

Deep machine learning in bridge structures durability analysis

Karolina Tomasziewicz, Tomasz Owerko

Department of Engineering Surveying and Civil Engineering, Faculty of Geo-Data Science, Geodesy and Environmental Engineering, AGH University of Science and Technology, Mickiewicza 30 Av., 30-059 Krakow, Poland, (tomaszki@agh.edu.pl; owerko@agh.edu.pl)

Key words: *deep machine learning; transfer learning; concrete cracks; bridge structure; durability*

ABSTRACT

According to Eurocode 0 structural durability is next to ultimate and serviceability one of the basic criteria in the structural design process. This article discusses the subject of concrete cracks observation in bridge structures, as one of the factors determining their durability. The durability of bridge structures is important due to both social, economic aspects and also the defense aspects of countries. Cracking of the reinforced concrete structures is a natural effect in concrete. The aim in the design and construction of structures is not to prevent the formation of cracks, but to limit their width to acceptable values. At the same time, there is a need for structure tests that allow for non-contact, fast measurements and algorithms that allow for efficient analysis of large amounts of measurement data. Deep machine learning algorithms can be used here. They can be used to analyse data which are acquired by means of photogrammetric methods (especially helpful during construction to inventory concealed works). Moreover, they can also be applied to standard data acquisition methods, consisting in photographing objects damage during works acceptance or periodic inspections. This paper discusses the application of deep machine learning to assess the condition of bridge structures based on photographs of object damage. The use of this method makes it possible to observe the rate and extent of damage development. Consequently, this method makes it possible to predict the development of damage in time and space in order to prevent failures and take structures out of service.

I. INTRODUCTION

According to Eurocode 0, structural durability is one of the basic criteria in the process of designing buildings, in addition to ultimate and serviceability. It has a significant influence on the adopted structural and material solutions. This paper concentrates on the subject of bridge structures damage observation, the reliability of which is crucial due to social and economic aspects as an element of transport and transit infrastructure. These objects also play an important role due to the aspects of national defense.

The cracks are occurring when the force in a given point in the structure exceeds concrete maximum tensile strength. The phenomenon of cracking in reinforced concrete structures is a natural phenomenon of concrete. The aim in the design and construction of structures is not to prevent the formation of cracks, but to limit their expansion to the values specified in national standards.

Identifying the location of the crack line in a concrete member is an equally important aspect that can give knowledge about the nature of the factor causing the crack in the structure. The progression of the crack over time must also be considered. Some cracks occurring on the surface of a structure do not increase in dilation over time and remain present on the surface of the structure throughout its life cycle. The conclusion is that not every crack may be a potential threat to the

durability of the structure. However, expert knowledge is needed to clearly and responsibly identify potentially dangerous locations.

The phenomenon of concrete cracking can be caused by many factors, which makes the problem complicated (Zhang, 2018). The problem of cracking concerns in particular the massive elements of the bridge such as foundations, abutments, pillars and pylons. Cracking occurs due to the increase of thermal-shrinkage stresses in the concrete of these elements. Under the influence of loads it may lead to the propagation of cracks through the entire thickness of the element, and as a consequence to the loss of monolithicity of the element and changes in the static scheme, or to the development of corrosion processes under the influence of water penetrating through cracks into the element.

At each stage of the life cycle of the structure, a number of factors can be listed that affect its durability (Germaniuk *et al.*, 2016; Zhang, 2018). It requires close cooperation between the designer, contractor, concrete technologist, inspector, and the investor or user of the facility to ensure durability. Examples of factors at different stages of the structural life cycle that influence the possibility of cracking are presented in Figure 1. All this makes the aspect of crack identification seem to be crucial for the durability of the structure.

