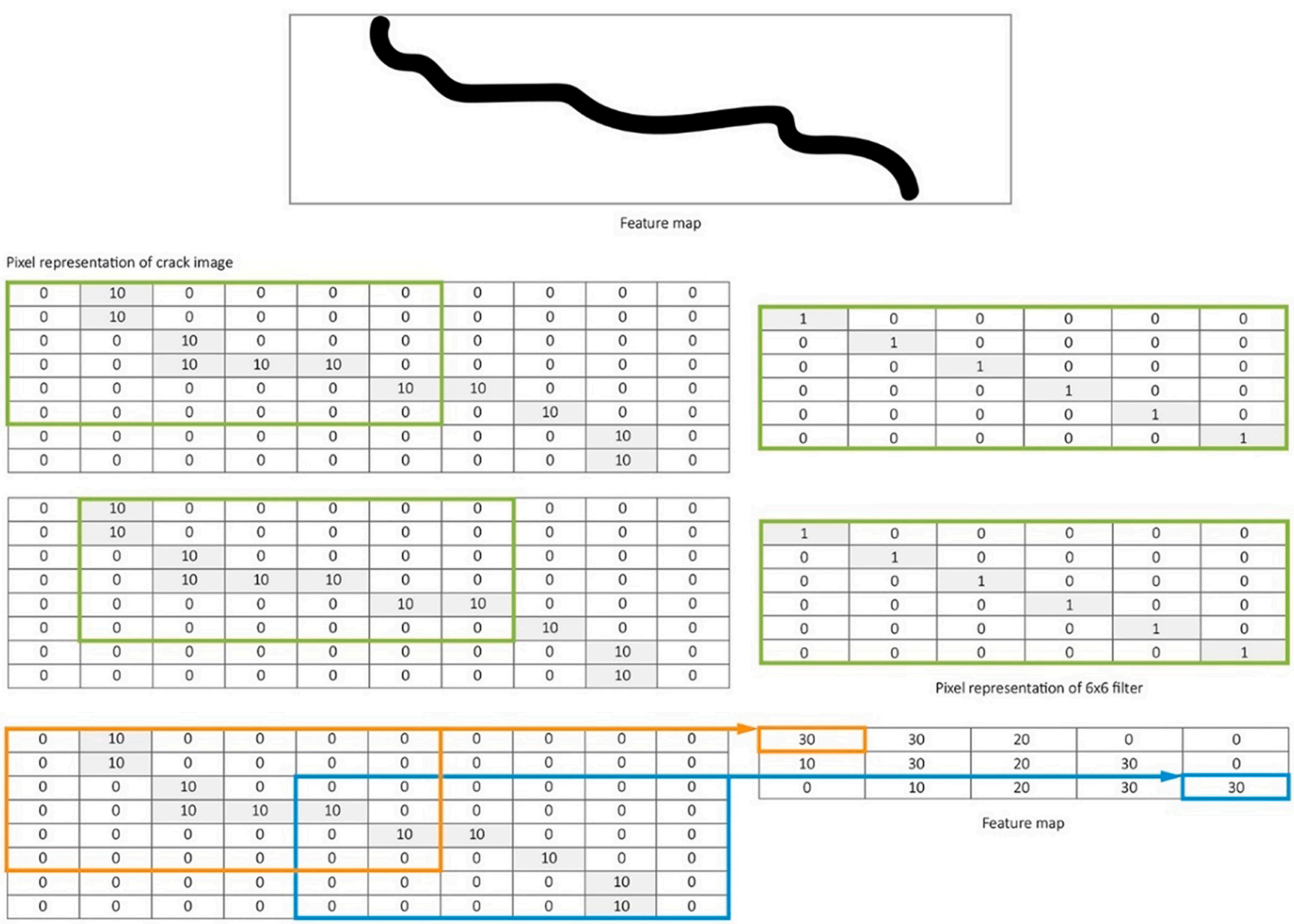

Fig. 4. Representation of convolutional operation of a crack image.

where the network weights correspond to the values in the filters. Then, each layer can have up to 3 dimensions: base, height, and depth. Where the depth refers to the number of channels. One of the problems with convolutional layers is that they can cause the image to shrink too much. In addition, corner and edge pixel information will be under-represented in the feature map. For this, the concept of padding is introduced, which aims to make the input and output image have the same dimensions. In other words, it is a matter of extending the input image by interpolating nearby pixels. For example, although the image has dimensions (n,n) and padding (p): $p = 1$ is applied, the image is magnified to dimensions $(n+2,n+2)$. Another important hyperparameter to configure is the stride which is just the size of the filter movement step along the receptive field mapping. Also, there are other types of layers that reduce the dimensionality of feature maps, max-pooling and average pooling, which reduce a region of the image to the maximum or average value, respectively. Analogous to traditional neural networks, activation functions are also applied between the convolutional layers.

Finally, while the compendium of convolutional layers handles feature extraction, the question is how the classificatory task is performed to detect crack types (Fig. 5). To do this, what is done is to flatten the last convolutional layer, recovering the structure of the traditional neural networks to perform the classification task (flatten or global average layer pooling). Fig. 6

When the state-of-the-art regarding pavement crack classification is shown, different studies propose different architectures and give a name to their proposal. Therefore, proposing an architecture is just indicating hyperparameters (kernel size, padding, stride, number of layers, types of layers, etc.), topology (connection mode between layers) and optimization algorithms (gradient descent methods).

In order to homogenize the different models presented above, Table 2 shows the algorithms, their advantages and disadvantages, and the types of data (imaging, GPR, laser and fiber optic) on which they are found or on which they could be optimally used. In addition, it will help the research community to choose study models based on the characteristics of their datasets.

3. Test pavement performance

3.1. Pavement condition indicators

Pavement condition assessment includes the technical characterization of the pavement, considering its physical characteristics (e.g., roughness, friction, distress type, etc.). Eventually, several indexes have been developed where some of the most applied are: Pavement Condition Index (PCI), Pavement Serviceability Index (PSI), Pavement Quality Index (PQI) and IRI [33].

In the next section, we will show the different data collection methods and how ML algorithms can solve a problem. Thus, when dealing with image related studies, ML metrics will be used to see how well the algorithm detects or classifies but no pavement condition indicators will be calculated. In fact, it is considered that this should be one of the next steps or efforts in future research (e.g., if there is an image problem, to determine the type of deterioration as well as its extent and frequency of occurrence, in order to build a new dataset that allows to extract a new standardized index).

However, in other studies, from field data or using databases such as those discussed later, ML metrics will show how well the model (regression problem) predicts the numerical value of the indicator (usually IRI or PCI). Therefore, the following is a brief description of pavement quality indicators and the various ML metrics for evaluation.

3.1.1. Pavement performance indicators

The IRI is most commonly obtained from measured longitudinal road profiles; i.e., vertical deviations or irregularities of the pavement surface from planar plane [34]. It is calculated using a quarter-car vehicle math model, whose response is accumulated to yield a roughness index with units of slope [35]. The IRI index can be estimated after performing profilometric measurements carried out on road pavements using specific laser devices.

The PCI, presented by the US Army's engineering department [36], is a numerical indicator that rates pavement condition according to a rating scale from 0 to 100. The mainly influencing factor for PCI calculation is surface distress and it depends on: distress type, level of severity and density of distress. Consequently, it would be a suitable index to extrapolate imaging problems to the calculation of the PCI.

The PSI estimates the serviceability rating from measurements of several physical parameters from surface distresses such as cracking, rut depth and roughness. It ranges from 0 (worst condition) to 5 (best condition).

The PQI computes the overall pavement condition combining pavement roughness and distress magnitudes. There are different formulations of the indicators, where the Minnesota Department of Transportation is the most widely applied.

3.1.2. ML metrics

To assess how optimal the developed model is, different metrics are used that relate the observed data to the predicted data. The most widely used metrics for classification and regression problems in ML algorithms as pavement performance evaluation techniques are shown below.

The classification problems related to pavement assessment are

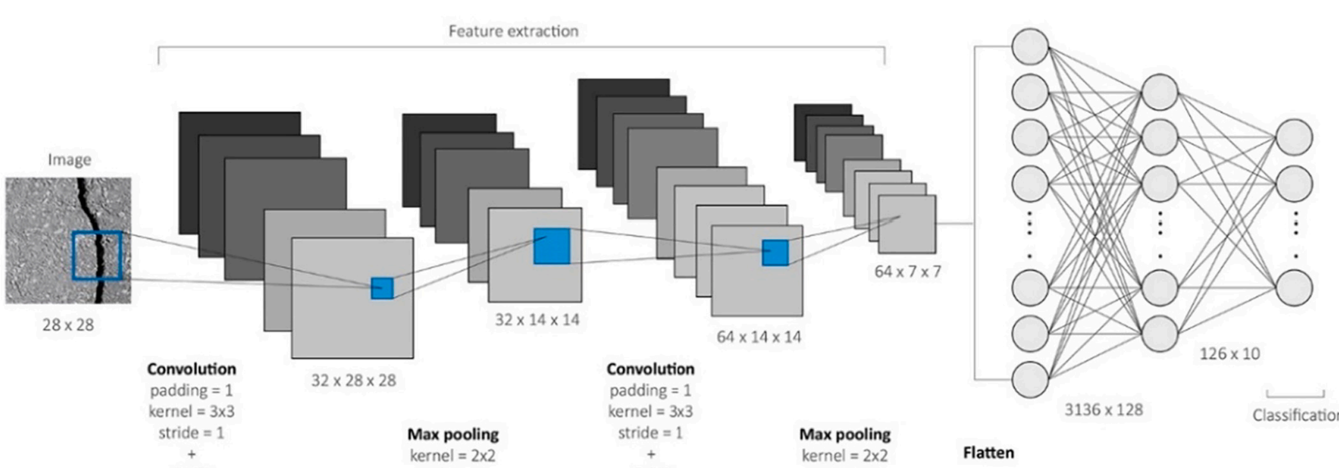

Fig. 5. Structural representation of a convolutional layered architecture with image classification capability.

Table 2
Brief description of the ML algorithms with benefits, drawbacks and association with data type.

