EPiC Series in Engineering
Volume 3, 2018, Pages 1478–1485
HIC 2018. 13th International Conference on Hydroinformatics

Automatic identification of sewer fault types using CCTV footage

Joshua Myrans[1], Zoran Kapelan[2], Richard Everson[1]
[1] University of Exeter, Harrison Building, North Park Road, Exeter, Devon, UK, EX4 4QF
jm494@exeter.ac.uk

Abstract

Water companies all over the world regularly perform inspections of their sewer networks. The data collected this way is then analysed by human technician which is time consuming and expensive. Previous work by the authors has developed methodology that can automatically detect faults in sewer pipes using standard CCTV footage. This paper presents a methodology to automatically identify types of detected faults aiming to further improve the efficiency and accuracy (i.e. consistency) of surveys. The methodology calculates a feature descriptor for individual frames of CCTV footage, before predicting the contents using a multi-class Random Forest classifier. Demonstrated on a comprehensive library of frames extracted from real-life CCTV footage of a UK water company, the methodology correctly identified the fault type in 71% of frames. Most common fault types were included in this experiment, covering a wide range of pipe sizes and materials, including vitrified clay, PVC and brick. Overall, this preliminary work shows promise for application in industry, proving an effective tool for analysing CCTV surveys.

Keywords: Automatic, Fault Identification, Machine Learning, Random Forest, Sewers

1 Introduction

Water companies in the UK must maintain thousands of kilometres of sewers, the condition of which is usually monitored using regular CCTV surveys. These surveys record the otherwise inaccessible interior of pipes using remote cameras, continuing to annotate the footage according to the WRc Manual of Sewer Condition Classification [1]. This annotation is usually performed after recording, by trained technicians, identifying the location, type and severity of faults. However, sewer surveys are extremely time consuming, prone to human error [2] and expensive. This paper presents a preliminary methodology to identify the type of a detected fault, using a machine learning classifier.

This is not the first work aiming to identify faults in sewer networks, multiple academics have approached this problem from many perspectives. Some of the earliest work by Duran et al. [3] suggested retrofitting the camera with a laser profiler. The laser's readings were used as inputs to a sequence of neural networks predicting the type of structural faults. Similarly, Sinha et al. [4] applied fuzzy neural networks to key features extracted from the footage to identify the presence of cracks in pipes. This work on crack detection has continued to flourish, with new techniques such as 3D scene reconstruction [5] and advanced pattern recognition tools [6] proving capable of the task. Outside the field of crack detection, Guo et al. [7] compared neighbouring frames of CCTV footage, identifying anomalies in the footage for further investigation. Most recently Halfawy and Hengmeechai [8] utilised optical flow to track the operator's camera motion to identify regions of interest within the footage. This work was continued, demonstrating the effectiveness of HOG features in combination with a SVM classifier to detect tree root intrusions [9].

The methodology presented here, builds on previous work, which identifies the presence of faults with at least 80% accuracy [10] [11]. Given the ability to detect faults, this new methodology identifies a fault's type based on its classification (i.e. Crack, Deposit, Root Intrusion etc.). Previous work in the field has largely concentrated on individual fault types, most notably cracks. This results in a large ensemble of extremely specialised classifiers, to accurately determine a fault type. This can perform well, as demonstrated by Hawari et al. [12], achieving an accuracy of >84% when detecting cracks, deformations, deposits and displaced joints in small scale experiments. In contrast, the technique presented here uses a single holistic methodology to determine fault type. Improving the efficiency of automatic fault identification, reducing the need for many specialised techniques, whilst maintaining a comparable level of performance and easy integration with other fault detection techniques.

2 Methodology

The fault type identification methodology, can be broken down into three stages: 'Pre-processing', 'Feature Extraction' and 'Classification' as demonstrated in Figure 1. This methodology is designed to identify faults within images, already identified as containing faults, as such the following assumptions are made:

- Images processed by the methodology contain at least one fault.
- Footage is of high enough quality for a human to identify the type of fault present in an image.
- Sufficient data, with a good representation of all sewer faults, is available to train the machine learning classifier.

As the methodology is designed to work alongside fault detection methodologies, the first assumption should be met with correct implementation. Similarly, if a human

technician is unable to discern the contents of an image, due to poor lighting, resolution etc., it is unreasonable to assume the methodology is able to. Finally, being a data driven technique, its effectiveness is limited by the data used during its calibration. If a fault type is underrepresented or dissimilar to the training examples, the methodology is likely to perform poorly on that fault type.