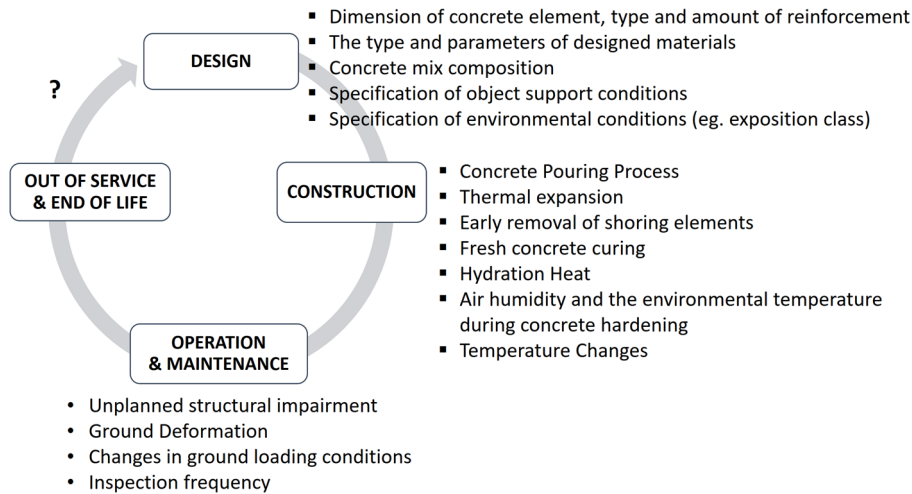


Figure 1. Examples of factors at different stages of the structural life cycle that influence the possibility of cracking.

The inspection, condition assessment and maintenance procedures for bridges are different in each country. Helmerich *et al.* (2008) describe bridge infrastructure maintenance requirements, inspection and condition assessment procedures, and ongoing testing programs. A review of bridge management system (BMS) and bridge inspection practices in China, Japan, Korea, and the US are presented in (Jeong *et al.*, 2018).

Currently, the still popular method for collecting data on concrete surface cracks is to obtain them from the results of an inspector's inspection. However, this process has many limitations (Kim *et al.*, 2015). Among other things, we can point out the costliness of this process, the time-consuming nature of the inspections performed, and the labor-intensive nature of the report preparation. Moreover, this process requires the use of specialized equipment to inspect surfaces located at height or in difficult terrain, thus posing a danger to the person performing the inspection. In addition, the process is dependent on the experience and subjective judgment of the inspector (Xu *et al.*, 2021).

It should be noted that current methods require the person performing the inspection to manually create an inspection map and mark damage locations. This further increases the labor intensity and cost of performing the inspection. Currently, in connection with the digitalization of construction, solutions have already been created to support the construction engineer and supervisor in their daily work. Kim *et al.* (2018) present an approach in which they use deep learning to simultaneously classify and locate cracks acquired with UAVs. In addition, they create a model of the bridge structure based on the point cloud and create a map of the structure on which the cracks identified by the network are marked. This makes it possible to link the results to the bridge management system (BMS) and process them automatically over time. Visualization of the detected damage on a 3D model based on photographs is also used by (Wu *et al.*, 2019).

Nowadays there is a need to use such construction surveys that allow for non-contact, quick measurements and algorithms that allow for efficient analysis of large amounts of measurement data. The algorithms of deep machine learning used to analyze data acquired by means of photogrammetric methods (especially helpful in the course of construction for inventory of disappearing works) as well as standard methods consisting in taking pictures of object damages with the use of cameras during works acceptance or periodic inspections are applicable here.

The aim of this work is to verify the possibility of using deep machine learning algorithms to assess the condition of structures on the basis of images of object damage taken with the camera, without prior preparation. The aim of the research is to show the possibilities of using this method to detect and locate cracks and to observe the rate and extent of processes causing damage. The goal is to be able to predict the development of cracks in time and space on the basis of damage images in order to prevent failures.

The use of machine learning algorithms can be a tool that will support the engineer in his daily work. Moreover, it can be particularly helpful for hard-to-reach sites, especially when combined with photogrammetric data acquisition methods. Well-crafted machine learning algorithms can help detect cracks at an early stage of their formation, where the human eye does not always perceive the danger anymore.

II. PROBLEM DESCRIPTION

The use of deep machine learning including transfer learning for structural condition assessment from images of object damage is a rising trend. The possibilities arising from using transfer learning and comparing the performance of several proposed deep convolutional neural networks in the problem of image-based automatic detection of concrete surface cracks on a small dataset are discussed in (Słowski, 2019). A comparison of the performance of pre-trained

networks on datasets of different sizes is presented in (Ali *et al.*, 2021). Network architecture selection in the context of limited computational resources is mentioned by Su *et al.* (2020). The impact of model parameter selection and network learning rate is presented in (Li and Zhao, 2019).