Algorithm	Pros	Cons	Data Type
CART [8]	Easy to interpret and explain the interaction between features; they can handle numeric, nominal and textual data; normalization or scaling not needed; irrelevant features won't affect.	It hardly handles high dimensional data (high computational requirement-data fragmentation problem); overfitting issue in case of correct pruning lack; sensitive to data changes.	FO, LS
SVM [9]	It usually provides high accuracy; it prevents theoretical guarantees regarding overfitting; accuracy and performance are independent of size; it deals correctly with high dimensional data and good generalization ability; outliers have less impact; it suits extremely well for binary classification and separable classes.	Accuracy dependent on the number of training cycles; slow processing for large datasets; poor performance with overlapped classes; dependent of the appropriate hyperparameters and kernel function.	ALL
k-NN [10]	Simple to understand and implement; there are no assumptions about data (e.g., dependency of variables); well suited for multi-modal classes; well evolving model to new data points.	Lower efficiency for large datasets (curse of dimensionality); performance dependent on selecting good value of "k" (cross validation but it is computationally expensive); performance varies according size of data (scaling is compulsory); it doesn't work well on imbalanced data.	FO, LS
Naïve Bayes [11]	Assuming independence correlations (class conditional independence), converges fast (useful for real-time predictions); lower computation training time; scalable with large datasets; insensitive to irrelevant features; adequate performance with high dimensional data.	Bad estimator; it assumes that all predictors are independent, rarely happening in real problems; inaccurate representation of data; it assigns zero probability to a categorical variable whose category in the test data set wasn't available in the training dataset (zero-frequency problem).	FO, LS
RF [12]	Fast; scalable; robust to noisy data; they do not over fit; easy to interpret; reduced prediction error; good performance on imbalanced datasets; it handles well high amounts of data and missing information; it suffers low impact of the outliers.	Slow for real-time prediction as the number of trees (estimators) increases; predictions need to be uncorrelated; it is difficult to understand the different parameters; they are found to be biased with categorical values; not suitable for linear methods with sparse features.	ALL
Logistic regression [13]	Nice probabilistic interpretation; easy to update with new data; small assumptions on the distributions of independent variables; simple to implement; feature scaling and hyperparameter tuning not needed.	It requires large sample size to achieve stable results; poor results of non-linear data and/or irrelevant, highly correlated features and/or when the number of observations is lesser than the number of features.	FO, LS
			ALL
Hierarchical Clustering [14]	No apriori information about the number of clusters required; easy to implement and to understand because the output dendrogram provides visual information.	The algorithm can never undo what was done previously; time complexity of at least o(n^2 log n) is required, where "n" is the number of data points; sensitive to noise and outliers; it hardly handles several sized clusters and convex shapes.	
k-means [15]	Relatively simple to implement; it scales to large datasets; it commonly guarantees convergence; easy adaptation to new instances; it generalizes to several shapes and sizes of clusters (e.g., elliptical).	Dependent of choosing the "k" value; it has troubles where clusters are of varying sizes and density; centroids can be dragged by outliers; it suffers with high number of dimensions.	ALL
k-medoids [16]	Simple to comprehend and easy to implement; it fast and converges in a fixed number of steps; it is less sensitive to outliers than other partitioning algorithms.	Not suitable for clustering non-spherical (arbitrary shaped) groups of objects; it may obtain different results for different runs on the same dataset because the first k medoids are chosen randomly.	ALL
ANN [17]	They are quite robust to noise in the training data, because the training examples may contain errors, which do not affect the final output; they can bear long training times depending on factors such as the number of weights in the network, the number of training examples considered, and the settings of various learning algorithm parameters.	They require processors with parallel processing power (hardware); it makes it very difficult for ANN to understand the problem statement; the ANN solution to the problem statements that we really don't know on what basis it will give the solution.	IM, LS
CNN [18]	CNN-based models achieve state-of-the-art results in classification, localisation, semantic segmentation and action recognition tasks; they minimize computation compared to regular ANN (convolutional operation); they are great at handling image classification and recognition and use the same knowledge across all image locations.	Classification of images with different positions (different angles, backgrounds or lighting conditions); they recognize similar images with different noise levels as the very same picture; CNNs do not have coordinate frames which are a basic component of human vision; GPUs are generally required.	IM, GPR, LS

IM = images, GPR = ground penetrating radar, LS = laser and FO = fiber optics, ALL = applicable for the 4 types of data.

mainly related to crack categories detection such as crack, non-crack, rutting, non-rutting, ravelling, etc. Also, depending on the study, other pavement indicators can be characterized in categories such as road friction estimates (RFE) or road surface condition (RSC). Since cracking can accelerate the deterioration process, crack evaluation is a demanding assignment to ensure public safety [37]. The most common classification metrics are shown below (Table 3):

Accuracy measures the ratio of correct predictions over the total number of instances, Sensitivity quantifies the fraction of positive

Table 3
Most popular classification metrics.

Metric	Formula	Metric	Formula
Accuracy	$\frac{TP+TN}{TP+TN+FP+FN}$	Precision	$\frac{TP}{TP+FP}$
True Positive Rate - TPR (Sensitivity)	$\frac{TP}{TP+FN}$	F1-score	$\frac{2}{\frac{1}{Precision}+\frac{1}{Recall}}$
True Negative Rate – TNR (Specificity)	$\frac{TN}{FP+TN}$	Recall	$\frac{TP}{TP+TN}$

TP = true positive; TN = true negative; FP = false positive; FN = false negative.

patterns that are correctly classified, Specificity represents the fraction of negative patterns that are correctly categorized, Precision measures the number of correct positive results divided by the number of positive results predicted, Recall estimates the ratio of correct positives over all samples that should have been identified as positive and f1-score is the harmonic mean between Recall and Precision [38].

For the evaluation problems with the objective of predicting continuous values such as IRI or PCI, regression metrics are utilized (Table 4) [39].

MSE measures the average of the squares of the errors (error means difference between observed and predicted value), RMSE is the square root of MSE, MAE provides the average of the absolute difference between the observed and predicted continuous variables and R^2 indicates the square of the correlation coefficient R which determines the strength of association between predicted and observed magnitudes [60].

3.2. Open pavement performance datasets using ML algorithms

The main purpose of this section is to mention some of the most relevant data sets used in the state of the art for pavement evaluation using ML algorithms. Such information banks enable the advancement of scientific progress according to FAIR principles [61]. Therefore, they will be mentioned below together with a brief description of their contents, as well as a link to their location on the web.

LTTP database: LTTP is a program that has the aim of collecting pavement performance data as one of the major research areas of the Strategic Highway Research Program (SHRP) in the US. The LTTP program includes two classes, General Pavement Study (GPS) and the Specific Pavement Studies (SPS). LTTP Information Management System contains datasets of data collected under the LTTP program (over 2500 tests). For example, Fathi[62] utilized historical LTTP data and made use of different MLAs in order to predict alligator deterioration index (ADI) using features like air voids (VA), voids in mineral aggregate (VMA) or voids filled with asphalt (VFA). Zeiada et al. [39] studied multiple ML techniques (CART, SVM, ensemble trees, GPR and ANN) to model asphalt pavement performance (IRI was adopted as the pavement performance indicator) from LTTP database of warm climate regions. Younos et al. [63] carried out a research on the impact of climatic conditions and traffic loading characteristics on pavement performance to predict PCI by applying ANNs to LTTP data.

Table 4
Regression popular metrics.

Metric	Acronym	Formula
Mean Square Error	MSE	$\frac{1}{N}\sum_{i=1}^{N}(y_i-\widehat{y}_i)^2$
Root Mean Square Error	RMSE	$\sqrt{MSE}$
Mean Absolute Error	MAE	$\frac{1}{N}\sum_{i=1}^{N}\lvert y_i-\widehat{y}_i\rvert$
R-Squared	R^2	$\left(\frac{\sigma_{y,\widehat{y}}}{\sigma_y\sigma_{\widehat{y}}}\right)^2$

y_i = observed value; $\widehat{y}_i$ = predicted value; σ = variance of a given variable; $\sigma_{y,\widehat{y}}$ = covariance of predicted and observed variables.

RDD2020: It is a large-scale heterogeneous road damage dataset comprising 26,620 images of different pavement distresses from India, Japan and the Czech Republic. Actually, a part of RDD2020 was utilized for Global Road Detection Challenge 2020. Road images were captured using a smartphone running a publicly available image-capturing application developed by Sekimoto Lab [64].

There are other popular datasets such as: SDNET2018, CrackTree, GeoPortale Lamma, RTK, KittiSeg, etc.

3.3. ML frameworks for pavement performance evaluation

All algorithms associated with pavement distress detection (PDD) and classification problems (frames with pavement deterioration) and pavement indicators prediction are programmed with code. For this purpose, the programming languages most commonly used in the state-of-the-art are Python, R and MATLAB. The most commonly used CV library for image processing (applying filters, resizing, normalizing, etc.) is OpenCV, and also Scikit-Image. The most widely applied open source ML frameworks are PyTorch, TensorFlow, Fastai and Caret. In addition, there are open repositories on platforms such as Github where you can find characteristic pre-trained models (interesting to apply transfer learning and alleviate the computational cost) that lately are being included in the articles through links as the research community is progressing towards open data.

4. Application of ML algorithms for different data extraction methodologies

In this section, the different ML algorithms used for different data extraction or data collection techniques will be presented. In terms of the structure, each section will contain a brief introduction and subsequently, each investigation will address if possible 4 concepts (see Tables 5-8): methodology (pre-processing techniques and ML models), type of problem (classification, regression), metadata (dataset and computational resources) and finally, the most interesting section, the discussion (results, difficulties and future investigations).

Table 5
Summary table of image-based analyses.

ID	Reference	MLA	Data source	Amount	Pixels resolution
A	[40]	SVM	Middle East Technical University campus rigid pavement images	109	4000 × 3000
B	[41]	Deep residual network	CrackForest	118	480 × 320
C	[25]	SqueezeNet	YouTube	5300	640 × 360–1280 × 720
D	[42]	DeepCrack	CrackTree260, CRKWH100, CrackLS315, Stone331	1006	512 × 512
E	[43]	FBI-LSSVC-FS	Da Nang city field trip imagery collection	2000	32 × 32
F	[44]	VGG16	SDNET2018	5200	256 × 256
G	[45]	Mask R-CNN	–	45	5184 × 3456
H	(G. [46])	PvmtTPNet	PaveVision3D imagery system	21,000	4096 × 2048
I	[47]	GoogleNet	–	2250	900 × 1000
J	[48]	LightGBM	–	98	5184 × 3456

MLA = Machine Learning algorithm, data source is the name of the dataset or the data collection system and amount is the number of images.

Table 6
Summary table of GPR analyses.

