Detecting Building Defects Using VGG16 & IBM Watson

CONTENTS

1.ABSTRACT

A growing number of clients require fast and effective methods of regularly surveying and communicating the condition of their buildings. This allows clients to conduct necessary repairs and maintenance work before it becomes too dangerous or expensive. The building surveyor is engaged to conduct a condition assessment, which includes a lengthy site visit to determine the physical condition of the building elements and estimate cost estimates for immediate and long-term renewal, repair, and maintenance. Current asset condition assessment processes are highly time-consuming, labor-intensive, and costly, as well as posing health and safety risks to surveyors, particularly at height and on roofs, which are difficult to access. This project aims to evaluate the use of convolutional neural networks (CNNs) for automated detection and localisation of key building defects, e.g., mould, decay, and stain, from images. A pre-trained CNN classifier using VGG-16 (later compared with ResNet-50, and Inception models) is used for object localization, along with class activation mapping (CAM). In real-life applications, the model poses challenges and has limitations. This model has proven to be robust and accurate for detecting and localizing defects in buildings. The approach is being scaled up and further advanced so that it can be used to detect defects and deterioration of buildings in real time using mobile devices and drones.

2. INTRODUCTION

In high-rise buildings, the detection of cracks and flaws on the surfaces of the walls is an important part of their maintenance. If left undetected and untreated, these defects can significantly affect the structural integrity and the aesthetic aspect of buildings, timely and cost-effective methods of building condition surveys are of practicing need for the building owners and maintenance agencies to replace the time- and labor-consuming approach of the manual survey.

As client expectations evolve, clients are looking for quick and effective ways to communicate the condition of their buildings so that essential repairs and maintenance can be performed in a proactive and timely manner before they become too costly and dangerous. This type of work commonly requires building surveyors to conduct a site inspection and produce a systematic record of the condition of the building elements, along with estimates for immediate and long-term costs of renewal, repair, and maintenance.

In this project detecting building defects such as cracks , flakes and roof defects, We are using CNN pre-trained model VGG16 to analyze the type of building defect on the given parameters. The objective of the project is to build an application to detect the type of building defect. The model uses an integrated webcam to capture the video frame and the video frame is compared with the Pre-trained model and the type of building defect is identified and showcased on the OpenCV window and emergency pull is initiated.

3. LITERATURE SURVEY

3.1 Existing problem:

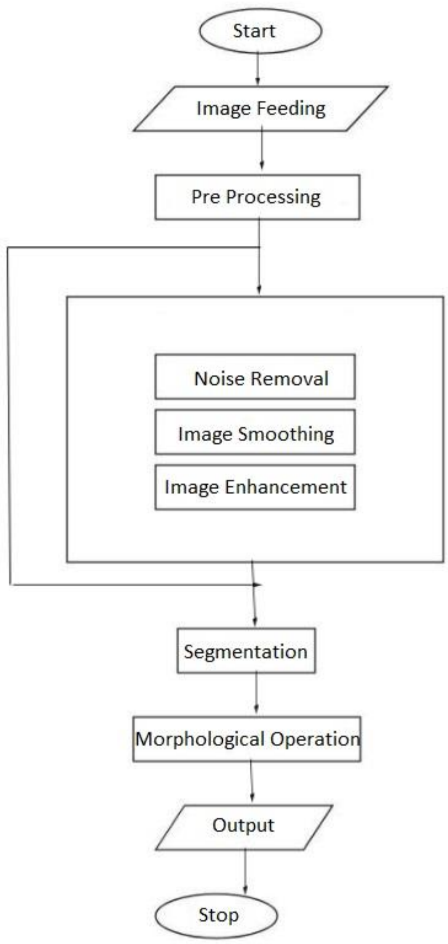
Traditionally, the type of work would consist of engaging builders to carry out a condition assessment which would include a thorough audit of the building elements, including estimations of the immediate and long-term costs of repair, replacement and maintenance. The current asset condition assessment procedures are time-consuming, labor-intensive, and expensive. They also pose health and safety risks to surveyors, particularly at heights and roof levels.