Classifying damage images as cracked and non-cracked is a binary problem. Solving the problem based on convolutional neural networks requires selecting the neural network architecture and model parameters such that the measures used to evaluate model accuracy are satisfactory. Convolutional networks are typically trained using a back propagation algorithm to compute the gradient of the loss function to iteratively update the model weights.

Based on the proposed network architecture and model parameters using a training dataset, which is part of the whole dataset, the network model is trained. Based on another validation subset extracted from the whole dataset, the solution is verified. The purpose of this activity is to validate the performance of the network against data that the model has not seen before. The testing and validation process is an iterative process until such an architecture and model parameters are selected that produce stable and satisfactory classification results. The network thus trained is tested on test sets that are either a subset of the dataset on which we have been working, or is a completely independent dataset.

The problem that occurs in deep machine learning is the need for very large datasets on which the neural network is trained. Moreover, the stage of architecture selection, model parameter tuning, as well as the network learning process itself are complicated, labor-intensive and costly issues. A frequently used practice in this case is transfer learning. It consists in using the neural network model, which was previously learned on a large dataset, to solve a new problem. This makes it possible to solve problems for which we have a relatively small amount of data.

In this paper, transfer learning was used. It is based on learning a convolutional base and retraining the classifier. The convolutional base was frozen, *i.e.*, for the layers included in the convolutional base, the weights adopted in the model were not updated in subsequent training steps. The images of concrete surface cracks were passed through the selected network, and by retraining the proposed classifier, it was possible to adapt the network to the dedicated crack detection problem. The model thus trained was validated on the test set and showed satisfactory prediction results.

III. METHODOLOGY AND COMPUTATIONAL IMPLEMENTATION

A. Training, validation and test datasets

The dataset used to train, validate and test the convolutional neural network model is the publicly

available dataset "Concrete Crack Images for Classification" (Özgenel, 2019). The dataset contains 40,000 images with 227 x 227 pixels with RGB channels. The dataset was based on 458 high-resolution images. These images have variance in terms of illumination conditions and surface finish. The images were divided into 2 classes: "Positive", containing images of cracked concrete surfaces, and "Negative", which is a dataset of images of non-cracked surfaces. Each class contains 20,000 images. Examples of images from this collection are shown in Figure 2 and Figure 3. This images were randomly divided into 3 subsets: train, validation, test. The images were divided in a percentage ratio of 60:20:20 (*i.e.* 24000, 8000, 8000 images). Each subset was equally divided into "Positive" and "Negative".



Figure 2. Examples of cracked surfaces ("Positive") (Özgenel, 2019).

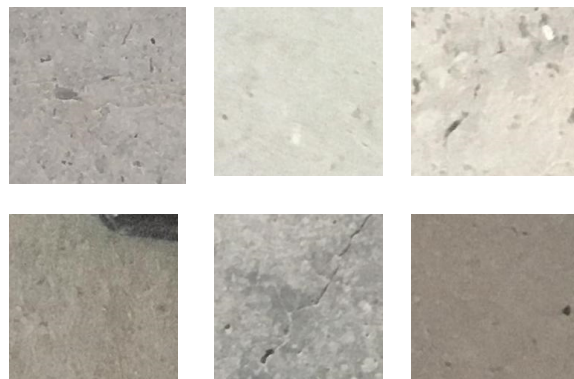


Figure 3. Examples of non-cracked surfaces ("Negative") (Özgenel, 2019).

The second dataset is the authors' own collection. The photographs of the damaged concrete surfaces were taken using a camera and a mobile phone camera and were taken without prior conditioning. The purpose of this form of data acquisition was to verify the possibility of using images taken without paying attention to the parameters and settings of the equipment used to acquire the image. This is particularly important because under construction conditions it is not always possible to have full access to the study area and to focus carefully on cropping and sharpening the image. Hence, one aspect of this work is to verify the feasibility of using deep machine learning for crack detection regardless of the quality of the

concrete surface image taken, the lighting conditions, and the type and parameter of the equipment used to capture the image.