ID	Reference	MLA	Data type	Data device	Samples	Resolution
A	[49]	OCSSVM	2D B-scan images	Finite-Difference Time Domain method and Accelerated (gprMax) Pavement Test	1174	–
B	[50]	YOLOv5	3D B-scan images	GeoScopeTM Mk IV	1750	320 × 320
C	[51]	Faster RCNN	2D B-scan images	Finite-Difference Time Domain method and Accelerated (gprMax)	30,000	–
D	[52]	SVM	2D B-scan images	LTD-2100 (China Electronics Technology Group Corporation)	100	–
E	[53]	Faster RCNN	2D B-scan images	Unknown GPR data collection device	1683	–

MLA = Machine Learning algorithm, data device is the name of the data collection system or software program and samples refer to the number of images.

Table 7
Summary table of optic fibers analyses.

ID	Reference	MLA	Data source	Data type (outcome)	Vehicles
A	[54]	SVM	3-D glass fiber-reinforced polymer packaged fiber Bragg grating sensors	Wavelength changes (vehicle type)	477
B	[55]	SpeedNet	Distributed fiber-optic sensing	Wavelength changes (average vehicle speed)	165,000
C	(T. [56])	CNN + SVM	Distributed Acoustic Sensing	Wavelength changes (sonic nap alert vibration)	–

MLA = Machine Learning algorithm, data source is the name of the data collection system, data type (outcome) is the obtained data (and the final calculated outcome) and vehicles mean the number of means of transports passing through the optic-fiber system.

Table 8
Summary table of laser analyses.

ID	Reference	MLA	Data source	Data type	Amount
A	[57]	ANN	LTTP database	Numeric dataset	300 instances
B	[58]	Random Forest	Georgia Tech Survey Vehicle	3D pavement surface images	–
C	[59]	SVM	Geoportale Lamma	2D SAR images	210 images

MLA = Machine Learning algorithm, data source is the name of the dataset or the data collection system and amount is the number of images or instances in case of numeric datasets.

4.1. *Image-based techniques*

4.1.1. *Fundamentals of image-based approaches*

Image-processing techniques to determine road conditions are considered as an encouraging non-destructive testing (NDT) method to quantify pavement distresses by evaluating pavement surface images [65]. Computer vision (CV) modules are becoming an integral component of contemporary Structural Health Monitoring (SHM) frameworks. In this respect, the present section explains the state-of-the-art CV methodologies, which are used to automate the process of defect and damage detection [66]. Thus, active safety systems and self-driving vehicles can unquestionably benefit from real-time prediction of drivable surface conditions.

Effective pavement rehabilitation polices can only be established with reliable prediction of future pavement cracking rates based on quantitative assessment of past and present pavement conditions. Image-based crack-recognition techniques have been employed to provide necessary quantitative measures of cracks in pavement surface images. In CV, crack can be defined as a group of low-intensity pixels compared to neighboring pixels. In fact, to deal with multi-level topological shapes of crack images, different image processing levels need to be employed for computer-aided crack recognition: crack extraction (non-crack background removal of input images), crack grouping (group fragmented crack pixels extracted in crack extraction by image segmentation), crack detection (image components) and crack classification.

Formerly, the labeling and quantification of the severity, type, and extent of surface cracking, was a challenging area for evaluating the asphalt pavements [67]. Image-based crack detection methods have been extensively studied due to their cost-effectiveness in terms of data acquisition and processing [68]. The most relevant issue is that automated pavement distress detection and classification has remained one of the high-priority research areas of transportation agencies. Image classification based on ML models is turning into the main application and study tool, in order to avoid preventive road maintenance. Actually, image-based models focus primarily on crack detection and classification. In addition, the state-of-the-art in recent years focuses on different image preprocessing methods and different architectures, where the most widely used are CNNs, with resources such as transfer learning.

Nonetheless, existing crack detection methods still have constraints due to their insufficiency to overcome inherent challenges associated with pavement images, such as background complexity, inhomogeneity of cracks or diversity of surface texture. Hence, regarding the state-of-the-art that implements the automated pavement detection by dint of ML models, a qualitative and quantitative description of the most recent models based on CV are shown below.

4.1.2. *Studies employing ML algorithms from images*

For a better understanding of the sub-section, a chronological and indexed descriptive table is shown below, followed by a further explanation of each study whose source of information is digital imagery.

A. In 2017, an unmanned aerial vehicle (UAV), DJI Inspire 1 Quadcopter, was developed as a crack identification system for monitoring rigid pavements [40]. The process accomplished can be divided into 2 stages: crack detection (image resizing, grayscale image transformation, thresholding, enhancement via median filtering and morphological operations) and crack identification on the basis of different properties such as extent, aspect ratio, eccentricity and circularity ratio. SVM algorithm was used as a binary categorization model (i.e., discerning between crack and non-crack classes, C&nC). The UAV provided 109 images from Middle East Technical University Campus at different altitudes ranging from 0.5 m to 3 m with 1–3 $m \bullet s^{-1}$ speed. It provides encouraging results obtaining 90% specificity and 97% accuracy. The proposed methodology serves as an alternative cost-effective solution for pavement monitoring but it involves some drawbacks. The restricted amount of data limits the model performance. Differently, other ingredients that cause performance failure are shadowy and low-resolution images.
B. A new approach was developed in 2018 using deep residual neural networks with transfer learning [41], to detect road crack at pixel-level. The dataset used, CrackForest, contains 118 images of the road surface in Beijing, China. Regarding computational resources, GTX 1080 Ti GPU i7–700 has been used. The investigation goal focuses exclusively on the model improvement

referring to previous analysis for crack detection, ergo model refinement. It evidences validation metric improvements: precision (93.57%), recall (84.90%) and f1-score (89.03%). The main reasons for the improvement are the use of transfer learning given the reduced number of images to train and the use of residual convolutional neural networks as a solution to CNNs with possible problems such as vanishing gradient.

C. Also, in 2018 an alternative investigation performs a comparison between CNN-based architectures for RSC and RFE determination [25]. The experimental part related to RSC classification compares several models: CNN, SqueezeNet [69] and feature-based model. Regarding RFE categorization, a rule-base method was considered. As pre-processing techniques, normalization was implemented in RSC categorization whereas patch segmentation and quantization for drivable surface determination was applied in RFE cataloguing. It is a 2-stage approach relative to multi-categorical classification of RFE and RSC. The RSC hierarchically detection serves as a previous step for RFE arrangement in order to estimate the patchiness in the vehicle's ego-lane. The images were collected from YouTube video sequences, ensuring variability across 5300 images (70–30% for train-test partition). SqueezeNet obtained better results: 97.36% accuracy, 97% precision and 97% recall and optimum computing speed, 0.69 train-hours and 4.0 milliseconds. Regarding the benefits, image resizing smooths model complexity, also during stage-1, training data is augmented to six-times by implementing histogram-based image equalization and contrast enhancement to decrease overfitting (small dataset) and it improves the state-of-the-art. The difficulties are mainly the lack of camera calibration information on public data (need of a high quality and calibrated dataset), the overfitting risk due to small datasets and misclassifications.

D. DeepCrack, is a CNN-based model built on the SegNet network which contains an encoder-decoder architecture blossomed in 2019 [42] in order to develop a pavement crack detection model. It is a binary classification problem (C&nC problem). A collection of 4 datasets have been deployed: CrackTree260 (260 road pavement images captured by area-array camera under visible-light illumination enlarged to 35,100 images thanks to DA), CRKWH100 (100 images with similar requirements and 1 mm of distance sampling), CrackLS315 (315 images under laser illumination) and Stone331 (331 stone images with analogous requirements). CrackTree260 was used for training and the rest, for testing procedure. The experiments were tested using GeForce GTX TITAN-X GPU. As for DeepCrack details, batch normalization accelerates convergence, is capable of detecting tiny cracks without producing false positives, runs efficiently due to the lack of fully-connected layers, fuses multiscale architecture showing metrics improvement, performs inadequate detection dealing with noisy crack images and could not detect bright cracks (the brightness of the training images was reversed in such a way that higher intensity than the background). DeepCrack achieved 0.87 f1-score on the test datasets in average.

E. Moreover, in 2019 an innovative model was analyzed, the forensic based investigation-least squares support vector classification-feature selection (FBI-LSSVC-FS). Initially, pavement texture features are extracted via Gabon filter and discrete cosine transform, serving as predictors for LSSVC for image binary classification (rutting and non-rutting categories) where model optimization is implemented by FBI [43]. Through Cannon EOS M10, 2000 balanced images were collected and the model was deployed via HP Z440 workstation. The suggested model achieves optimal results in terms of 0.994 precision, 0.984 recall and 0.989 F1-score. On the interpretation of the results, it is confirmed that the efficiency of the model decreases in the presence of images with strong shadow texture and irregular patterns.

F. Another study contemporary to the previous one for binary concrete crack classification (C&nC problem) compares 4 different CNNs (small CNN with/without DA and large pre-trained CNNs with DA and with/without hyperparameter fine-tuning [44]. A public labeled dataset was used, SDNET2018 [37], but just a balanced sample of 5200 images was deployed. Dell Inspiron 15 with 64-bit, 32GB RAM, i7 and NVIDIA GeForce GTX 1060 Ti GPU was used for experiments. In terms of quantitative comparison, large CNN (VGG16, [70]) with DA and fine-tuning obtained greater validation metric results achieving 95% and 93% training and validation accuracy. Nevertheless, the biggest challenge posed by the simpler model was the overfitting, solved with DA inclusion. In order to improve training/validation metrics, large CNNs and hyperparameters fine-tuning produced superlative results.

G. In 2020, a model was presented which implements semantic segmentation through deep learning for crack detection owing to Mask Residual CNN, also applying hyperparameter fine-tuning [45]. It is indeed a binary categorization problem (C&nC study) but with extra labels, traces of tie-rods and form works, in order to avoid misclassification errors (false detections). Firstly, 45 images were collected, notwithstanding poor accuracy was retrieved. Consequently, each image was cut into smaller portions, improving validation metrics. The proposed method reduces the influence of false positives considering extra labels and it enhances the classification of adverse images with shadows or dirt. Furthermore, a quantitative increase is obtained: from 0.9915 to 0.9921 accuracy, from 0.7881 to 0.7847 sensitivity, from 0.9927 to 0.9933 specificity, from 0.3924 to 0.4044 and f1-score went from 0.4862 to 0.4994. Detailed object detection leads to poor accuracy recognition (shrunk image) but portion cut provided learning accuracy improvement due to higher detail capacity.

