Using Image Processing to Classify Roofs based on Damage Level.

K. Mostafa and T. Hegazy

Department of Civil and Environmental Engineering, University of Waterloo, Canada
E-mail: ktmostafa@uwaterloo.ca, tarek@uwaterloo.ca

Abstract:

Roofing systems are considered one of the items that in most need of frequent inspection and rehabilitation due to its ongoing exposure to the elements. Manual roof inspections are time consuming and subjective. This study uses Convolutional Neural Network (CNN), an image-processing technique, to classify roofs according to their damage level. The proposed model analyzes images showing general views of roofs to determine (on a macro level) whether the roof has sustained no (or low), moderate, or severe damage. Based on this analysis, more detailed roofing inspection can be conducted if needed. The study was applied on more than 200 images of roofs of the University of Waterloo campus, collected using a drone. Different CNN architectures were examined where the number of convolution kernels (i.e., the depth of the CNN layer) has been the main variable. This experiment has revealed that complicating the model by changing the depth of the layer causes the model to overfit with no performance improvements on the validation dataset. The proposed model, using only 5 kernels in the first convolution layer, has achieved 90% accuracy level. The developed model serves as the initial step of a larger roofing inspection framework, reducing the need for more time and resource intensive assessments. The end goal of the proposed inspection framework is to provide a fast, objective, and reliable roofing inspection and assessment, helping asset managers of large portfolios better assign the allocated rehabilitation funds.

Keywords:

Facility Management, Capital Renewal, Rehabilitation, Inspection, Roofing, Image-based Analysis, Convolutional Neural Network (CNN)

1 Introduction

Roofs are frequently exposed to extreme temperatures, rain, snow, and other severe weather events. As such, roofs are considered to be among the most vulnerable building components [1, 2]. While the average life expectancy of a roof is a function of its type and material, having an adequate maintenance program is essential to ensure the roof would live up to the expectations and maintain an adequate level of service throughout.

Currently, most inspections are done manually. Manual inspections suffer from a plethora of shortcomings. First of all, manual data collection and processing is a time-consuming process [3]. Kamarah [4] estimated that an average inspection site visit takes about 3 hours, and Abou shaar [5] estimates that for every hour spent on site for data collection, three more hours are spent in analysis. Second, manual inspections are subjective as they rely on the inspector's training and experience [6]. This often means that two inspectors can produce two different reports for the same asset [6]. Third, roofs are sometimes difficult to access and, if accessed, often pose safety risks to inspectors [7]. This puts extra overhead on the inspection company in terms of special training and/or insurance. Hence, automation of inspection is a must.

Computer vision (i.e., image-based analysis) allows computers to extract and analyze information directly from pictorial data such as images and videos [8]. Hence, and because images are easy to collect (e.g., using a cell phone camera), computer vision has been under investigation as a way to automate supervision and inspection tasks in the construction domain. For example, [9] developed a computer-vision-based pavement crack inspection model capable of achieving 92% accuracy using a commercial grade GoPro as its data collection source. [10] developed a convolutional neural network (CNN) based classification model capable of detecting concrete cracks as well as distinguish them from handwritten markings on the concrete surface. As a step towards a more comprehensive assessment, [11] was capable of detecting multiple defects such as cracks, concrete delamination and rebar exposure. Other Examples of computer vision applications towards automating asset inspections include defect detection in concrete structures [12], pavements [13], and sewers and pipelines [14, 15]. Unfortunately, however, little to no work was done on roofing condition assessment despite roofs being one of the most vulnerable building components. As such, this paper aims to fill this research gap by utilizing CNN technology towards roofing condition assessment.

As a step toward automating roofing inspections, this paper proposes a computer vision approach that utilizes a Convolutional Neural Network (CNN) for preliminary classification of roof condition. The proposed system relies on images that show the roof general view to categorize the roof into one of three categories; Clear/Minor, Moderate, or Severely Damaged. Based on the classification, more detailed inspection can be conducted as needed. This aims to reduce the time and resources spent on the inspection process because, using this method, not all the roofs will need to undergo the same detailed inspections. The remaining part of the paper explains the model development and testing process. starting with the motivation, followed by the data collection process and the proposed CNN architecture, and concluding with analysis results and discussion.

