ISSN Online: 2821-1693

Transactions on Machine Intelligence

Journal Homepage: https://tmachineintelligence.ir/index.php/TMI

Detection of Pavement Damage Using Smartphones

A. Nouri[1] , V. Toufigh[2,*]

[1] M.Sc. in Civil Engineering, Department of Civil Engineering, Sharif University of Technology, Tehran, Iran
[2] Ph.D. in Civil Engineering and Associate Professor, Department of Civil Engineering, Sharif University of Technology, Tehran, Iran

ARTICLE INFO

Article History:
Received 14 March 2021
Received in revised form 4 April 2021
Accepted 3 June 2021
Available online 3 June 2021

Keywords:
Machine Learning, Deep Learning, Pavement, Diagnostics

ABSTRACT

Early detection and repair of pavement damage can significantly reduce the associated maintenance and repair costs. In recent years, efforts have been made to use digital hardware and software to streamline the inspection and diagnostic process. However, in a real-world scenario, the resource limitations, high cost, and time-consuming nature of these digital units have diminished their efficiency. In the past decade, smartphones have gained remarkable hardware capabilities. Mobile phones, aided by GPS, record location information and capture high-quality images with powerful lenses. This research aims to utilize machine learning algorithms to employ smartphones for pavement inspection. Deep learning, a method capable of pattern recognition and solving complex problems, has been chosen as the machine learning technique. The learning process of this algorithm is conducted using samples collected from pavement surfaces via smartphones. The study further aims to enhance the speed and processing power of the learning method through new parameters. Additionally, a defined framework is provided to assess the quality of damage, facilitating effective action by route managers.

1. INTRODUCTION

The existing road network in any country is a key element for its growth. In the past two decades, Iran's infrastructural road network has expanded significantly. Pavement is considered a crucial part of transportation infrastructure and must be constructed to withstand expected traffic loads. Over time, pavements deteriorate due to exposure to traffic loads and environmental conditions. Consequently, many of these pathways have begun to deteriorate, and it is expected that the number of infrastructures to be inspected will increase rapidly in the coming decades. The pavement sector will be at the forefront of evaluation and inspection for transportation infrastructure. Until now, the discovery of old and damaged parts of the infrastructure has relied solely on the expertise of veteran field engineers. However, due to the increased demand for inspections, the shortage of field technicians (experts), and financial resources, there have been consequences in many areas. The risks associated with aging infrastructure, instability, and deterioration are among these consequences [1]. In Iran, the number of operators neglecting proper inspections due to resource or expert shortages is increasing. Moreover, maintenance and management problems of infrastructure are likely experienced by countries worldwide. Given this negative trend in infrastructure maintenance

* Corresponding Author: toufigh@sharif.edu
M.Sc. in Civil Engineering, Department of Civil Engineering, Sharif University of Technology, Tehran, Iran

http://dx.doi.org/10.47176/TMI.2021.54

and management, it is evident that efficient and sophisticated methods for infrastructure maintenance are essential [2].

In response to the aforementioned problem, numerous methods for the efficient inspection of infrastructure, particularly road conditions, have been studied, with damage detection consistently being a hot topic of research. Over the past 40 years, significant and promising results have been achieved in this field. Digital hardware and software have received considerable attention from researchers. Techniques such as laser technology and image processing are among the methods referenced in the literature. For instance, three-dimensional laser imaging-based sensors can be used to inspect pavement issues. This technology creates a longitudinal profile for roughness and a transverse profile for rutting, while also evaluating and predicting surface defects for safety analysis. Additionally, a mobile measurement system can obtain highly accurate spatial information using a moving vehicle. This system includes a Global Positioning System (GPS) unit, an internal measurement unit, measurable digital images, a digital camera, a laser scanner, and an omnidirectional video recorder. Although quantitative inspections are highly accurate, conducting such comprehensive inspections is very costly, especially for small operators lacking necessary financial resources. Due to these issues, operators without sufficient resources do not perform proper and frequent infrastructure inspections, thereby increasing the risks associated with structural deterioration.

