Concrete crack detection using Convolutional Neural Networks

[1]S. Uma Maheswari, [2] F. Owino, [3] A. Anirudh, [4] VS. Praganesh
[1]Assistant professor, [2]Student, [3]Student,[4]Student
Department of Computer Science and Engineering,
Sri Shakthi Institute of Engineering and Technology, Coimbatore, India

Abstract: The purpose of this paper is to explore the different methods used to detect concrete cracks. It reviews the non-destructive crack detection methods, their strengths and weaknesses. The goal is to analyze their application and feasibility to find the most optimal and practical methodology that will be most suitable for implementation. Crack detection has been manually over a long period of time, this presents major security risks and it is also time-consuming. Other implementations require using dedicated hardware which in turn may be expensive to acquire. Using a readily available device which is the smartphone, it is possible implement the detection system by processing the video stream, even a low-cost smartphone has all the sensors that can be used for the system making it is more accessible for usage by many people. Due to the smartphone compactness, lighter weights and smaller form factors, by tapping into readily available devices such as drones and small remote-controlled cars people could attach their smartphone and reach places they could not analyze before such as high walls, dam walls and underground tunnels and get the precise location of the cracks

Index Terms **- image processing, convolutional neural networks, small unmanned vehicles, concrete crack detection, supervised learning, transfer learning**

I. INTRODUCTION

Concrete cracks as simple as they are, play a major role in signalling concrete structure strength and signal structure deterioration when seen. Implementing a cost-efficient automated crack detection method ensures both safety and efficiency. Convolutional neural networks are receiving a boom in utilization majorly due to the improvement in the computational technology's improvement in the calculation algorithms and largely helped by the availability of big data. The computational algorithms such as the Winograd's algorithm have greatly fastened the major bottleneck of the networks which was in the massive matrix multiplication operations. Using convolutional neural networks, people are able to achieve a high state of the art performance levels. The CNN being an area of growing interest There is massive research being undertaken thus making technologies such as transfer learning available for use. Machine learning libraries also play a major role in getting people started with machine learning by providing a good amount of abstraction while still providing high performance. By utilizing these advancements, classification images more easily than utilizing traditional image processing techniques while being novice-friendly at the same time.

II. REVIEW OF LITERATURE

Cracks are the earliest, most common and reliable signs signalling the beginning of structural degradation. This makes crack detection an important task to be done in order to allow engineers to take early preventive measures. These early preventive measures help to prevent extensive damage that may be cost intensive to repair and also to prevent accidents that would occur due to extensive structural damage. In order to make it feasible, it is needed to ensure that the method is a cost-efficient method such that the proactive approach costs don't exceed the reactive approach costs. For complex structures, it is difficult to get the manual evaluation of the cracks and it is a risky, time consuming and a budget intensive approach. Due to the high costs of manual human-based evaluation, the following will be a review of the various image-based evaluation systems. The image below shows the overall flow of the image-based crack detection systems.

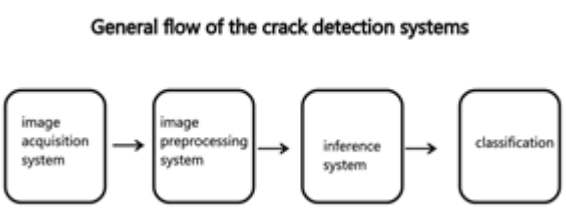


Fig 1 general crack detection system

A. Image Acquisition System

The image acquisition system is used to acquire the images that will be subjected to testing. With the growth and availability of cameras, people will have an extensive list of options to use in the system form dedicated cameras, web cameras or even the ubiquitous mobile phone cameras. Shimamoto et al. [1] have shown how the camera sensors have been advanced with higher resolutions and ability to record at higher frame rates. The developed camera has a spatial resolution approaching human vision's ability is an advancement that will enable the systems to get as many details as they possibly can get out of the scene. These high resolutions and higher frame rates will, in turn, lead to more data to be analyzed. This creates room for the developers and researchers to optimize the inference engines to keep up with a large number of frames availed for analysis. Chaudhury et al. [2] suggest a pixel-wise crack detection method that scans the frames in a video sequence to detect cracks, optimization will ensure faster evaluation of the frames as they are delivered.

