Concrete Cracks Detection Based on Deep Learning Image Classification [†]

Wilson Ricardo Leal da Silva [1,*] and Diogo Schwerz de Lucena [2]

[1] Danish Technological Institute, Gregersensvej 4, 2630 Taastrup, Denmark

[2] Biorobotics Lab—Dept. of Mechanical and Aerospace Engineering, Univ. of California, Irvine, CA 92697, USA; dschwerz@uci.edu

* Correspondence: wrls@teknologisk.dk; Tel.: +45-7220-1034

[†] Presented at the 18th International Conference on Experimental Mechanics (ICEM18), Brussels, Belgium, 1–5 July 2018.

Abstract: This work aims at developing a machine learning-based model to detect cracks on concrete surfaces. Such model is intended to increase the level of automation on concrete infrastructure inspection when combined to unmanned aerial vehicles (UAV). The developed crack detection model relies on a deep learning convolutional neural network (CNN) image classification algorithm. Provided a relatively heterogeneous dataset, the use of deep learning enables the development of a concrete cracks detection system that can account for several conditions, e.g., different light, surface finish and humidity that a concrete surface might exhibit. These conditions are a limiting factor when working with computer vision systems based on conventional digital image processing methods. For this work, a dataset with 3500 images of concrete surfaces balanced between images with and without cracks was used. This dataset was divided into training and testing data at an 80/20 ratio. Since our dataset is rather small to enable a robust training of a complete deep learning model, a transfer-learning methodology was applied; in particular, the open-source model VGG16 was used as basis for the development of the model. The influence of the model's parameters such as learning rate, number of nodes in the last fully connected layer and training dataset size were investigated. In each experiment, the model's accuracy was recorded to identify the best result. For the dataset used in this work, the best experiment yielded a model with accuracy of 92.27%, showcasing the potential of using deep learning for concrete crack detection.

Keywords: artificial intelligence; concrete cracks; deep learning; image classification

1. Introduction

The maintenance and repair of large concrete infrastructure is rather costly, e.g., in Europe, €4–6 billion are spent annually on maintenance of concrete infrastructure [1], let alone the associated costs that derive from traffic disruptions. It is reasonable to assume that the number of required inspections is expected to increase in the next few decades as concrete structures from the 60's and 70's are presumably exceeding their expected service life.

Solutions to monitor structures usually consists in embedded sensors to detect structural stiffness changes or corrosion initiation, e.g., vibrating wires [2] and optical fibres [3], to mention a few. The results from these are then associated with the presence of damages in concrete. However, such monitoring systems are common to modern infrastructure with long service life, i.e., much of the existing ageing concrete structures ought to be inspected in the traditional way, i.e., visual inspection followed by non-destructive testing.

The presence of cracks in concrete has a major impact on the durability of reinforced concrete since they represent an easy path to aggressive agents to reach the reinforcement and trigger the onset

of corrosion. Thus, the detection of cracks is a key aspect in the damage assessment. The earlier a crack is detected, the better the chance to counter the side effects that derive from them. When performing damage assessment of infrastructures, a visual inspection provides an easy mean to detect damages, especially concrete cracks since they are apparent. However, for large infrastructures, visual inspections are rather time consuming and difficult, let alone the safety aspects when dealing with areas that are hard to reach, e.g., the lower part of slab decks in bridges [4]. To facilitate that, several companies are making use of unmanned aerial vehicles (UAV) for visual inspection of structures; especially because the UAV industry is providing reliable, easy to use, and affordable UAV that can help inspectors to improve their efficiency.

Although commercial UAV are becoming easier to operate, it does not necessarily translate to an easier damage assessment workflow, i.e., from data collection to assessment. The biggest challenge relies on the latter, since UAV enables the collection of a larger dataset (at least a few orders to magnitude higher) than that from traditional visual inspection using cameras. Extracting meaningful insights from such extensive amount of data can be time consuming and cumbersome, so the time saved to collect data from a structure can easily be spent on data analysis. Hence, the provision of new paradigms for the assessment of concrete structures, e.g., to automatically identify concrete cracks, is a step towards improving UAV-based inspection workflow, thus providing a safer, low cost, and more objective analysis.

