

ROAD DRAINAGE SYSTEM LOCALISATION AND CONDITION DATA CAPTURE

M. Gkovedarou*, I. Brilakis

Department of Engineering, University of Cambridge, Cambridge, United Kingdom

* Corresponding author

ABSTRACT Inspection of road networks is time and cost consuming. Over half of the money allocated for road maintenance are spent on this area. Visual inspection is still the most commonly employed way of inspection for most parts of the road network. The sheer amount of lane miles of the road network renders this type of road monitoring a costly process. The ultimate goal of this paper is to reduce the time and cost needed to perform routine drive-by, visual and conditional data capturing. The road inspector devotes effort to visually capture the great number of the road assets as it consists of capturing of both the geometry and condition of multiple road assets. In this paper, we propose a method to capture road drainage covers through images and assess if they could potentially be blocked or not. This proposed novel framework focuses on two main tasks: a) localisation of road drainage b) assessment of drainage condition (if they are clean or not). This solution uses Speeded up Robust Features (SURF) and Scale Invariant Feature Form (SIFT) detector, the Bag of Visual Words (BoVW) enhanced with k-means algorithms for grouping the features, and a Support Vector Machine classifier for classifying the data to their respective categories.

1. Introduction

Highway Operational Standards 2018-2028 refer to the importance of healthy highway infrastructure as the key for a robust economy stating the need to use all the available resources (Highway Operational Standards 2018). A publication by the United Kingdom Royal Society for the Prevention of Accidents states the importance of healthy road assets (pavement, signs, markings etc.) as they constitute a huge factor for the safety of road users (International Road Federation 2006). Their maintenance is essential (Levik 2001) and cost effective compared to complete road reconstruction (Medd 2009). As a result, the identification and the measurement of the road defects (as part of maintenance policy) is quite challenging, rendering the use of suitable advanced technologies essential. There are three maintenance categories (Burningham and Stankevich 2005):

Routine Maintenance, which is conducted on a regular basis and entails debris removal, minor repairs etc.

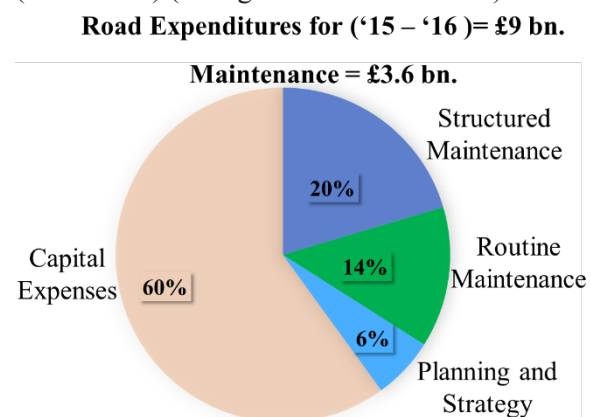
Periodic Maintenance, which is also conducted on a regular basis but is more costly than the routine maintenance, as it includes pavement reconstruction, planning and strategy

Urgent Maintenance, which includes repairs that could not have been foreseen

The United Kingdom (UK) road expenditures for the year 2015-2016 were £9 bn. and 40% of them were allocated for the maintenance (£3.6 bn.). From this amount, 20% (£0.72 bn.) constituted the routine maintenance and the planning and strategy of the maintenance policy, which also represents the

area of interest of this paper (Figure 1) (George and Kershaw 2016).

Figure 1: Allocation of road expenditures for the UK (2015-2016) (George and Kershaw 2016).



The maintenance policy of the authorities in the UK is both expensive (Figure 1) and time consuming (Table 1). Road maintenance monitoring occurs once or twice per year for the motorways and once per two years for the minor roads (Highway Operational Standards 2018). This inspection is mainly visual for road assets like lights, barriers and the drainage system (Royal Borough of Windsor and Maidenhead 2017). A trained inspector needs in total 30 days for both the visual inspection and the defect investigation. There are also online council report systems that give road users the opportunity to report any possible road network problems (Roads Safety Inspection Manual 2017).