B. Image pre-processing system

The image pre-processing system performs the necessary alterations to make the inference easier to be performed and also enhance the accuracy levels of the inference system. Most common pre-processing techniques are usually focused on reducing the noise and unnecessary data points from the image. Here the necessary transformations are performed to make the image ready for evaluation in a better state. Transforming the images into grayscale and resizing the images by compression are the most common form of image pre-processing performed on images.

Noise is a common factor as these images are taken in an uncontrolled environment with a wide variety of camera setups, different lighting conditions and variable weather conditions. Mean filter is one of the used noise reduction methods. It is used to clean salt and pepper noise. Mean filter is mostly faced with a blurring effect as its biggest disadvantage. Vallepalli and Rajendran [4] show how this blurring can be reduced by making use of Non-linear mean filters. They show minimal variance from the images showing how they suppress noise in the presence of Gaussian noise and also its performance in the presence of edges by applying the harmonic mean filter.

The median filter was developed to supplement the tasks that could not be done using the low pass filters or the high pass filter separately. It performs better than the mean filer by eliminating the blurring and maintaining sharpness. This is mainly because the median filter does not introduce any new pixel values.

George et al. [4] have shown the various methods of implementation of the median filtering technique and shows how the forwarded approach is effective in salt and pepper noise reduction. They also show how two-phase filters enhance the performance by employing parallelism and still managing to use only a fraction of the standard CPU usage. By utilizing a smaller amount of the CPU, it is able to be implemented in various types of low powered systems such as the wireless sensors and IoT devices that are power and processing power constrained. This will greatly contribute boost the lifetime especially of the wireless sensors which have been a major limiting factor.

Another way to tackle the uncertainty of the image inputs is to add some noise to the images during the training process the images are often exposed to some types of transformation to introduce some noise. The transformations can also be useful to help increase the size of the dataset used for training especially when the dataset is limited. The transformations that can possibly be done include random cropping, rotation, reflection etc. Just by applying two different transformations the benefit obtained is the abiltiy to increase the data set by three folds which greatly increases the data set size. Seo et al. [9] have shown how the lack of data leads to overfitting during the training process. Overfitting makes the model learn all the features of the training data and also learns the noise. This greatly lowers the performance of the models when it is given new data to classify when it is to be working on actual problems.

C. Inference System

The inference system is composed of the engine that is responsible for the segmentation of the image and determining whether the cracks are present in the evaluated frame or not. Choi and Kim [6] presented a crack detection system that makes use of UAV to capture the footage which is then subjected to edge detection algorithms. The edge detection algorithms will be practical for application in a limited and controlled environment free from massive obstructions because it would bring up a lot of false positives. The advantage of using unmanned vehicles reduces the risk associated with the analysis in risky places and also increases the range of places that could be analyzed. To supplement the shortcomings of the pure edge detection mechanism the usage of fuzzy inference and the SOM algorithm to better supplement the accuracy.

Heng et al. [7] have suggested a modification of the beamlet tree methodology by enabling it to also connect fragments. This plays an impressive role especially when the length of the cracks is needed in order to be able to tell the extent of the crack in thus enabling to estimation the lengths. It can also help in concealing the effects of noise in the images. This will be of great importance in cases where it is of great importance to classify the precise type of crack. The implementation showed great performance in representing the actual crack which most skeletonization algorithms could not display properly by eliminating some portions of the cracks.It was able to detect even the very thin portions of the crack that the other approaches could not represent. From his experiment, it can be clearly seen how the pure canny edge detection is susceptible to noise interference with the canny image inference having a lot of noise. Estimating the length of the crack is the largest application of their implementation it adds a new dimension in the utility. By adding the linking operation to the normal beamlet tree method the new implementation managed to score higher in completeness and quality than the pure beamlet tree implementation which was already ahead of the other skeletonization mechanisms. Also, by being able to classify the different types of cracks people can end up finding the root cause of the problem. people can be able to tell precisely what led to the crack formation thus helping civil engineers make proper mixture proportions and also foster faster research and development of new construction materials.

The advancement of machine learning algorithms has made it possible for a more complex set of images to be classified. They have been widely used across many fields especially in the medical field to analyze the outputs from medical imaging systems and

they have had a massive success rate of accurate diagnosis. Neural networks have been widely used in the medical field because they have been largely influenced by the interaction and impressive collaboration between developers and the neuroscientists.

