

DEEP LEARNING ON UAV VIDEOS FOR VISUAL INSPECTION OF STRUCTURAL AUDIT

ABSTRACT:	3
INTRODUCTION:	3
LITERATURE REVIEW:	3
GAP ANALYSIS:	13
PROBLEM STATEMENT:	13
PROPOSED SOLUTION:	14
REFERENCES:	14

Abstract:

Keywords: neural network, crack, dampness, algae, fungi, paint peel-off, unmanned aerial vehicle, drone, structural audit

Introduction:

Construction of different infrastructures is a necessity for human life, as it bring value to the land, fulfill human needs and also develop the area around it. As the infrastructures get old, the risk living/working in the building increases. We can see cracks, paint peel off and dampness on the walls of old infrastructures. The infrastructure should be examined after 30 years. The process is called structural audit. For structural audit the surveyor has to go the location and do visual inspection of the infrastructure, different chemical test, etc. While performing visual inspection for detection of cracks, dampness, paint peel off, the process totally depends on the surveyor experience and knowledge. Sometimes there may be human error. With the help of machine learning and UAVs we can reduce the human error caused during inspection and monitor almost the whole structure.

Many systems are developed for crack detection using image processing but it takes lot of time and resources. There are many systems developed using deep learning for detecting cracks on pavement, bridges. But only cracks are not the cause for damaging the infrastructure. For visual inspection a well-trained surveyor is required. But with the help of drone technology and deep learning a person with drone operating skills can have the job done. With UAVs, capturing images of higher grounds will be much easier and developing a neural network model for detection of the defects on captured images of walls will help us reduce human error rate and speed up the process. This process will be automated and reduce human work.

Literature review:

Real Time Image Processing System for Detecting Infrastructure Damage: Crack [1]

The paper states a crack detection algorithm using a small camera and an embedded board that is Jetson TX2. There are three methods for crack detection. They are laser sensor based, acceleration sensor based and camera based. The laser method has high accuracy but it is very slow in scanning and the acceleration sensor based method also has low accuracy. The camera based method is cost effective and is very good in terms of accuracy. The algorithm is divided in two parts, the image preprocessing and crack detection. In the preprocessing part color image is converted to gray scale and after that preprocessing methods are applied to remove noise from the image. In the detection part the outline of the crack is obtained and then filtering methods are applied to the image and finally the crack is obtained. When the algorithm was tested it was 89.33% accurate and it correctly classified nine out of ten images.

Image Processing Based Analysis of Cracks on Vertical Walls [2]

The paper proposed an approach to identify various type of cracks and for calculating the width of the crack on walls. For crack detection the binarization algorithm is used. The image undergoes binarization and gets converted to a grayscale image and then to binary image by setting a threshold value. The obtained binary image undergoes edge detection and the width of the crack is

calculated. Due to issues in binarization caused by changing user parameters (sensitivity and window size) the cracks sometime remain undetected. For increasing the accuracy of the binarization, preprocessing is done for removing the distortion in the image and then the image is converted to grayscale. A hybrid image processing algorithm is applied on the preprocessed image with two new user parameters, they are P_w and P_l . By using these parameters two binary images will be obtained and they will undergo the edge detection algorithm. This approach is been tested and compared with Savola's method and it gives better results than traditional binarization. Cracks visible to human eyes are still not classified by the hybrid approach.

Inspection, Identification and Repair Monitoring of Cracked Concrete structure –An application of Image processing [3]

The paper proposed an image processing algorithm for detecting cracks and using UAVs for capturing images. The algorithm is as follows: after capturing the image 1) The RGB image is converted to grayscale image to reduce the intensity of the image and to represent the image in single channel. 2) Void detection is done to find any voids present on the surface cause by air bubbles or water droplets between the concrete. This is done using MATLAB. 3) Image stitching is done to merge the different images captured by UAVs from different positions so as to get a clear vision of the crack. 4) Next step is median filtering followed by edge detection for which CANNY and SOBEL filters are used. It is done to reduce the pixels to process the image faster. 5) Texture analysis is carried out using SOBEL and CANNY to filter out a specific color from the image. 6) Then a binary image is formed and then thresholding is done to remove skeleton pixels from the image. This method can be used to detect cracks in chimney, etc.

Detection and analysis of large-scale WT blade surface cracks based on UAV-taken images [4]

Traditional method for defect detection on the wind turbine is by using telescope and sensors which are very costly and have low efficiency and take lots of time. This paper suggest an approach using image processing algorithm and UAVs. UAV is just used for capturing the images of WT. The algorithms is as follows: Wiener filter is applied on the image remove the motion blur. An adaptive median filtering method is used for removing the noise from the image. After these steps are performed the image losses some quality, so to regain the quality enhancement algorithms are used to the image will be better for analysis. Using this approach there is an improvement in the crack detection as that of the traditional method. Comparing the proposed method with the extended Haar cascade classification method, the proposed method analysis crack with more detail.