3.2 Proposed solution:

Here we aim at evaluating the application of convolutional neural networks (CNN) towards an automated detection and localisation of key building defects.The proposed model has proven to be robust and able to accurately detect and localize building defects.With further advance to support automated detection of defects and deterioration of buildings can be done in real-time using mobile devices and drones

4. THEORETICAL ANALYSIS

4.1 Block diagram:



4.2 Hardware / Software designing:

Software Requirements:

Python 3 - We have used Python which is a statistical mathematical programming language like R instead of MATLAB due to the following reasons:

1. Python code
2. The python data structure is superior to MATLAB
3. It is an open source and also provides more graphic packages and data sets Keras (with TensorFlow backend 2.3.0 version) - Keras is a neural network API consisting of TensorFlow, CNTk, Theano etc. Python packages like Numpy, Matplotlib, Pandas for mathematical computation and plotting graphs. OpenCV (Open Source Computer Vision) is a library of programming functions aimed at real time computer vision i.e. used for image processing and any operations relating to image like reading and writing images, modifying image quality, removing noise by using Gaussian Blur, performing binary thresholding on images, converting the original image consisting of pixel values into an array, changing the image from RGB to grayscale etc. It is free to use, simple to learn and supports C++, Java, C, Python. It **Hardware Requirements:**

Processor: Intel® Core™ i3-2350M CPU @ 2.30GHz

Installed memory (RAM):4.00GB

System Type: 64-bit Operating System

5. EXPERIMENTAL INVESTIGATIONS

The team examined how to build models using AL and ML and how to process images while working on the solution. And mainly we had studied about the CNN because our solution mainly needed this so we worked on these aspects.

Machines and especially computer systems can simulate human intelligence processes (AI), enabling it to mimic even human behavior. The software finds application in areas such as Computer Vision, Natural Language Processing, Robotics, Speech Recognition, and much more.

Operating Neural Networks: Neural Networks (NN) form the basis of deep learning, a subfield of machine learning where the algorithms are inspired by brain structure. Data is fed into NN, which train themselves to recognize patterns and then predict outcomes for new, similar data. There are layers of neurons in NN. These neurons are the core processing units of the network.

Transfer Learning: A major assumption in many machine learning and data mining algorithms is that the training and future data must be in the same feature space and have the same distribution. However, in many real-world applications, this assumption may not hold. For example, we sometimes have a classification task in one domain of interest, but we only have sufficient training data in another domain of interest, where the latter data may be in a different feature space or follow a different data distribution.

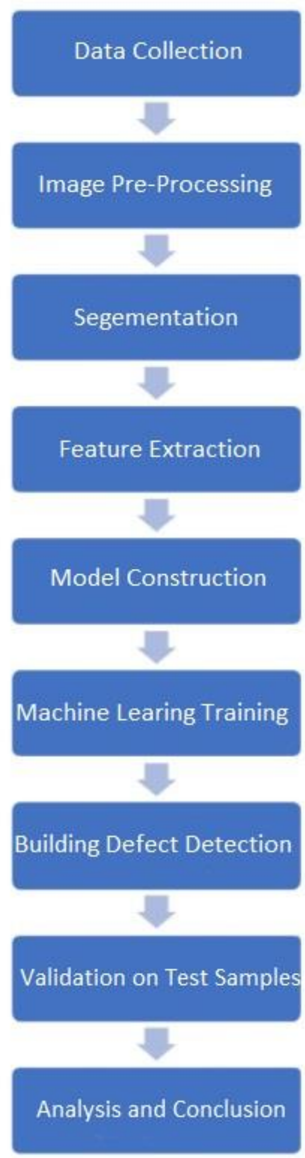
Convolutional Neural Network:

A classification model can be classified into two categories in principle, which are generative models and discriminative models based on traditional learning such as support vector machines (SVMs), Random Forests (RFs), and Convolutional Neural Networks (CNNs). As a result of the computation of a large number of features, methods based on hand-crafted features can have difficulty being accurate when combined with traditional machine learning techniques. These methods are often slow to compute and memory-intensive.