Though the development of automated crack detection system is not entirely new, with a few deployments on UAV [5,6], but it still represents a challenging task that has been explored over the last decades. Most of the development relies on a combination or improvement of conventional digital image processing techniques such as thresholding, mathematical morphology, and edge detection, while utilising photometric (e.g., pixel value) and geometric assumptions (e.g., continuity and local orientation) about properties of crack images to detect cracks—an extensive review can be found in [7]. While reliable in some applications, these methods are based on shallow abstractions (as in traditional rule-based artificial intelligent systems) with limited learning capabilities. Thus, they do not encompass the complexity of conditions that a concrete surface might exhibit, e.g., different light, surface finish, roughness, etc. In fact, such complexity makes it virtually impossible to hard code a rule-based method that is capable to account for all features in concrete surfaces.

Deep learning algorithms, such as Convolutional Neural Networks (CNN), offer means to overcome the existing limitations in crack detection using image processing. Specifically, CNN have successfully been applied to image classification, while featuring a great level of abstraction (generalisation) and learning capabilities, a few examples can be found in [8]. These features are key to detect damages such as cracks in concrete in a robust and reliable manner; modern CNN-based automatic crack detection system under development for pavements are a proof to that [9].

With this vision in mind, we focus on developing a CNN-based model to detect damages on concrete surfaces, thus providing new paradigms for the assessment of structures. At present, the developed system is limited to detecting concrete cracks using a binary classification method, i.e., the system identifies whether or not a crack is present on the concrete surface. The reference image dataset for development has 3500 images of concrete surfaces.

The basis for CNN development relies on transfer-learning, i.e., we build upon a pre-trained open-source model VGG16 and use an experimental study to evaluate the influence of training parameters such as learning rate, number of neurons (or nodes) in the last fully connected layer, and training dataset size on the accuracy of the classification model. The obtained results provide an insight for researchers working with machine learning-based algorithms for crack detection, while the developed model represents the first-generation concrete crack detection tool that can be scaled to more complex models, including multi-label classification.

2. Experimental Program

The experimental program comprises four phases: (1) create a reference classified image dataset; (2) establish the reference artificial intelligent system; (3) implement transfer-learning approach; and (4) run the training experiments. The details on each phase are presented as follows.

2.1. Image Dataset

The dataset comprised 851 pictures with size of 756 × 756 pixels. These pictures were taken from various concrete specimens after mechanical testing at the Concrete Centre at the Danish Technological Institute. The main idea was to collect images of concrete services at different surface appearance to increase the diversity of the dataset and, consequently, of the AI system that learns from this dataset.

To augment the dataset without compromising the resolution, the pictures were sliced into images of 256 × 256 pixels—composing a final dataset with 3500 samples, which were then manually classified in two categories: concrete surfaces with and without cracks. The classified dataset is composed of 2336 and 1164 images with and without cracks, respectively. Examples are shown in Figure 1. The dataset is divided into a training and validation dataset at an 80/20 ratio.
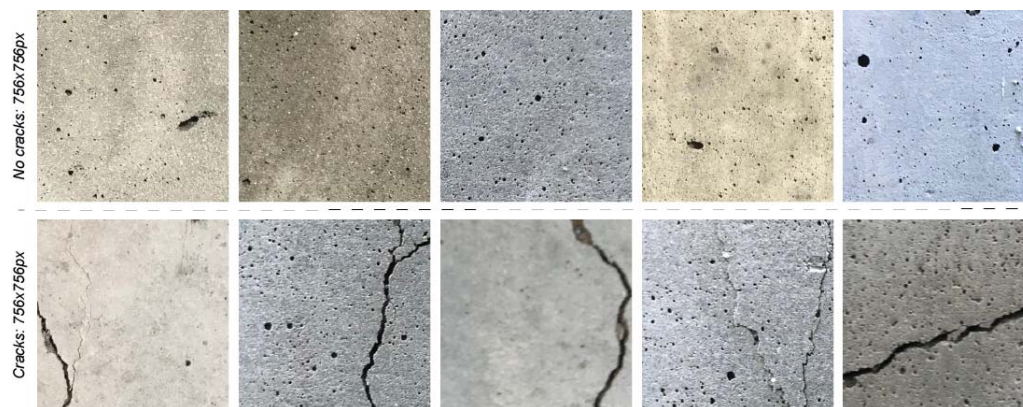

Figure 1. Examples of the images that compose the reference concrete image dataset.