Contrary to these techniques, studies employing various machine learning techniques to evaluate the surface and identify types of pavement damage have been conducted over the past decade. However, inspection methods focused on artificial intelligence still appear to suffer from certain weaknesses. Automating pavement crack detection generally requires robust algorithms with a high level of intelligence. One of the problems relates to human limitations in developing mathematical models that simulate cognitive capabilities. Furthermore, there seems to be no common dataset for comparing results. In each study, the proposed method is evaluated using its own dataset of road damage images. This issue pertains to the complex nature of imaging pavement surfaces. Although modern object detection methods utilize end-to-end deep learning techniques, such a method for road damage detection does not exist. End-to-end learning in artificial intelligence and machine learning is a technique where the model learns all the steps between the initial input phase and the final output result. It is a deep learning process where all different parts are trained simultaneously rather than in sequence. Although road surface damages are categorized into several types (eight categories based on road maintenance and repair guidelines), many studies have limited themselves to identifying or classifying damage only longitudinally and transversely. Therefore, it is challenging for road operators to directly apply these research findings to practical scenarios [3].

Within this challenging context, governments worldwide have recently begun leveraging information technology based on smartphones to enhance public service delivery. For instance, some local municipalities in Iran have established communication channels (such as websites) to share information about minor issues reported by citizens. Another example is the use of simple and lightweight applications to collect citizens' reports [4]. Smartphones possess significant capabilities such as precise geolocation data collection and high-quality imaging. Reporting systems enable citizens to report and share information about infrastructural issues using images and location data, allowing local governments to track and address these issues. Consequently, a substantial amount of information about damaged infrastructure can be collected at minimal cost based on citizen reports. However, these assessments are not considered expert evaluations. It is challenging for citizens without specialized knowledge to accurately and reliably assess the extent of road damage. As a result, reports submitted by citizens may include insignificant damages that are deemed unimportant by professional road managers. Furthermore, there is bias in the data collected by citizens. Manual validation of these reports, including minor ones, would be burdensome for local governments.

This research aims to utilize a machine learning algorithm to prepare smartphones for lightweight analysis of raw initial data. The proposed algorithm is lightweight with a simple learning process. The method presented not only simplifies the validation of reports but also addresses the issues related to identifying types of damage by focusing on pavement quality.

2. ARTIFICIAL NEURAL NETWORK

The artificial neural network (ANN) is modeled based on the biological neural network. Similar to the biological neural network, an ANN is an interconnected network of nodes that resemble neurons. Each neuron is a cell that performs a simple task, such as responding to an input signal. However, when a network of neurons is interconnected, they can perform complex tasks such as speech and image recognition with astonishing speed and accuracy. Neural networks are a set of algorithms modeled after the human brain and designed to recognize patterns. In other words, their goal is to simulate the behavior of biological systems composed of neurons. They interpret sensory data through a kind of machine perception, labeling, or clustering of raw input. The patterns they recognize are numerical and reside in vectors, requiring all real-world data, whether images, sound, text, or time series, to be translated into them.

Every neural network has three vital components: node character, network topology, and learning rules. The node character determines how the node processes signals, including the number of inputs and outputs associated with the node, the weight related to each input and output, and the activation function. Network topology determines how nodes are organized and connected. Learning rules dictate how weights are initialized and adjusted. ANNs use a learning process to train the network. During training, weights are adjusted to optimal values. Various learning schemes for ANNs have been devised to achieve different learning objectives. The most common learning methods are error correction methods and nearest neighbor methods.

Acceptance of the paper will be communicated to the corresponding author. However, the current status of the papers can be tracked at any time through the congress website. Upon acceptance, authors must incorporate the reviewers' requested revisions in the final version and submit the final version through the congress website within the specified time frame [5].