2 Existing Roof Management Systems

While buildings, in general, are considered to be durable and expected to last for decades, a key for that durability is continuous and effective maintenance and repair programs. This is because periodic maintenance and rehabilitation increases the building's lifespan and improves the level of service of its components [16]. As such, the last few decades have witnessed the development of a variety of building management and inspection systems that cater to a variety of needs. Examples include de Brito et al. [17] concrete bridge management system, Curt et al. [18] dam safety assessment, and Bortolini and Forcada [19] and Faqih and Zayed [20] building inspection systems. However, these systems still rely on manual inspections as the main source of data.

As seen in Fig. 1, most roofing systems used in non-residential buildings such as schools and hospitals are considered to be low-slope roofs. Among those roofs, the built-up roof (BUR) system is the most common type of roofing in [21]. BURs are generally composed of alternating layer of bitumen and reinforcing fabric that create a finished membrane (often referred to as "roofing felt") which is finally covered by gravel to reduce its exposure to the weather elements.

Originally developed by the US army, ROOFER [22] is one of the most common roofing-based asset management software. Despite its capabilities, it still relies on manual inspections as its sole source of data. To reduce the subjectivity of manual inspections, [21] developed a pictorial database of common roof defects, ranked by their severity. Other examples of roofing rehabilitation systems include the works of [1] and [23], who defined the most common roofing defects, along with possible causes and remedies. However, these models, like their predecessors, still rely on manual data input. As such, computer vision technology provides the opportunity for automation of roof inspections, leading to cost, time, and subjectivity reductions.