Automatic Pavement Crack Detection Using HMRF-EM Algorithm [5]

The paper proposed an algorithm for crack detection on pavement using the adaptive line detector and hidden Markov random field model and its expectation-maximization algorithm. First step of the proposed method is to preprocess the original image using band pass filter method to reduce the noise in the image and then an adaptive line detector is used due to changing region size and direction. The second step is to use the HMRF-EM algorithm to increase the accuracy to label the data. The labels obtained in this step are also consisting of noise. Third step uses start and end points of the labeled data and cluster them on relative position so to get the true end points. When compared with other methods the proposed method is much more accurate and has a Recall of

94.54%, Precision of 90.89% and F-measure of 93.04% which is an improvement over other method.

Crack Detection in Concrete Elements from RGB Pictures using Modified Line Detection Kernels [6]

This paper proposed a modified line detection kernel algorithm to detect crack. Main aim of the paper is to find cracking patterns in concrete from the image captured. Necessary steps are digital image processing to reduce the noise and image sharpening, spatial filtering, feature detection, hypothesis and then the algorithm which is as follows: 1) Edge detection is carried out to reduce the pixels. 2) The pixels obtained from edge detection are yet not the crack pixels, to get the crack pixels we have to perform mean detection. 3) In the orientation kernel step, the group of pixels which are dark but not crack are removed by passing it through modified line kernel. 4) In angle detection the pixel matrix is multiplied element wise with kernel. 5) Calculating the width of the crack. After apply this algorithm the results are not good. This method works very well on plain surface but not on rough surfaces.

Track Crack Detection Method in Complex Environment [7]

This paper states an algorithm for crack detection in rail slabs using image processing. The sample image is preprocessed to remove the noise, uneven brightness. Then the image size is scaled to 25% of the original size. There are 3 types of cracks they are deformation, crack and surface damage. To classify them edge detection algorithm is used, but normal edge detection is not enough so CANNY edge detection algorithm is used to extract information from the preprocessed image. After performing these steps we get three characteristics: 1) crack is long, 2) relatively large length and width and 3) center ratio is large. These characteristics of the image are calculated by the algorithm. With the values of the three characteristics, we determine which crack it is. When the experiment was performed the accuracy was 87.73%. Total 220 images were captured and used for testing and 193 images were correctly identified.

Robust Image-Based Crack Detection In Concrete Structure Using Multi-scale Enhancement And Visual Features [8]

This paper states an algorithm for detection of crack using image processing. As the captured image has low contrast and noise. For making the image more enhance a multi-scale enhancement method is used. The method uses a MESG algorithm. After this step proposed algorithm is applied as follows: 1) Taking this enhance image apply SOBEL filter on the image. 2) Then obtain the binary image. 3) Morphological processing is done to remove noise and join parts of cracks. 4) refine the obtained image so that cracks are clearer. The proposed algorithm is compared with OSMA and PMM algorithm. The dataset selected for the testing purpose has 200 concrete images of dimension 4000×6000 pixels captured at NSH of CMU. It performed very well against OSMA and slightly well against PMM and got a score of 94.22% on average TRP.

Deep Learning For Semantic Segmentation Of UAV Videos [9]

The traditional approaches use image segmentation deep learning methods and extend it to frames of video. This does not take into account the temporal features of subsequent frames. This paper states video segmentation method using a combination of FCN (Full Convolutional Network) and Conv-LSTM (Convolutional Long Short Term Memory). The architecture stated in this paper's methodology consist of FCN instead of CNN (Convolutional Neural Network) because FCN can

be used for arbitrary sized inputs and it generates pixel level segmentation output. The input videos are captured using UAV (Unmanned Aerial Vehicle). The input video frames are fed into FCN and the output segmentations of FCN are grouped into blocks of 4 frames with 3 overlapping frames in consecutive blocks. These blocks are then fed to Conv-LSTM. Conv-LSTM has a loop in its network which allows previous state features to be used in the next frames. The results of this paper show that using a combination of FCN and Conv-LSTM increases the accuracy of almost all classes.

Effective Object Detection From Traffic Camera Videos [10]

This paper states the work done by a group of researchers for the NVIDIA AI Challenge. This challenge consisted of video recording of a traffic intersection. The objective was to classify different kinds of four wheelers, 2 wheelers, pedestrians, etc. This paper compared 2 deep learning methods namely, R-FCN (Region-based Full Convolutional Network) and Faster RCNN (Region-based Convolutional Neural Network). This paper used the dataset provided by NVIDIA which consisted of 8 hour length video with annotated frames for training and testing. The transformations this paper used was flipping of frames. The architecture used in this paper consists of a RPN (Region Proposal Network) which gives score to regions in the frames. These region scores are then passed to the above 2 deep learning networks. The results show that R-FCN performed better than Faster RCNN. Also, R-FCN has pixel level classification design which can be easily applied to any general detection task. The results were not able to differentiate similar appearing classes (like variants of four wheelers). This results can be improved by using temporal features and using more dataset and variance.