H. PvmtTPNet is a CNN subvariant to learn feature from pavement categories, flourished in 2021 (G. [46]). It automatically recognizes different pavement types: asphalt concrete, jointed plain concrete and reinforced concrete. For this purpose, PaveVision3D system collected 23,000 images of 1-mm resolution covering Oklahoma fields. Development stages used NVIDIA TITAN V GPU card services. The proposed methodology deals favorably with overfitting and performs training and testing accuracy of 91.27% and 96.66%. The model's operability was impaired by the misclassification of bridge deck, probably due to the similarity of concrete sections of bridge decks and rigid pavement sections and they should be included in future related investigations.

I. SqueezeNet, GoogleNet, Demesne, AlexNet and Inception (pre-trained CNN-based architectures) were compared in terms of multi-classification (linear, non-linear and surface cracking), computational speed, model complexity, feature extraction and validation metrics in 2021 [47]. Previously, a step-by-step pre-processing techniques compendium was applied: histogram equalization for image contrast enhancement, gaussian filter for image smoothing, wavelet transform for defects improvement and thresholding and morphological filtering for noise removal. About computational resources, Intel Core i7-4710HQ running GeForce GTX 850 M GPU performed analysis procedure. SqueezeNet and GoogleNet have better performance than the other pre-trained models in terms of simpler structure, less complexity and adequate validation metrics when dealing with a small dataset achieving a train-test accuracy of 0.991–0.989, a precision of 0.986–0.984 and f1-score of 0.986–0.984. What is clear is that the use of pre-trained models gives better results for small datasets. It also highlights the efficiency of the road detection system thanks to the image processing procedure.

J. Eventually, LightGBM is an efficient gradient boosting decision tree developed in 2021 for C&nC classification [48]. Regarding methodology sequential steps: grayscale conversion to reduce calculation time through maximum RGB (red-green-blue) value, correction of contrast differences resulting from shadows or dirt via median filter considering that filter size affects the adequate precision, use of the crack target pixel and pixels surrounding as features to avoid false detections due to dark appearance through gaussian filter and non-square kernels for crack extraction and noise removal looking upon geometric characteristics. A total amount of 98 images of 5184 × 3456 pixels size were analyzed using NVIDIA Quadro RTX 8000, Intel Xeon W-2195 CPU and 512 GB RAM. The pre-processing techniques lead to higher validation values obtaining an accuracy of 0.9935, precision of 0.4040, sensitivity of 0.7880, specificity of 0.9945 and f-measure of 0.1565. Nevertheless, it also involves some obstacles such as the wrong recognition of the exact boundaries detecting fine cracks, plastic corn and formwork. The addition of geometric features is proposed as an achievable solution. In the future, including Principal Component Analysis (PCA) for reducing calculation time is another improvement.

4.2. Ground penetrating radar (GPR)

4.2.1. Fundamentals of GPR approach

GPR has recently been considered as a pavement quality control and quality assurance method. Notwithstanding, GPR has been used for asphalt concrete pavement density prediction for the past two decades [71]. GPR is a non-destructive, cost and time-effective testing method, widely applied for monitoring civil structures. Utterly, GPR data are used to predict pavement density and layer thickness, and the detection of anomalies underneath pavement surfaces, estimating its remaining service life and pavement performance, etc. Among available non-destructive test techniques, GPR has contrasting benefits: performing potential high-accuracy measurements, a large-coverage area, and relatively high-speed surveys.

GPR is a multidisciplinary non-destructive evaluation technique that requires knowledge of EM wave propagation, material properties and antenna theory [72]. GPR data interpretation can be provided thanks to different ML algorithms, mainly different CNN topologies for feature extraction, selection and damage classification. Nonetheless, GPR poses some adversities such as the difficult interpretation of the data, since it is not physically competent to detect layers unless there is sufficient dissimilarity in their dielectric constants [73].

The different studies based on ML models from GPR survey data are shown below, with analogous structure to the previous section, except that in this case the detailed research will be realized in 2021.

4.2.2. Studies employing ML algorithms from GPR data

The following is a structured representation of the studies associated with the use of pavement GPR data for subsequent analysis using ML algorithms.

A. In order to detect horizontal stratified thin debondings or delaminations (inter-layer cracks) from B-scan images focusing on air-voids defects from B-scan images, One-class SVM (OCSSVM) was applied [49] to classify between 2 classes (debonding and non-debonding). Different pre-processing techniques were applied via the Sensitive Analysis which is the study of uncertainties between input and expected output, studying noise, data size and debonding thickness. Training data was generated by an EM simulation software based on Finite Time Domain method, *GprMax*, and testing samples from an Accelerated Pavement Test at the University of Gustave Eiffel, Nantes. Both datasets were acquired in 2 configurations: ground-coupled and air-launched GPR. In terms of results, the noiseless learning data is maintained invariant in the feature distribution, thus, the learning data size can be decreased. Regarding debonding thickness, a priori knowledge about 2 classes is required during the learning step, the imbalance of the dataset produced lower performance of the model and some misclassifications occurred. However, time resolution increased by increasing frequency, thus providing a better performance (just some false alarms were detected probably close to the hyper-sphere boundary). Validation metrics were not provided. The proposal limitations are the debonding data presence in the learning partition leading to false detections and the impossibility to apply the model for multiclass classification.
B. Looking upon another contemporary investigation, a CNN architecture-based which is a subvariant of YOLO (You Only Look Once) model (YOLOv5) was used to scrutinize internal defects from GPR images and compares traditional GPR detection via maintenance benefits [50]. Cracking and void are the two categories (settlement excluded due to different actuation scale). The inverse discrete Fourier Transform, DA and background removal were used for data preprocessing. Regarding data collection, B-scan images (1397/179/174 images for training/validation/test of 320 × 320 pixels) reflect basic internal features through GPR. The labelling process of those images was performed manually and with LabelImg software. Computational features are AMD Ryzen 52,600× CPU 16 GB memory. The results demonstrate that maintenance cost is reduced by $49.398/km, and the energy consumption and carbon emissions are reduced by 16.94% and 16.91%, respectively. Also, relevant metrics are obtained with 0.76 of precision, recall of 0.94 and 0.82 of f1-score.
C. Referring to another approach, a Faster Residual CNN (Faster RCNN) was optimized using GPR B-scan images in which the characteristic visual patterns of defeat can be leveraged for detection using visual descriptors [51]. It is a typical two-stage detection neural network (series proposal region and classification). It is a solution for automatic detection of roadbed subgrade defect by Faster RCNN. The aim was predicting just one class, defect, which includes all the different subgrade defects types. 30,000 roadbed defect GPR B-scan data have been simulated by the simulation software gprMax (also applying DA), which are labeled automatically in terms of 4 magnitudes: maximum/minimum values in x/y direction. Faster RCNN (pre-trained model on Imagenet database) is a compromise between accuracy and ease of comparison, achieving an average precision of 0.8067. Future investigations will concern the addition of detecting more labels (pixel level) to develop more complex classifications.
D. With regard to the study of automatic detection of road hazards (cracking, hollowing and subsidence) based on GPR imaging, LTD-2100 system, a novel analysis was carried out using a modified SVM [52]. Firstly, the original images (100 images) suffered from environmental issues such as reflection and clutter types. Therefore, to extract those shortcomings, an image pre-processing based on the following sequential methods was matured: image filtering (gaussian filter to minimize noise effects and interference), image segmentation (Canny operator for foreground object detection), feature extraction (to satisfy several requirements like translation, rotation and scale-transformation insensitivity) and feature selection (application of K-L algorithm under MSE criterion). SVM was the classifier but three hyperparameter optimization methods were addressed: (1) grid search (time-consuming and discrete combinations of hyperparameters), (2) Particle Swarm Optimization (PSO) algorithm (good convergence exhibition and optimization performance but local minimal drawback was detected) and (3) modified PSO (introducing mutations into PSO, the convergence problem was solved). Consequently, SVM produced different results in terms of hyperparameter optimization approach: (1) accuracy of 88.33% and image recognition time (IRT) of 0.630 s, (2) 86.667% of accuracy and similar IRT and (3) accuracy of 91.667% and IRT of 0.615 s. Then, time computation consumption and accuracy were highest using SVM with variance PSO improvement. Notwithstanding, some defect-related information was detected in spite of hyperparameter

optimization and imaging procedure; it could be solved using zero-distortion image processing. Another enhancement is the inclusion of more road features.

E. GPR images combined with deep learning techniques effectuate a solution to time consumptions dealing with GPR data processing and manual judgement. Accordingly, a Faster RCNN was proposed to capture 2 road defects: underground pipelines and uneven settlement [53]. The dataset combines simulated and real pre-processed images (DA mainly and redundancy and noise removal). In order to examine detection performance, a mean accuracy of 0.8595 was obtained. It is only noted that the inclusion of depth of feature selection provides better accuracy results.

4.3. Fiber optics

4.3.1. Fundamentals of fiber optics approach

Optical fiber is a transmission medium commonly used in data networks; a very thin wire of transparent material, glass, or plastic materials, through which light pulses representing the data to be transmitted are sent (H.-N. [74]). The light beam is completely confined and propagates inside the fiber with a reflection angle above the limit angle of total reflection, according to Snell's law.

Fiber optics is a current case of study that offers certain advantages over other conventional sensors, such as low electromagnetic (EM) interference or easy maintenance. In addition, it allows coupling sensors such as stress-strain, pressure, piezoelectric, temperature, displacement and humidity sensors. Thus, it is applied to pavement structure health and traffic information monitoring. The most used ML classification techniques are SVMs, ANNs and RF.

Fiber optics can be used to measure magnitudes such as light intensity, phase, polarization state or light frequency, when external vibration is applied. As a matter of fact, the measurement and monitoring of vibration is essential for the detection of abnormal events and pre-warning of infrastructure damage. Moreover, there are different distributed fiber-optic vibration sensing-technologies, such as interferometric, back-scattering based, phase-sensitive optical time domain reflectometer (Φ-OTDR), Brillouin optical time domain analysis (BOTDA) and Brillouin optical correlation domain analysis (BOCDA), etc. [75]. Nevertheless, various traditional vibration sensors are available, but they suffer from electromagnetic interference, short monitoring distance and high maintenance cost. For that reason, optic fibers have attracted research attention due to their capabilities: light weight, flexible length, high accuracy, signal transmission security, easy installation, cost-effective, immunity to electromagnetic interference and corrosion resistance. In the past decades, distributed fiber-optics vibration sensors have found a wide range of applications such as perimeter security protection or borehole seismic implementation [76]. Following, the different Machine Learning techniques will be shown from data produced by fiber optics with the objective of monitoring roads (vehicular traffic monitoring) and evaluating pavement performance (infrastructure health monitoring) [77].