Table 1: Type and Inspection Frequency on the UK roads (Highway Operational Standards 2018).

Hierarchy Description	Inspection frequency and type from City Councils
Motorway	1-2 per year (from Highways England)
Strategic Route	12 times per year (monthly)
Main Distributor	12 times per year (monthly)
Secondary Distributor	12 times per year (monthly)
Link Road	4 times a year (3 monthly)
Local Access Road	Annually (once per year)
Minor Roads	1 per two years (24 monthly) (standard is that they are passable with care)
Soft Roads (Green Lanes)	No formal inspection regime. Inspected on a reactive basis (standard is that they are passable in a 4 wheel drive vehicle)

Towards the goal of helping the time and cost of the routine drive-by visual inspection, machine-learning methods are used extensively in road assessment through image classification, object recognition or detection. The following methods are mostly used in pavement defect detection. For example, a research method uses SVM that detects the patches of the pavement followed by their measurement with accuracy over 80% (Hadjidemetriou et al. 2017). The novelty of this method is that it uses besides an external camera (during day and night), it also uses a simple smartphone camera as a sensor to detect and measure a specific pavement defect (Figure 2). Another method proposes the detection of the quality of the drainage system. This method (Figure 3) records the pavement surface with a camera, the temperature as well as the surface saturation by water jets in order to assess the drainage system. In this method, they saturate the pavement with the water jet, calculate the temperature and then capture images from the drainage system in order to assess through image processing, the quality of the drainage (Mataei et al. 2018). However, this method cannot be used in a drive-by, non-stop inspection process and is very slow considering how long it takes to follow its process per drainage given the number of drainages in the road network. This is likely why it is not used on a commercial basis. Another method detects the drainage covers

and feeds a classifier with images from Unmanned Aerial Vehicles (de Vitry et al. 2018). The novelty of this method is that it was the first to detect drain covers from UAV imagery. The method achieved 73% accuracy, and 50% recall. This performance is too low for practical use, as inspectors would need to manually review all the data to find the missing 50% of the drainages and discard the 27% of the imagery that does not contain drainages.

Figure 2: Detection of roadway pavement patches (yellow area) (Hadjidemetriou et al. 2017)

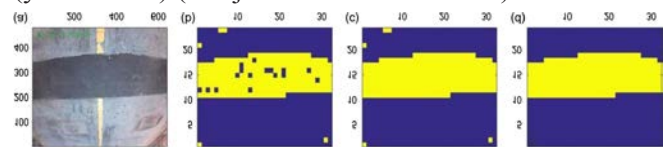
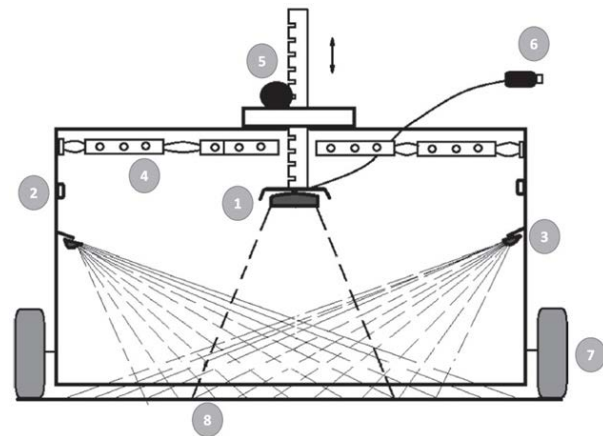


Figure 3: The proposed device (1) camera; (2) thermometer; (3) waterjet; (4) LED lamps; (5) height adjustment tools; (6) USB cable; (7) wheels; and (8) pavement surface. (Mataei et al. 2018)



2. Research Questions and Proposed Solution

This paper aspires to contribute to reducing the time and cost needed to perform routine drive-by, visual condition data capturing. A set of questions that derive from the gaps need to be addressed to achieve the aforementioned goal. Out of these gaps, this paper aims to address these three:

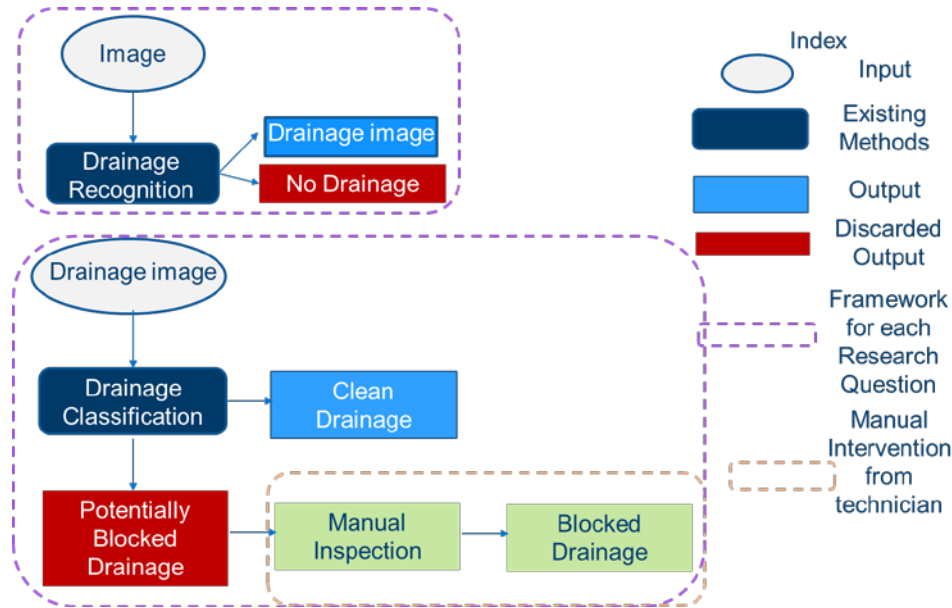
1. Automate the capture of drainage systems locations.
2. Automate the capture of drainage systems condition data (either blocked or not).

The research questions that arise are:

1. How to recognise drainage systems automatically using cameras?
2. How to differentiate clean from blocked (unclean) drainage?

In my case supervised methods will be used as they provide classes, which aligns with the paper's scope. Figure 4 introduces the framework of the proposed solution. Each part of the flowchart addresses one of the research questions.

Figure 4: Framework of the proposed solution



Step 1

Step 1 addresses Research Question 1. The input of this step is an image of the edge of the road pavement, captured by camera, where the drainage covers are usually located. One assumption is that the images captured are localised from an on-board data collection system for localisation (e.g. GPS coordinates), so the focus is only on detecting whether a drainage cover is present in the image or not. This appears to be a classic machine-learning problem for image classification. The process comprises of feature extraction, feature codification and finally image classification. For feature extraction, both the SIFT and SURF descriptor are suitable for the scope of this paper so both of them will be experimentally examined so as to understand which one is more appropriate by using the proposed parameters listed in the literature review. The author suggests for the SIFT detector 3 scales per octave as 1 to 4 is recommended in the literature. Increasing number of octaves leads to increase of the computational time as the algorithm detects both smaller and larger blobs. The number of octave layers affects the ability to detect finer elements. The recommended values are between 1 to 6 so we choose 4 as an initial value. The recommended number of σ is $\sqrt{2}$. Increase of σ leads to smoother images (Yu et al. 2007). Similar to SIFT the parameters that the author suggests are: threshold 1000 and number of octaves 3, as 1 to 4 is recommended, and 4 octave layers, as 1 to 6 is recommended (Juan and Gwun 2009).

For feature encoding, we use Bag of Visual Words (BoVW) with 0.8 clustering threshold since this is the optimal according to the literature. Empirical studies try to determine the best vocabulary size, while recommend the optimal clustering threshold 0.8 (Hou et al. 2010). Small vocabulary size means that there might not be many words to represent the features of the image, whereas, large vocabulary size might cause

overfitting problems. The result of this algorithm is a frequency histogram of the vocabularies versus the frequencies of them. The output of BoVW algorithm is a feature vector, which derives from image encoding (Tsai 2012). The number of vocabularies is under research hence, the size 250 (256 is suggested by the literature) is selected and we monitor the histogram of occurrences to check if our choices create underfitting or overfitting issues.