Neural Network-based Vehicle And Pedestrian Detection For Video Analysis System [11]

This paper compares 3 neural network architectures for vehicle and pedestrian detection. The three neural network architectures compared are R-CNN (Region-based Convolutional Neural Network), YOLO (You Only Look Once), and SSD (Single Shot multibox Detector). The parameters used for comparison were AUC (Area Under precision-recall Curve), mAP (mean Average Precision), and processing time. Different variants/networks of these 3 architectures were used testing (Faster R-CNN + Resnet-101, Faster R-CNN + InceptionResnet-2, YOLOv3, SSD + MobileNet-1, SSD + MobileNet-2). Results show that Faster-RCNN performs better in terms of AUC and mAP. The drawback of this architecture is that RCNN uses grids of region whereas other 2 methods process the entire image at once. Thus the processing time of Faster-RCNN method is higher than other 2 methods. This processing time is crucial for real time processing where the accuracy requirements cannot always be met.

Object Detection And Tracking Based On Convolutional Neural Networks For High-resolution Optical Remote Sensing Video [12]

This paper states a method for object detection in high resolution remote sensing videos using neural networks. The considered architectures were YOLO (You Only Look Once) and Faster R-CNN (Region-based Convolutional Neural Network). Faster R-CNN was selected for this paper because it produces better result for smaller sized objects which are present in remote sensing video. The dataset used in this paper is satellite videos capturing aircraft and other similar objects.

This paper uses sub sampling strategy which uses sliding window mechanism. This sub sampling strategy allows for adapting to different resolution images. The proposed architecture uses temporal features of previous frame. The RoI (Region of Interest) generated for the previous frame is fused with current frame's RoI. In this way, the current frame's RoI is passed on to the next frame. This temporal feature has improved the results as compared to Faster R-CNN.

Object Detection On Video Images Based On RFCN And GROWCUT Algorithm [13]

This paper focuses on automation of picking objects from inventory in the logistics industry due to decrease in the productive age population in Japan. As there are a lot of different objects in logistics industry, the object detection algorithm is required in the robot vision. The dataset used by this paper is PASCAL Visual Object Classes 2007 dataset consisting of large number of classes. The algorithm this paper uses is R-FCN (Region-based Full Convolutional Network) for object detection. After object detection, the paper's methodology uses Grow cut algorithm for segmentation. This segmentation refines the bounding box produced by R-FCN to actual shape of the object using background and foreground seeding method. This paper uses Kinect for Windows v2 (Kinect v2) for capturing RGB images and depth images. This RGBD channels acts as the input. And thus, by using refined segmentation and depth information, robot vision can measure the depth of the object for picking. The mean accuracy precision obtained by this paper is 67%.

Object Detection, Classification And Tracking Methods For Video Surveillance: A Review [14]

This paper states various methods of object detection, classification and tracking of objects in videos. The detection methods are Motion-based Object Detection (including Background Subtraction, Frame Differencing, and Optical Flow Method) and Appearance-based Object Detection (including Template Matching and Feature Extraction). The classification of the detected objects can be done based on 4 attributes: shape and size, color, texture, and motion. This paper also states methods of tracking of the classified objects. This paper compares the different methods and sub-methods mentioned above with their advantages and disadvantages.

Panorama Stitching, Moving Object Detection And Tracking In UAV Videos [15]

This paper states methodology for panorama stitching of frames obtained by UAV (Unmanned Aerial Vehicle). The subsequent frames are compared and merged to form a continuous panoramic image. The paper also states object detection methodology using lq-estimator which is the normalization function comparing features of the image. This method suffers from outliers and hence tuning is required for normalizing function. The detected object is then tracked in the subsequent frames using thresholding. The results show that the methods mentioned in this paper works well for both smooth textured and rough textured images.

Video Object Detection For Tractability With Deep Learning Method [16]

This paper states object detection methods in video dataset using deep learning methods for traceability. Annotated database is constructed first marking the object of interest and training the model off line. The architecture used in this paper is CNN (Convolutional Neural Network). Once trained, the model is deployed online and objects are detected and traced using Gaussian

background modelling and other non-parametric modelling. The results show that this deep learning method is suitable for downloaded videos as well as real time video analysis.

Video Predictive Object Detector [17]

This paper states a methodology for Predictive Object Detection (POD) using deep learning techniques. The architectures used in this paper uses YOLOv2 (You Only Look Once version 2) and LSTM (Long Short Term Memory). The YOLO architecture has time benefits as it processes the image only once. The LSTM architecture is used to incorporate temporal features of previous frames in the video. The paper suggests 2 different architectures PODv1 and PODv2 which differs in the feature mapping stage. The PODv2 gives better results.

Video Stream Analysis In Clouds An Object Detection And Classification Framework For High Performance Video Analytics [18]

This paper proposes a method for video stream analysis using cloud computing for object detection and classification. The traditional methods are time consuming even after using GPU cores as the resolution of today's cameras are increasing. This paper states that saving the input video data on the cloud, decoding the videos, and moving operator function code to the computing nodes on the cloud can give scalability and time benefits. The nodes on the cloud have GPU cores thereby collectively improving the time required for analysis of videos.