3. CONVOLUTIONAL NEURAL NETWORK

A Convolutional Neural Network (CNN) is a specialized type of multilayer neural network inspired by the visual system of living organisms. Introduced in 1980, multilayer neural networks capable of hierarchical visual pattern recognition through learning served as the theoretical inspiration for CNNs. By using a deep architecture to mimic the multilayered, neuromorphic natural networks, CNNs can adaptively learn hierarchical representations of models, from low-level to high-level functions, and subsequently identify the most critical features for a specific model. A deep CNN typically refers to a structure that includes convolutional layers, pooling layers, and a fully connected network. The convolution operation in this structure is used for feature extraction, while the fully connected network classifies these features [6].

The Convolution Layer is a structure with a number of filters of fixed size that enables the application of complex functions to the input image. This process is performed by sliding the locally trained filters over the image. Each filter has the same weights and bias values across the entire image during this process. This mechanism is called weight sharing, which allows the same feature to be represented throughout the image. The local receptive field of a neuron represents the area to which the neuron in the previous layer is connected. The size of the receptive field is determined by the size of the filters. Suppose m × n and c × c are the dimensions of the input image and the kernel size, respectively. i denotes the image, w and b represent the weights and bias values of the filter, respectively. The output at position (0,0) can be calculated as in equation (1), where f is the activation function. ReLU or sigmoid functions can be used as the activation function in this process. The behavior of the ReLU activation function is also shown in equation (2).

$$a_{0.0} = f(b + \textstyle\sum_{t=0}^{c}\sum_{r=0}^{c} w_{t.r} i_{0+t.0+r}) \quad (1)$$

$$f(x) = \begin{cases} x & x > 0 \\ 0 & else \end{cases} \quad (2)$$

The pooling process is applied to feature maps that have passed through the convolution layer and activation function. This results in smaller feature maps that summarize the input feature maps. Pooling is performed by sliding a window over the image to apply the chosen operation. The main advantages provided by the pooling operation include reducing the image size and extracting visual features independently across the image. After the convolution and pooling layers, the data is transformed into a one-dimensional vector. This vector will be the input to a fully connected network. The fully connected structure may contain one or more hidden layers. Each neuron multiplies the connection weight by the data from the previous layer and adds a bias value. The computed value passes through an activation function before being transferred to the next layer. The computations performed by a neuron in this layer can be seen in equation (3).

$$fc_1 = f(b + \sum_{q=1}^{M} w_{1.q} o_q) \quad (3)$$

where f is the activation function, w is the weight vector, o is the input vector of the q the neuron, and b is the bias value [7].

4. FUZZY LOGIC

Fuzzy systems were introduced by Dr. Lotfi Zadeh at Berkeley in the 1960s. The primary inspiration behind the introduction of fuzzy set theory was the need to model real-world phenomena that are inherently vague and ambiguous. Human knowledge about complex issues can be successfully represented using the imprecise terms of natural language. Fuzzy set theory and fuzzy logic provide formal tools for the mathematical representation and efficient processing of such information. Fuzzy systems are structures based on fuzzy techniques directed towards information processing. In the literature, terms such as fuzzy system, fuzzy model, fuzzy rule-based system, fuzzy controller, or fuzzy associative memory are used interchangeably depending on the type of application. Fuzzy logic depends on the degree of truth or falsehood [8].

The typical structure of a fuzzy system consists of four functional blocks: a fuzzifier, a fuzzy inference engine, a knowledge base, and a defuzzifier. Both linguistic values (defined by fuzzy sets) and crisp (numerical) data can be used as inputs for a fuzzy system. If crisp data are applied, the inference process begins with fuzzification, which assigns the appropriate fuzzy set to the non-fuzzy input. The values of the input variables are mapped to the linguistic values of the output variable using an appropriate approximate reasoning method (inference engine) with the aid of special knowledge, which is represented as a set of fuzzy conditional rules (knowledge base). Besides linguistic values, numerical data may also be required as the output of the fuzzy system. In such cases, defuzzification methods are used to assign representative crisp data to the output fuzzy set.

