Recent Trends in Civil Engineering and Built Environment Vol. 4 No. 3 (2023) 499-505

RTCEBE

Homepage: http://publisher.uthm.edu.my/periodicals/index.php/rtcebe
e-ISSN :2773-5184

Development of Computational Algorithm for Concrete Leakage Detection Using Deep Convolutional Network

Dyeanna Donnis[1]**, Nickholas Anting Anak Guntor**[2]*
Mariana Dina Anak Malong[1]

[1]Faculty of Civil Engineering and Built Environment,
Universiti Tun Hussein Onn Malaysia, Batu Pahat, 86400, MALAYSIA

[2]Big Data and Advance Analytics Research
Faculty of Civil Engineering and Built Environment, Universiti Tun Hussein Onn Malaysia, Batu Pahat, 86400, MALAYSIA

*Senior Lecturer, Faculty of Civil Engineering and Built Environment,
Universiti Tun Hussein Onn Malaysia

DOI: https://doi.org/10.30880/rtcebe.2023.04.03.053
Received 06 January 2022; Accepted 15 May 2023; Available online 31 December 2023

Abstract: Traditional methods for detecting leaks in concrete are very expensive, time-consuming, and require human intervention. Since leakage inside a concrete building cannot be seen, another method for detecting leakage is to use a thermal camera. In this paper, an approach was made to develop a data processing python algorithm using a convolutional neural network (CNN) for predicting the presence of leakage based on 60 thermal images that had been collected from published literature and internet resources. The model performance was evaluated using metrics measured based on the accuracy of the trained model. Python is used as the main programming language to develop the algorithm. Furthermore, four hidden layers will transform the input data to its output which go through the sub-functions from input, convolutional layer, pooling layer, and then output. The testing data makes up to 30% of the datasets, while the training data makes up 70% of the datasets. The accuracy of the system was measured for the detection of the leak and no leak area by increasing the epochs value. The results show that the accuracy of the model is not significantly improved when the number of epochs increases. However, when the number of hidden layer increase there is a slightly different in the accuracy of the model. Herein, the model train to classify leak and no leak images have been proven successful. This is proved from the accuracy results, which showed that 95%. The findings indicated that based on the results obtained, CNN can classify the collected images based on the accuracy of the trained model. The findings imply that increasing the number of images can increase data accuracy, emphasizing the specificity and sensitivity of the output results.

Keywords: Thermal image, CNN, Python, Accuracy

Corresponding author: nickholas@uthm.edu.my

1. Introduction

Building leakage is most common at the main building element. Thermal imaging is a useful tool nowadays because it is one way of determining the structural health of a structure [1]. Leakage inside a building is generally caused by a few possible factors. Water leakage will occur as a result of all of the factors, causing dampness in the building [2]. There are many proposed procedures and techniques to detect leakage in building, but many of them are very expensive, time-consuming and require human involvement. In the traditional leak detection method, maintenance personnel need to inspect periodically. This periodical inspection does not provide real-time monitoring. When the leak is not detected at the time of the incident, environmental pollution and economic loss will happen. To upgrade into technology system, ultrasonic water leakage detection had been widely used. The risk of noise interference is one of the most serious issues with ultrasonic leak detectors.

Since leakage inside a concrete building's cannot be seen, the thermal camera will be used to capture an image where leakage happens inside the wall. The thermal camera would then produce images of infrared or "heat" radiation that is invisible. Machine learning is thus needed to detect leakage area at early stage which can recognize pattern. The field of machine learning known as CNN is based on learning representation levels. Monitoring a building's structural health is useful to minimize the building's life cycle cost from construction to maintenance [3]. Thus, thermal images that are not evaluated using a machine learning model will not be able to spot future leaks in an area.