UAV Autonomous Target Search Based on Deep Reinforcement Learning in Complex Disaster Scene [19]

The proposed system is about designing UAV autonomous path using deep reinforcement learning in order to rescue people in post disaster. This paper proposed a snake game to automate path for drone. Snake body is used to represent drone and hotspot is used to denote target to be searched. Main focus is to create a strategic agent that will be able to play as many game as possible. System uses Markov decision to represent agent with state action and reward state, Q-learning, Markov Decision process, Depth Q-Learning (DQL). DQL consist of multiple state as input and actions as output that need to be perform. State include raw picture frames (240×240 pixel images) performing prepossessing to reduce dimension, RGB image is converted to grayscale, used Q learning with neural network in order to find small paths autonomously.

UAV-Based Situational Awareness System Using Deep Learning [20]

The proposed system mainly focuses on situational awareness system implemented using UAV which is very useful in many areas like search and rescue, surveillance, disaster response etc. Designed a Personal-Action-Locator (PAL) system uses Visual camera mounted to UAV, intelligent processing to detect people, their action, map their location on PAL. Main component is deep learning trained to detect multi-person and action recognition. The Pixel2GPS converter that estimates the GPS position of persons by image processing. CNN is used for object detection and recognition. PAL deployed on Jetson TX2 (8 GB RAM), that we use as the onboard computer for the DJI Matric 100. Used DJI Phantom 4 and its built-in gimbal equipped camera to scan a target area. As a result human detection, we obtained a map of 60.9% in the surveillance test, as compared to 85.0% in the target.

Deep Learning-Based Super-Resolution Reconstruction and Marker Detection for Drone Landing [21]

This system proposed a deep learning based SR reconstruction and marker detection for drone landing. In SR reconstruction low resolution image frame 320×240 pixels is capture then this feed to the module that combined the SR reconstruction and Marker detection CNN for detecting landing area. The inputted image feed to CNN with DCSCN which produced high resolution image. Combining and training multiple CNN network with is difficult so the proposed system focuses on two network separately. DCSACN achieve more efficient computation, uses 1×1 CNN layer, input feature map of $80 \times 80 \times 410$. R-CNN exhibit high performance in terms of object detection, process 200 frame per second. YOLO is alternative that exhibit 10 high performance than R-CNN. Snapdragon 835 and Jetson board TX2 tool used. The self-collected dataset of low-resolution marker images and our trained models for SR reconstruction and marker detection are made available to other researchers.

Autonomous Navigation Of UAV In Large-scale Unknown Complex Environment With Deep Reinforcement Learning [22]

The proposed system focuses on incrementally building a map of the unknown environment and use it to localize. The model include deep Reinforcement learning (DRL), reinforcement learning (RL), Markov decision processes (MDP), partially observable Markov decision processes (POMDP) which helps in navigation in complex environment. Recurrent deterministic policy gradient (RDPG) and fast RDPG algorithm helps in navigation. There are certain limitation associated with proposed system firstly we need to test this model in real environment so it include complex obstacle like wires, and tress that is challenging.

Real-time UAV Detection based on Deep Learning Network [23]

The proposed system mainly focuses on deep learning based YOLO (You Only Look Once) model for detecting UAV. YOLO model takes input image, resize input image to 416×416 , divide image into 5×5 grid, object center fall on grid, if yes then responsibility of grid to detect object, predict B's bounding boxes, calculate bounding box x, y, z, h confident, predict conditional probability, determine bounding box confidence score. Else undefined object. YOLOv2 is 92.10% accurate and YOLOv3 is 95.20% which is better but YOLOv3 took longer training time. Mean average precision (MAP) is also calculated for both YOLOv2 and YOLOV3. Deep learning used in YOLO model for real UAV detection.

A Deep Learning Approach to UAV Image Multilabeling [24]

This paper mainly focuses on generating multilabeling image using deep learning approach. Input frame passed to CNN module which divide image into equal frame size further multilayer perceptron (MLP) is used for classification. GoogLeNet used to generates a feature vector of size equal to 1024. Otsu's Thresholding Algorithm is used to find hypothesis between 2 classes. A RBFNN model is trained for the classification task. AlexNet-RBFNN and GoogLeNet-RBFNN used for feature extraction. GoogLeNet-RBFNN and AlexNet-RBFNN score 79.3% and 78.4% of

average accuracy respectively for data set 1, and 77.4% and 76.9% of average accuracy for data set 2 for given data set.

Strength Analysis of Buildings using Image Processing and SHM Principles [25]

The system proposed include inspection methods that is vital for diagnosis and maintenance of structures but we cannot depend upon conventional strategies of detection. Manual methods are that often used is error prone and need considerable time and effort. Conventional damage detection approaches does not produce better results in damage detection. So the need of automation of visual inspections increased efficiency in damage assessment.

Crack Detection and Classification in Asphalt Pavement Images Using Deep Convolution Neural Network [26]

This paper is about study proposed a deep convolution neural network (DCNN) as a detection of asphalt pavement crack that capable to detect and classify the pavement crack robustly when dealing with complexity background image. A method to detect and classify asphalt pavement crack using deep CNN is proposed. This proposed deep CNN architecture composed of eight layers consist of three convolution layer, three pooling layer and two fully connected layer. The network is trained using stochastic gradient (SGD) training algorithm.

Crack detection and classification comprises steps – 1) Image acquisition, 2) image processing and labeling 3) crack detection and 4) crack classification. In order to explore how the size of input image affects deep CNN with respect to detection and classification performance, different sizes and grid scale are adopted in the architecture. The network achieved overall recall, precision and accuracy on the 1000 testing images are 98.0%, 99.4% and 99.2% respectively.