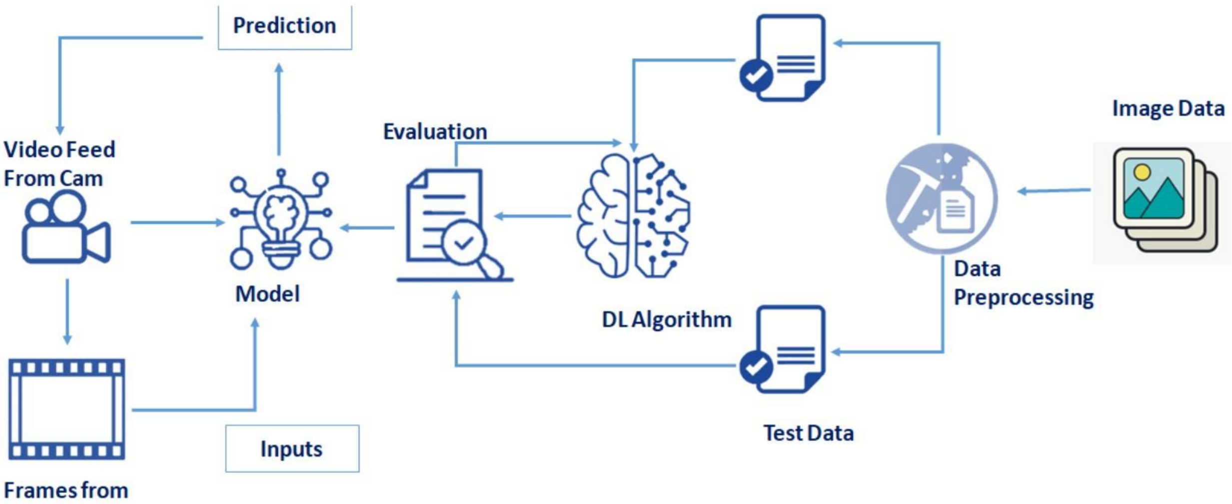
Activation Function: Sigmoid function ranges from 0 to 1 and is used to predict probability as an output in case of binary classification while Softmax function is used for multi-class classification. tanh function ranges from -1 to 1 and is considered better than sigmoid in binary classification using feed forward algorithm. ReLU (Rectified Linear Unit) ranges from 0 to infinity and Leaky ReLU (better version of ReLU) ranges- from -infinity to +infinity. ReLU stands for Rectified Linear Unit for a non-linear operation.

The output is $f(x) = \max(0,x)$.ReLU's purpose is to introduce non-linearity in our ConvNet. Since, the real world data would want our ConvNet to learn would be non-negative linear values. There are other nonlinear functions such as tanh or sigmoid that can also be used instead of ReLU.

6. FLOW CHART


Data Collection
Image Pre-Processing
Segementation
Feature Extraction
Model Construction
Machine Learing Training
Building Defect Detection
Validation on Test Samples
Analysis and Conclusion

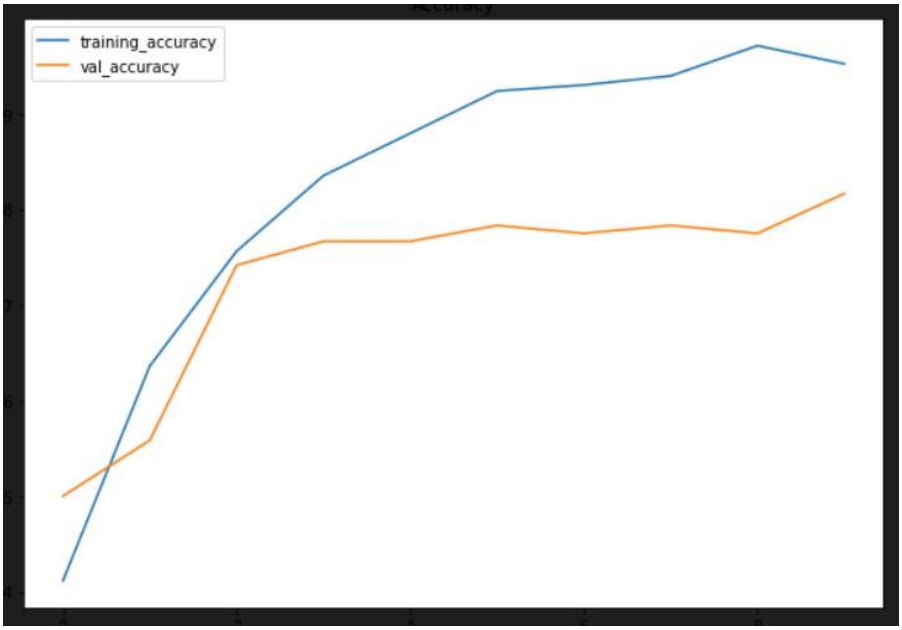
Technical Architecture:



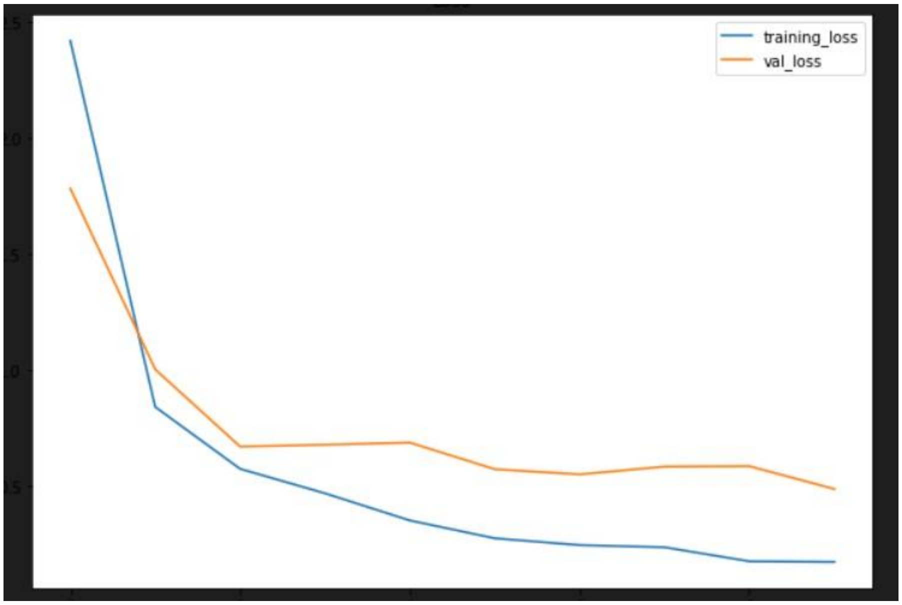
7. RESULT

Plotting the learning Curves

Plot Accuracy



Plot Loss



Final Result

```
# Train the model
history_1 = model_1.fit(
    train_data,
    validation_data = test_data,
    epochs = 10,
    steps_per_epoch = len(train_data),
    validation_steps = len(test_data)
)
```

```
Epoch 1/10
10/10 [==============================] - 7s 647ms/step - loss: 2.4197 - accuracy: 0.4114 - val_loss: 1.7825 - val_accuracy: 0.5000
Epoch 2/10
10/10 [==============================] - 6s 608ms/step - loss: 0.8394 - accuracy: 0.6361 - val_loss: 1.0014 - val_accuracy: 0.5583
Epoch 3/10
10/10 [==============================] - 7s 633ms/step - loss: 0.5721 - accuracy: 0.7563 - val_loss: 0.6684 - val_accuracy: 0.7417
Epoch 4/10
10/10 [==============================] - 7s 640ms/step - loss: 0.4665 - accuracy: 0.8354 - val_loss: 0.6761 - val_accuracy: 0.7667
Epoch 5/10
10/10 [==============================] - 7s 634ms/step - loss: 0.3493 - accuracy: 0.8797 - val_loss: 0.6854 - val_accuracy: 0.7667
Epoch 6/10
10/10 [==============================] - 7s 634ms/step - loss: 0.2724 - accuracy: 0.9241 - val_loss: 0.5706 - val_accuracy: 0.7833
Epoch 7/10
10/10 [==============================] - 7s 655ms/step - loss: 0.2430 - accuracy: 0.9304 - val_loss: 0.5481 - val_accuracy: 0.7750
Epoch 8/10
10/10 [==============================] - 7s 643ms/step - loss: 0.2336 - accuracy: 0.9399 - val_loss: 0.5817 - val_accuracy: 0.7833
Epoch 9/10
10/10 [==============================] - 7s 647ms/step - loss: 0.1730 - accuracy: 0.9715 - val_loss: 0.5835 - val_accuracy: 0.7750
Epoch 10/10
10/10 [==============================] - 7s 657ms/step - loss: 0.1705 - accuracy: 0.9525 - val_loss: 0.4850 - val_accuracy: 0.8167
```

8. ADVANTAGES & DISADVANTAGES

Advantages:

1. It is considered as the best ml technique for image classification due to high accuracy.
2. Image pre-processing required is much less compared to other algorithms.
3. It is used over feed forward neural networks as it can be trained better in case of complex images to have higher accuracies.
4. It reduces images to a form which is easier to process without losing features which are critical for a good prediction by applying relevant filters and reusability of weights
5. It can automatically learn to perform any task just by going through the training data i.e. there no need for prior knowledge
6. There is no need for specialised hand-crafted image features like that in case of SVM, Random Forest etc.

Disadvantages:

1. It requires a large amount of training data.
2. It requires an appropriate model.
3. It is time consuming.
4. It is a tedious and exhaustive procedure.
5. While convolutional networks have already existed for a long time, their success was limited due to the size of the considered network.

9. APPLICATIONS

The main application of this model is to predict whether the provided image is a Crack,Falke or Roof and find whether it is damaged or not. It is well trained so that it will predict the correct data.