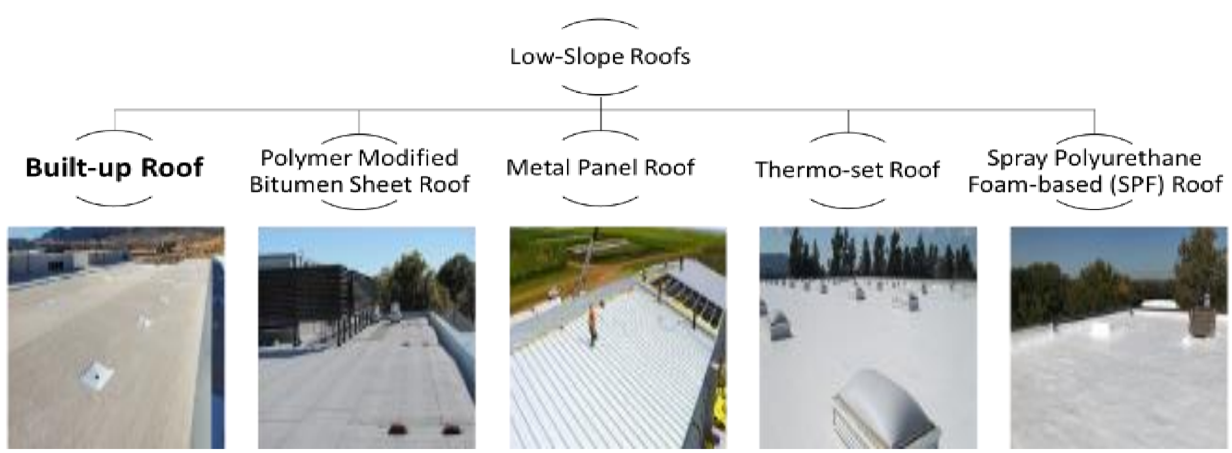

Fig. 1: Types of Low-Slope Roofs

3 Data Collection

University of Waterloo is one of the largest universities in Canada, having a total of 80 buildings. The pictorial dataset used in this study comes from 20 buildings that are highlighted in Fig. 2. Two different data collection methods were used. First, the authors were allowed to physically access the building roofs and take photographs using their personal cameras (16 Megapixel phone camera). The second method was only used for a group of five buildings (highlighted by the red star in Fig. 2) as the roof was inaccessible. In that case, the images were collected by a commercial drone. The drone was set to record a video of its flight, then one of each 10 frames were extracted. The images collected were subject to the weather conditions (e.g., lighting) prevalent at the time of the data collection, which took place over the span of multiple days. The camera angle differed depending on the method of collection. Images collected by the commercial drone had, on average, a 45-degree angle relative to the roof surface (subject to wind conditions and extra tilting of the drone for navigation purposes), while images collected manually had a 60-degree angle. While the camera angle was kept constant relative to the surface being captured, the images collected were subject to the weather conditions prevalent at the time of the data collection, which took place over the span of multiple days. As such, the total number of images used in this study is 346 representing three damage categories: low, moderate, and severe. Examples of the images representing the three different categories are in Fig. 3. One third of the data set (114 images) was used for validation and testing purposes, while the remaining two-thirds were used for training.

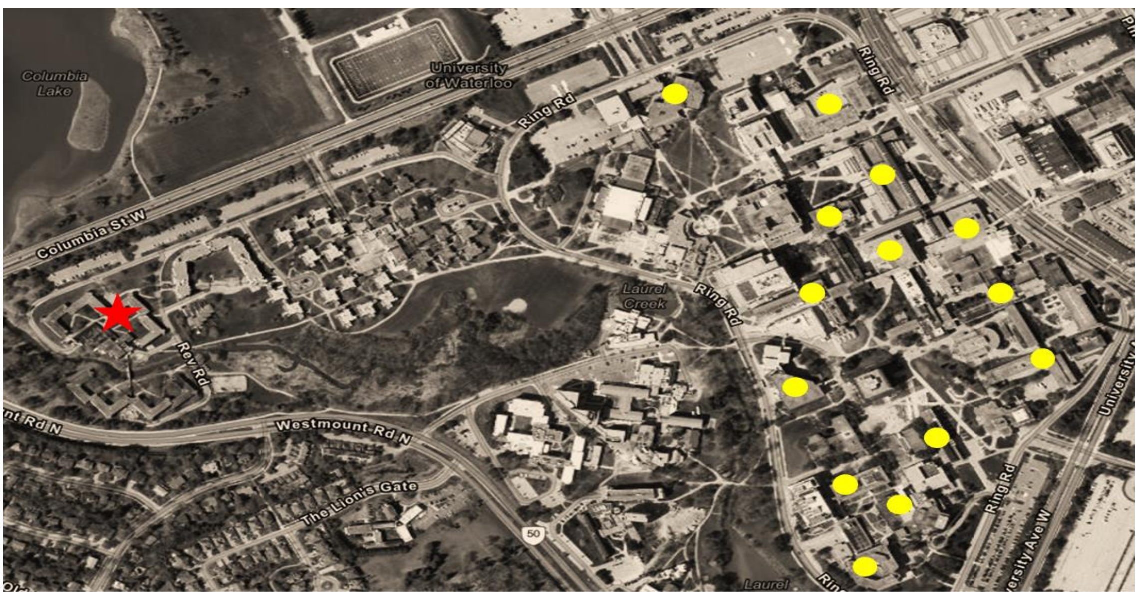

Fig. 2: Image Locations - University of Waterloo (Red Star-Images Collected by a drone)

Fig. 3: Examples of the Collected Images Representing the Three Different Damage Categories

4 Proposed CNN architecture

The composition of CNN elements is an important task to achieve maximum performance level. While increasing the complexity of the CNN by adding more layers or convolution kernels can improve performance, over complicating the model might lead to performance degradation and overfitting. Inspired by AlexNet [24] and LeNet [25], three CNN architectures are tested, all followed the architecture demonstrated in Fig. 4. The only difference was in the number of channels in each layer (i.e., the value of **N**). **N** took three different values; 5, 6, and 8. For notational convenience, the three different CNNs that were experimented will be referred to as CNN_**N**. For all CNNs, RELU was used as the activation function for the fully connected layer. Other studies reported that changing the kernel size has a minimal effect on accuracy [26]. Hence, the size of the kernel was fixed at 10x10 with zero padding and a stride of two.