An Extraction and Classification Algorithm for Concrete Cracks Based on Machine Vision [27]

To solve the problem of large errors in extraction and the difficulty in classifying crack images in health monitoring of civil engineering structures, a new classification algorithm of concrete crack extraction based on machine vision is proposed in this paper.

Nonlinear grayscale transformation, improved OSTU threshold segmentation and then bifurcation points are all used in algorithm. . The obtained features are used as input to train a support vector machine classifier, which is then used to perform crack classification This algorithm had a good ability to extract cracks with obvious contours, but had difficulty with fine cracks, used a multiscale linear filtering approach based on the Hessian matrix to enhance the crack area and then determined the threshold interval according to the maximum entropy of the image histogram to obtain the crack binarization segmentation result.

Crack and Noncrack Damage Automatic Classification from Concrete Surface Images using Broad Network Architecture [28]

In this paper, an automatic crack damage classification method using broad network architecture, it extracts feature nodes from input raw data by convolution function or other mapping transformation, flatted enhancement nodes are combined with feature nodes to construct broad network architecture, finally a decision threshold is set to obtain a binary classification output. Methodology -The Principle of Broad Learning Algorithm a broad network is flatted in the width instead of the deep structure by multiple feature nodes and enhancement nodes. Dataset – It consists of 40000 concrete images with crack (positive) and noncrack

(negative) .Each class has 20000 images with $227 \times 227 \times 3$ RGB pixels [L. Zhang, F. Yang, and D. Zhang, et, al. Road crack detection using deep convolutional neural network. In 2016 IEEE International Conference on Image Processing (ICIP). 2016:3708-3712. this method avoids hyper parameter adjustment and complicated deep structure.

Application of Internet of Things Technology and Convolutional Neural Network Model in Bridge Crack Detection [29]

This paper defines the damage structure of the bridge structure as a comprehensive perception of the damage situation of the structure through the information sensing equipment, the ubiquitous interconnection of structural security impact factors (load, displacement, service life, use environment, etc.). Methodology – Convolution neural Network (CNN) the set of small neurons in the convolution neural network is connected with a small region of the input image. Application of convolution neural network in classification of bridge cracks consist image preprocessing, the establishment of convolution neural network model, and the example analysis. Studied a digital and intelligent bridge crack detection method to improve the efficiency of bridge safety diagnosis and reduced the risk factor. Firstly, the collected bridge crack pictures were preprocessed, the bridge crack convolution neural network classification model was established, and the model was simulated and trained using MATLAB. The bridge crack classification was obtained. This method could effectively solve the problems of low fracture diagnosis efficiency and high risk factor in domestic fractures.

A Deep Convolutional Neural Network for Semantic Pixel-Wise Segmentation of Road and Pavement Surface Cracks [30]

In this paper propose a deep fully convolutional neural network to perform pixel-wise classification of surface cracks on road and pavement images. The network consists of an encoder layer which reduces the input image to a bank of lower level feature maps. This is followed by a corresponding decoder layer which maps the encoded features back to the resolution of the input data using the indices of the encoder pooling layers to perform efficient up-sampling. This paper presents an algorithm for semantic segmentation of road and pavement surface cracks using a Convolutional Neural Network, namely U-Net. The algorithm is trained, validated and tested on the publicly available CrackForest. Dataset which consists of 118 images of surface cracks on pavement and road surfaces, taken with a hand-held camera. The patch based training method proposed here will be extended to include augmentations of the input data in an attempt to equally represent cracks running at multiple angles.

Road Crack Detection Using Deep Convolutional Neural Network and Adaptive Thresholding [31]

Crack detection is performed by either certified inspectors or structural engineers. This task is, time-consuming, subjective and labor-intensive. In this paper, a novel road crack detection algorithm which is based on deep learning and adaptive image segmentation is proposed. Dataset [<http://dx.doi.org/10.17632/5y9wdsg2zt.1>]. In this experiments, the proposed deep neural network is trained on an NVIDIA GTX 1080 Ti GPU1, which has 3584 CUDA cores and 11 GB GDDR5X memory. The GPU memory bandwidth is 484 GB/s. The dataset utilized for training the proposed network was created by the researchers from Middle East Technical University. The dataset contains 40000 RGB images (resolution: 227×227). The number of positive and negative images are both 20000. The experimental results illustrate that our network can classify

images with an accuracy of 99:92%, and the cracks can be successfully extracted from the images using our proposed thresholding algorithm.

Design Application of Deep Convolutional Neural Network for Vision-Based Defect Inspection [32]

In this paper, a design application of DCNN (Deep Convolutional Neural Network) is considered and developed for vision-based defect inspection. As a trial test, three kinds of DCNNs are designed, implemented and tested to inspect small defects, such as, crack, burr, protrusion, chipping and spot phenomena seen in the manufacturing process of resin molded articles. An image generator is also implemented to systematically generate range of relevant deformed version of similar images for training.