2. Literature Review

Machine learning, specifically CNN, has become increasingly popular in engineering applications in recent years [4]. In the computational materials science field, the number of CNN applications is rapidly growing. As a result of the increased availability of data, machine learning algorithms may now be trained on a vast number of instances. The system's analytical skills have been boosted by increasing computer processing capability. Within the discipline, there have also been algorithmic advancements that have given machine learning additional strength. CNN model architecture of neural network is extremely effective at dealing with image data. 60 images will be collected to construct a prediction model using CNN model, which will allow the genuine leakage to be easily classified based on temperature differences.

2.1 Innovation in Thermal Images

The primary commercial thermal imaging camera was sold in 1965 for tall voltage control line assessments. Since at that point, the utility of thermal imaging cameras for mechanical applications has ended up a significant showcase portion for FLIR. The thermal imaging technology has extensively developed seeing that then, and thermal imaging cameras have developed to come to be compact in dimension and seem to be like a digital photograph camera. Thermal imaging cameras are used to determine the renovation requirements for electric and mechanical installations as they have a tendency to generate unusual warmness earlier than they fail.

2.2 Machine Learning Approach

Machine learning is split into four categories of learning based on the learning style: supervised learning, unsupervised learning, semi-supervised learning, and reinforcement learning. The utmost common technique is supervised learning, in which the trained machine model's sample data has associated target values [5]. A dataset must be gathered before machine learning approaches can be used. The dataset will then go through many stages, including pre-processing, data training, and application of a learning model, as well as an evaluation phase.

3. Research Methodology

There are several methods applied to get the required information. It refers to the most suitable and effective methods to carry on with the study. This study emphasizes of data collection, data preparation, data training and testing and model evaluation.

3.1 Research Design

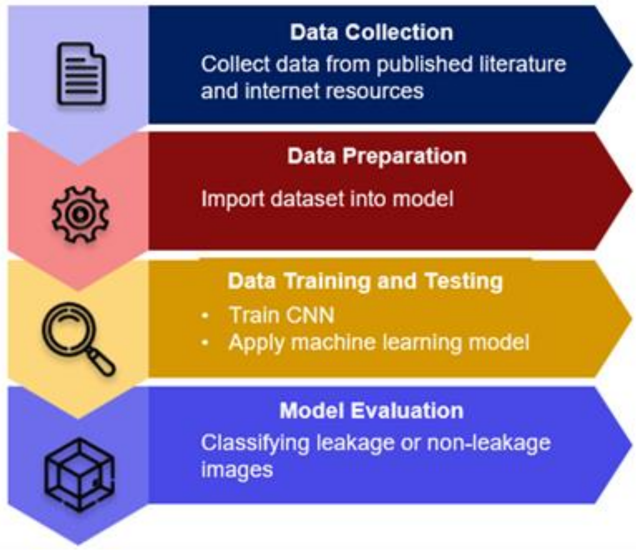

Figure 1: Methodology Framework

3.2 Data Collection

A dataset of roughly 60 thermal pictures of leak and no leak in building structures was acquired through internet resources and a prior study for this investigation. The accuracy and precision of the method used for the experiment are examined in this phase to ensure that the data is of high quality.

3.3 Data Preparation

The step of data preparation is the phase to reconsider the data selection criteria again. At this point, the data is cleaned by deleting any missing value data that could cause redundancy in the dataset. Additionally, visualizing data will aid in the discovery of specific correlations between variables. Clean and well-curated data is generated through good data preparation, which results in more realistic and accurate model outcomes.

3.4 Data Training

Data training or testing is a means of determining the model's accuracy. Because data must be spitted into two sets, a training set and a testing set, this procedure is called train or test. The training set should consist of a random selection of 80% of the original data, whereas the testing set should consist of 20% of the data.

3.5 Model Evaluation

Model evaluation metrics are needed to measure the model performance. The evaluation metrics used are determined by the machine learning process such as classification, regression, ranking, clustering, topic modeling, and so on. The model will be evaluated in this study by classifying the image for each dataset obtained from data training. Leakage classification is defined as a method for identifying specific leakage types using machine learning algorithms. This leakage is divided into leakage non-leakage images categories. At this stage, the result of accuracy and precision are summarized and the method with the highest accuracy is determined.