The authors' own dataset was built from images of 16 bridge structure elements (abutments, pillars) that were segmented into 4665 sub-images with 227 x 227 pixels. The images were manually classified and labeled as "Positive" if there was a crack in the image or "Negative" if there was no crack - to maintain analogy with the dataset described above (Özgenel, 2019). Finally, 309 images of the cracked concrete surface (*i.e.*, "Positive") and 4356 images of the non-cracked surface (*i.e.*, "Negative") were obtained. The authors' own dataset consists of cracks of various dilation. It contains images with various obstructions including shadows, varying illumination, surface roughness, holes, and background contamination. Examples of images from the authors' dataset are shown in Figures 4 and 5.

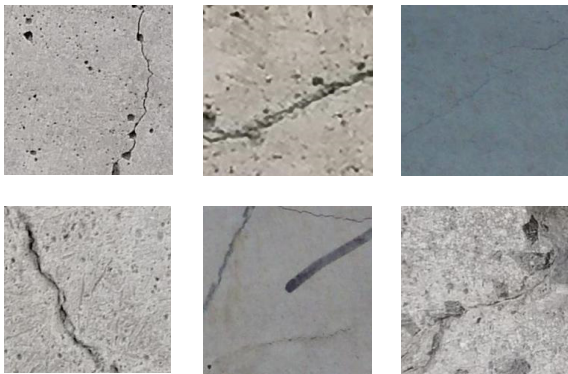


Figure 4. Examples of cracked surfaces ("Positive") - authors' own dataset.



Figure 5. Examples of non-cracked surfaces ("Negative") - authors' own dataset.

B. Architecture of convolutional neural network

The computer implementation was programmed in Python using the open-source Keras library. This library runs on the open-source TensorFlow machine learning library developed by Google.

This paper uses transfer learning using the VGG-16 network proposed by (Simonyan and Zisserman, 2015), trained on the ImageNet dataset. The VGG-16 model consists of 16 layers, including 13 convolutional layers for feature extraction and a classifier composed of three fully connected layers with a filter size of 3 x 3.

In the approach used in this study, the frozen convolutional basis of the VGG-16 model was used for feature extraction and a custom classifier was proposed. This procedure was intended to adapt the network to the dedicated problem of concrete surface crack detection. The overall architecture of the proposed CNN is shown in Figure 6.

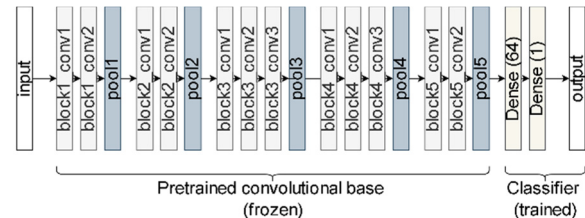


Figure 6. Proposed architecture of a convolutional neural network.

The image size at the output of the VGG16 network was (512, 7, 7), and the total number of network parameters at this stage was 14 714 688.

The proposed classifier consisted of a deep layer with ReLu activation function. Before entering the data into this layer, the vector was flattened to a one-dimensional form, because this is the form of data required by the Dense layer. Additionally, Dropout layer with probability 0,4 was used for regularization. This technique, which involves random freezing of individual neurons in the network during the learning process, allows for more accurate matching of the model to the data. In particular, this is useful for small data sets that tend to over-fit, and using dropout allows the network to be taught in a more generalized way. The second deep layer uses the Sigmoid activation function to determine the probability that the analyzed image is outlined.

As described earlier, the network learning process used a frozen convolutional base, *i.e.*, the parameters included in it were not subject to the learning process. The weights obtained for the network learned on the ImageNet dataset were adopted. Only the parameters included in the proposed classifier were trained. As part of the classifier selection, the performance of the network was tested as a function of the number of filters used in the first Dense layer, and thus the number of parameters learned in training the network. The use of 16, 32 and 64 filters in this dense layer was verified.

The total number of parameters depending on the number of dense layer filters is shown in Table 1. In addition, the table shows the number of trainable and non- trainable parameters.