The DCNN based design application is described to evaluate the usefulness for image defect detection. Setting parameters for the training of the DCNN. The training was conducted using single PC with a Core i7 CPU and a GPU (NVIDIA GeForce GTX 1060).

The designed DCNNs are trained using the generated images and then evaluated through classification experiments. The proposed DCNN design application is planned to be applied to actual physical inspection processes.

A Bridge Crack Image Detection and Classification Method Based On Climbing Robot [33]

Develop a bionic climbing robot which can climb on rough surface due to its smart structure and bionic design. After loading simple camera, wall images can be acquired wirelessly in real time, which is suitable for health detection of bridge structure. But, due to the small load of the robot, both the size and precision of its camera are limited, which leads to the lower quality of obtained pictures. The goal of this system is to obtain crack pictures of bridge surface, use algorithms to make up for the deficiency of hardware precision through a series of image processing methods, divide complete crack samples. Based on the visual and geometrical characteristics, a decision-tree-based multiclass support vector machine algorithm is applied to classify crack target.

Grid-based Pavement Crack Analysis Using Deep Learning [34]

In this paper segment the pavement crack images into different scales of grids. We choose suitable scale of grids to segment image. We design the structure of CNN to detect pavement crack. To classify the type of crack, we utilize principal component analysis (PCA) to calculate the distribution of grids. After analysis the distribution of grids, we achieve the pavement crack classification an image is cut into many non-overlapping grids, and then we use CNN to detect the existence of crack. We only keep the grids that containing crack so that the skeleton of crack can be preserved. The pavement images are collected by author along the Youyi Avenue in Wuhan city. There were 510 pavement images which including three kinds of crack pictures and non-crack pictures. Took 30000 grid images as training set images. The rest of images were testing set images. The correct rate of classification is increased compared with pavement crack classification using neural network.

A Genetic Algorithm for Convolutional Network Structure Optimization for Concrete Crack Detection [35]

GA has been applied to a variable depth CNN. This means that in the proposed method, the GA evolves the network depth, the layer size, and the hyper parameters of the network. The size of layers in the CNN affects the level of details that are recognized by the CNN, a case where GAs

have excelled previously. The proposed method utilizes a GA for CNN structure optimization. This method evolved several parameters that dictate the structure of a CNN, including: number of convolution layers, size of convolution filters, and number of convolution filters used in each layer. Evolving CNN structures produces high-performance networks that have higher classification accuracy than the state-of-the-art network when tested on images of concrete containing cracks. This process allows for a small training set, since mass training data is sometimes difficult to obtain in real-world applications. The process also performs the search for network structures automatically, which removes the need for a deep knowledge of the features being described and the neural network design process. Visual results indicate that the generated networks perform.

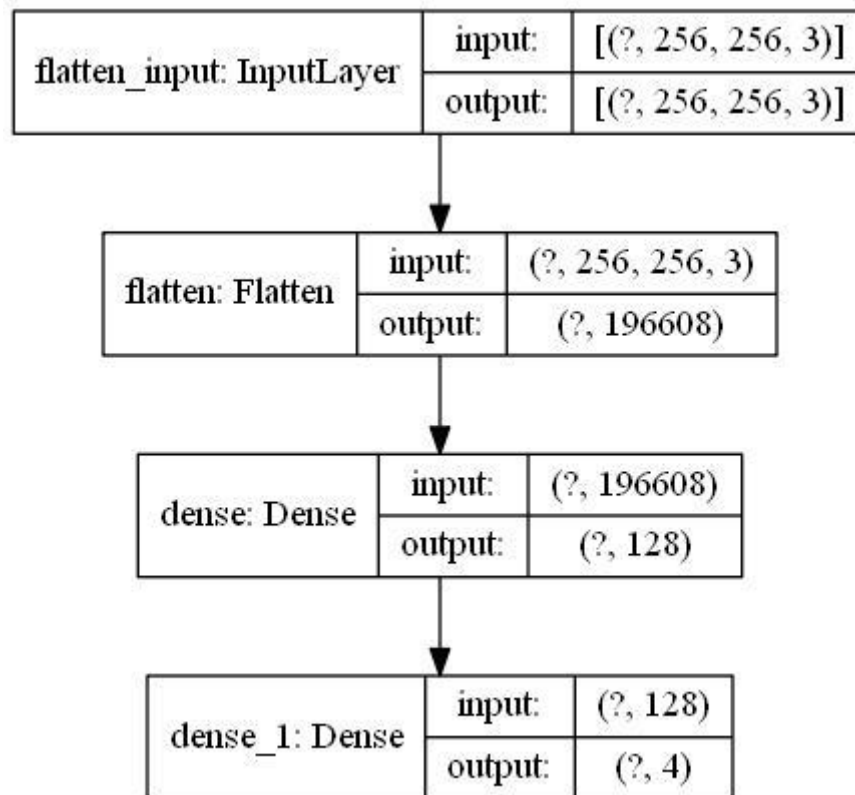
Gap analysis:

[1-8, 14-15, 25, 27, 33] have stated digital image processing technique for crack detection. Instead of using image processing, using deep learning techniques can automatically learn complex features that need not be hardcoded and requires less computation once the model is trained. [1-2, 5-8, 10-12, 14, 16-17, 25-32, 34-35] papers' methods require manual capturing of images or automatic capturing by fixing the camera in one location. Using Unmanned Aerial Vehicle (UAV) for capturing input data would be useful to capture dynamic data from human inaccessible areas. [1-8, 20-35] papers are working on image data. We can extend these work to video data. [1-8, 26-31, 33-35] papers are detecting only cracks in their work. But there are other structural defects in the infrastructure that can be detected visually including dampness of wall and peeling off of paint. [9-24] papers are working on other fields of image and video segmentation. We can extend their work in the field of structural audit.