5 Implementation Details and Results

All experiments were performed using the python programming language (CUDA 7.5) on a laptop with Core i7-10750H@2.6GHz CPU, 16GB RAM, and 4GB NVIDIA GeForce GTX 1650 Graphical Processing Unit (GPU). Learning rate was set to be 0.0002 while the batch size was set to 22 images. All experiments were run for 500 epochs. Accuracy (i.e., number of correct classifications) was used as the evaluation metric.

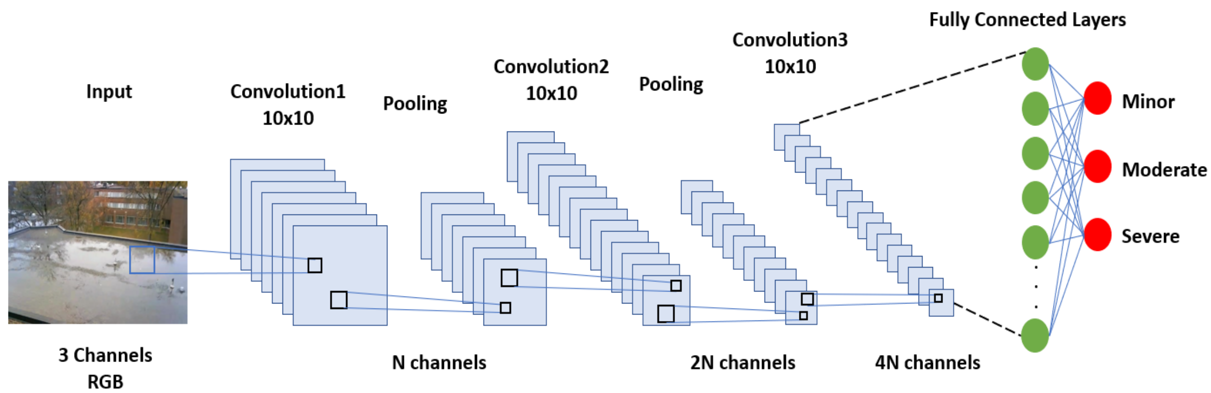

Fig. 4: Proposed CNN Architecture

5.1 Comparing the different models

Table 1 shows the performance of the different CNN models in terms of correct classifications for the training and testing datasets. Since the was no function to terminate the training process when convergence or performance degradation takes place, the epoch where the optimal performance is reached is reported in Table 1 as an indication for the speed of the training process. It is important to note that all three models took approximately the same time per epoch (approx. 18 seconds).

Table 1: Performance Comparison of different models

	CNN_5	CNN_6	CNN_8
Training Accuracy (%)	91.8	91.4	96.5
Testing Accuracy (%)	89.5	86.8	88.6
Epoch	469	326	331

Table 1 shows that CNN_5 took the longest time (i.e., highest number of iterations) to reach an optimal performance, followed by CNN_8 and CNN_6. However, the performance of CNN_5 is considered to be the best of the three for two reasons. First, CNN_5 is reporting the highest testing accuracy, which speaks to its generalization ability. Second, comparing the training and testing accuracy reveals that CNN_5 has the smallest variation (2.3%). Unlike CNN_8 which, despite having the best performance on the training dataset, has the highest difference between the training and testing accuracies (8%) which raises some overfitting concerns. From the data shown in Table 1, it can be deduced that, for this particular application, creating a simpler model by using less number of channels yields models with higher generalization abilities and more resilience against overfitting.

The number of channels refers to the number of convolutional kernels used in each layer. Each of these convolutional kernels work independently and aim to capture significant features of the image to help reach a correct classification. As such, changing the number of channels correspond to changing the number of significant features captured by the CNN to guide its classification activity. The data in Table 1 suggests that the distinctive features that differentiate between images showing different degrees of roof damage are not many, and therefore require a relatively simple model to avoid "going deep in the weeds" and losing generalization ability. This is supported by the sample images shown in Fig. 3, where the differences between the damage levels are noticeable even to non-inspection experts. As such, macro-level roof inspection (i.e., holistically analyze roofing conditions from images that capture the entire roof surface) can be quickly and effectively used to triage the different roofs before investing in relatively more resource-intensive analyses for micro-level inspections.