10. CONCLUSION

The work is concerned with the development of a deep learning-based method for the automated detection and localisation of key building defects from given images. This research is part of work on condition assessment of built assets. The developed approach involves classification of images into Crack, Flake and Roof. For our classification problem, we applied transfer learning to a VGG-16 network pre-trained on ImageNet.In order to obtain sufficient robustness, we applied different augmentation techniques to generate a larger dataset.The evaluation test showed a overall accuracy of 81.67% of images. For a large dataset, Dice loss is preferred over Accuracy. For small data, we should use simple models, pool data, clean up data, limit experimentation, use regularisation/model averaging ,confidence intervals and single number evaluation metric. To avoid overfitting, we need to ensure we have plenty of testing and validation of data i.e. dataset is not generalised. This is solved by Data Augmentation. If the training accuracy is too high, we can conclude that the model might be over fitting the dataset. To avoid this, we can monitor testing accuracy, use outliers and noise, train longer, compare variance (=train performance-test performance).

11.FUTURE SCOPE

For the future works, the challenges and limitations that we were facing in the project will be addressed. The presented paper had to set a number of limitations, i.e., firstly, multiple types of the defects are not considered at once. This means that the images considered by the model belonged to only one category. Secondly, only the images with visible defects are considered. Thirdly, consideration of the extreme lighting and orientation, e.g., low lighting, too bright images, are not included in this study. In the future, these limitations will be considered to be able to get closer to the concept of a fully automated detection. Through fully satisfying these challenges and limitations, our present work will be evolved into a software application to perform real-time detection of defects using vision sensors including drones. The work will also be extended to cover other models that can detect other defects in construction such as structural movements, spalling and corrosion. A long-term vision includes plans to create a large, open source database of different building and construction defects which will support world-wide research on condition assessment of built assets.

12.BIBILOGRAPHY

[1]. Mohseni, H.; Setunge, S.; Zhang, G. M.; Wakefield, R. In Condition monitoring and condition aggregation for optimised decision making in management of buildings, Applied Mechanics and Materials, 2013; Trans Tech Publ: 2013; pp 1719-1725.

[2]. Agdas, D.; Rice, J. A.; Martinez, J. R.; Lasa, I. R., Comparison of visual inspection and structuralhealth monitoring as bridge condition assessment methods. Journal of Performance of Constructed Facilities 2015, 30, (3), 04015049

[3]. Mimura, T.; Mita, A., Automatic estimation of natural frequencies and damping ratios of building structures. Procedia Engineering 2017, 188, 163-169.

[4]. . Davoudi, R.; Miller, G. R.; Kutz, J. N., Structural load estimation using machine vision and surface crack patterns for shear-critical RC beams and slabs. Journal of Computing in Civil Engineering 2018, 32, (4), 04018024.

[5]. Hoang, N. D., Image Processing-Based Recognition of Wall Defects Using Machine Learning Approaches and Steerable Filters. Comput. Intell. Neurosci. 2018, 2018.

[6]. . Jo, J.; Jadidi, Z.; Stantic, B., A drone-based building inspection system using software-agents. In Studies in Computational Intelligence, Springer Verlag: 2017; Vol. 737, pp 115-121

[7]. Valero, E.; Forster, A.; Bosché, F.; Renier, C.; Hyslop, E.; Wilson, L. In High Level-of-Detail BIM and Machine Learning for Automated Masonry Wall Defect Surveying, ISARC. Proceedings of the International Symposium on Automation and Robotics in Construction, 2018; IAARC Publications: 2018; pp 1-8.

[8]. Valero, E.; Forster, A.; Bosché, F.; Hyslop, E.; Wilson, L.; Turmel, A., Automated defect detection and classification in ashlar masonry walls using machine learning. Autom Constr 2019, 106.

[9]. Radopoulou, S.C.; Brilakis, I. Automated detection of multiple pavement defects. J. Comput. Civ. Eng. 2016, 31, 04016057.

[10]. Sinha, S.K.; Fieguth, P.W. Neuro-fuzzy network for the classification of buried pipe defects. Autom. Constr. 2006, 15, 73–83.

[11].German, S.; Brilakis, I.; DesRoches, R. Rapid entropy-based detection and properties measurement of concrete spalling with machine vision for post-earthquake safety assessments. Adv. Eng. Inform. 2012, 26, 846–858.

[12]. Ji, M.; Liu, L.; Buchroithner, M. Identifying Collapsed Buildings Using Post-Earthquake Satellite Imagery and Convolutional Neural Networks: A Case Study of the 2010 Haiti Earthquake. Remote Sens. 2018, 10.

[13].Organisation (ISO), I.S. ISO 19208:2016- Framework for specifying performance in buildings; ISO, 2016;

[14].Maguire, M.; Dorafshan, S.; Thomas, R. SDNET2018: A concrete crack image dataset for machine learning applications. Browse Datasets 2018.