Problem statement:

The proposed system should be able to capture video images using UAV, which will be processed for detection of cracks, dampness and paint peel-off using neural network as a part of visual inspection while doing structural audit of infrastructures.

Proposed Solution:



We are building a neural network whose architecture is given in the above image. The input image is of dimension (256, 256) having RGB channels. We had generated huge amount of data (~160k) from small set of data (80 images). The model overfitted with training accuracy of 99% and testing accuracy 25%. Currently we are adding more variety to the input data and applying data augmentation techniques like rotation, flipping and similar operations on the data to increase the data with increase in variety in training data.

References:

- [1] S. Kang, C. Chun, S. Shim, S. Ryu and J. Baek, "Real Time Image Processing System for Detecting Infrastructure Damage: Crack," 2019 IEEE International Conference on Consumer Electronics (ICCE), Las Vegas, NV, USA, 2019, pp. 1-3.
- [2] J. Vora, M. Patel, S. Tanwar and S. Tyagi, "Image Processing Based Analysis of Cracks on Vertical Walls," 2018 3rd International Conference On Internet of Things: Smart Innovation and Usages (IoT-SIU), Bhimtal, 2018, pp. 1-5.
- [3] R. K. Meghana, S. Apoorva, Mohana and Y. Chitkara, "Inspection, Identification and Repair Monitoring of Cracked Concrete structure –An application of Image processing," 2018 3rd International Conference on Communication and Electronics Systems (ICCES), Coimbatore, India, 2018, pp. 1151-1154.

- [4] L. Peng and J. Liu, "Detection and analysis of large-scale WT blade surface cracks based on UAV-taken images," in *IET Image Processing*, vol. 12, no. 11, pp. 2059-2064, 11 2018.
- [5] M. Lin, R. Zhou, Q. Yan and X. Xu, "Automatic Pavement Crack Detection Using HMRF-EM Algorithm," 2019 International Conference on Computer, Information and Telecommunication Systems (CITS), Beijing, China, 2019, pp. 1-5.
- [6] L. S. Calderón and J. Bairán, "Crack detection in concrete elements from RGB pictures using modified line detection Kernels," 2017 Intelligent Systems Conference (IntelliSys), London, 2017, pp. 799-805.
- [7] W. Li, M. Zhang, Z. Shen, W. Hu and P. Li, "Track Crack Detection Method in Complex Environment," 2018 11th International Symposium on Computational Intelligence and Design (ISCID), Hangzhou, China, 2018, pp. 356-359.
- [8] X. Liu, Y. Ai and S. Scherer, "Robust image-based crack detection in concrete structure using multi-scale enhancement and visual features," 2017 IEEE International Conference on Image Processing (ICIP), Beijing, 2017, pp. 2304-2308.
- [9] Y. Wang, Y. Lyu, Y. Cao and M. Y. Yang, "Deep Learning for Semantic Segmentation of UAV Videos," *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*, Yokohama, Japan, 2019, pp. 2459-2462.
- [10] H. Shi, Z. Liu, Y. Fan, X. Wang and T. Huang, "Effective object detection from traffic camera videos," 2017 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computed, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI), San Francisco, CA, 2017, pp. 1-5.
- [11] P. V. Babayan, M. D. Ershov and D. Y. Erokhin, "Neural Network-Based Vehicle and Pedestrian Detection for Video Analysis System," 2019 8th Mediterranean Conference on Embedded Computing (MECO), Budva, Montenegro, 2019, pp. 1-5.
- [12] B. Hou, J. Li, X. Zhang, S. Wang and L. Jiao, "Object Detection and Trcacking Based on Convolutional Neural Networks for High-Resolution Optical Remote Sensing Video," *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*, Yokohama, Japan, 2019, pp. 5433-5436.
- [13] K. Mouri, H. Lu, J. K. Tan and H. Kim, "Object Detection on Video Images Based on R-FCN and GrowCut Algorithm," 2018 International Conference on Information and Communication Technology Robotics (ICT-ROBOT), Busan, 2018, pp. 1-4.
- [14] R. C. Joshi, M. Joshi, A. G. Singh and S. Mathur, "Object Detection, Classification and Tracking Methods for Video Surveillance: A Review," 2018 4th International Conference on Computing Communication and Automation (ICCCA), Greater Noida, India, 2018, pp. 1-7.