5.2 Analyzing the Optimal Model (CNN_5)

After analyzing the three different models as shown in Table 1, CNN_5 was selected as the "optimal model" and the results were analyzed in more detail. Table 2 shows the confusion matrix of the model's result on a 110-image dataset collected by the authors. Two observations can be drawn from Table 2. First, the overall accuracy of the model is 90%, which is acceptable compared to other models in the literature. Second, the model is not exhibiting any major biases, as the chances of "under classifying" and "over classifying" a roof are equally likely to happen.

Table 2: Confusion Matrix for CNN_5 predictions

	True Label			
Predicted Label	No/Minor	Moderate	Severe	Total
No/Minor Damage	**40**	2	2	44
Moderate Damage	4	**8**	2	14
Severe Damage	1	--	**51**	52
Total	45	10	55	110

The performance of the proposed model (CNN_5) is compared with models from the literature in Table 3. Due to the absence of models that address roofing condition assessments, the model was compared with other defect detection models that are intended for other purposes such as pavements and concrete elements. Based on Table 4, the performance of the proposed model is on par with the state of the art. However, this model has the advantage of the lower data collection burden. While all models in the literature perform micro level assessments and thus require scanning a relatively small are of the structure to detect the defects. The proposed model only looks at a few snapshots that show the entire roof. This means that the analysis can (and should) take place frequently to produce updated assessments.

Table 3: Comparison with Models in the Literature

Model	Purpose	Accuracy
[9]	Pavement Crack Detection	92%
[27]	Building Defect Detection	89.1%
[28]	Concrete Defect Detection	89.7%
Proposed	**Roofing Defect Detection**	**90%**

6 Conclusion

In this study, a CNN-based model is proposed for holistic roofing condition assessment. The proposed model relies on a few snapshots showing the general view of the roof and classifies the roof accordingly into one of three categories (No damage, Moderate, Severe). The proposed model is part of a three-step roofing inspection and rehabilitation framework that aims to aid asset management practitioners with prioritizing assessment and rehabilitation events under limited budgetary constraints. The proposed model addresses the first step which is prioritizing the inspection efforts. Using only a few snapshots, the model can triage the roofs according to their condition and suggest more rigorous inspections for the ones that deemed worthy, thus saving on the total time spent on inspection site visits and their accompanied office work (four hours per building according to [4] and [5]). First, three different CNN architectures were examined, and one was selected as the better performer due to its higher accuracy on the testing dataset as well as its lower susceptibility for overfitting. While the model addresses an important research gap which is the lack of image-based roof inspection frameworks, its performance is on par with other models in the literature that are intended for other purposes such as concrete or pavement defect detection. All without the need for detailed images like the other models in comparison.

Regarding the proposed model, future work includes automating the data collection process by maximizing the drone usage as well as, if possible, testing the validity of live satellite images. The proposed model is the first of a three-step assessment and prioritization framework (Fig. 5). First, the proposed model is used to holistically assess the roofs and determine their need for more detailed assessment. More detailed assessment can then be conducted for roofs that are deemed worthy, which leads to a more granular assessment for the roof condition highlighting the different damage types and sizes [29]. Using cost and crew information (e.g., RS means) work packaging estimates for the required rehabilitation can be developed. Finally, this information is aggregated with other text-based data mining information (e.g., building age, description) to optimize the use of the limited rehabilitation budget.