Regarding the state-of-the-art on the study of pavement performance by means of fiber optics applying ML techniques, this is a new field. Nevertheless, it has been applied in other engineering fields, such as monitoring track slab deformation using fiber optic sensing technology and identifying track slab deformation using RF model [78,79].

Another application is highway monitoring system based on distributed fiber-optic sensor sensing (DFOS) like in [55] where DFOS measures the vibration amplitude of passing vehicles every 250 m/s and uses the deep neural network based on VGG-16 to predict average speed. There are other approaches for vehicle classification to improve pavement management and maintenance. The following are ML-based applications based on this novel methodology according to 2 different approaches: traffic and structural monitoring.

4.3.2. Studies employing ML algorithms from fiber optics data

A. In 2018, some embedded 3D glass fiber-reinforced polymer packaged fiber Bragg grating sensors (3D GFRP-FBG system) provided speed of passing vehicles in order to estimate wheelbase and classify cars according to Federal Highway Administration (FHWA) standard via One-Against-One SVM (OAO SVM) into 3 categories [54]. The classification of vehicles is relevant for surveillance, access control, traffic demand planning, traffic congestion prevention and accidents avoidance. The occurrence of wavelength changes (individual peaks) of the 3D GFRP-FBG system identifies the occurrence of a passing vehicle. To derive wheelbase, it is important to measure vehicle speed (Eq. (9)) from distance between two sensors (D) and time interval between strain peaks (t). Once the vehicle speed is calculated, the vehicle's wheelbase can be estimated (Eq. (10)) A posteriori, multi-categorical SVM classifies vehicles according to FHWA requirements. By means of 477 samples (passing vehicles), an accuracy above 0.94 was obtained by using OAO SVM (a video camera was used as reference to validate the proposed optic fiber system). Future investigations will focus on vehicle classification label extension and the inclusion of more physical parameters to gain categorization accuracy.

$$v = \frac{\frac{D}{t_1} + \frac{D}{t_2}}{2} \tag{9}$$

$$wb = \frac{v \bullet t_1 + v \bullet t_2}{2} \tag{10}$$

B. Maintaining the dynamics of traffic monitoring ML-based approaches based on fiber optic, a distributed fiber-optic sensing (DFOS) system was analyzed to gather vibrations of passing vehicles [55] to predict average traffic speeds. Firstly, a pre-processing stage was applied: normalization (to solve intensity gains with distance from the sensing system and the type of roads) and localization (inability to assign traffic due to additional fiber segments and junctions of roadway). Then a neural network was trained, SpeedNet, with labeled data generated synthetically to ensure uniformly distributed information (e.g., traffic congestion). Congregating 150,000 and 15,000 synthetic samples for training and validation stage, and accuracy of 97% and 935 respectively. Finally, SpeedNet was compared with already installed loop detectors (90% of accuracy). Intel 4 Core i5 processor with a 16GB was used. DFOS is proposed as a wide-area cost-effective system to improve camera system which suffers from weather conditions or loop detector difficult installations. Future investigations will concern SpeedNet refinement, the correlation study between vibration intensity and thickness and the vehicle weight and length, to diagnose over-weight and zone-restricted means of transport.

C. Finally, another similar approach using DFOS for traffic monitoring considers vehicle run-off-road events detection using ML algorithms (T. [56]). The fiber sensing system, Distributed Acoustic Sensing (DAS) measures Rayleigh scattering modifications through interferometric phase beating in fiber considering factors such as ground type, weather conditions, vehicle types and sensing distance. Considering spatiotemporal sensing signals as images to classify whether it is Sonic Nap Alert Pattern (SNAP) or normal vibration, a CNN determines if the vibration patterns are caused by the same event and SVM identifies the previous verified images (DA was also deployed during learning stage). The following validation metrics were obtained for a reduced dataset: accuracy of 96.9%, Area under ROC (Receiver Operating Characteristic) Curve (AUC) of 97.7% and 95.3% for Area Under Precision-Recall Curve (AUPRC) for sunny weather and accuracy of 96.4%, AUC of 98.2% and AUPRC of 96.4% for rainy situation. Some drawbacks were observed like high damping coefficient considering under grass (ground type) with long sensing distance which can degrade CNN-SVM performance. This study is just the starting point for future field deployment with a ML approach and it should be engaged to existing management system of transports to relieve traffic congestions and accidents.

4.4. Laser systems

4.4.1. Fundamentals of laser approach

The deviation of a pavement surface from a true planar surface is known as the texture. The texture is a component that can be subject to different scales of investigation, where the discriminant is represented by the wavelength (minimum distance between periodical repeated parts of the curves). Subsequently, laser scanning techniques are commonly used to describe asphalt texture characterization. In contrast, laser systems constitute expensive equipment [80].

Vehicle-mounted laser profilers constitute practical systems to sense pavement profile data. Road Surface Profilers (RSP) consist of laser sensors, odometer, and accelerometer, and operate at high speeds, detecting and analysing long wavelengths and providing pavement profile.

The advancement of 3D laser technology with line-laser imaging and triangulation range computation has become a mainstream technology to collect high-resolution, 3D pavement surface data. Thus, 2D imaging systems are combined with 3D laser systems for collecting 3D pavement surface data.

Laser systems are NDT methods which are the fundamental supply of Pavement Management Systems (PMS). As a matter of fact, many current studies assess NDT-based techniques to develop predictive models, decreasing road surveys to enhance public safety. The following are the most current papers related to laser-based methodology using ML algorithms.

4.4.2. Studies employing ML algorithms from laser data

A table summarizing the studies subsequently developed with the main characteristics will be implemented.

A. An ANN architecture was used for IRI prediction for flexible pavements based from only climate and traffic data (LTTP database). Some of the dataset features were: average daily traffic, humidity, freezing index, annual average temperature or daily truck traffic [57]; from different years and climatic regions. A back-propagation ANN was designed and optimized providing a testing RMSE of 0.01 and consisting of: *tansig* activation function, 7–9–9-1 architecture (e.g., 4 layers of 7,9,9,1 neurons). The main limitation of the study was the lack of available data for flexible pavements. To overcome this issue, a synthetic dataset was created based on statistical distribution of the current climatic data. Future research will focus on developing an ANN model that could be trained more frequently and adaptatively tuned using available data online, then, IRI could be predicted more accurately using a massive online database as a state network level.
B. In 2020, the Georgia Tech Survey Vehicle, equipped with 2D imaging system and 3D line laser imaging system for gathering 3D pavement surface images and an Inertial Measurement Unit and Differential Global Positioning System (GPS) for location references was unveiled [58]. First, a series of pre-processing techniques were claimed to 3D data: invalid depth points removal, pavement marking needs removal due to reflectivity and range data rectification to supress the curvature of pavement which can lead to false detections via high-pass filtering. The objective is to detect ravelling distresses and categorize ravelling severity levels according to Georgia Department of Transportation. AdaBoost with decision trees, SVM and RF were analyzed to classify ravelling severity level. RF performed best results achieving a mean precision value of 86.6% and a recall of 91.55%. Actually, the analyzed system has been deployed into Georgia's highway. Future investigations consider the sophistication of different distresses impact, the refinement of ML algorithms and the development of a quantitative ravelling index.
C. A novel analysis also proposed in 2021, studied the correlation between multitemporal Synthetic Aperture Radar (SAR) interferometry or Persistent Scatter Interferometric SAR (PS-InSAR) and profilometric measurements of road roughness by laser profiler [59]. Especially, PS-InSAR will provide ML algorithms (MLAs) input features to predict the average velocity of each PS (in terms of mm/year); contrarily, a profilometric survey will compute IRI. Therefore, to compare both quantities in order to find out a correlation because MLAs units are mm/year and IRI's units are mm/m, a normalized weighted sum of the absolute value of ML predictions was computed. In relation to MLAs, a backward wrapper was added to extract the most relevant features and also, a Bayesian Optimization algorithm for hyperparameter tuning. CART, RF, SVM and Boosted Regression Trees were developed and SVM obtained the highest performance in terms of Taylor diagram and lowest difference between standard deviation (2.05 mm/year). The objective is to automate pavement performance surveys by replacing time-consuming and expensive laser profilometer surveys with MLAs from PS-InSAR (free SAR open data from Sentinel-1, *Geoportale Lamma*). Regarding correlation interpretations, when road regularity is driven by natural occurrences (external or exogenous factors), IRI and normalized MLA output are similar. In contrast, if the samples are governed by endogenous or local factors, the IRI and MLA output are weakly correlated. Then, future investigations will dissert the calibration of MLAs including endogenous magnitudes to develop non-destructive surveys by MLAs.

4.5. Devices to enhance PMS components

This section aims to manifest the incorporation of IoT devices to the previous methodologies for real-time data collection and the combination of some of those previous technologies in hybrid systems. Previously, in order to build a global idea of the 4 methods for monitoring pavement performance using ML algorithms, the following figure is shown, with the objective, valuable points, limitations and future research lines (see Fig. 6).

4.5.1. IoT monitoring system

Internet of Things (IoT) is an open and comprehensive network of intelligent objects that have the capacity to auto-organize; share information, data, and resources; as well as reacting and acting in face of situations and changes in the environment [81]. Hence, to improve durability and efficiency of the road-embedded monitoring system, IoT provides a connection of real-time monitoring and acquisition data between physical sensors and databases via Internet. As for its advantages, it is important to highlight low cost, low energy, ease to install and suitable compatibility and scalability.