In a fuzzy classification system, an item or an object can be classified by applying a set of fuzzy rules based on its characteristics. Each rule has a weight, a number between 0 and 1, which is applied to the value given by the antecedent. The process involves two distinct parts. The first part includes the evaluation of the antecedents, fuzzifying the input, and applying any necessary fuzzy operators. The second part involves applying the result to the consequent, known as inference. The most challenging task in building a fuzzy classification system is finding a set of fuzzy rules pertinent to the specific classification problem. A fuzzy inference system is a rule-based system that uses fuzzy logic instead of Boolean logic to reason about the data.

5. PAVEMENT DISTRESSES

Surface distresses occur when the serviceability of the pavement is compromised due to surface degradation, without necessarily losing the structural load-bearing capacity of the pavement system. Surface-distressed pathways can disrupt usage without complete pavement failure. There are four main categories of surface distresses: cracks, surface deformations, disintegration, and slipperiness. However, not all surface distresses pose the same level of danger to users. For example, distresses that lead to slipperiness or surface unevenness present greater risks to users.

Making decisions regarding pavement maintenance and rehabilitation is a significant challenge for highway operators worldwide. To overcome this challenge, many road operators have developed procedures and practices aimed at maintaining their pavement networks by making appropriate maintenance decisions at the right time. Evaluating pavement performance using pavement condition indices is also an essential component of any pavement management system. Most cost-effective maintenance and rehabilitation strategies developed using pavement management systems are the result of precise pavement evaluations. Various indices, such as the Pavement Condition Rating (PCR), Pavement Condition Index (PCI), Present Serviceability Rating (PSR), and International Roughness Index (IRI), are typically used to determine maintenance strategies for existing pavements [9].

The Pavement Condition Index (PCI) method is the most commonly used index for evaluating pavement conditions across the United States and Canada. It provides a comprehensive measure of the current pavement condition based on observed surface distresses and statistically accurate analysis for pavement sampling. It also reflects the structural integrity and operational status of the pavement surface. However, its application is challenging as it involves assessing 19 different distresses with varying levels of severity (low, medium, and high). Data for determining the PCI are collected through visual inspections or image-based survey methods. In the PCI calculation method, different types of distress with varying severities are combined into a single PCI value. Each distress contributing to pavement deterioration is measured in units of length or area with different severities (i.e., low, medium, and high). The PCI ranges from 100 to 0, where 100 represents newly constructed pavement and 0 represents the worst possible condition. The PCI calculation method for roads and parking lots is thoroughly detailed in international standards (ASTM 2003, D6433-03). Although manually calculating the PCI for a single sample unit might not be tedious, a database collected from a survey is generally very large, and the PCI calculation process for an entire database can be time-consuming. MicroPAVER is a commonly used software that can automatically calculate the PCI value once distress information is entered [10].

6. PROPOSED NEW NETWORK STRUCTURE

The architecture of the proposed fuzzy convolutional neural network model consists of four types of layers: convolutional layer, pooling layer, fuzzy layer, and fully connected layer. To form a complete architecture of the fuzzy convolutional neural network, we combine three sections: a convolutional network, a fuzzy layer, and a classifier.