Ertam et al. [8] show how artificial neural networks have grown and their applications. The convolutional neural networks were developed to advance the neural networks. Here they also present the performance of different activation functions with ReLu the most preferred activation function having the highest accuracy followed by a softplus activation function. The convolutional neural networks growth is greatly been boosted by the availability of open data sets that people can use in order to directly compare the performance of their algorithms and implementations. This was not possible earlier due to lack of a standard data set, the data people used was individually collected and was therefore flawed to utilize their results to compare with others.

The ease of access of implementation has been greatly influenced by the availability of many open source machine learning libraries such as TensorFlow, Kera's etc. [8] The ease and quality of the results in term of accuracy, completeness and quality of the classification makes the convoluted neural networks a go-to solution in many image processing systems. The quality of the classifications is greatly dependent on the training data and the layer architecture and the selection of the different functions as they are all suited for different use cases.

Overfitting is a very undesirable thing in the machine learning world. Implementing a dropout function helps to greatly reduce the overfitting and also helps in the convergence and reduces the error rate greatly.

D. Transfer Learning

Classification is an interesting part of machine learning that has gained a lot of interest among many people. It has been the centre of focus for many and it brought about the ImageNet competition. Training the classification models from scratch tend to take a large amount of time even on high-end computing devices with multiple top of the line graphical processing units. The models from the ImageNet competitions along with their weights are made public. This is a very useful reference resource for those beginning their journey in machine learning and majorly classification and object detection problems.

The ImageNet competition has led to revolutionization and a major refinement of the perfect arrangements of layers in convolutional neural networks with many models attaining state of the art levels of accuracy. These models present a learning model for people who want to begin their journey with machine learning making use of transfer learning networks can be built on top of some already available models and find the most efficient for a particular type of problem.

By utilizing transfer learning, good classification results can be attained and the training is not as resource intensive as it is in the training from the ground up. Transfer learning only retrains the last classification layer in the fully connected layer to be able to classify a given set of classes in a very short time. This training can be done using commodity computer and can also be sped up using GPU support. They can also be done in the cloud such as GCP and Amazon for a better GPU but at a smaller rate since you pay per use and the training time has been shortened massively. By making use of these models the machine learning beginners are able to get their products for testing up and running in a very short amount of time and also enables them to test different models to find the different layer architectures that are best suited for a particular set of problems. These models are a good reference point for people to start their journey by learning the good layering architectures since they greatly affect the performance of the models.

III. PROPOSED SYSTEM

The proposed system is designed to make use of all the above benefits to build an easily accessible cost-efficient and safe system without limiting the accuracy obtained by using actual humans doing the observation. The suggested approach here is building an android application to cut cost since most of the android phones have all the sensors needed for the task. The users will only need to install the application to get started. With the application. In order to get full utility of the system automation, the users will only have the flexibility to purchase a small remote-controlled vehicle or a drone of their choice and use the phone to perform the inference and store the GPS location when the crack is detected. This ensures the users have an extra level of quality control on their hand by selecting their drones since many companies might cut costs in fully dedicated systems. This gives quality control as well as flexibility while cutting the overall costs.

A. Supervised learning

The training was done with a dataset from SDNET2018 which is an annotated image dataset. The dataset has 56000 images of both cracked and uncracked images. The training used 40000 images from which 80% of the data for training and 20% of it for testing.