2.2. Network Implementation

As the reference dataset described in Section 2.1 is relatively small to enable a robust training of a complete deep learning model, a transfer-learning methodology was applied. Specifically, an open-source model, namely VGG16 [10], was used as basis for the development of the concrete crack identification model. The VGG16 is a CNN with a total of 16 layers, with 13 convolution layers and 3 fully connected layers, as shown in Figure 2. This model was proposed in the ImageNet Large Scale Visual Recognition Challenge in 2014, and it was trained on the ImageNet Dataset that consists of millions of images divided into thousands of categories (none of which is related to concrete cracks).
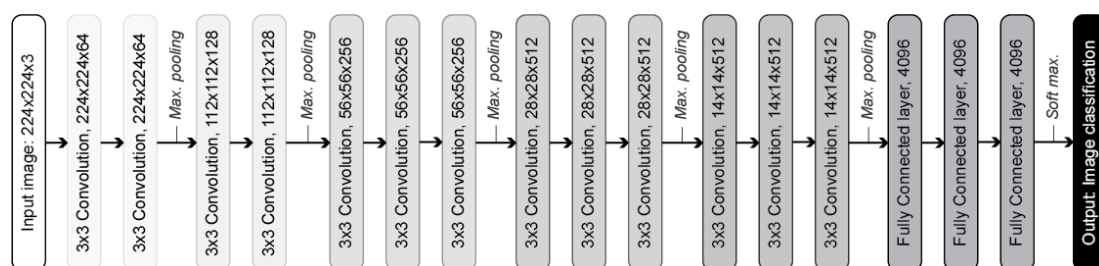

Figure 2. VGG16 architecture - the last layer was re-trained based on the transfer-learning approach.

In our work, the training experiments were carried out using Google Cloud's Computer environment. In this environment, a virtual machine with 8 vCPUs Intel Broadwell, 30GB memory, and 1× NVIDIA Tesla K80 was setup, allowing experimental runs with various values of learning rate, number of nodes, and training sample-size as described in the Section 2.3.

2.3. Training Experiments

The training experiments correspond to a factorial study of three variables, namely learning rate (l_r), number of nodes in the fully connected layer (N_d) and training image dataset size (S_z). Specifically,

the ranges selected for each variable are l_r = [0.01, 0.10, 1.00]; N_d = [32, 64, 128, 256] and S_z = [25%, 50%, 75%, 100%] of the original image dataset. As such, the training experiment of the crack detection network comprises size of l_r × size of N_d × size of S_z = 48 experiments. For each experiment, the value of accuracy (A_c) in each epoch was recorded and used to identify the best model.

It is important to bear in mind that a global optimisation of the parameters from the crack detection model is beyond the scope of this publication. Our primary focus is to showcase the potential of using deep learning with transfer-learning for the development of a crack detection model and identifying the effects of model parameters. Hence, each training experiment was performed only once, and the standard deviation of the results was not computed.