Table 1. The number of model parameters

Number	Filters		
	16	32	64
Total	15 116 129	15 517 569	16 320 449
Trainable	401 441	802 881	1 605 761
Non-trainable	14 714 688		

Results showed that using a Dense layer with 64 filters produces satisfactory results. In the following section we present results for a classifier consisting of this Dense layer with 64 filters.

C. Network training, validation and testing

Calculations were performed on a laptop with a GPU. The computations were performed on a platform with the parameters: CPU - Intel(R) Core(TM) i7-8750H, GPU - NVIDIA Quadro P1000, 16GB RAM. The use of such hardware has shown that it is not necessary to use specialized hardware to perform such analysis. Therefore, such analyses are within the range of equipment available to all the engineers.

Due to the binary nature of the problem, a binary cross-entropy loss function was used. The use of different optimizers in the training process was considered. Results show that the highest accuracy was obtained using the Adam type optimizer. Finally the Adam optimizer was used for parameter updates. A validation set was used to monitor during learning the accuracy of the model on data that the network had never seen before. This consisted of images previously extracted from the output set (Özgenel, 2019), as described in paragraph III.A The performance of the network was tested by prediction on two datasets. The first consisted of samples separated from the output set (Özgenel, 2019). The second consisted only of samples from the authors' own dataset.

Various metrics were used to evaluate the network, such as accuracy, precision, sensitivity and specificity (König *et al.*, 2022). This metrics are explained in more detail below. The symbols in the formulas denote elements of the confusion matrix, where: TP - True Positive - number of true cracked images correctly predicted as cracked, FP - False Positive - number of true non-cracked images predicted as cracked, TN - True Negative - number of true non-cracked images predicted as non-cracked, FN - False Negative - number of true cracked images predicted as non-cracked.

Accuracy can be defined as the ratio of correct model predictions to the total number of tested images, as shown in Equation 1:

$$\text{accuracy} = \frac{TP+TN}{TP+FP+TN+FN} \quad (1)$$

Precision (Eq. 2) defines how many of the images predicted as positive were predicted by the model correctly:

$$\text{precision} = \frac{TP}{TP+FP} \quad (2)$$

Sensitivity, shown in Equation 3, defines how many of all true positive images were correctly predicted by the model:

$$\text{sensitivity} = \frac{TP}{TP+FN} \quad (3)$$

Specificity (Eq. 4) defines how many of all true negative images were correctly predicted by the model:

$$\text{specificity} = \frac{TN}{TN+FP} \quad (4)$$

F1 score can be defined as the harmonic mean of precision and recall, as shown in Equation 5:

$$\text{F1 score} = \frac{2 \times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \quad (5)$$

Because the authors' test dataset is imbalanced, a balanced accuracy metric (Eq. 6) was used to evaluate the model for this dataset. This metric is defined as the arithmetic mean of sensitivity and specificity:

$$\text{balanced accuracy} = \frac{\text{sensitivity} + \text{specificity}}{2} \quad (6)$$

The values of each metric for both test datasets are presented in the following section.

D. Results

Verification of damage classification shown in images of cracked and non-cracked concrete surfaces for a test dataset extracted from (Özgenel, 2019) showed a 99% accuracy. The confusion matrix for this set is shown in Table 2. Using the network to classify cracks on the authors' own dataset showed a accuracy of over 91%, but because the authors' own dataset is an imbalanced dataset the balanced accuracy metric was used as more adequate. The value of balanced accuracy for the authors' own dataset is over 55%. The confusion matrix for this set is shown in Table 3. The metrics for both test datasets are shown in Table 4.

Table 2. Confusion matrix for publicly available dataset (Özgenel, 2019)

	Predicted non-cracked	Predicted cracked
True non-cracked	3996	4
True cracked	6	3994

Table 3. Confusion matrix for authors' own dataset

	Predicted non-cracked	Predicted cracked
True non-cracked	4315	41
True cracked	295	14

The analysis showed that the presented network architecture allows for high accuracy, but the architecture is not very universal. For the test dataset of images with crack character corresponding to the cracks on which the model was trained, high precision is obtained (over 99%). This is shown by the analysis

performed using the test set extracted from the public data (Özgenel, 2019). However, as the analysis performed on the authors' own data set shows, the network is able to detect mainly such cracks whose nature is similar to that of the images that were used to train the model. This is confirmed by an analysis of classification results examples for cracked images from the authors' own dataset which is shown in Figure 7.

