Machine learning and computer vision for urban drainage inspections

Meijer, D.W.J.

6 Discussion and Conclusions

This thesis has explored applications of machine learning and computer vision to automate and enrich urban drainage inspections. This chapter will provide a conclusion to the thesis by answering the six research questions posed in chapter 1 and ending with some closing remarks.

Q1

What knowledge can be obtained from available inspection data without the utilization of expert classification, which might be inconsistent or unavailable?

In chapter 3, unlabeled sewer CCTV images were analysed with unsupervised learning. Image patches were classified as anomalous or non-anomalous, based on how common elements of the image patches were in the larger dataset. The use of image feature extractors and PCA decomposition allows us to detect anomalies based on the variance in the dataset they explain.

Urban drainage inspection is in fact a problem that lends itself well to unsupervised learning: because defects are very uncommon, they can be treated as anomalies in an anomaly detection problem more easily. The extreme class imbalance [1] works in our favour in this instance.

[1] Less than 1 % of images contain defects

It must be noted that these anomalies are not all defects, and that it is not trivial to separate defect and non-defect anomalies. The way we extract knowledge from unlabeled data with this approach is negative classification: images with no anomalous patches are very unlikely to contain defects. Because of this, such knowledge can be used as preselection for a later classification stage, regardless of whether that classification will be performed by humans or other algorithms.

How can the data collected with current inspection practices be analysed with machine learning techniques in order to improve processing efficiency and accuracy?

Q2

In chapter 4, we have trained a convolutional neural network to perform classifications of defects as human operators would. We have demonstrated that it may be possible to adequately perform future classification in an automated manner, provided enough varied training data is available. The problem is not a trivial one however. There is a rather extreme class imbalance and the human-labeled training data is known to have some errors.

The class imbalance leads to a dilemma: training a classification model without regard for the imbalance means the model might not adequately learn how to classify the extremely uncommon defects as they make up such a small portion of the variance present in the dataset. At the same time, adapting the training set or classification algorithm to make sure the under-represented class is properly classified introduces a bias that will decrease performance on the majority class in future, unclassified data.

And because the inspection quality is lacking, [2] the labels are not entirely reliable. The fact that the labels themselves are known to have errors leads to a limitation of a model's capabilities: we can scarcely expect the model to outperform its training data. The answer to the research question then, is that supervised machine learning could at best reach human parity with current data collection practices, which would provide an improvement to processing efficiency without (too much) reduction in quality. This in itself could be a more important improvement than it seems to be at first glance. Not only does a menial, repetitive task not have to be performed manually anymore, the consistency with which it is performed when automated is also increased.

[2] Dirksen, J., Clemens, F., Korving, H., Cherqui, F., Le Gauffre, P., Ertl, T., Plihal, H., Müller, K., and Snaterse, C. 2013. The consistency of visual sewer inspection data. *Structure and Infrastructure Engineering 9*, 3, 214–228

Q3

How do we assess the quality and operational impact of (partial) automation of the current inspection practices?

First and foremost, meaningful assessment requires that the data used to test the model be as realistic as possible: no rebalancing of datasets, no images of pipes that were present in the training set, no zooming and panning of the camera to better frame suspected defects. This might sound obvious in the context of this thesis, but published and peer-reviewed works have missed such crucial details in the past.

Secondly, and as discussed at length in chapter 4, commonly used quality metrics such as accuracy are of limited use for this problem. On the one hand, the extreme class imbalance makes some metrics difficult to interpret: an accuracy value of 99 % might look impressive in most cases, but we easily could achieve this by classifying every image as not containing a defect. On the other hand, because properly operating urban drainage systems are essential to public health and infrastructure, there are limits to the amount of false negative classifications that can be acceptable, regardless of the metric that is being used.

We have put forward two quality metrics specifically for this use case, that may provide more insight into the usefulness of classification models: the precision-at-recall and the specificity-at-recall. These metrics allow the user to define a minimum acceptable recall, and report the precision or specificity achieved if we tune the model output to achieve at least that recall value. Which of the two to use depends on the context: are we interested in knowing the amount of false positive detections as a fraction of all detections (use precision-at-recall), or as a fraction of all positives (use specificity-at-recall)?

To estimate operational impact, we need a clear picture of how the model will be used in practice. Assuming that false positive detections will be relatively common, as is to be expected with the extreme class imbalance, we might use the specificity-at-recall to estimate a portion of data that may be discarded, such that we can still achieve the necessary true positive detections. In our research, we estimated that 60.5 % of images may be discarded as not containing defects for us to still be able to achieve 90 % detection of defects with the trained convolutional neural network. In practice this will mean a sizeable reduction in workload for this task.

To what extent are the current inspection practices automatable?