3. Results and Discussion

The results for each experiment are listed in Figure 3, which indicates that a model with l_r = 0.10, N_d = 32, and S_z = 100% yields an accuracy of 92.27%, i.e., the best result encountered in our study.
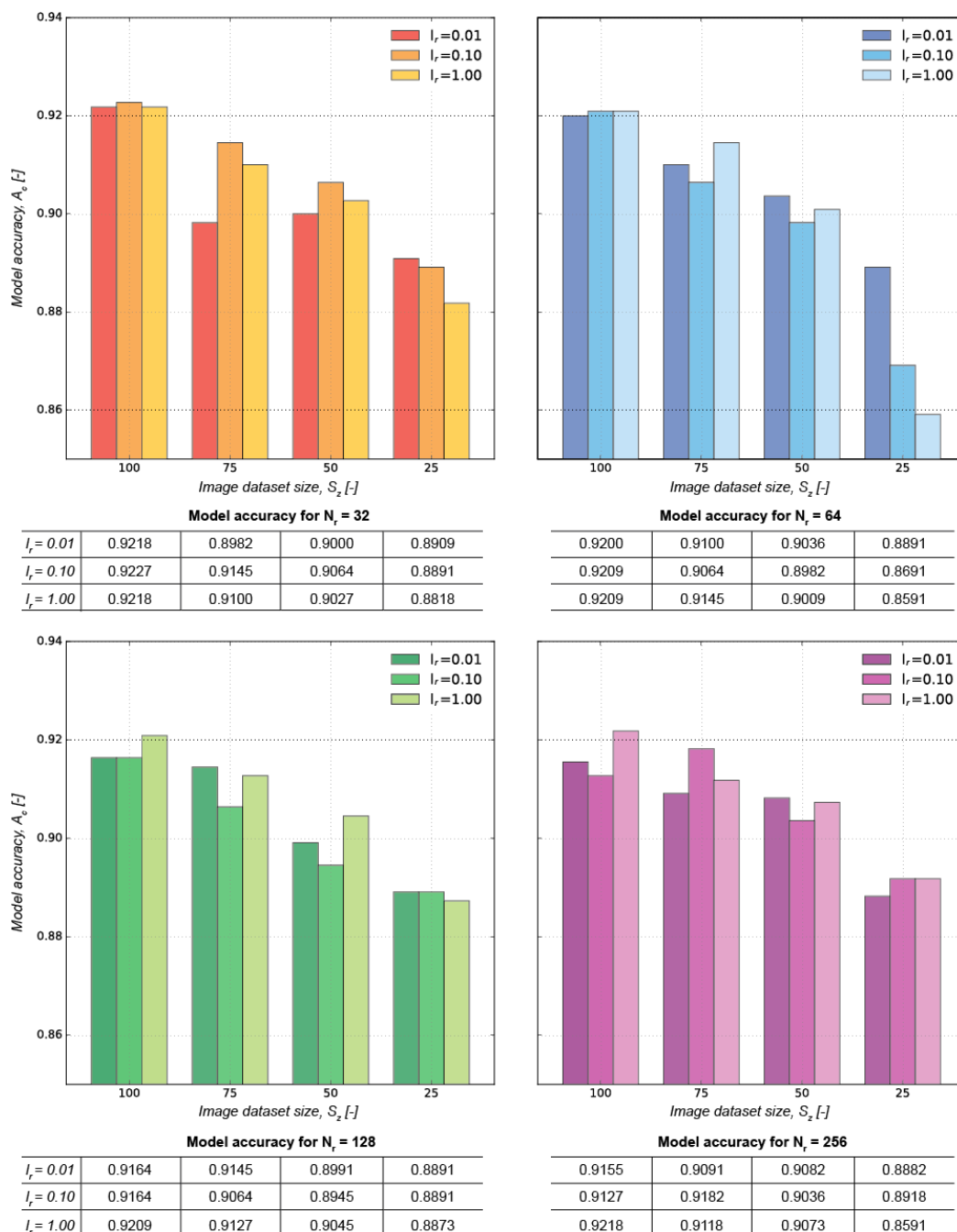

Model accuracy for N_r = 32

	100	75	50	25
$l_r = 0.01$	0.9218	0.8982	0.9000	0.8909
$l_r = 0.10$	0.9227	0.9145	0.9064	0.8891
$l_r = 1.00$	0.9218	0.9100	0.9027	0.8818

Model accuracy for N_r = 64

	100	75	50	25
$l_r = 0.01$	0.9200	0.9100	0.9036	0.8891
$l_r = 0.10$	0.9209	0.9064	0.8982	0.8691
$l_r = 1.00$	0.9209	0.9145	0.9009	0.8591

Model accuracy for N_r = 128

	100	75	50	25
$l_r = 0.01$	0.9164	0.9145	0.8991	0.8891
$l_r = 0.10$	0.9164	0.9064	0.8945	0.8891
$l_r = 1.00$	0.9209	0.9127	0.9045	0.8873

Model accuracy for N_r = 256

	100	75	50	25
$l_r = 0.01$	0.9155	0.9091	0.9082	0.8882
$l_r = 0.10$	0.9127	0.9182	0.9036	0.8918
$l_r = 1.00$	0.9218	0.9118	0.9073	0.8591

Figure 3. Summary of the experiments results—model accuracy.