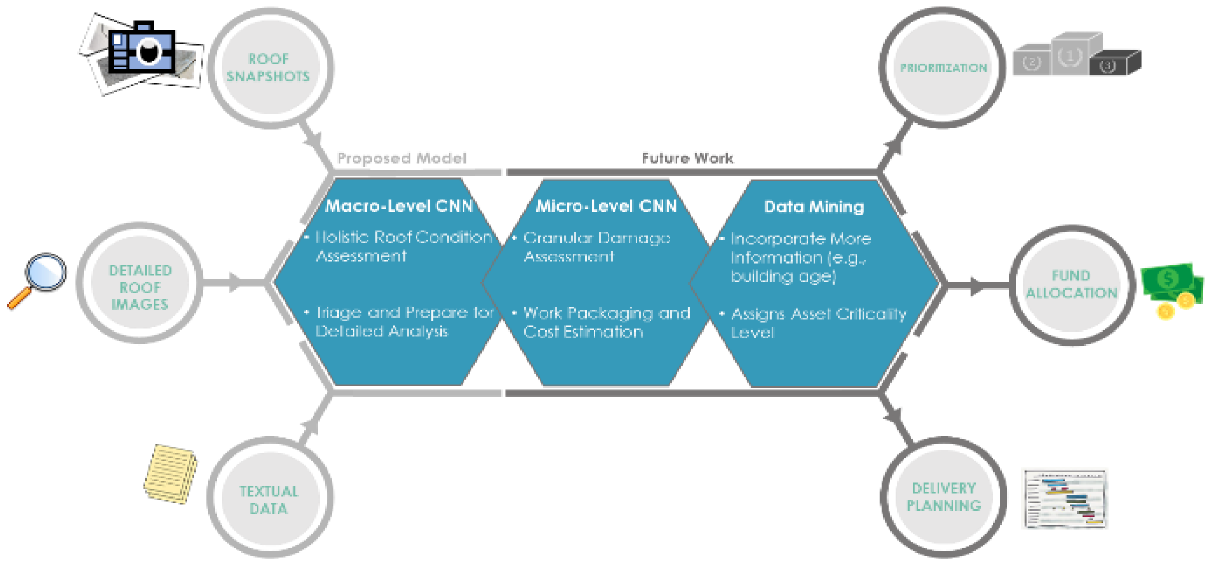

Fig. 5: Proposed Roofing Assessment Framework

References

[1] Morgado, J., Flores-Colen, I., Brito, J. D., and Silva, A. (2017). "Maintenance Planning of Pitched Roofs in Current Buildings." *Journal of Construction Engineering and Management*, 143(7), 05017010.

[2] Garcez, N., Lopes, N., Brito, J. D., and Silvestre, J. (2012). "System of inspection, diagnosis and repair of external claddings of pitched roofs." *Construction and Building Materials*, 35, 1034–1044.

[3] Radopoulou, S. C., and Brilakis, I. (2017). "Automated Detection of Multiple Pavement Defects." *Journal of Computing in Civil Engineering*, 31(2), 04016057.

[4] Kamarah, E. (2019). *Framework for Scheduling, Controlling, and Delivery Planning for Scattered Repetitive Infrastructure Rehabilitation Projects*. Dissertation. Department of Civil and Environmental Engineering, University of Waterloo.

[5] Abou Shaar, B. (2012). *Adaptable Three-Dimensional System for Building Inspection Management*. Dissertation. Department of Civil and Environmental Engineering, University of Waterloo.

[6] Hoang, N.-D., Nguyen, Q.-L., and Bui, D. T. (2018). "Image Processing–Based Classification of Asphalt Pavement Cracks Using Support Vector Machine Optimized by Artificial Bee Colony." *Journal of Computing in Civil Engineering*, 32(5), 04018037.

[7] Mimura, T., and Mita, A. (2017). "Automatic estimation of natural frequencies and damping ratios of building structures." *Procedia Engineering*. 188, 163–169.

[8] Mostafa, K., and Hegazy, T. (2021). "Review of image-based analysis and applications in construction." *Automation in Construction*, 122, 103516.

[9] Mei, Q., and Gül, M. (2020). "A cost effective solution for pavement crack inspection using cameras and deep neural networks." *Construction and Building Materials*, 256, 119397.

[10] Deng, J., Lu, Y., and Lee, V. C. S. (2019). "Concrete crack detection with handwriting script interferences using faster region-based convolutional neural network." *Computer-Aided Civil and Infrastructure Engineering*, 35(4), 373–388.

