Automation in Construction 95 (2018) 155–171

Contents lists available at ScienceDirect

Automation in Construction

journal homepage: www.elsevier.com/locate/autcon

# Automated detection of sewer pipe defects in closed-circuit television images using deep learning techniques

Jack C.P. Cheng, Mingzhu Wang*

*Department of Civil and Environmental Engineering, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong*

ARTICLE INFO

*Keywords:*
Sewer pipe inspection
Defect detection
Deep learning
Computer vision
Faster region-based convolutional neural network (faster R-CNN)
Object detection

ABSTRACT

Sanitary sewer systems are designed to collect and transport sanitary wastewater and stormwater. Pipe inspection is important in identifying both the type and location of pipe defects to maintain the normal sewer operations. Closed-circuit television (CCTV) has been commonly utilized for sewer pipe inspection. Currently, interpretation of the CCTV images is mostly conducted manually to identify the defect type and location, which is time-consuming, labor-intensive and inaccurate. Conventional computer vision techniques are explored for automated interpretation of CCTV images, but such process requires large amount of image pre-processing and the design of complex feature extractor for certain cases. In this study, an automated approach is developed for detecting sewer pipe defects based on a deep learning technique namely faster region-based convolutional neural network (faster R-CNN). The detection model is trained using 3000 images collected from CCTV inspection videos of sewer pipes. After training, the model is evaluated in terms of detection accuracy and computation cost using mean average precision (mAP), missing rate, detection speed and training time. The proposed approach is demonstrated to be applicable for detecting sewer pipe defects accurately with high accuracy and fast speed. In addition, a new model is constructed and several hyper-parameters are adjusted to study the influential factors of the proposed approach. The experiment results demonstrate that dataset size, initialization network type and training mode, and network hyper-parameters have influence on model performance. Specifically, the increase of dataset size and convolutional layers can improve the model accuracy. The adjustment of hyper-parameters such as filter dimensions or stride values contributes to higher detection accuracy, achieving an mAP of 83%. The study lays the foundation for applying deep learning techniques in sewer pipe defect detection as well as addressing similar issues for construction and facility management.

## 1. Introduction

Sewer pipe systems form an important component of civil infrastructure which is designed to collect and transport wastewater, stormwater and groundwater to treatment facilities. In the United States, there are over 800,000 miles of sewers and 500,000 miles of private lateral sewers that connect private property to public sewer pipes [1]. It was estimated that there are 19,500 sewer systems for handling an average daily flow of roughly 50 billion gallons of raw sewage but most of the sewer systems are between 30 and 100 years old [2], making them susceptible to structure failure, blockage and overflows. It was reported that there are at least 23,000 to 75,000 sewer sanitary overflows (SSOs) in the United States every year, which can release 3 billion to 10 billion gallons of untreated wastewater [3], leading to water quality problems, property losses and threats to public health. Over 50% of the reported SSOs are attributed to pipe blockage, which are mainly caused by grease and debris. Other reasons include root intrusion, line breaks and imbalanced amount of inflow and infiltration. In addition, large amount of budget is needed for the maintenance of sewer pipe systems. For example, it was estimated that US $297 billion will be required for wastewater infrastructure over the next 25 years [1]. According to the defect severity and distribution characteristic, there may be different sub-types under each category. For example, cracks can be divided into longitudinal cracks (with the same orientation of the pipeline), vertical cracks (whose orientation is vertical to pipeline) as well as complex cracks consisting of both longitudinal and vertical ones [4]. Deposits have different types such as attached deposits (which are attached to the pipe wall) and settled deposits (which are settled on the pipe ground) [5]. As for the tree root intrusion, it can belong to mass root (with high density or large occupying area) or minor root (which are scattered and with low density). Considering the serious consequences caused by those defects such as

* Corresponding author.
*E-mail address:* mwangaz@connect.ust.hk (M. Wang).

https://doi.org/10.1016/j.autcon.2018.08.006
Received 3 April 2018; Received in revised form 29 June 2018; Accepted 11 August 2018

cracks, deposits, tree root intrusions as well as water infiltrations, it is important to detect and evaluate the defects at an early stage such that further pipe deterioration can be avoided, and existing defects can be repaired to maintain normal sewer operations.

Currently, visual inspection techniques such as closed-circuit television (CCTV) have been commonly utilized for underground sewer pipe inspection. A CCTV usually consists of a camera and an illumination device mounted on a tractor. During the inspection, the CCTV unit moves along the interior pipe wall and transmits the inspection video to an external monitor on the ground. When encountering a pipe defect or pipe lateral, the inspector would stop the unit and zoom the camera into the abnormal part to check if there are potential defects. After the inspection, the inspector needs to watch the captured images or videos to identify the defect type and location. Such manual interpretation of the inspection images or videos is time-consuming, labor intensive and the results can be subjective and inaccurate.