Table 4. Metrics for test datasets [%]

	Publicly available dataset (Özgenel, 2019)	Authors' own dataset
Accuracy	99,88	91,55
Balanced accuracy	-	57,44
Precision	99,90	28,43
Sensitivity	99,85	18,12
Specificity	99,90	96,76
F1 score	99,88	22,13

The analyzed network correctly classifies images with similar crack types, but is limited in its ability to detect damage of a different nature. To obtain higher levels of accuracy, the type of damage shown in the analyzed image must belong to the hypothesis space of the model.

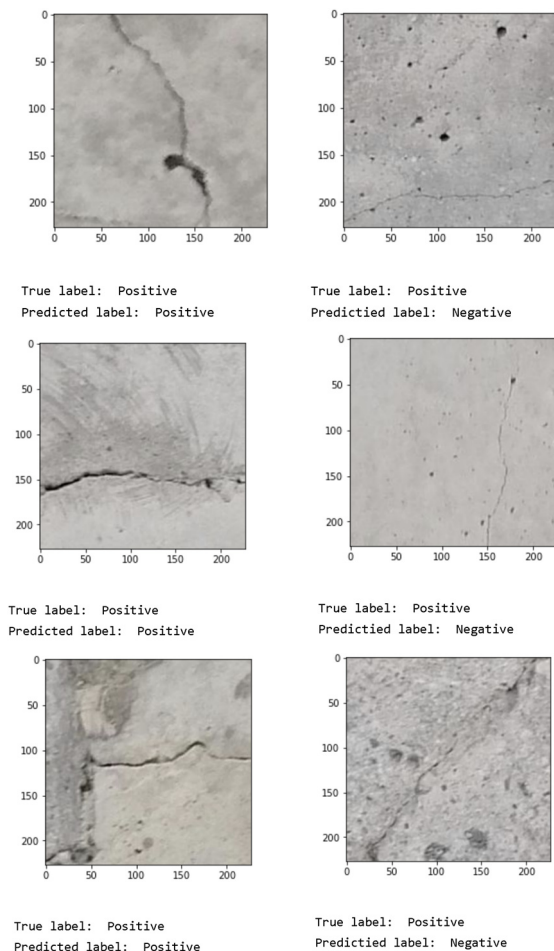


Figure 7. Examples of classification results for cracked images from the authors' own dataset.

IV. CONCLUSION

The crack detection is critical in the diagnosis of bridge structures. Early detection of defects allows for faster response to damage and planning of necessary repair work. This is expected to lead to a situation where the bridge structure does not need to be taken out of service. There is a need for research that allows for non-contact, high-speed measurements of objects and algorithms that allow for efficient analysis of large amounts of measurement data. It is also important to be able to store this data in a single facility management system so that you have a complete database and can observe the development of damage over time.

The application of deep machine learning to assess the condition of bridge structures based on images of structure damage is a response to these needs and one of the better developing trends today. This paper presents an example of using transfer learning to train a network based on publicly available data. The model trained on these data was used to predict damage on the authors' own dataset. The result of the research is one of the stages of work aimed at presenting a solution that would make it possible not only to detect the presence of cracks, but also to distinguish those cracks that constitute an excess of permissible norms based on their dilation value.

The results of the authors' research confirm the feasibility of using deep machine learning for cracks detection from photos. Importantly, as the results show, it is also possible to use photos that were taken with a camera or phone camera without configuring its settings or preparing the shot. This is extremely important in construction conditions or during the inspection of the object, where it is not always possible to have full access to photograph the element of interest and the possibility to take an accurate shot.

As the analysis for the applied network architecture shows it can be concluded that despite obtaining high accuracy this architecture is not very universal and the obtained results depend on whether the analyzed crack belongs to the model hypothesis space. In future steps, the authors plan to use the proposed model architecture to identify pavement cracks as having a crack character similar in size to that presented in the used publicly available dataset. In addition, other architectures will be tested to verify the possibility of finding a more universal architecture. It should allow the use of transfer learning to classify cracks regardless of their nature.