IoT technology is not a technique for obtaining information about the state of the pavement, but an extra tool that allows linking physical devices capable of extracting information (e.g., accelerometers) and connecting them to an environment or database via Internet. The crucial objective is to receive data in real time and perform the appropriate analysis based on MLAs. ML techniques and IoT applications are the main axes for establishing the so-called Intelligent Transportation Systems. Accordingly, smart transportation challenges are divided into six categories in [82]: route optimization-navigation, parking, lights, accident detection, road anomalies and infrastructure. Based on the proposal of the current review, only aspects related to road anomalies that mix IoT devices with applications using machine learning techniques are discussed. To give some examples, [83] proposed the Pothole Detection System (PDS). This system uses accelerometer sensors of Android smartphones for the detection of potholes through ANN and GPS for plotting the location of potholes on Google Maps. In 2021, a real-time traffic information via pavement vibration IoT Monitoring Systems

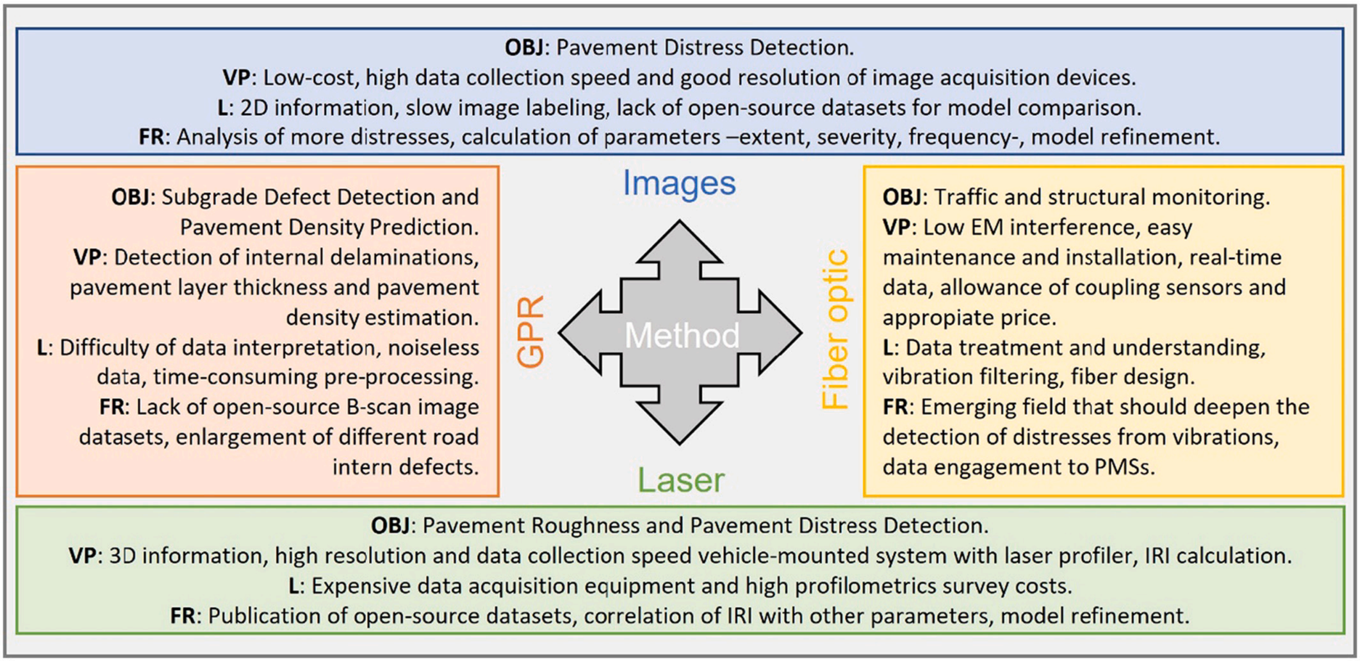

Fig. 6. General description of objective (OBJ), valuable points (VP), limitations (L) and future research (FR).

was proposed by [84]. The pavement vibration IoT monitoring system consists of multi-acceleration sensing nodes, a gateway, and a cloud platform. The acceleration sensing node is used to collect pavement vibration induced by the moving vehicle load; the gateway makes possible the communication between the sensor network and the remote server (4G communication module); and the cloud platform, stores data. This study can be viewed in a web platform, and it can be analyzed: speed and wheelbase, number of axles and vehicle type, location of vehicle load, driving direction and traffic volume.

Consequently, the incorporation of IoT devices would enable automated surveys where data collection would be predicted in the cloud regardless of methodology. For example, it would be interesting to manage the state of aging and road condition through data received in real time combined with fiber optics, increasing maintenance costs.

4.5.2. Hybrid systems

A hybrid system is a compendium of the previously mentioned methodologies that allows linking the information provided by different sensors (GPR, laser profilometers, cameras or fiber optics). By obtaining data via NDT methods, an efficient and low-cost data acquisition is achieved for further processing-based MLAs. To provide some context by exemplifying some hybrid systems, a couple of recent projects are shown below.

Asfault is a low-cost system to monitor road pavement real-time conditions using smartphone sensors and MLAs [80]. An Android mobile app is developed which gathers accelerometer and geolocation data while driving. Moreover, using signal processing techniques, SVM performs a real-time evaluation of road sections classifying asphalt labels. Future works intend to include a model refinement including coefficients such as IRI.

Landmark project [85] is monitored with a low-cost inertial measurement unit (IMU) and a GPS, connected to a Raspberry and the comparison between the comfort index is obtained with accelerations on the 2 different sensors and correlated this index with pavement indicators, IRI and ride number (RN). The correlation studies were developed via linear regression models obtaining $R^2 = 0.98$ with the 2 devices, $R^2 = 0.83$ for IRI and $R^2 = 0.90$ for RN.

5. Conclusions

The main purpose of this article is to detail the different approaches based on ML algorithms for monitoring pavement condition assessment. To this end, an introduction of the problematic and the need for research have been provided. Next, the models used for pavement evaluation in state-of-the-art studies have been theoretically traced, where their performance, benefits, drawbacks and metrics have been explained. Also, the open-source datasets useful for the ML algorithms for pavement evaluation and the different programming frameworks have been mentioned. Then, for the various data collection methods highlighted –Digital Images, Ground Penetrating Radar (GPR), Fiber Optics and Laser– the most recent up-to-date articles characterizing pavement condition using ML models have been analyzed, presenting the methodology, characteristics (data source, data acquisition devices, data volume, computational resources), discussion (solutions, evaluation metrics and drawbacks) and future research directions.

Image-based investigations primarily apply CNN-based architectures for pavement distress detections (mainly cracks and potholes) and future research should focus on solving the current limitations, which are: preprocessing strategies to deal with shadowy and shrink images, improvement of camera calibration and resolution, lack of large image datasets and overfitting. GPR studies, from the B-scan images, focus on detecting interlayer defects via CNN topologies where future investigations should scrutinize the inclusion of more internal road features and the lack of presence of defects to detect more labels with complex algorithms. Fiber optics' target is traffic monitoring through vibrations from CNN or SVM models, thus obtaining parameters such as vehicle categories. The main problems of this technique are the inclusion of more physical features and that no attempt has been made to link the surface condition of the pavement by means of fiber optics. Laser scanner studies result in high-resolution 3D images to predict IRI and detect pavement distress. Unfortunately, beyond its possible improvements such as lack of available data or quantitative pavement indicator development, its main disadvantage is the cost of the data collection system and surveys. Finally, IoT and hybrid data collection systems are explained for a better comprehension of devices to enhance pavement data collection. In conclusion, monitoring or characterizing the pavement condition has a beneficial impact for road administrations, road users and the environment. Consequently, the research community should focus on data collection techniques and ML algorithms to develop services that allow to accomplish an acceptable pavement condition, thus promoting an optimal road infrastructure.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] C. Koch, K. Georgieva, V. Kasireddy, B. Akinci, P. Fieguth, A review on computer vision based defect detection and condition assessment of concrete and asphalt civil infrastructure, Advanced Engineering Informatics 29 (2) (2015) 196–210, https://doi.org/10.1016/j.aei.2015.01.008.

[2] R. Victor, G. Baskir, J. Bennett, J. Camp, R. Capka, S. Curtis, G. Davids, L. Frevert, H. Hatch, A. Herrmann, others, 2013 report card for America's infrastructure, in: American Society of Civil Engineers, American Society of Civil Engineers, 2013, https://doi.org/10.1061/9780784478837.

[3] H. Ziari, J. Sobhani, J. Ayoubinejad, T. Hartmann, Prediction of IRI in short and long terms for flexible pavements: ANN and GMDH methods, International Journal of Pavement Engineering 17 (9) (2016) 776–788, https://doi.org/10.1080/10298436.2015.1019498.

[4] P. Marcelino, M. de Lurdes Antunes, E. Fortunato, M.C. Gomes, Machine learning approach for pavement performance prediction, International Journal of Pavement Engineering 22 (3) (2021) 341–354, https://doi.org/10.1080/10298436.2019.1609673.

[5] J. Yang, J.J. Lu, M. Gunaratne, Application of Neural Network Models for Forecasting of Pavement Crack Index and Pavement Condition Rating. https://trid.trb.org/view/643210, 2003.

[6] M. Ghadge, D. Pandey, D. Kalbande, Machine learning approach for predicting bumps on road, in: 2015 International Conference on Applied and Theoretical Computing and Communication Technology (ICATccT), 2015, pp. 481–485, https://doi.org/10.1109/ICATCCT.2015.7456932.

[7] R. Justo-Silva, A. Ferreira, G. Flintsch, Review on machine learning techniques for developing pavement performance prediction models, Sustainability 13 (9) (2021) 5248, https://doi.org/10.3390/su13095248.

[8] L. Breiman, J.H. Friedman, R.A. Olshen, C.J. Stone, Classification And Regression Trees, Routledge, 2017, https://doi.org/10.1201/9781315139470.

[9] H. Ziari, M. Maghrebi, J. Ayoubinejad, S.T. Waller, Prediction of pavement performance: application of support vector regression with different kernels, Transportation Research Record 2589 (1) (2016) 135–145, https://doi.org/10.3141/2589-15.

[10] D. Wettschereck, D.W. Aha, T. Mohri, A review and empirical evaluation of feature weighting methods for a class of lazy learning algorithms, Artificial Intelligence Review 11 (1–5) (1997) 273–314, https://doi.org/10.1007/978-94-017-2053-3_11.

[11] S. Inkoom, J. Sobanjo, A. Barbu, X. Niu, Pavement crack rating using machine learning frameworks: partitioning, bootstrap Forest, boosted trees, Naïve Bayes, and K -nearest neighbors, Journal of Transportation Engineering, Part B: Pavements 145 (3) (2019) 04019031, https://doi.org/10.1061/jpeodx.0000126.

[12] S. Ren, X. Cao, Y. Wei, J. Sun, Global refinement of random forest, in: Proceedings of the IEEE computer society conference on computer vision and pattern recognition, 07-12-June-2015, 2015, pp. 723–730, https://doi.org/10.1109/CVPR.2015.7298672.

[13] C. Corcoran, C. Mehta, N. Patel, P. Senchaudhuri, Computational tools for exact conditional logistic regression, Statistics in Medicine 20 (17–18) (2001) 2723–2739, https://doi.org/10.1002/sim.739.