Image classification can be achieved with either supervised or unsupervised methods. Supervised methods use evidence (data) to predict a model. In our case, road data is easily collected, while the objective of this paper is to assign labels to each image and not simply find patterns in the collected data, so supervised methods are preferred. Potential methods for image classification are: Artificial Neural Networks (ANNs), the Support Vector Machine classifier and the Semantic Texton Forest classifier. The disadvantages of ANNs are: 1) they need justified underlying assumptions otherwise they perform poorly and 2) their performance in noisy data is low as there is danger of overfitting (Kotsiantis, S. B., Zaharakis, I., and Pintelas 2007; Zhang 2000). Overfitting is when the predictor corresponds too close to a dataset making it difficult to fit new data to predict new results (Hawkins 2004). Deep Neural Network (category of ANNs) tools are more appropriate for complex data like hyperspectral images or large datasets (over 1000 images per category or over 100 classes) (Hu et al. 2015; Oquab et al. 2015), while, STFs are mostly used for per-pixel classification. Hence, we choose to start with the SVM classifier, and try all the kernels to find the most appropriate one. We will test several kernels of the SVM classifier (Linear, Gaussian, Quadratic and Cubic) to benchmark their performance. The output of this step is labelled images (yes/no drainage cover).

2.1 Assumptions

There are various types of drainage systems in the road network as shown in Figure 5. The data set contains the (c) and (d) drainage type, so these types will be the ones to look for.

Step 2

Step 2 approaches Research Question 2. In this step, the input consists of images that contain a drainage cover. There is a variety of road debris, so it is difficult to classify the images based on the debris. Hence, in this step, the hypothesis is that the clean drainage is the one with no debris on the surface, and that the presence of debris may indicate varying levels of potential existing or soon to appear blockage. The first step of this hypothesis is the manual distinguishing of the clean from the non-clean drainage covers and then the recognition of only the clean ones. This is considered as another machine-learning problem; thus, similar approach is followed to Step 1. The outcome of this step is the separation of images of clean drain covers from those with debris (i.e. potentially blocked).

Hypothesis 1: This method will lead to accurate and precise classification.

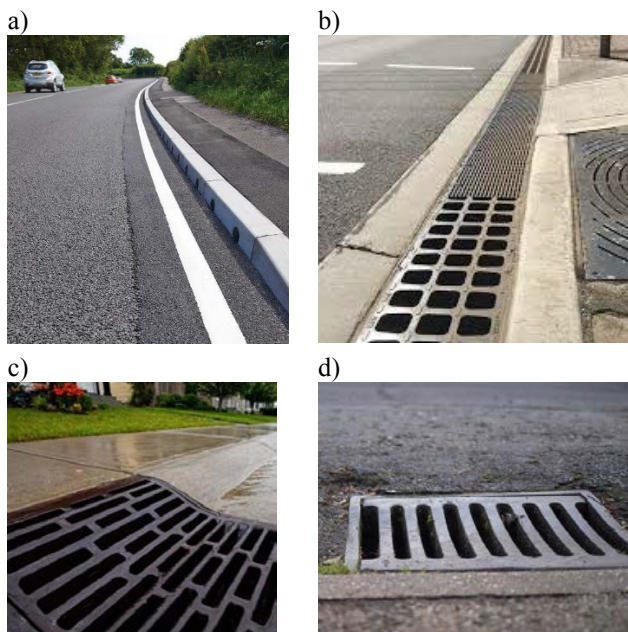
Hypothesis 2: The proposed solution will accurately distinguish between blocked and unblocked drainage.

One assumption is that the drainage system is always placed at the edge of the road to account for basic road design principles.