3.6 Python Programming Language

Python is an interpreted language with a declarative syntax that has been likened to executable pseudocode by some [6]. Python may be just a language for user but it is a real programming language which offer much more structure and support for large programs than the shell has. On the other hand, Python also offers much more error checking than C. Being a very high-level language, Phyton has high-level data types built-in, such as flexible arrays and dictionaries that would cost user days to implement efficiently in C. Because of its more general data types, Python applies to a much larger problem.

4. Results and Discussion

This chapter will discuss the details about the research design and implementation of the research. This chapter also presents the experimental evaluation of the purposed algorithm which is the classifying leak and no leak image. It also will explain the detail of the experimental setup and will be showing the results of the two classification methods. The project was evaluated with the CNN algorithm. The performance between the two classifiers had been measured using loss and accuracy performance. The analysis results of the two classifiers are discussed in this chapter.

4.1 Implementation

The implementation of the project begins with combining all of the datasets and software code into a single file. Before the augmented imaging process can be loaded, the required libraries must be imported, and the dataset must be trained and modelled using training data. About 60 thermal photos, 30 leak and 30 non-leak images, were acquired via literature review and internet resources and taken using a thermal imaging camera in this investigation.

This model is made with a 2D Convolutional layer (Conv2D), which generates a convolutional kernel that is convolved with the input layer to get a tensor of outputs. The model is programmed to train for two to four layers. In order to complete the model, a 2 x 2 filter with a stride of 2 will be employed. The ReLu activation function in Keras transforms the node's summed weighted input into the node's activation or output for that input. The epochs value is used to train the data accuracy pattern. In this investigation, the epoch values employed were 10, 20, 20, 40, and 50. The testing dataset and the training model are used in the model assessment output to estimate the loss and accuracy of the datasets. The metrics are used to assess the trained model's performance. The model's final result is a classification of the dataset into the leak and no leak sets.

4.2 Analysis of Datasets based on Accuracy

The performance of various epoch values on the dataset is shown in Table 1. The epoch values in this study are 10, 20, 30, 40, and 50.

Table 1: Loss, accuracy and validation accuracy percentage for different epochs value

Epochs	Loss (%)	Accuracy (%)	Validation Accuracy (%)
10	0.070	98.33	98.33
20	0.090	98.33	98.33
30	0.090	98.33	98.33
40	0.050	98.33	98.33
50	0.006	100.00	100.00

(a) Epochs 10

Figure 2 depicts the data accuracy throughout ten epochs. The greatest and minimum data accuracy values are 97.06 percent and 58.82 percent, respectively. In this study, data accuracy is used to identify whether a dataset leaks or not when obtained with a thermal camera. The accuracy of the data reveals that the temperature of leak and no leak photos from thermal imaging differs dramatically.

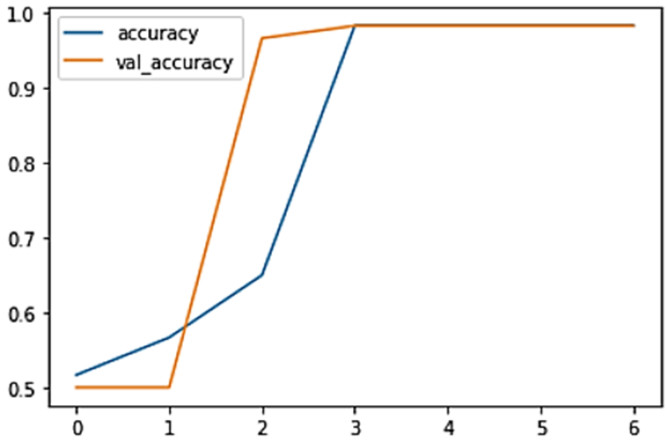

Figure 2: Accuracy graph for 10 epochs

(b) Epochs 20

The accuracy value for epochs 20 in Figure 3 is the same as the accuracy value for epochs 10. The data is trained more than epochs 10 at this point. The datasets are 97.06 percent correct, as evidenced by the accuracy of the data. Because the datasets were trained more in epochs 10 than in previous epochs, all of the trends had the same accuracy percentage value, but the trend was consistent.