[14] B. Andreopoulos, A. An, X. Wang, M. Schroeder, A roadmap of clustering algorithms: finding a match for a biomedical application, Briefings in Bioinformatics 10 (3) (2009) 297–314, https://doi.org/10.1093/bib/bbn058.

[15] P.O. Olukanmi, B. Twala, K-means-sharp: Modified centroid update for outlier-robust k-means clustering, in: 2017 Pattern Recognition Association of South Africa and Robotics and Mechatronics International Conference, PRASA-RobMech 2017, 2018-January(November), 2017, pp. 14–19, https://doi.org/10.1109/RoboMech.2017.8261116.

[16] T. Velmurugan, T. Santhanam, Computational complexity between K-means and K-medoids clustering algorithms for normal and uniform distributions of data points, Journal of Computer Science 6 (3) (2010) 363–368, https://doi.org/10.3844/jcssp.2010.363.368.

[17] D. Svozil, V. Kvasnieka, J. Pospichal, Chemometrics and intelligent laboratory systems introduction to multi-layer feed-forward neural networks, Chemometrics and Intelligent Laboratory Systems 39 (1997) 43–62, https://doi.org/10.1016/S0169-7439(97)00061-0.

[18] Z. Zhang, Y. Lei, X. Mao, P. Li, CNN-FL: An effective approach for localizing faults using convolutional neural networks, in: SANER 2019 - Proceedings of the 2019 IEEE 26th International Conference on Software Analysis, Evolution, and Reengineering, 2019, pp. 445–455, https://doi.org/10.1109/SANER.2019.8668002.

[19] H. Schnoor, E. Horak, Possible method of determining structural number for flexible pavements with the falling weight deflectometer, in: 31st Annual Southern African Transport Conference, July, 2012, pp. 94–109, https://doi.org/10.13140/2.1.1459.6165.

[20] C. Le Bastard, V. Baltazart, Y. Wang, J. Saillard, Thin-pavement thickness estimation using GPR with high-resolution and superresolution methods, IEEE Transactions on Geoscience and Remote Sensing 45 (8) (2007) 2511–2519, https://doi.org/10.1109/TGRS.2007.900982.

[21] A.L. Straub, F.J. Wegmann, Determination of freezing index values, Highway Research Record 68 (1965) 17–30. https://onlinepubs.trb.org/Onlinepubs/hrr/1965/68/68.pdf#page=21.

[22] H. Blockeel, L. De Raedt, Top-down induction of first-order logical decision trees, Artificial Intelligence 101 (1) (1998) 285–297, https://doi.org/10.1016/S0004-3702(98)00034-4.

[23] S. Karamizadeh, S.M. Abdullah, M. Halimi, J. Shayan, javad Rajabi, M., Advantage and drawback of support vector machine functionality, in: 2014 International Conference on Computer, Communications, and Control Technology (I4CT), 2014, pp. 63–65, https://doi.org/10.1109/I4CT.2014.6914146.

[24] A.T. Olowosulu, J.M. Kaura, A.A. Murana, P.T. Adeke, Classification of surface condition of flexible road pavement using Naïve Bayes theorem, IOP Conference Series: Materials Science and Engineering 1036 (1) (2021), 012036, https://doi.org/10.1088/1757-899X/1036/1/012036.

[25] S. Roychowdhury, M. Zhao, A. Wallin, N. Ohlsson, M. Jonasson, Machine learning models for road surface and friction estimation using front-camera images, in: 2018 International Joint Conference on Neural Networks (IJCNN), 2018, pp. 1–8, https://doi.org/10.1109/IJCNN.2018.8489188.

[26] A.M. Andrew, Reinforcement learning: an introduction, Kybernetes 27 (9) (1998) 1093–1096, https://doi.org/10.1108/k.1998.27.9.1093.3.

[27] Z. Boger, H. Guterman, Knowledge extraction from artificial neural network models, in: 1997 IEEE International Conference on Systems, Man, and Cybernetics. Computational Cybernetics and Simulation 4, 1997, pp. 3030–3035, https://doi.org/10.1109/ICSMC.1997.633051.

[28] L. Hardesty, Explained: Neural networks Ballyhooed Artificial-Intelligence Technique Known as "Deep Learning" Revives 70-Year-Old Idea, 2017-04-14, https://news.mit.edu/2017/explained-neural-networks-deep-learning-0414, 2017.

[29] J. Starmer, Neural Networks Part 8: Image Classification with Convolutional Neural Networks, Access date: 01/03/2021, https://statquest.org/video-index/, 2022 (n.d.).

[30] K. Gurney, An Introduction to Neural Networks, CRC Press, 2018, https://doi.org/10.1201/9781315273570.

[31] K. Janocha, W.M. Czarnecki, On Loss Functions for Deep Neural Networks in Classification, ArXiv:1702.05659, 2017, https://doi.org/10.48550/arXiv.1702.05659.

[32] K. Gopalakrishnan, S.K. Khaitan, A. Choudhary, A. Agrawal, Deep convolutional neural networks with transfer learning for computer vision-based data-driven pavement distress detection, Construction and Building Materials 157 (2017) 322–330, https://doi.org/10.1016/j.conbuildmat.2017.09.110.

[33] P. Marcelino, M. de Lurdes Antunes, E. Fortunato, Comprehensive performance indicators for road pavement condition assessment, Structure and Infrastructure Engineering 14 (11) (2018) 1433–1445, https://doi.org/10.1080/15732479.2018.1446179.

[34] M.W. Sayers, On the calculation of international roughness index from longitudinal road profile, Transportation Research Record 1501 (1995) 1–12. http://onlinepubs.trb.org/Onlinepubs/trr/1995/1501/1501-001.pdf.

[35] S. Madeh Piryonesi, T.E. El-Diraby, Using machine learning to examine impact of type of performance indicator on flexible pavement deterioration modeling, Journal of Infrastructure Systems 27 (2) (2021) 4021005, https://doi.org/10.1061/(ASCE)IS.1943-555X.0000602.

[36] N.G. Gharaibeh, Y. Zou, S. Saliminejad, Assessing the agreement among pavement condition indexes, Journal of Transportation Engineering 136 (8) (2010) 765–772, https://doi.org/10.1061/(ASCE)TE.1943-5436.0000141.

[37] S. Dorafshan, R.J. Thomas, M. Maguire, Comparison of deep convolutional neural networks and edge detectors for image-based crack detection in concrete, Construction and Building Materials 186 (2018) 1031–1045, https://doi.org/10.1016/j.conbuildmat.2018.08.011.

[38] M. Hossin, M.N. Sulaiman, A review on evaluation metrics for data classification evaluations, International Journal of Data Mining & Knowledge Management Process 5 (2) (2015) 01–11, https://doi.org/10.5121/ijdkp.2015.5201.

[39] W. Zeiada, S.A. Dabous, K. Hamad, R. Al-Ruzouq, M.A. Khalil, Machine learning for pavement performance modelling in warm climate regions, Arabian Journal for Science and Engineering 45 (5) (2020) 4091–4109, https://doi.org/10.1007/s13369-020-04398-6.

[40] A.B. Ersoz, O. Pekcan, T. Teke, Crack identification for rigid pavements using unmanned aerial vehicles, IOP Conference Series: Materials Science and Engineering 236 (1) (2017), 012101, https://doi.org/10.1088/1757-899X/236/1/012101.

[41] S. Bang, S. Park, H. Kim, Y. Yoon, H. Kim, A deep residual network with transfer learning for pixel-level road crack detection, in: ISARC 2018 - 35th International Symposium on Automation and Robotics in Construction and International AEC/FM Hackathon: The Future of Building Things, July, 2018, https://doi.org/10.22260/ISARC2018/0103.

[42] Q. Zou, Z. Zhang, Q. Li, X. Qi, Q. Wang, S. Wang, DeepCrack: learning hierarchical convolutional features for crack detection, IEEE Transactions on Image Processing 28 (3) (2019) 1498–1512, https://doi.org/10.1109/TIP.2018.2878966.

[43] N.D. Hoang, Image processing based automatic recognition of asphalt pavement patch using a metaheuristic optimized machine learning approach, Advanced