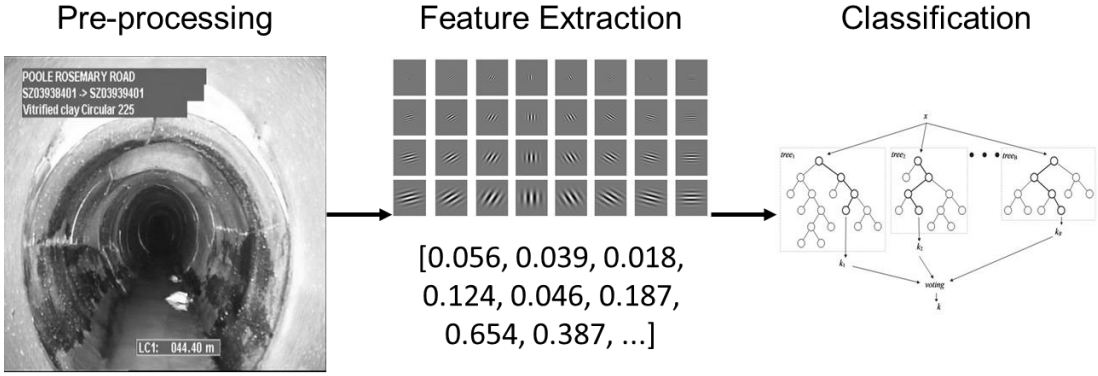

Figure 1. The three-stage structure if the fault type identification methodology

The first ‘Pre-processing’ stage of the methodology, extracts frames, identified as faulty from the raw CCTV footage, re-sizes them, and finally converts the image from colour to grayscale. In doing so each frame of CCTV footage is treated as an individual image, receiving its own classification. Experiments showed that colour to be noisy, dependant on varying illumination within the footage. Similarly, a lower resolution of 128 x 128 pixels is used, as experiments showed minimal improvement and increased processing times for higher resolutions.

Continuing to the second ‘Feature Extraction’ stage, a processed frame’s GIST feature descriptor is calculated [13]. The GIST descriptor represents the image in a much lower dimensional space, simplifying the classification problem. Initial experimentation showed the GIST descriptor to perform well, comparable to the HOG descriptor used by Halfawy and Hengmeechai [8], and much faster than other feature descriptors, such as SIFT and SURF. A GIST descriptor is calculated by convolving the pre-processed image with a series of 32 Gabor wavelets and summing the convolved values over 16 equally sized grid cells. The resulting 512 values are appended to form the final feature descriptor, ready for classification.

The final ‘Classification’ stage applies a trained multi class Random Forest [14], to the extracted GIST descriptor. The workhorse of the methodology, this machine learning classifier identifies the type of fault in the image. A Random Forest is an ensemble classifier, containing a collection of randomly ‘grown’ decision trees. Each tree in the forest makes an independent prediction of the frame’s contents, before the entire ensemble vote on the assigned fault type. A Random Forest was chosen over

other multi-class machine learning classifiers for it's easy to understand structure and its previous successes in the fault detection methodology [10]. The key to an effective Random Forest, like all machine learning classifiers, is effective training. In this methodology we apply the Extra Trees algorithm [15], which grows the internal decision trees by randomly selecting a parameter of GIST descriptor and splitting it randomly. This is dissimilar to standard Random Forests, which splits parameters based on the maximum normalised information gain.

3 Case study

3.1 Data

The testing and development of this methodology has been performed on real-world data provided by UK water company Wessex Water. The footage covered over 30km of pipe ranging in diameter (150 -1500mm), shape (circular, egg, horseshoe) and material (vitrified clay, PVC and brick). From this footage a selection of 2424 faults distributed as shown in Table 1.