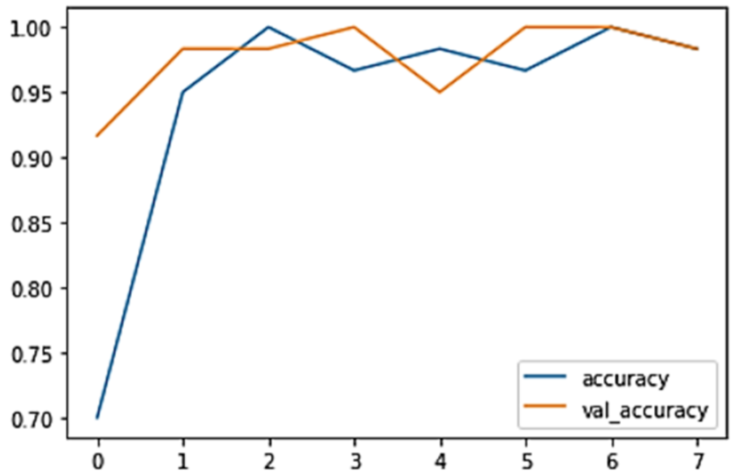

Figure 3: Accuracy graph for 20 epochs

(c) Epochs 30

The graph is shown in Figure 4 when epochs 30 are used. The graph demonstrates that the proportion increases to 100% in the first step of image training. The epochs training has an impact on the highest accuracy %. As the datasets became more accurate, this percentage increased. Because the epochs values are 10 and 20, respectively, the accuracy for Figure 1 and 2 is less than 100%. This is also the rationale for demonstrating that the datasets are accurate enough to categories photos as having a leak or not having a leak.

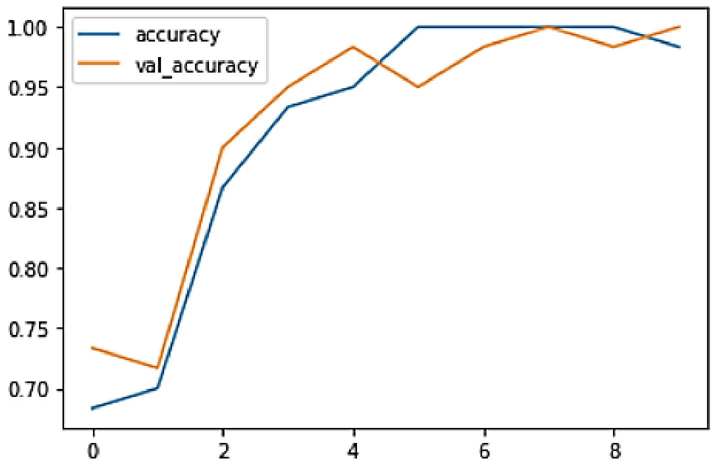

Figure 4: Accuracy graph for 30 epochs

(c) Epochs 40

The curve in Figure 4 demonstrates that accuracy is improving, with the same accuracy % as in Figure 3. The minimum percentage accuracy is reduced to 0.003% in this graph, making the data more accurate to see.