Q4

With the results obtained from the experiments outlined in chapters 3 and 4, we conclude that automation of current inspection practices is limited mostly by the available data and its quality. While we had significant amounts of image data available, these images pertained, as expected, mostly to pipes with few visible defects. In addition, the quality of data and accompanying metadata is inherently limited by the current inspection practices: not all defects can be accurately captured in CCTV footage; the defect registration standards are contentious; there seems to be limited consensus on defect severity when multiple experts independently review the same CCTV footage. [3]

[3] van der Steen, A. J., Dirksen, J., and Clemens, F. H. 2014. Visual sewer inspection: detail of coding system versus data quality? *Structure and infrastructure engineering 10,* 11, 1385–1393

Based on results obtained, we conclude that it may be possible to entirely automate current inspection practices, with a more sophisticated neural network model, provided enough images of each defect type are provided and enough time and energy is spent on the network hyperparameter optimization. Such automation would still be limited to achieving human parity in terms of quality, which is to say, less than perfect.

A proper followup question would then be: *"Should we aim to automate current inspection practices?"* Our answer to this question is that the short-term benefits of doing so might be too short-lived. The current inspection practices are outdated in terms of methodology and lagging decades behind in their use of available technology. [4] Energy, time,

[4] Tscheikner-Gratl, F., Caradot, N., Cherqui, F., Leitão, J. P., Ahmadi, M., Langeveld, J. G., Le Gat, Y., Scholten, L., Roghani, B., Rodríguez, J. P., et al. 2019. Sewer asset management–state of the art and research needs. *Urban Water Journal 16,* 9, 662–675

and money may be better spent developing new inspection workflows that are by design adaptable to future innovations such as machine learning pipelines or inspection techniques.

Q5 Does introducing depth information through computer stereovision improve the data quality and analysis capabilities?

In chapter 5 we outlined a method to recreate a three-dimensional point cloud of a sewer pipe through computer stereovision. It is immediately clear that *human* analysis capabilities of these point clouds, as compared to the images it was constructed from, are drastically increased. Inspecting the three-dimensional geometry in an interactive environment is a much easier task than inspecting a pipe based on two-dimensional images, especially for those not trained in recognizing defects from a two-dimensional image.

What we are of course more interested in, is the analysis capabilities with machine learning techniques. In the same chapter, we outlined a possible method of quantifying the degree of 'anomalousness' in a pipe, by using a robust regression method to reconstruct the original shape of the pipe. We found that this anomaly detection worked well on its own: the positions of the points in the point clouds gave us an anomaly score with a moderate, positive correlation with human-graded quality assessment of the pipes.

While the amount of data gathered in the chapter does not lend itself to a machine learning approach, we assume that the information present in the point cloud does not overlap entirely with the information in a single image. That is to say, if we would augment the data as used in chapter 4 or similar research to contain depth information from a second camera, this has a high chance to improve detection capabilities.

Q6

How can we employ machine learning and computer vision to improve the efficiency and quality of urban drainage inspections?

This thesis has provided a collection of *possible* methods to employ machine learning and computer vision to improve efficiency and quality of urban drainage inspections. We have provided examples of unsupervised learning, supervised learning, and 'classical' computer vision techniques to automate parts of the inspection process. As noted in the answer to research question 4, full automation is a while off, but applying the advances made in machine learning and computer vision to this specific problem can lead to short-term improvements in efficiency of inspections: large amounts of data may not need human classification, and for the parts that do, it might be possible to aid the human inspectors with data obtained from the machine learning and computer vision algorithms and improve the quality of their assessments.

6.1 Future Work

To speculate on possible future work that could stem from ours, we might consider a combination of the different techniques described in this work.

Unsupervised anomaly detection (as described in chapter 3) could be used as a pre- or post-processing step for convolutional neural network classification (as described in chapter 4). As a post-processing step, it might be used to estimate locations of defects within an image that was classified as containing defects. As a pre-processing step, it might be part of a semi-supervised learning approach, to select samples for active learning for example.

As mentioned in the answer to research question 5, introducing geometry information into a deep learning pipeline has the potential to greatly improve results. While we did not have enough data to do so, adding a second camera to an inspection vehicle is inexpensive compared to more sophisticated 3D-scanning devices, and can in time provide enough data for neural network training. In this way we can imagine combining the techniques described in chapters 4 and 5.

We feel that a combination of all three techniques may be a significant step forward in the processing possibilities of urban drainage inspection data.

6.2 Closing Remarks

What this thesis has touched on is only a fraction of the possibilities for enrichment of urban drainage inspections with machine learning and computer vision. The field of urban drainage is only recently catching up on novel digital and virtual innovation, and the space for innovators to explore is virtually endless. Applying convolutional neural networks may have been an obvious first step that we and other researchers are collectively taking, but the next steps should be taken in terms of data collection and making this data accessible to researchers.