Figure 3 highlights that the image dataset size (S_z) has the strongest influence of the accuracy of the model. Such trend is expected since a larger dataset enables a greater number of iterations, contributing to a more accurate output. Also, the ascending trend when moving from S_z = 25% to 100% in Figure 3 suggests that an even larger training dataset has the potential to improve the network accuracy further, as there is steady increase in accuracy with no indication of convergence to a maximum accuracy. The evolution of the model's accuracy in each epoch for models with l_r = 0.10, N_r = [32, 256] and S_z = [25%, 50%, 75%, 100%] is depicted in Figure 4.
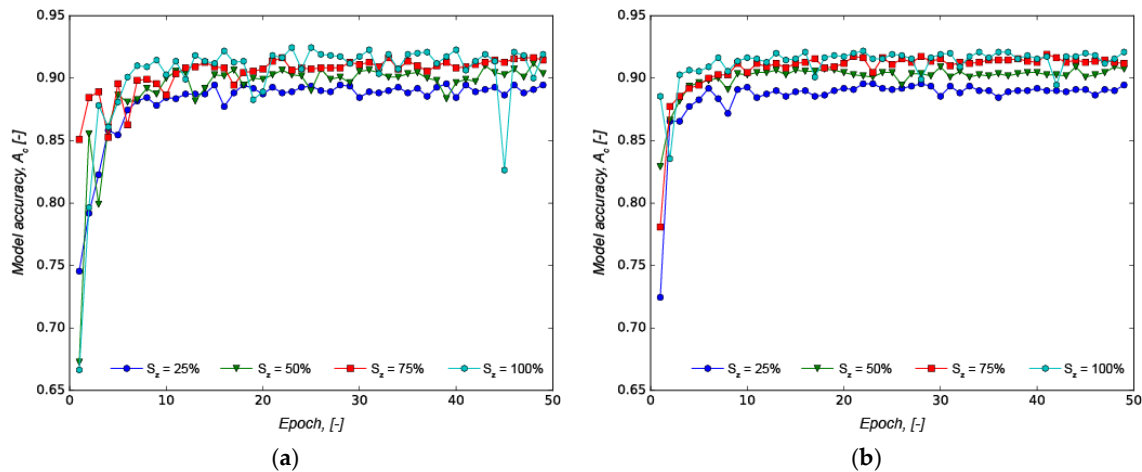

Figure 4. Effect of the image dataset size ($\underline{S_z}$) on the model accuracy: (**a**) N_r = 32 and (**b**) N_r = 256.

Differently from the image dataset size, the effect of N_r and l_r is rather neglectable. Such results are likely because the fully connected layer—i.e., the last layer of the network—relies on binary cross entropy, which is a relatively simple classification approach that can be computed with less complex networks. Furthermore, smaller N_r requires fewer parameters and, therefore, enables a likely faster optimal convergence of the network. Hence, even with N_r = 32 nodes, the neural network yielded great accuracy when compared to more complex models. Notice that a further reduction of $N_r < 32$ is likely to yield a model with prediction accuracy at the same level as the ones listed in Figure 3.

The validation of the abovementioned assumption and the development of a comprehensive parametric study focused on optimising the network variables is currently under investigation and will be published elsewhere. Another aspect for a complementary study relates to the recognition of cracks at the pixel level to enable the creation of a map of cracks and computing the crack width on an evaluated surface. At present, our model is only capable of finding patch level cracks, e.g., for large images that are split into several blocks of 256 × 256 pixels, without considering the pixel level.

Finally, our efforts towards developing a machine learning-based crack identification system have a goal to automate the damage assessment workflow of inspectors dealing with large infrastructure—increasing their productivity when the model is combined with UAV system. The results listed in this article are a step towards that. Notice, however, that while a fully-automated inspection seems far-fetched in a short-term scenario—especially because a much greater dataset of images comprising several damage types and a more complex network is required—the utilisation of machine learning-based tools to assist experts in concrete damage assessment tasks is likely to have a greater chance of success. The future of building and infrastructure inspection workflow envisioned by the authors is depicted in Figure 5.
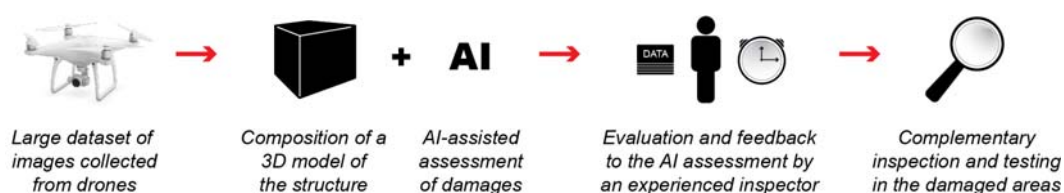

Figure 5. Future workflow of UAV and AI-assisted inspection of concrete infrastructure.