The second assumption is that the camera looks always to the edge of the pavement. Hence: the driver drives on the left side of the road where the drainage covers are located.

The last assumption is that this work will be embedded in a greater system that will use GPS and can easily attach coordinates to its collected images, for localising the detection of drainage covers.

Figure 5: Various road drainage systems (Department of the Environment Heritage and Local Government 2004)



3. Conclusions

The identification of the geometrical characteristics and the possible defects of the various parts of the road network is quite challenging. Visual inspection (as part of the road maintenance) is a commonplace first-level inspection for the majority of the parts of the road network. It is also time and cost consuming; thus advanced methods need to be implemented. The goal of this paper was to devise a solution for detecting drainage covers and which of them could be potentially blocked in drive-by image data.

The proposed solution of this paper aimed to: a) localise the drainage system covers captured by images, and b) separate the clean ones from the potential blocked ones. Both the SIFT and SURF detector could be examined. SIFT as it is expected by the literature review (Yu et al. 2007), detects mostly edges which is the goal of the first experiment (drainage covers edges). There is much research related to the number of clusters (words) in the Bag of Visual Words. The diagrams of Word Occurrences produced by the BoVW model indicate that the number of vocabulary size (250) is adequate as almost each visual word index appears in both “Drainage” and “Non-Drainage” image category. The classification training could be conducted using all the available SVM kernels with both 5 and 10-fold cross validation. Different kernels can be examined using A good indicator is the ROC curve which if it follows the left and the right hand border closely which can be interpreted that as long as the specificity increases, the sensitivity increases in a lower grade. The results of this method can be compared to another one from the literature review. They reported accuracy 65% and 73% recall. In this research, the authors used Unmanned Aerial Vehicles (UAV). These devices could not capture objects on the side of the street since trees, cars, or other objects hide them. In addition, there was no much diversity in the trained objects as the image capturing had a fixed vertical angle of 90°. Also in this research, the authors mentioned that better results might be possible if the images were taken from lower (de Vitry et al. 2018).

The second part of the proposed solution will be based on the hypothesis that there is a variety of road debris, thus the clean drainage covers were detected based on the assumption that the rest will be covered by debris (potentially blocked). Similar methodology as the one in the first part is followed. The difference in this experiment is that SURF algorithm that is faster and more like a blob detector might be the preferable option. The equation for sample sufficiency indicated that the test data (66) are not sufficient (68 were required). There is, however, published research where the number of test data is not sufficient but it gives high results (de Vitry et al. 2018). The same number of clusters (250) was used for the BoVW algorithm as the one in Experiment one, despite the fact that the number of features in experiment two is half the number of features in experiment one. However, the histogram of BoVW indicates that this number is sufficient as almost each visual word index appears in both “Clean” and “Non-Clean” image category. The classification training conducted in all the available SVM kernels both with 5 and 10-fold cross validation. Cubic kernel achieved better results in both 5 and

10-fold cross validation with 61.8% and 66.4% respectively. 10-fold cross-validation was again a better validation.

3.1 Contributions

This research project contributes to the knowledge spectrum of road network asset monitoring. In academic terms, this research advanced road object localization and assessment. The former is achieved by automatically recognizing drainage covers in drive-by road imagery. The latter is achieved by automatically separating clean from potentially blocked drain covers. Consequently, this research contributes to road design and maintenance by assisting road designers and inspectors in their work. In addition, the automated drain cover localisation and assessment research solution can reduce the time and cost devoted to inspection. Road designers could also benefit from the asset classification, as they will be aware of the parts that they should pay attention to during design, for maintenance purposes. Ultimately, this research also has a contribution in road safety. It is a stepping stone towards achieving healthier road conditions. It can be embedded in a greater system that will conduct routine road maintenance more often than the current ones. Conventional vehicles equipped with camera or an autonomous vehicle, which is already equipped with these type of sensors can serve as the foundation for such a system. As a result, road monitoring will occur with greater frequency and help catch problems like blockage causing road floods well before they become serious.