- [15] Q. Wei, S. Lao and L. Bai, "Panorama Stitching, Moving Object Detection and Tracking in UAV Videos," 2017 International Conference on Vision, Image and Signal Processing (ICVISP), Osaka, 2017, pp. 46-50.
- [16] B. Tian, L. Li, Y. Qu and L. Yan, "Video Object Detection for Tractability with Deep Learning Method," 2017 Fifth International Conference on Advanced Cloud and Big Data (CBD), Shanghai, 2017, pp. 397-401.
- [17] M. H. Gasmallah and F. Zulkernine, "Video Predictive Object Detector," 2018 IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), Vancouver, BC, 2018, pp. 365-371.
- [18] A. Anjum, T. Abdullah, M. F. Tariq, Y. Baltaci and N. Antonopoulos, "Video Stream Analysis in Clouds: An Object Detection and Classification Framework for High Performance Video Analytics," in IEEE Transactions on Cloud Computing, vol. 7, no. 4, pp. 1152-1167, 1 Oct.-Dec. 2019.
- [19] C. Wu et al., "UAV Autonomous Target Search Based on Deep Reinforcement Learning in Complex Disaster Scene," in IEEE Access, vol. 7, pp. 117227-117245, 2019.
- [20] R. Geraldies et al., "UAV-Based Situational Awareness System Using Deep Learning," in IEEE Access, vol. 7, pp. 122583-122594, 2019.
- [21] N. Q. Truong, P. H. Nguyen, S. H. Nam and K. R. Park, "Deep Learning-Based Super-Resolution Reconstruction and Marker Detection for Drone Landing," in IEEE Access, vol. 7, pp. 61639-61655, 2019.
- [22] C. Wang, J. Wang, X. Zhang and X. Zhang, "Autonomous navigation of UAV in large-scale unknown complex environment with deep reinforcement learning," 2017 IEEE Global Conference on Signal and Information Processing (GlobalSIP), Montreal, QC, 2017, pp. 858-862.
- [23] S. A. Hassan, T. Rahim and S. Y. Shin, "Real-time UAV Detection based on Deep Learning Network," 2019 International Conference on Information and Communication Technology Convergence (ICTC), Jeju Island, Korea (South), 2019, pp. 630-632.
- [24] A. Zeggada, F. Melgani and Y. Bazi, "A Deep Learning Approach to UAV Image Multilabeling," in IEEE Geoscience and Remote Sensing Letters, vol. 14, no. 5, pp. 694-698, May 2017.
- [25] S. Koshy, B. Radhakrishnan and L. P. Suresh, "Strength analysis of buildings using image processing and SHM principles," 2016 International Conference on Emerging Technological Trends (ICETT), Kollam, 2016, pp. 1-6.
- [26] N. A. M. Yusof, M. K. Osman, M. H. M. Noor, A. Ibrahim, N. M. Tahir and N. M. Yusof, "Crack Detection and Classification in Asphalt Pavement Images using Deep Convolution

Neural Network," 2018 8th IEEE International Conference on Control System, Computing and Engineering (ICCSCE), Penang, Malaysia, 2018, pp. 227-232.

[27] S. Liang, X. Jianchun and Z. Xun, "An Extraction and Classification Algorithm for Concrete Cracks Based on Machine Vision," in IEEE Access, vol. 6, pp. 45051-45061, 2018.

[28] L. Guo, R. Li, X. Shen and B. Jiang, "Crack and Noncrack Damage Automatic Classification from Concrete Surface Images using Broad Network Architecture," 2019 Chinese Control And Decision Conference (CCDC), Nanchang, China, 2019, pp. 1966-1971.

[29] L. Zhang, G. Zhou, Y. Han, H. Lin and Y. Wu, "Application of Internet of Things Technology and Convolutional Neural Network Model in Bridge Crack Detection," in IEEE Access, vol. 6, pp. 39442-39451, 2018.

[30] M. David Jenkins, T. A. Carr, M. I. Iglesias, T. Buggy and G. Morison, "A Deep Convolutional Neural Network for Semantic Pixel-Wise Segmentation of Road and Pavement Surface Cracks," 2018 26th European Signal Processing Conference (EUSIPCO), Rome, 2018, pp. 2120-2124.

[31] R. Fan et al., "Road Crack Detection Using Deep Convolutional Neural Network and Adaptive Thresholding," 2019 IEEE Intelligent Vehicles Symposium (IV), Paris, France, 2019, pp. 474-479.

[32] F. Nagata, K. Tokuno, K. Watanabe and M. K. Habib, "Design Application of Deep Convolutional Neural Network for Vision-Based Defect Inspection," 2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Miyazaki, Japan, 2018, pp. 1705-1710.

[33] Y. Chen, T. Mei, X. Wang and F. Li, "A bridge crack image detection and classification method based On climbing robot," 2016 35th Chinese Control Conference (CCC), Chengdu, 2016, pp. 4037-4042.

[34] X. Wang and Z. Hu, "Grid-based pavement crack analysis using deep learning," 2017 4th International Conference on Transportation Information and Safety (ICTIS), Banff, AB, 2017, pp. 917-924

[35] S. Gibb, H. M. La and S. Louis, "A Genetic Algorithm for Convolutional Network Structure Optimization for Concrete Crack Detection," 2018 IEEE Congress on Evolutionary Computation (CEC), Rio de Janeiro, 2018, pp. 1-8.