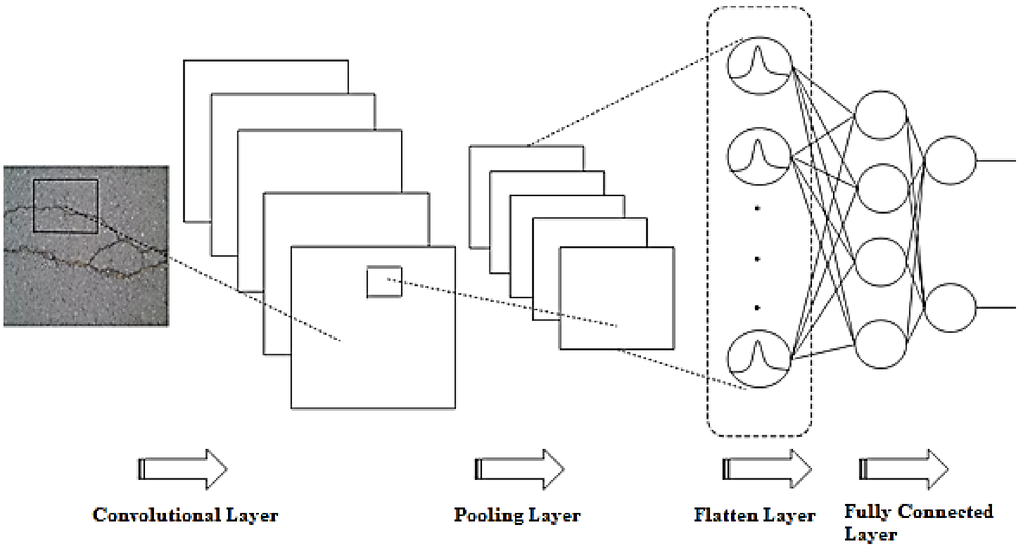

Fig. 1. Proposed new network structure.

Initially, the image is fed into the stack corresponding to the convolutional network. The first layer in the convolutional network is the convolutional layer, which is the fundamental building block and performs most of the computational work. The image is convolved using filters or kernels. Filters are small units applied to the data through a sliding window. The output of a convolutional layer with a three-dimensional color filter is a two-dimensional matrix. The next component is the activation layer, which applies the Rectified Linear Unit (ReLU). At this stage, we apply the activation function to introduce non-linearity into the convolutional network. Generally, images of various objects are not linearly related to each other. Training a network composed solely of linear activation functions is straightforward, but it cannot learn complex mapping functions. Third, the pooling layer samples the features. The pooling layer eliminates any redundant features captured during convolution. A pooling layer typically uses a 2×2 max filter with a stride of 2, resulting in a non-overlapping filter. There are numerous pooling methods, but the most popular is max pooling.

Unlike a conventional convolutional neural network, the proposed fuzzy convolutional neural network includes a fuzzy layer (a self-organizing layer), which acts as a type of preprocessor. It is situated between the convolutional network and the classifier (a type of postprocessor). The fuzzy layer initially distributes the input data into a predetermined number of clusters. These clusters do not correspond to the output classes, and the number of clusters and target classes can differ. The outputs of the neurons in the fuzzy layer represent the membership function values for the fuzzy clusters of the input data. These membership degrees indicate the extent to which data points belong to each cluster. These values are then fed into a classifier. Finally, a fully connected classifier layer exists at the end. A fully connected neural network consists of a series of fully connected layers. In fully connected layers, every input from one layer is connected to every activation unit of the next layer. The last few layers are fully connected layers that aggregate the data extracted by the previous layers to form the final output.

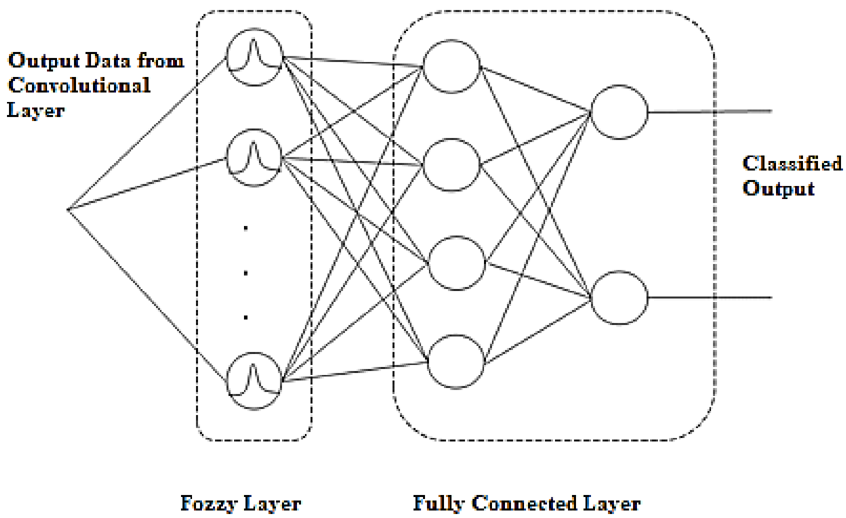

Fig. 2. Structure of the final layer of the proposed network.