[11] Rubio, J. J., Kashiwa, T., Laiteerapong, T., Deng, W., Nagai, K., Escalera, S., Nakayamad, K., Matsuo, Y., and Prendinger, H. (2019). "Multi-class structural damage segmentation using fully convolutional networks". *Computers in Industry*. 112, 103121.
[12] Kumar, B., and Ghosh, S. (2020). "Detection of Concrete Cracks Using Dual-channel Deep Convolutional Network." *2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*.
[13] Hou, Y., Li, Q., Han, Q., Peng, B., Wang, L., Gu, X., and Wang, D. (2021). "MobileCrack: Object Classification in Asphalt Pavements Using an Adaptive Lightweight Deep Learning." *Journal of Transportation Engineering, Part B: Pavements*, 147(1), 04020092.
[14] Yin, X., Chen, Y., Bouferguene, A., Zaman, H., Al-Hussein, M., and Kurach, L. (2020). "A deep learning-based framework for an automated defect detection system for sewer pipes." *Automation in Construction*, 109, 102967.
[15] Wang, M., Luo, H., and Cheng, J. C. P. (2021). "Towards an automated condition assessment framework of underground sewer pipes based on closed-circuit television (CCTV) images." *Tunnelling and Underground Space Technology*, 110, 103840.
[16] Amaral, S., and Henriques, D. (2013). "Inspection and diagnosis: A contribution to modern buildings sustainability." Proc., Portugal SB13, Univ. do Minho, Guimarães, Portugal, 75–82.
[17] de Brito, J., Branco, F. A., Thoft-Christensen, P., and Sørensen, J. D. (1997). "An expert system for concrete bridge management." *Engineering Structures*, 19(7), 519–526.
[18] Curt, C., Talon, A., and Mauris, G. (2011). "A dam assessment support system based on physical measurements, sensory evaluations and expert judgements." *Measurement*, 44(1), 192–201.
[19] Bortolini, R., and Forcada, N. (2018). "Building inspection system for evaluating the technical performance of existing buildings." *Journal of Performance of Constructed Facilities*, 32(5), 04018073.
[20] Faqih, F., and Zayed, T. (2021). "Defect-based building condition assessment." *Building and Environment*, 191, 107575.
[21] Ahluwalia, S. S., and Hegazy, T., (2010). "Roof deterioration and impact: A questionnaire survey." *Journal of Retail and Leisure Property*, 9, 337–348.
[22] https://digonsystems.com/roofer/
[23] Conceição, J., Poça, B., de Brito, J., and Flores-Cohen, I. (2019). "Inspection, Diagnosis, and Rehabilitation System for Flat Roofs." *Journal of Performance of Constructed Facilities*, 31(6), 04017100.
[24] Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). "ImageNet classification with deep convolutional neural networks". *International conference on neural information processing systems*, 1097–1105.
[25] Lecun, Y., Bottou, L., Bengio, Y., and Haffner, P. (1998). "Gradient-based learning applied to document recognition". *Proceedings of the IEEE*, 86 (11), 2278–2324.
[26] Wu, H., & Gu, X. (2015). "Towards dropout training for convolutional neural networks." Neural Networks, 71, 1–10.
[27] Perez, H., Tah, J. H., and Mosavi, A. (2019). "Deep Learning for Detecting Building Defects Using Convolutional Neural Networks." *Sensors*, 19(16), 3556.
[28] Cha, Y-J., Choi, W., Suh, G., and Mahmoudkhani, S. (2018). "Autonomous Structural Visual Inspection Using Region-Based Deep Learning for Detecting Multiple Damage Types." *Computer-Aided Civil and Infrastructure Engineering*, 33, 731-747.
[29] Mostafa, K., Hegazy, T., Hunsperger, R., and Elias, S. (2021) "Image Analysis Using a Two-Phase Convolutional Neural Network to Identify and Classify Defects in Roofing Elements." *Automation in Construction*, Under Review