The aim of all works is to create such a tool which will be helpful for civil engineer in everyday work and will not require him/her to train network, special preparation of data. This solution will allow real-time assessment of the tested concrete surfaces with indication of the location of those damages which, in accordance with the regulations in force in a given country, are beyond the serviceability limit state and are a danger to the structure.

References

- Ali, L., Alnajjar, F., Jassmi, H.A., Gocho, M., Khan, W., and Serhani, M.A., (2021). Performance Evaluation of Deep CNN-Based Crack Detection and Localization Techniques for Concrete Structures. *Sensors* 21, 1688. DOI: 0.3390/s21051688
- Germaniuk, K., Gajda, T., Sakowski, A., Wierzbicki, T., and Kamiński, P., (2016). Bridge Structures Cracks – What Made that Phenomena so Common? *Transportation Research Procedia* 14, pp. 4030–4039. DOI: 10.1016/j.trpro.2016.05.500
- Helmerich, R., Niederleithinger, E., Algernon, D., Streicher, D., and Wiggenhauser, H., (2008). Bridge Inspection and Condition Assessment in Europe. *Transportation Research Record* 2044, pp. 31–38. DOI: 10.3141/2044-04
- Jeong, Y., Kim, W., Lee, I., and Lee, J., (2018). Bridge inspection practices and bridge management programs in China, Japan, Korea, and U.S. *Journal of Structural Integrity and Maintenance* 3, pp. 126–135. DOI: 10.1080/24705314.2018.1461548
- Kim, H., Sim, S.H., and Cho, S., (2015). Unmanned Aerial Vehicle (UAV)-powered Concrete Crack Detection based on Digital Image Processing. Presented at the 6th International Conference on Advances in Experimental Structural Engineering 11th International Workshop on Advanced Smart Materials and Smart Structures Technology, Urbana-Champaign, United States.
- Kim, I.-H., Jeon, H., Baek, S.-C., Hong, W.-H., and Jung, H.-J., (2018). Application of Crack Identification Techniques for an Aging Concrete Bridge Inspection Using an Unmanned Aerial Vehicle. *Sensors* 18, 1881. DOI: /10.3390/s18061881
- König, J., Jenkins, M., Mannion, M., Barrie, P., and Morison, G., (2022). What's Cracking? A Review and Analysis of Deep Learning Methods for Structural Crack Segmentation, Detection and Quantification. DOI: 10.48550/ARXIV.2202.03714
- Li, S., and Zhao, X., (2019). Image-Based Concrete Crack Detection Using Convolutional Neural Network and Exhaustive Search Technique. *Advances in Civil Engineering* 2019, 1–12. DOI: 10.1155/2019/6520620
- Özgenel, Ç.F., (2019). Concrete Crack Images for Classification. DOI: 10.17632/5Y9WDSG2ZT.2
- Simonyan, K., and Zisserman, A., (2015). Very Deep Convolutional Networks for Large-Scale Image Recognition. CoRR.
- Słoński, M., (2019). A comparison of deep convolutional neural networks for image-based detection of concrete surface cracks. *Computer Assisted Methods in Engineering and Science* 26, pp. 105–112. DOI: 10.24423/CAMES.267
- Su, C., and Wang, W., (2020). Concrete Cracks Detection Using Convolutional NeuralNetwork Based on Transfer Learning. *Mathematical Problems in Engineering* 2020, 1–10. DOI: 10.1155/2020/7240129
- Wu, Z.Y., Kalfarisi, R., and Systems, B., (2019). Deep Learning-based Defect Detection and Assessment for Engineering Structures. Presented at the 5th International Conference on Smart Monitoring, Assessment and Rehabilitation of Civil Structures (SMAR 2019), Potsdam, Germany.
- Xu, Z., Qian, S., Ran, X., and Zhou, J., (2021). Application of Deep Convolution Neural Network in Crack Identification. *Applied Artificial Intelligence* 1–22. DOI: 10.1080/08839514.2021.2014188
- Zhang, T., (2018). Analysis on the Causes of Cracks in Bridges. *Journal of Construction Research* 1. DOI: 10.30564/jcr.v1i1.83