The proposed fuzzy neural network is divided into three stages: the input data (image) undergoes a series of transformations, resulting in a feature vector. Subsequently, the fuzzy layer performs an initial distribution of the input data into fuzzy clusters. Finally, the fully connected layers execute classification and assign the class label outcome to each group of clusters.

7. DATASETS

Figure 3 illustrates four common types of crack images in pavement surfaces: (a) a densely distributed crack network, where the cracks segment the pavement into multiple blocks of varying sizes and shapes, posing a danger to traffic. Therefore, it is crucial to identify and classify these cracks before significant propagation occurs. However, detecting and classifying them is challenging. (b) A crack in a relatively uniform background, where detection and classification are easier. (c) An obscure crack network, characterized by uneven illumination and some noise inclusion, making image processing difficult and network identification challenging. (d) An image with fine, thin cracks that are not easily discernible. Various noises are present in the image, and some noise bands are much broader than the cracks, thus complicating the application of existing algorithms to manage the image. Due to the texture and variable intensity of the image, accurately detecting the type of pavement damage is relatively difficult. Most methods are based on good image quality and clear cracks that are inconsistent with the actual environment, making it difficult to meet practical needs. In this study, labeling is conducted based on the Pavement Condition Index (PCI) method, focusing on the pavement surface degradation status rather than the type of damage. The labeling categorizes images into three conditions: poor, fair, and good.

Fig. 3. Four types of cracks: (a) scattered cracks, (b) network cracks, (c) obscure cracks, (d) fine cracks.

Until now, the captured images have been clear and noise-free, focused on the crack area, and possessing a uniform texture. Applying this pattern in practice to find cracks faces certain limitations. The collected data here represent the general condition of road pavement in Tehran. The total number of images for training the model is 600. Additionally, to reduce computational costs, the number of longitudinal and latitudinal pixels has been halved from the initial value.

8. DISCUSSION ON RESULTS

The training of the proposed convolutional fuzzy neural network involves three independent stages corresponding to the three components of the network. Firstly, we train the convolutional neural network (comprising convolutional layers and pooling layers) to form some abstract features of the input image. Utilizing pre-trained models with large datasets from a related domain is a common strategy. Therefore, we can train and prepare this layer with a rich dataset. The second part involves tuning the parameters of the fuzzy layer, a process known as self-organization. This layer is trained in an unsupervised manner using a competitive learning scheme. The self-organization of the fuzzy layer means selecting the positions of cluster centers (choosing the parameters of membership functions). Various fuzzy clustering algorithms can be applied, such as the C-means algorithm or the Gustafson-Kessel algorithm. Competition means that, given an input, processing elements in a neural network will compete for the output. This competition fosters specialization within the network. For each input, the processing elements produce an output. Only the most appropriate output is used, and only the winning processing element is updated.

The third part involves training the classifier. The parameters of the convolutional and fuzzy layers remain stable. Only the weights of the fully connected layers are adjusted. The classifier is trained using a standard backpropagation algorithm. After completing these three training phases, the proposed model is ready for operation.

The starting point involves using a three-layer convolution-pooling model for the architecture of the proposed convolutional neural network. This lightweight and simple architecture demonstrates the capabilities of the proposed model. Training a convolutional neural network requires substantial resources, so we only use 2 training epochs. Subsequently, self-organization of the fuzzy layer is carried out using fuzzy c-means clustering. The dataset is clustered several times with different numbers of clusters. We then select the number of clusters that maximizes the Fuzzy Partition Coefficient (FPC), which is defined in the range of 0 to 1, with 1 being the best. The FPC is a metric that indicates how clearly our data is described by a specific clustering model.