4. References

- Highway Operational Standards. (2018). "Highway Operational Standards 2018-2028." 1–9.
- International Road Federation. (2006). "Maintenance of Road Signs, Markings and Other Safety Assets: A COST EFFECTIVE MEANS OF OPTIMISING ROAD SAFETY."
- Levik, K. (2001). "How to sell the message 'Road maintenance is necessary' to decision makers." *First Road Transportation Technology Transfer ...*, 460–467.
- Medd, D. G. P. A. S. R. I. A. I. (2009). *Road safety audit*. (T. I. O. H. & TRANSPORTATION, ed.).
- Burningham, S., and Stankevich, N. (2005). "Why road maintenance is important and how to get it done." *Transport*, (June), 1–10.
- George, N., and Kershaw, K. (2016). "Road Use Statistics Great Britain 2016." *Department for Transport*, (April), p.23.
- Royal Borough of Windsor and Maidenhead. (2017). "Royal Borough of Windsor and Maidenhead HIGHWAY SAFETY INSPECTION MANUAL." 1–16.
- Royal Borough of Windsor and Maidenhead. (2017). "Royal Borough of Windsor and Maidenhead HIGHWAY SAFETY INSPECTION MANUAL." 1–16.
- Roads Safety Inspection Manual. (2017). "Roads Safety Inspection Manual April 2012." (April), 1–38.
- Hadjidemetriou, G., Vela, P. A., and Christodoulou, S. E. (2017). "Automated Pavement Patch Detection and Quantification Using Support Vector Machines." 32(1), 1–11.
- Mataei, B., Moghadas Nejad, F., Zahedi, M., and Zakeri, H. (2018). "Evaluation of pavement surface drainage using an automated image acquisition and processing system." *Automation in Construction*, Elsevier, 86(November 2017), 240–255.
- de Vitry, M. M., Schindler, K., Rieckermann, J., and Leitão, J. P. (2018). "Sewer inlet localization in UAV image clouds: Improving performance with multiview detection." *Remote Sensing*, 10(5), 1–18.
- Yu, S. J., Sukumar, S. R., Koschan, A. F., Page, D. L., and Abidi, M. A. (2007). "3D reconstruction of road surfaces using an integrated multi-sensory approach." *Optics and Lasers in Engineering*, 45(7), 808–818.
- Juan, L., and Gwun, O. (2009). "A comparison of sift, pca-sift and surf." *International Journal of Image Processing (IJIP)*, 3(4), 143–152.
- Hu, W., Huang, Y., Wei, L., Zhang, F., and Li, H. (2015). "Deep convolutional neural networks for hyperspectral image classification." *Journal of Sensors*, 2015.
- Oquab, M., Bottou, L., Laptev, I., and Sivic, J. (2015). "Is Object Localization for Free? - Weakly-Supervised Learning With Convolutional Neural Networks." *Cvpr '15*, (iii), 685–6
- Hou, J., Kang, J., and Qi, N. (2010). "On Vocabulary Size in Bag-of-Visual-Words Representation." *Advances in Multimedia Information Processing - PCM 2010*, 6297(September 2010).
- Tsai, C.-F. (2012). "Bag-of-Words Representation in Image Annotation: A Review." *ISRN Artificial Intelligence*, 2012, 1–19.
- Kotsiantis, S. B., Zaharakis, I., and Pintelas, P. (2007). *Supervised machine learning: A review of classification techniques*. (J. B. R. D.-K. N. G. J. N. K. J. L. R. L. de M. R. M. M. M. and N. Zhong, ed.), IOS Press, Amsterdam.
- Zhang, G. P. (2000). "Neural networks for classification: a survey." *IEEE Transactions on Systems, Man and Cybernetics, Part C (Applications and Reviews)*, 30(4), 451–462.
- Hawkins, D. M. (2004). "The Problem of Overfitting." *Journal of Chemical Information and Computer Sciences*, 44(1), 1–12.
- Department of the Environment Heritage and Local Government. (2004). "Guidelines for road drainage." 1–62.