Engineering Informatics 40 (809) (2019) 110–120, https://doi.org/10.1016/j.aei.2019.04.004.
[44] M. Słoński, A comparison of deep convolutional neural networks for image-based detection of concrete surface cracks, Computer Assisted Methods in Engineering and Science 26 (2) (2019) 105–112, https://doi.org/10.24423/cames.267.
[45] T. Yamane, P. Chun, Crack detection from a concrete surface image based on semantic segmentation using deep learning, Journal of Advanced Concrete Technology 18 (9) (2020) 493–504, https://doi.org/10.3151/jact.18.493.
[46] G. Yang, K.C.P. Wang, J.Q. Li, Y. Fei, Y. Liu, K.C. Mahboub, A.A. Zhang, Automatic pavement type recognition for image-based pavement condition survey using convolutional neural network, Journal of Computing in Civil Engineering 35 (1) (2021) 04020060, https://doi.org/10.1061/(asce)cp.1943-5487.0000944.
[47] S. Ranjbar, F.M. Nejad, H. Zakeri, An image-based system for pavement crack evaluation using transfer learning and wavelet transform, International Journal of Pavement Research and Technology 14 (4) (2021) 437–449, https://doi.org/10.1007/s42947-020-0098-9.
[48] P. Chun, S. Izumi, T. Yamane, Automatic detection method of cracks from concrete surface imagery using two-step light gradient boosting machine, Computer-Aided Civil and Infrastructure Engineering 36 (1) (2021) 61–72, https://doi.org/10.1111/mice.12564.
[49] S.S. Todkar, V. Baltazart, A. Ihamouten, X. Dérobert, D. Guilbert, One-class SVM based outlier detection strategy to detect thin interlayer debondings within pavement structures using ground penetrating radar data, Journal of Applied Geophysics 192 (2021), https://doi.org/10.1016/j.jappgeo.2021.104392.
[50] Z. Liu, W. Wu, X. Gu, S. Li, L. Wang, T. Zhang, Application of combining YOLO models and 3D GPR images in road detection and maintenance, Remote Sensing 13 (6) (2021) 1081, https://doi.org/10.3390/rs13061081.
[51] Z. Fang, Z. Shi, X. Wang, W. Chen, Roadbed defect detection from ground penetrating radar B-scan data using faster RCNN, IOP Conference Series: Earth and Environmental Science 660 (1) (2021) 12020, https://doi.org/10.1088/1755-1315/660/1/012020.
[52] Xinyu Liu, P. Hao, A. Wang, L. Zhang, B. Gu, X. Lu, Non-destructive detection of highway hidden layer defects using a ground-penetrating radar and adaptive particle swarm support vector machine, PeerJ Computer Science 7 (2021), https://doi.org/10.7717/peerj-cs.417.
[53] Y. Gao, L. Pei, S. Wang, W. Li, Intelligent detection of urban road underground targets by using ground penetrating radar based on deep learning, Journal of Physics: Conference Series 1757 (1) (2021) 12081, https://doi.org/10.1088/1742-6596/1757/1/012081.
[54] M. Al-Tarawneh, Y. Huang, P. Lu, D. Tolliver, Vehicle classification system using in-pavement fiber Bragg grating sensors, IEEE Sensors Journal 18 (7) (2018) 2807–2815, https://doi.org/10.1109/JSEN.2018.2803618.
[55] C. Narisetty, T. Hino, M.-F. Huang, R. Ueda, H. Sakurai, A. Tanaka, T. Otani, T. Ando, Overcoming challenges of distributed fiber-optic sensing for highway traffic monitoring, Transportation Research Record 2675 (2) (2021) 233–242, https://doi.org/10.1177/0361198120960134.
[56] T. Li, Y. Chen, M.-F. Huang, S. Han, T. Wang, Vehicle run-off-road event automatic detection by fiber sensing technology, in: 2021 Optical Fiber Communications Conference and Exhibition (OFC), 2021, pp. 1–3, https://doi.org/10.1364/OFC.2021.Th1A.26.
[57] M.I. Hossain, L.S.P. Gopisetti, M.S. Miah, International roughness index prediction of flexible pavements using neural networks, Journal of Stomatology 145 (1) (2019) 1–10, https://doi.org/10.1061/JPEODX.0000088.
[58] Y.-C.J. Tsai, Y. Zhao, B. Pop-Stefanov, A. Chatterjee, Automatically detect and classify asphalt pavement raveling severity using 3D technology and machine learning, Int. J. Pave Res. Technol. 14 (4) (2021) 487–495, https://doi.org/10.1007/s42947-020-0138-5.
[59] N. Fiorentini, M. Maboudi, P. Leandri, M. Losa, Can machine learning and PS-InSAR reliably stand in for road profilometric surveys? Sensors 21 (10) (2021) 3377, https://doi.org/10.3390/s21103377.
[60] M.Z. Naser, A.H. Alavi, Error metrics and performance fitness indicators for artificial intelligence and machine learning in engineering and sciences, Architecture, Structures and Construction (2021) 1–25, https://doi.org/10.1007/s44150-021-00015-8.
[61] M.D. Wilkinson, M. Dumontier, J. Aalbersberg, G. Appleton, M. Axton, A. Baak, N. Blomberg, J.-W. Boiten, L.B. da Silva Santos, P.E. Bourne, J. Bouwman, A. J. Brookes, T. Clark, M. Crosas, I. Dillo, O. Dumon, S. Edmunds, C.T. Evelo, R. Finkers, A. Gonzalez-Beltran, A.J.G. Gray, P. Groth, C. Goble, J.S. Grethe, J. Heringa, P.A.C. Hoen, R. Hooft, T. Kuhn, R. Kok, J. Kok, S.J. Lusher, M. E. Martone, A. Mons, B. Mons, A.L. Packer, B. Persson, P. Roca-Serra, M. Roos, R. van Schaik, S. Sansone, E. Schultes, T. Senstag, T. Slater, G. Strawn, M.A. Swertz, M. Thompson, J. van der Lei, E. van Mulligen, J. Velterop, A. Waagmeester, P. Wittenburg, K. Wolstencroft, J. Zhao, The FAIR guiding principles for scientific data management and stewardship, Scientific Data 3 (1) (2016), 160018, https://doi.org/10.1038/sdata.2016.18.
[62] A. Fathi, M. Mazari, M. Saghafi, A. Hosseini, S. Kumar, Parametric study of pavement deterioration using machine learning algorithms, in: Airfield and Highway Pavements 2019: Innovation and Sustainability in Highway and Airfield Pavement Technology, American Society of Civil Engineers, Reston, VA, 2019, pp. 31–41, https://doi.org/10.1061/9780784482476.004.
[63] M.A. Younos, R.T. Abd El-Hakim, S.M. El-Badawy, H.A. Afify, Multi-input performance prediction models for flexible pavements using LTPP database, Innovative Infrastructure Solutions 5 (1) (2020) 27, https://doi.org/10.1007/s41062-020-0275-3.
[64] D. Arya, H. Maeda, S.K. Ghosh, D. Toshniwal, A. Mraz, T. Kashiyama, Y. Sekimoto, Deep learning-based road damage detection and classification for multiple countries, Automation in Construction 132 (August) (2021), 103935, https://doi.org/10.1016/j.autcon.2021.103935.
[65] L. Wu, S. Mokhtari, A. Nazef, B. Nam, H.-B. Yun, Improvement of crack-detection accuracy using a novel crack defragmentation technique in image-based road assessment, Journal of Computing in Civil Engineering 30 (1) (2016) 04014118, https://doi.org/10.1061/(ASCE)CP.1943-5487.0000451.
[66] R. Zaurin, F.N. Catbas, Integration of computer imaging and sensor data for structural health monitoring of bridges, Smart Materials and Structures 19 (1) (2010), 015019, https://doi.org/10.1088/0964-1726/19/1/015019.
[67] S. Mokhtari, Analytical Study of Computer Vision-Based Pavement Crack Quantification using Machine Learning Techniques, Access date: 13/07/2021, https://stars.library.ucf.edu/etd/1156, 2015.
[68] H. Zakeri, F.M. Nejad, A. Fahimifar, Image based techniques for crack detection, classification and quantification in asphalt pavement: a review, Archives of Computational Methods in Engineering 24 (4) (2017) 935–977, https://doi.org/10.1007/s11831-016-9194-z.
[69] F.N. Iandola, S. Han, M.W. Moskewicz, K. Ashraf, W.J. Dally, K. Keutzer, SqueezeNet: AlexNet-Level Accuracy with 50x Fewer Parameters and <0.5MB Model Size, 2016, pp. 1–13. ArXiv:1602.07360, 10.48550/arXiv.1602.07360.
[70] Y. Qu, W. Tang, B. Feng, Paper defects classification based on VGG16 and transfer learning, Journal of Korea Technical Association of the Pulp and Paper Industry 53 (2) (2021) 5–14, https://doi.org/10.7584/JKTAPPI.2021.04.53.2.5.
[71] Q. Cao, I.L. Al-Qadi, Development of a numerical model to predict the dielectric properties of heterogeneous asphalt concrete, Sensors 21 (8) (2021) 2643, https://doi.org/10.3390/s21082643.
[72] X.L. Travassos, S.L. Avila, N. Ida, Artificial neural networks and machine learning techniques applied to ground penetrating radar: a review, Applied Computing and Informatics 17 (2) (2021) 296–308, https://doi.org/10.1016/j.aci.2018.10.001.
[73] I.L. Al-Qadi, S. Lahouar, Measuring layer thicknesses with GPR–Theory to practice, Construction and Building Materials 19 (10) (2005) 763–772, https://doi.org/10.1016/j.conbuildmat.2005.06.005.
[74] H.-N. Li, D.-S. Li, G.-B. Song, Recent applications of fiber optic sensors to health monitoring in civil engineering, Engineering Structures 26 (11) (2004) 1647–1657, https://doi.org/10.1016/j.engstruct.2004.05.018.
[75] X. Liu, B. Jin, Q. Bai, Y. Wang, D. Wang, Y. Wang, Distributed fiber-optic sensors for vibration detection, Sensors 16 (8) (2016) 1164, https://doi.org/10.3390/s16081164.
[76] Y. Aono, E. Ip, P. Ji, More than communications: Environment monitoring using existing optical fiber network infrastructure, in: Optics InfoBase Conference Papers, Part F174, 2020, pp. 1–3, https://doi.org/10.1364/OFC-2020-W3G.1.
[77] X. Chapeleau, J. Blanc, P. Hornych, J.L. Gautier, J. Carroget, Use of distributed fiber optic sensors to detect damage in a pavement, in: 7th European Workshop on Structural Health Monitoring, EWSHM 2014 - 2nd European Conference of the Prognostics and Health Management (PHM) Society, 2014, pp. 1847–1854, https://doi.org/10.1201/b17219-60.
[78] G. Guo, X. Cui, B. Du, Random–Forest machine learning approach for high–speed railway track slab deformation identification using track-side vibration monitoring, Applied Sciences 11 (11) (2021) 4756, https://doi.org/10.3390/app11114756.
[79] S. Wang, F. Liu, B. Liu, Research on application of deep convolutional network in high-speed railway track inspection based on distributed fiber acoustic sensing, Optics Communications 492 (2021), 126981, https://doi.org/10.1016/j.optcom.2021.126981.
[80] V.M.A. Souza, R. Giusti, A.J.L. Batista, Asfault: A low-cost system to evaluate pavement conditions in real-time using smartphones and machine learning, Pervasive and Mobile Computing 51 (2018) 121–137, https://doi.org/10.1016/j.pmcj.2018.10.008.
[81] S. Madakam, V. Lake, V. Lake, V. Lake, others, Internet of things (IoT): a literature review, Journal of Computer and Communications 3 (05) (2015) 164, https://doi.org/10.4236/jcc.2015.35021.
[82] F. Zantalis, G. Koulouras, S. Karabetsos, D. Kandris, A review of machine learning and IoT in smart transportation, Future Internet 11 (4) (2019) 94, https://doi.org/10.3390/fi11040094.
[83] A. Kulkarni, N. Mhalgi, S. Gurnani, N. Giri, Pothole detection system using machine learning on android, International Journal of Emerging Technology and Advanced Engineering 4 (7) (2014) 360–364, https://doi.org/10.14257/astl.2018.150.35.
[84] Z. Ye, G. Yan, Y. Wei, B. Zhou, N. Li, S. Shen, L. Wang, Real-time and efficient traffic information acquisition via pavement vibration IoT monitoring system, Sensors 21 (8) (2021) 2679, https://doi.org/10.3390/s21082679.
[85] G. Loprencipe, F.G.V. de Almeida Filho, R.H. de Oliveira, S. Bruno, Validation of a low-cost pavement monitoring inertial-based system for urban road networks, Sensors 21 (9) (2021) 3127, https://doi.org/10.3390/s21093127.