4. Conclusions

This work introduced the implementation of a machine learning-based model to detect cracks on concrete surfaces for infrastructure inspection. Considering the analysis carried out in our experiments, the use of the transfer-learning approach proved suitable to train a model with a limited dataset size. The developed model is limited to a binary classification at a patch level of 256 × 256 pixels—and complementary studies are being carried out to implement a model capable of identifying cracks at a pixel level. From a practical perspective, it is important to mention that, while AI constitutes a path towards automated inspection of concrete structures, the classification model mirrors the knowledge used during its training. Hence, human expertise is key in the development of an appropriate tool. Also, to enable AI models to evolve, a feedback system with corrections from an expert in the field of inspections is crucial. Thus, machine learning-based models are likely to be initially deployed as a tool that assists experts to provide a safer, faster and more productive inspection, creating new possibilities for increased effectiveness in infrastructure asset management by making unbiased, periodic structure monitoring and/or damage assessment feasible.

Author Contributions: W.R.L.d.S. and D.S.d.L. conceived and designed the experiments as well as performed the experiments; analysed the data, and wrote the paper.

Conflicts of Interest: "The authors declare no conflict of interest" and "The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results".

References

1. HealCon: Self-Healing Concrete to Create Durable and Sustainable Concrete Structures. EU FP7-NMP, Project ID: 309451; 2013-16. Available online: https://cordis.europa.eu/project/rcn/106380_en.html (accessed on 15 March 2018).
2. Aktan, A.E.; Farhey, D.N.; Helmicki, A.J.; Brown, D.L.; Hunt, V.J.; Lee, K.L.; Levi, A. Structural Identification for Condition Assessment: Experimental Art. *ASCE J. Struct. Eng.* **1997**, *123*, 1674–1684, doi:10.1061/(ASCE)0733-9445(1997)123:12(1674).
3. Lau, K.-T. Fibre-optic sensors and smart composites for concrete applications. *Mag. Concr. Res.* **2003**, *55*, 19–34, doi:10.1680/macr.2003.55.1.19.
4. Chan, B.; Guan, H.; Jo, J.; Blumenstein, M. Towards UAV-based bridge inspection systems: A review and an application perspective. *Struct. Monit. Maint.* **2015**, *2*, 283–300, doi:10.12989/smm.2015.2.3.283.
5. Jong-Woo, K.; Sung-Bae, K.; Jeong-Cheon, P.; Jin-Won, N. Development of Crack Detection System with Unmanned Aerial Vehicles and Digital Image Processing. In Proceedings of the World Congress on Advances in Strcutural Engineering and Mechanics (ASEM15), Incheon, Korea, 25–29 August 2015; pp. 1–11.
6. Aldea, E.; Hégarat-Mascle, S.L. Robust crack detection for unmanned aerial vehicles inspection in an a-contrario decision framework. *J. Electr. Imaging* **2015**, *24*, 1–16, doi:10.1117/1.JEI.24.6.061119.
7. Wang, K.C.P. Basis and Design of a Deep Learning System for Automated Cracking Survey. In European Road Profile Users' Group (ERPUG2017), Copenhagen, Denmark, 19–20 October 2017.
8. Ciregan, D.; Meier, U.; Schmidhuber, J. Multi-column deep neural networks for image classification. In Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition, Providence, RI, USA, 16–21 June 2012; doi:10.1109/CVPR.2012.6248110.
9. Fan, Z.; Wu, Y.; Lu, J.; Li, W. Automatic Pavement Crack Detection Based on Structured Prediction with the Convolutional Neural Network. 2018. Available online: https://arxiv.org/pdf/1802.02208.pdf (accessed on 15 March 2018).
10. Simonyan, K.; Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition, 2014. Available online: https://arxiv.org/abs/1409.1556v6 (accessed on 15 March 2018).