Machine learning and computer vision for urban drainage inspections

Meijer, D.W.J.

English Summary

Sewer pipes are an essential infrastructure in modern society and their proper operation is important for public health. To keep sewer pipes operational as much as possible, periodical inspections for defects are performed. Instead of repairing sewer pipes when a problem becomes critical, such inspections allow municipalities to plan maintenance. This means the disruptions of the service can be planned for by users of the pipe in question, and there is less chance that a problem slips by unnoticed.

Sewer pipe inspections are generally performed visually with the aid of a *pipe inspection gadget*, or PIG. The PIG is a remote-controlled vehicle equipped with cameras and possibly other sensors. The PIG is lowered into a manhole to inspect a stretch of pipe, after which it is returned to the surface. A trained human operator inspects the footage recorded by the cameras, often while controlling the PIG from a vehicle at ground level.

Inspection reports are made according to a European classification norm. This norm groups defects of a similar nature together and has guidelines for what constitutes ratings from 1 ("no intervention necessary") to 5 ("immediate intervention necessary"). Problematically, these guidelines consider defects in a vacuum. Take a fissure in the wall of a pipe for example, the guideline assigns a rating 1 to 5 to different ranges of fissure sizes. The actual consequences depend on many more factors, such as whether the pipe is above or below the groundwater level, the zoning district the pipe is located in, etc. As a result, operators have learned to assign ratings not according to the guidelines, but according to an intuitive assessment of severity. This, in turn, means that severity ratings can vary wildly between operators, and even between inspections by the same operator.

This makes sewer pipe inspections an attractive target for automation. While a potential improvement in terms of assessment quality and processing efficiency is generally promised by automation, in this case we would also decrease the variability which is a current problem. Besides the reasons for automating, the methods for automating are also attractive: a lot of (visual) data has been gathered over the past decades which may be used to train algorithms.

This thesis compiles the results of five years of research into the possible automation of sewer pipe inspections with the tools of machine learning and computer vision. In this thesis, three distinct, yet complementary approaches to automating sewer pipe inspections are described.

Chapter 3, *Image-Based Unsupervised Anomaly Detection*, describes an approach based on anomaly detection of the contents of the images. At this stage, the data that was available to us consisted of images from inspections performed in two Dutch municipalities. The inspection reports themselves were not available at that time, meaning it was unclear which images were showing defects and which were not. While more complete data became

available at a later date, at this stage we decided to leverage the image data that we did have.

The structure of the different images is very similar: the pipes were photographed with the same equipment, and the pipes from the same municipality were often installed in the same year, from the same manufacturer, and had seen similar use. This resulted in a delineation of two image sets, one of images of pipes made of smooth concrete, one of images of pipes made of granulate. Within either of the image sets, the images look mostly uniform, meaning that anomalies—both expected (such as pipe joints) and unexpected (such as defects)—stand out.

We applied principal component analysis to the images and extracted features from the images, to detect the most common elements in an image set. Then, when we express an image in those most common elements, we obtain a faithful reconstruction for the images that do not contain any anomalies, and a less perfect reconstruction for the images that do contain anomalies. Leveraging this reconstruction error, we compare the reconstruction to the original image to estimate how likely it is to contain an anomaly.

In addition, we trained a convolutional autoencoder, a type of artificial neural network, to perform a function similar to the principal component analysis, without enforcing a linear relation of the common elements.

The results of these experiments were promising for the images of pipes made from smooth concrete, but less so for the images of pipes made from the rougher granulate.

Chapter 4, *Convolutional Neural Network Classification*, describes an approach based on supervised classification with a convolutional neural network. Convolutional neural networks are artificial neural networks that are particularly suited to handle images, audio and video. We were provided with sewer pipe images like the ones used in chapter 3, but a much larger volume and including machine-readable classifications as assigned by human operators. A total of 2.2 million images were available and the classification data allowed us to estimate what defects should be visible in any given image. A single neural network was trained to detect the twelve most common defect types in the dataset.

The problem of sewer pipe defect detection a strongly unbalanced one: only approximately 1 % of images actually contain defects. Most of the existing literature at the time was assessing the performance of their models in terms of accuracy, the fraction of correctly classified images, both as having, or not having, a defect. On a realistic dataset, an accuracy of 99 % is then to be expected if we classify every image as not having any defects, which is clearly not the intention of defect detection. To counter this, many works rebalanced the dataset to contain about 50 % images with defects. While this is not per se a bad idea, nearly every one of them also rebalanced the test set that was used to assess the performance of the model, making the assessed performance not at all indicative of actual, real-world performance. Many also treated false positive and false negative detections identically, while these have very different results in a realistic scenario: the former costs time, the latter might pose a public health hazard.

Many earlier works randomly divide images of pipes into training and test set, meaning that images of the same pipe at locations close to one another might end up in both training and test set. This introduces a danger of data leakage: a high performance on the test set might not necessarily mean that the defects themselves are being detected, but could rather mean that the pipe is being detected.

To have any real-world meaning, we asserted that the test set used to assess the model must be as realistic as possible, including having a realistic ratio of images with and without defects, and only containing pipes that were not also used to train the model. We also approach the problem from a more context-sensitive perspective, noting that accuracy is not a useful metric in realistic situations, and introducing metrics that can be more meaningfully interpreted by a human, as well as translate more directly to operational impact.

Chapter 5, *Stereovision and Geometry Reconstruction*, extends beyond the current sewer pipe inspection process and investigates the added value of a second camera, allowing us to reconstruct the three-dimensional geometry of the sewer pipe. Much like how human beings perceive depth only with both eyes open, a second camera allows us to estimate the positions of objects in relation to the viewpoint.

In collaboration with Eindhoven University of Technology, we photographed 26 sewer pipes in various conditions with a set of two side-by-side cameras. We built upon existing stereovision techniques and adapted them for this unique use case to reconstruct a three-dimensional point cloud of the pipe's inner surface.

A pipe surface model is constructed under the assumption that the cameras are aligned approximately along the pipe center. The model is powerful enough to capture the geometry of any of the pipes we have used, but also based on human understanding of the shape of a pipe, making it very interpretable.

The model is fit to the point cloud to estimate the original pipe geometry, without taking into account minor deviations that are visible in the pipes after years of use. This allows us to easily detect the portions of the pipe where the surface wall is deviating from the expected shape. The detected deviating portions of the surface correlate with the presence of actual defects in this small-scale experiment. The end result is an interpretable computer vision technique that can be used to assist human-guided inspections.