Table 1. Distribution of fault types in extracted CCTV frames

Fault type	Subtype	Percentage (%)
Joint	Displaced, Open	33.7
Deposits	Attached, Settled	18.6
Surface	-	12.9
Crack	Longitudinal, Circumferential, Multiple, Spiral	12.0
Roots	Fine, Tap, Mass	10.2
Infiltration	Running, Gushing	4.5
Obstacles	Intruding Junctions, Masonry, Protrusion	3.9
Broken / Collapsed	-	2.3
Hole	-	1.1
Brickwork	Missing mortar, Displaced bricks, Missing bricks	0.5
Deformation	-	0.3

To make best use of the available dataset, 25-fold cross validation is used to separate the frames into training and testing sets [16]. Cross validation splits the randomly shuffled dataset into 25 equally sized groups (folds). Each of the 25 folds is in turn set aside to form the test dataset, whilst the remaining 24 folds are used to train the Random Forest classifier; the generalisation accuracy of the method is then estimated by averaging the accuracy over the 25 validation sets. In addition, any frames containing multiple fault types were excluded from the experiment, reducing the number of usable frames to 1972. Finally, initial experiments showed faults with less than 100 examples to be underrepresented, as such they are grouped into an 'other' class. This left a classifier capable of detecting: 'joints', 'deposits', 'surface', 'crack', 'roots', 'infiltration' and 'other'.

3.2 Results & Discussion

Applied to the collection of still images, the methodology achieved an overall accuracy of 71%, correctly assigning the type of fault in each case. The breakdown of the methodology's success by fault is demonstrated in the Table 2. It is also worth noting that each prediction took less than $1/25^{th}$ of a second on a standard laptop computer, enabling the methodology to run in real-time.

Table 2. Confusion rate matrix for the fault type identification methodology over 7 classes using *non-stratified* data.

		Actual						
		Joint	Crack	Deposits	Roots	Infiltration	Surface	Other
Predicted	Joint	0.86	0.27	0.21	0.26	0.24	0.28	0.15
	Crack	0.04	0.57	0.02	0.03	0.06	0.03	0.05
	Deposits	0.05	0.06	0.69	0.11	0.02	0.10	0.28
	Roots	0.00	0.01	0.01	0.54	0.02	0.02	0.01
	Infiltration	0.00	0.01	0.01	0.02	0.65	0.00	0.02
	Surface	0.03	0.04	0.04	0.04	0.00	0.56	0.04
	Other	0.01	0.01	0.03	0.01	0.01	0.02	0.45

Given the above results it is clear to see the methodology is working. This is best demonstrated by the highlighted diagonal in Table 2 showing the True Positive Rate (TPR) for each fault type. TPR is the rate at which, given an image with a specific fault type, the methodology correctly identifies the fault's type. Given 'joints' are the best represented fault type (roughly a third of the dataset, see Table 1), it is unsurprising the methodology is best at identifying them with a TPR of 86%. However, this saturation of 'joints' appears to have led to many other fault types being misclassified. In most other classes the highest False Positive Rate (FPR) is attributed to being misclassified as a 'Joint'.

To address the above issue, the experiment was re-run, stratifying the dataset to contain only 150 examples of each fault, reducing the size of the dataset to 1,050 images. This second experiment achieved a reduced overall accuracy of 68% but, at the same time, it has improved the TPR of most other fault types, as shown in Table 3.

Table 3. Confusion rate matrix for the fault type identification methodology over 7 classes using the *stratified* dataset.

		Actual						
		Joint	Crack	Deposits	Roots	Infiltration	Surface	Other
Predicted	Joint	0.61	0.07	0.08	0.08	0.09	0.09	0.04
	Crack	0.14	0.67	0.03	0.05	0.03	0.07	0.05
	Deposits	0.08	0.03	0.62	0.07	0.00	0.04	0.07
	Roots	0.03	0.04	0.07	0.69	0.04	0.06	0.02
	Infiltration	0.03	0.03	0.02	0.03	0.82	0.01	0.02
	Surface	0.03	0.03	0.06	0.01	0.01	0.61	0.04
	Other	0.08	0.13	0.12	0.07	0.01	0.12	0.76

Comparing Tables 2 and 3, the only fault types to achieve a reduced TPR after stratification of observed data are 'joint' and 'deposits', whilst all other fault types were better classified. 'joints' are now most likely to be misclassified as 'cracks', likely due to their similar physical appearance. On the other hand, 'deposits' and most other faults are likely to be mislabelled into the 'other' category, likely due to the diversity of the 'other' class. The 'other' fault category itself has seen a marked improvement, from a TPR of 45% to 76%. Similarly, 'infiltration' faults have achieved the highest TPR (82%), this could be attributed to their unique appearance, with most forms of infiltration ('seeping', 'running', 'gushing') appearing to be very similar. It is worth noting that 'Infiltration' faults are most commonly misclassified as 'joint' faults, likely because 'infiltration' faults often occur around pipe joints.