In the final part, training of the classifier is performed using a stochastic optimization method. Only the weights of the fully connected layers are adjusted, while the parameters of the convolutional and fuzzy layers remain stable. The result of testing the proposed model on the test data shows an accuracy of 63.7%. The result of testing the data with a convolutional network with the same number of layers is 58.3%. This outcome indicates that the inclusion of the fuzzy layer in the convolutional layer allows for an improvement in quality and accuracy.

9. CONCLUSION

A convolutional fuzzy neural network model for image classification is presented in this paper. Fuzzification is integrated into the structure of the convolutional network. The proposed model combines the strengths of convolutional neural networks and fuzzy logic, making it capable of handling uncertainty and learning more effectively. An empirical comparison was conducted to measure the accuracy of the convolutional fuzzy neural network against a standard convolutional neural network. The results demonstrate that the proposed model can achieve better accuracy with reduced training time.

Transparency Statement

The data supporting this study are available upon reasonable request to the corresponding author, subject to ethical and confidentiality considerations.

Acknowledgments

We would like to express our gratitude to all individuals who contributed to this project.

Declaration of Interest

The authors declare that they have no competing interests.

Funding

This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

REFERENCES

[1] Elhadidy, A. A., El-Badawy, S. M., & Elbeltagi, E. E. (2021). A simplified pavement condition index regression model for pavement evaluation. International Journal of Pavement Engineering, 22(5), 643–652. https://doi.org/10.1080/10298436.2019.1633579

[2] Maeda, H., Sekimoto, Y., & Seto, T. (2016). Lightweight road manager: Smartphone-based automatic determination of road damage status by deep neural network. In Proceedings of the 5th ACM SIGSPATIAL International Workshop on Mobile Geographic Information Systems (MobiGIS 2016) (pp. 37–45). https://doi.org/10.1145/3004725.3004729

[3] Maeda, H., Sekimoto, Y., Seto, T., Kashiyama, T., & Omata, H. (2018). Road damage detection using deep neural networks with images captured through a smartphone. Computer-Aided Civil and Infrastructure Engineering, 33(4–6), 1–4. https://doi.org/10.1111/mice.12387

[4] Zhang, A., Wang, K. C. P., Fei, Y., Liu, Y., Chen, C., Yang, G., Li, J. Q., Yang, E., & Qiu, S. (2019). Automated pixel-level pavement crack detection on 3D asphalt surfaces with a recurrent neural network. Computer-Aided Civil and Infrastructure Engineering, 34(3), 213–229. https://doi.org/10.1111/mice.12409

[5] Nath, V., & Levinson, S. E. (2014). Machine learning. https://doi.org/10.1007/978-3-319-05606-7_6

[6] Džeroski, S. (2008). Data mining. In Encyclopedia of Ecology (pp. 857–863). https://doi.org/10.1016/B978-008045405-4.00153-1

[7] Studies, W. C., Singh, H., & Lone, Y. A. (2020). Deep neuro-fuzzy systems with Python: With case studies and applications from the industry.

[8] Gaweda, A. E. (2006). Flexible neuro-fuzzy systems: Structures, learning and performance evaluation. IEEE

Transactions on Neural Networks, 17(2), 270. https://doi.org/10.1109/TNN.2005.864047

[9] Karim, F. M. A., Rubasi, K. A. H., & Saleh, A. A. (2016). The road pavement condition index (PCI) evaluation and maintenance: A case study of Yemen. Organization, Technology and Management in Construction: An International Journal, 8(1), 1446–1455. https://doi.org/10.1515/otmcj-2016-0008

[10] Shahnazari, H., Tutunchian, M. A., Mashayekhi, M., & Amini, A. A. (2012). Application of soft computing for prediction of pavement condition index. Journal of Transportation Engineering, 138(11), 1495–1506. https://doi.org/10.1061/(ASCE)TE.1943-5436.0000454