There is recently a trend of applying computer vision for interpreting the inspection images or videos automatically. However, conventional computer vision techniques require designing complex feature extractors and images used for training need a large amount of pre-processing. In addition, the training process is tedious and inefficient. In recent years, deep learning has obtained promising performance in various computer vision tasks such as image classification and object detection. Compared with conventional computer vision techniques, approaches based on deep learning are capable of extracting image features automatically and there is not much requirement of image pre-processing, which improves the accuracy and efficiency greatly. Therefore, an automated defect detection approach is proposed in this study based on faster region-based convolutional neural network (faster R-CNN), which is a deep learning model for object detection, for identifying and locating sewer pipe defects from CCTV images.

This article is organized as follows. Related works are reviewed in Section 2. The components of the developed defect detection approach including the applied model, model training and evaluation are introduced in Section 3. The setup and results of three experiments for validating the proposed approach are elaborated in Section 4, followed by related discussions in Section 5. Conclusions, limitations and future work are provided in Section 6.

## 2. Related work

An automated approach is proposed in this study for identifying and locating multiple sewer pipe defects in CCTV images using deep learning to address the limitations of conventional computer vision methods. Therefore, related studies on conventional computer vision techniques for inspecting civil infrastructure and deep learning based approaches are reviewed in Section 2.

### *2.1. Conventional computer vision techniques for civil infrastructure inspection*

With the wide development and application of visual inspection techniques for infrastructure, such as CCTV robots and unmanned aerial vehicles (UAVs), large number of inspection images or videos are generated during the inspection. Manual interpretation to obtain inspection results is required, which is inefficient and ineffective. Computer vision is the capability of a computer or a machine to obtain understandings like human beings from digital images or videos [6]. Therefore, there are potentials of computer vision techniques for addressing the limitation of manual interpretation of images or videos from visual inspections of civil infrastructures. Image processing techniques such as edge detection and morphological operations were mainly studied previously [7]. Although the noise reduction methods [8] were studied, prior knowledge of the images is still required for image processing, which is complex especially with images taken in different environment.

Generally, applications of conventional computer vision methods include image pre-processing, image segmentation, feature extraction, object recognition and structural analysis [9,10]. Various vision-based tasks have been studied for civil infrastructure inspection including automated detection and dimension measurement of concrete cracks [11–13], recognition of damage pattern changes and 3D visualization of cracks [14], precise crack extraction with less model computation cost [15]. In addition, 3D reconstruction of road surface distresses [16] and power line corridors as well as automatic recognition of obstacles [17] were also investigated based on UAV images. Computer vision techniques have also been applied for efficient underground sewer pipe inspection, such as the segmentation and classification of sewer pipe images [18–20] to obtain joints, laterals and defects for condition assessment. Another application of image processing techniques for underground utility inspection is the automatic detection and segmentation of cracks [21–23]. Several challenges were encountered and discussed, such as dealing with images with low resolution and noisy data [24,25], image distortion and structure motion [26] and the influence of illumination and shooting distance [27].

However, one main limitation of the most conventional computer vision techniques applied in the aforementioned studies is the need of designing complex feature extractors which is only suitable for one particular task. Another problem is that large amount of image pre-processing is required when preparing the training dataset and the training process is tedious. In addition, previous studies mainly focus on identification and property retrieval of single defect (e.g. cracks) and there has been limited research on the automated identification and localization of other common defects of underground sewer pipelines, such as tree root intrusion and water infiltration, which can lead to severe consequences for sewer pipe networks.

### *2.2. Deep learning based approaches for image classification and object detection*

#### *2.2.1. Deep learning for image classification*

Deep learning has been widely developed and applied in various areas such as computer vision, speech recognition and natural language processing through various deep learning architectures, among which convolutional neural networks (CNNs) are commonly applied [28]. As shown in Fig. 1, a CNN model typically consists of feature extraction through a stack of layers on the input image such as convolution, activation and poling, and classification through fully connected layers for outputting the scores for each class. Each layer is responsible for different functions and uses the result from the previous layer as the input. For supervised computer vision tasks, CNNs (1) extract features from raw images, (2) feed the features forward using filters assigned with initial random weights and bias to predict the classes, (3) calculate the loss between predicted scores and the ground truth, (4) apply backpropagation to adjust the filter weights and bias continuously to finally obtain an optimized model. Details of the convolution process are introduced in Section 3.1.2.

Compared with conventional approaches, CNNs require less image pre-processing, and image features are extracted through learning. Therefore, there is no requirement of expertise for manual design of complex feature extractors. The promising performance of CNNs in various computer vision tasks such as image classification, object recognition and localization in recent years indicates the great potentials of CNNs for automated defect detection for civil infrastructures. Currently, there have been some studies attempting to apply deep learning approaches for civil infrastructure. For example, CNN based approaches are proposed for detecting concrete cracks and the advantages of the proposed model over conventional image processing techniques have been proven [29,30]. In addition, CNNs have also been applied for efficient tunnel inspection [31] and road crack detection [32]. However, the application of CNNs for infrastructure inspection is still in its infancy and there has been limited research using CNNs for