4 Conclusion

This paper demonstrates a preliminary methodology for identifying the type of sewer faults detected in CCTV footage. The methodology calculates a frames GIST descriptor [13] before using a multi class Random Forest [14] to identify the type of a given fault. In summary, the methodology shows promise for further investigation, achieving a 71% accuracy on this complex multi-class problem. The work was demonstrated on a comprehensive dataset, covering most common types of sewer pipe and fault as described by the WRc Manual of Sewer Condition Classification [1]. The case study shows that the methodology performs best on well represented faults, which are dissimilar to others in physical appearance, although over-represented faults often bias the methodology's classifications. As such the methodology was applied to a stratified dataset, containing 150 examples of each fault type. This reduced the overall accuracy to 68%, but balanced the TPR of most fault types, as shown by Table 3.

Future work aims to perform more comprehensive tests on the methodology, aiming to train the classifier on a larger, balanced collection of fault types. The methodology will also be compared to other similar techniques, such as an ensemble of binary

classifiers, to better understand the methodology's performance. Finally, the new methodology will be combined with previous work [10] [11] and expanded to cover fault location and severity in continuous CCTV footage. These technologies will be combined to create a decision support tool capable of assisting engineers in the field, improving the speed and quality of CCTV surveys.

References

[1] WRc plc, 2013 Manual of Sewer Condition Classification. 5th Revised edition. WRc Publications.

[2] 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), pp.1385-1393.

[3] Duran, O., Althoefer, K. and Seneviratne, L.D., 2007. Automated pipe defect detection and categorization using camera/laser-based profiler and artificial neural network. IEEE Transactions on Automation Science and Engineering, 4(1), pp.118-126.

[4] Sinha, S.K. and Fieguth, P.W., 2006. Neuro-fuzzy network for the classification of buried pipe defects. Automation in Construction, 15(1), pp.73-83.

[5] Jahanshahi, M.R. and Masri, S.F., 2012. Adaptive vision-based crack detection using 3D scene reconstruction for condition assessment of structures. Automation in Construction, 22, pp.567-576.

[6] Chen, F.C., Jahanshahi, M.R., Wu, R.T. and Joffe, C., 2017. A texture- Based Video Processing Methodology Using Bayesian Data Fusion for Autonomous Crack Detection on Metallic Surfaces. Computer- Aided Civil and Infrastructure Engineering, 32(4), pp.271-287.

[7] Guo, W., Soibelman, L. and Garrett, J.H., 2009. Automated defect detection for sewer pipeline inspection and condition assessment. Automation in Construction, 18(5), pp.587-596.

[8] Halfawy, M.R. and Hengmeechai, J., 2014. Integrated vision-based system for automated defect detection in sewer closed circuit television inspection videos. Journal of Computing in Civil Engineering, 29(1), p.04014024.

[9] Halfawy, M.R. and Hengmeechai, J., 2014. Automated defect detection in sewer closed circuit television images using histograms of oriented gradients and support vector machine. Automation in Construction, 38, pp.1-13.

[10] Myrans, J., Kapelan, Z. and Everson, R., 2016. Automated Detection of Faults in Wastewater Pipes from CCTV Footage by Using Random Forests. Procedia Engineering, 154, pp.36-41.

[11] Myrans, J., Kapelan, Z., Everson, R. and Britton, J., 2016, Using Support Vector Machines to identify faults in sewer pipes from CCTV surveys. CCWI 2016, electronic proceedings

[12] Hawari, A., Alamin, M., Alkadour, F., Elmasry, M. and Zayed, T., 2018. Automated defect detection tool for closed circuit television (cctv) inspected sewer pipelines. Automation in Construction, 89, pp.99-109.

[13] Oliva, A. and Torralba, A., 2001. Modelling the shape of the scene: A holistic representation of the spatial envelope. International Journal of Computer Vision, 42(3), pp.145-175.

[14] Breiman, L., 2001. Random forests. Machine learning, 45(1), pp.5-32.
[15] Geurts, P., Ernst, D. and Wehenkel, L., 2006. Extremely randomized trees. Machine learning, 63(1), pp.3-42. of structures. Automation in Construction, 22, pp.567-576.
[16] Kohavi, R., 1995, August. A study of cross-validation and bootstrap for accuracy estimation and model selection. In Ijcai (Vol. 14, No. 2, pp. 1137-1145)