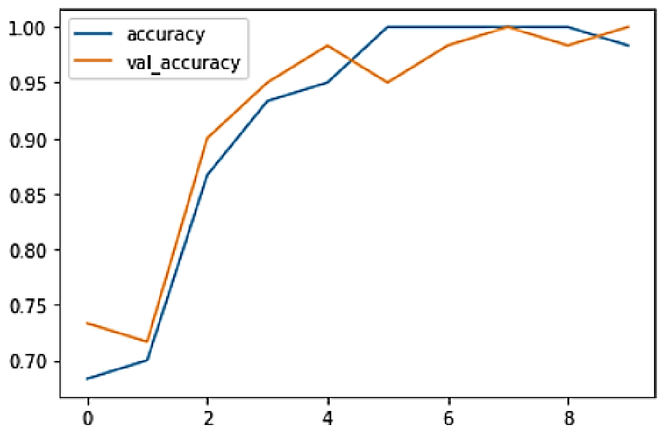

Figure 5: Accuracy graph for 40 epochs

(c) Epochs 50

At the end of the 50th epochs, the data accuracy is 100% and the validation accuracy was also 100%. The evaluation test revealed consistent overall accuracy for all the epochs value with 30 leak and 30 no leak images which than can be classified correctly.

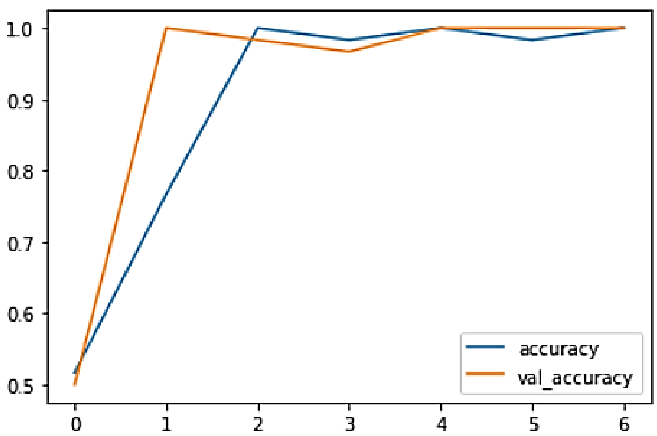

Figure 5: Accuracy graph for 50 epochs

The accuracy of the datasets was demonstrated by all the five graphs, which showed that all datasets that were trained using the CNN model were accurate, with consistent accuracy percentage of 98.33% at 10 to 40 epochs and 100% accuracy at the 50 epochs.

5. Conclusion

In line with these findings, the objectives of the study were achieved. The dataset which contains 60 images was collected from literature study and google search are listed in Chapter 3. The dataset consists of 30 leak images and 30 no leak images. A machine learning model was developed using CNN algorithm to classify leak and no leak images. The process of data collection, data preparation, data training and testing. Model training and model performance evaluation was implemented in this study.

To perform determine the metrics performance of the model, the accuracy of the testing data performance was conducted by making the epoch value as the variable in this stage. The number of epochs will control the number of complete passes through the training dataset. The epoch used in this study is 10, 20, 30, 40 and 50. The machine learning model is then performed by determine the dataset to be classified in leak or no leak classes.

Acknowledgement

The writer would like to thank her supervisor, family and friends for assistant and support in the accomplishment and completion of research. The authors would also like to thank the Faculty of Civil Engineering and Built Environment, Universiti Tun Hussein Onn Malaysia for its support.

References

[1] Moon, H. G., & Kim, J. H. (2011). Intelligent crack detecting algorithm on the concrete crack image using neural network. Proceedings of the 28th International Symposium on Automation and Robotics in Construction, ISARC 2011, 1461–1467.

[2] Polat, K., Guenes, S., 2009. A novel hybrid intelligent method based on C4.5 decision tree classifier and one-against-all approach for multi-class classification problems. Expert Syst. Appl. 36, 1587–1592.

[3] Gulshan V, Peng L, Coram M et al (2016) Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA 316:2402–2410

[4] Schmidhuber, J. (2015). Deep Learning in neural networks: An overview. Neural Networks, 61, 85–117. https://doi.org/10.1016/j.neunet.2014.09.003

[5] Lu, Z., 2010. The elements of statistical learning: data mining, inference, and prediction, 2nd edition. J R Stat. Soc. A Stat. 173, 693–694.

[6] Faust, A., Palunko, I., Cruz, P., Fierro, R., Tapia, L., 2017. Automated aerial suspended cargo delivery through reinforcement learning. Artif. Intell. 247, 381–398.