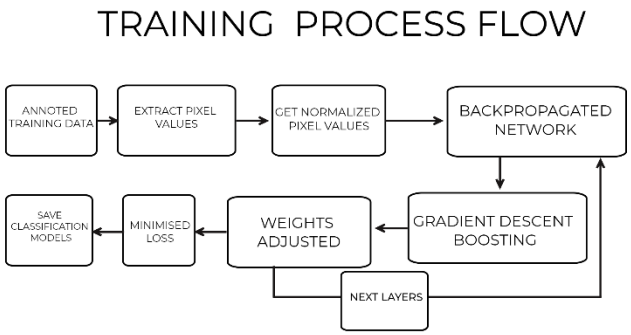


Fig 2 training process

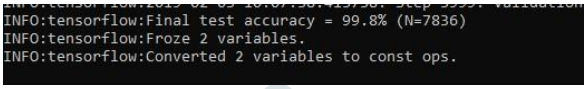


Fig 3 training results

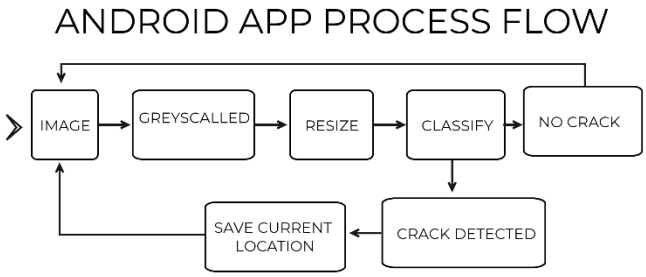
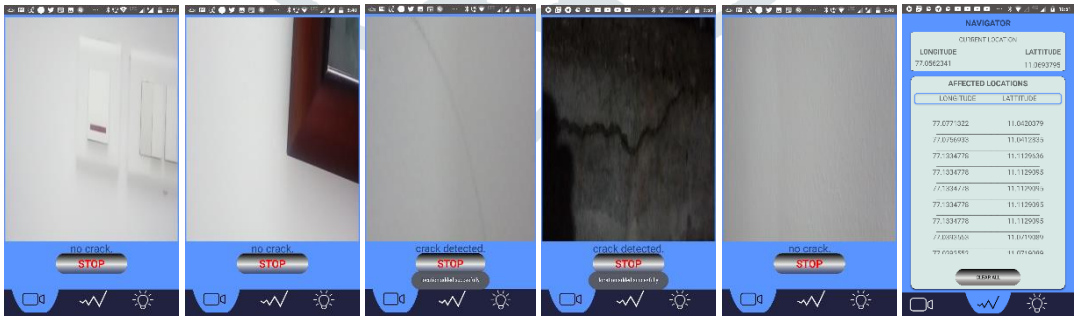


Fig 4 android application flow

IV. RESULTS AND DISCUSSION

As a result, the end system is a system that is able to correctly classify the cracks in the live video feed in a very reliable manner and an efficient way by eliminating the need for storing and later reviewing the image feed. By classifying the frames in real time, the system is only required to store the GPS locations of the cracks locations thus being both storage and processing power efficient as there is an elimination of the worry about compressing the data or setting it up to transmit the data for evaluation after it has been recorded



(a) (b) (c) (d) (e) (f)

Fig (a-f) resultant system in action

From the results in figures a-f, it can be clearly seen how efficient the system is from by eliminating other components such as paintings and switches n not registering them as cracks which would be a common problem if using edge detection methodologies.

s

V. CONCLUSION AND FUTURE SCOPE

The network offered a reliable method of detecting the presence of cracks in the images. This is a good starting ground in obtaining different characteristics by coupling it with other wireless sensors to get more information such as the depth of the cracks etc. These additional sensors could be synchronized with the application to add more useful data for inference of some more useful information

Due to the ease and convenience prospects of the project the next versions should be trying to implement the estimation of crack length and also group different types of the cracks detected. By utilizing this classification, it will be able to help civil engineering research teams in their material research work. By classifying the types of cracks it could also be able to set up the trigger mechanism to help engineers put more focus on more risky types of cracks that need more immediate attention. This will be made possible by following works such as those done by Heng et al. [7] and building on top of these to come up with a more reliable mode of accurate classification.

REFERENCES

[1] Road Crack Detection Based on Video Image Processing-Gang Shen, China Transport-Telecommunications & Information Center

[2] Thyroid Segmentation and Volume Estimation in Ultrasound Images ITBE page 1348-1357.

[3] Image De-noising using mean pixel algorithms corrupted with photocopier noise-Sri Srujan Vallepalli Madan Mohan Rajendran2012 IWSSIP 530-535

[4] A Survey on Various Median Filtering Techniques For Removal of Impulse Noise From Digital Image-Ginu George; Rinoy Mathew Oommen; Shani Shelly; Stephie Sara Philipose; Ann Mary Varghese2018 Conference on Emerging Devices and Smart Systems (ICEDSS)

[5] Building Crack Inspection using Small UAV - Sung-suk Choi*, Eung-kon Kim*

[6] Accurate extraction of cracks on the underside of concrete bridges-Heng Zhang, Huai Yu, Haijian Zhang, Wen Yang

[7] Data Classification with Deep Learning using TensorFlow- Faith Ertam, Galip Aydinc

Image Classification of Fine-grained Fashion Image Based on Style Using Pre-Trained Convolutional Neural Network -Yian Seo and Kyung-shik Shin page 387-390