MANHOLE COVER LOCALIZATION IN AERIAL IMAGES WITH A DEEP LEARNING APPROACH

B. Commandre[a,e], D. En-Nejjary[a,b], L. Pibre[b,c], M. Chaumont[d], C. Delenne[a,e]*, N. Chahinian[f]

[a] HSM, Univ. Montpellier, IRD, CNRS, Montpellier, France
[b] LIRMM, Univ. Montpellier, CNRS, Montpellier, France
[c] Berger-Levrault, Montpellier, France
[d] LIRMM, Univ. Nîmes, Univ. Montpellier, CNRS, Montpellier, France
[e] Lemon, Inria, Montpellier, France
[f] HSM, IRD, Univ. Montpellier, CNRS, Montpellier, France

KEY WORDS: Deep learning, high resolution imagery, urban object detection, Convolutional Neural Network

ABSTRACT:

Urban growth is an ongoing trend and one of its direct consequences is the development of buried utility networks. Locating these networks is becoming a challenging task. While the labeling of large objects in aerial images is extensively studied in Geosciences, the localization of small objects (smaller than a building) is in counter part less studied and very challenging due to the variance of object colors, cluttered neighborhood, non-uniform background, shadows and aspect ratios. In this paper, we put forward a method for the automatic detection and localization of manhole covers in Very High Resolution (VHR) aerial and remotely sensed images using a Convolutional Neural Network (CNN). Compared to other detection/localization methods for small objects, the proposed approach is more comprehensive as the entire image is processed without prior segmentation. The first experiments using the Prades-Le-Lez and Gigean datasets show that our method is indeed effective as more than 49% of the ground truth database is detected with a precision of 75%. New improvement possibilities are being explored such as using information on the shape of the detected objects and increasing the types of objects to be detected, thus enabling the extraction of more object specific features.

1. INTRODUCTION

Urban areas are undergoing fast and continuous growth leading to the expansion of underground utility networks. As a result, network management and monitoring problems may rise, especially in the case of unplanned city growth (Wang et al., 2014). In the context of Smart Cities, planners and decision makers require information about the actual state of the urban infrastructure (Rajendra and Chandraskaran, 2014). There is thus a strong expectation from authorities for technical solutions that may lead to the automation of the modeling and monitoring of urban infrastructure.

Wastewater and stormwater networks are a perfect example of the mislocalisation of buried utilities both in industrialized and developing countries. Over the past century it was common practice for each service provider and district to install, operate and repair its network separately (Rogers et al., 2012). Maps and databases were often not well archived or centralized and as a result it is difficult nowadays to obtain accurate information on the localization and characteristics of the buried pipes and hydraulic equipment. The manhole covers and inlet grates which are surface elements and thus indicators of the location of these networks, can be observed on aerial images. Hence, we can rely on the automatic detection of these objects to provide an estimation of the underground utility networks' position (Pasquet et al., 2016).

Some works have aimed to automatically detect manhole covers based on MLS (Mobile Laser Scanning). This has the advantage of providing a feedback about the current road conditions and thus avoiding accidents whether it is for intelligent transportation systems, moving vehicles or pedestrians (Yu et al., 2015).

But these approaches are extremely expensive, need considerable post-processing and are restricted to objects located on roads. In previous works, we have attempted to detect manhole covers using aerial images and have thus demonstrated the potential of these low cost and efficient methods (Pasquet et al., 2016, Bartoli et al., 2015).

In this paper, we present an automatic recognition and localization method for manhole covers using aerial images and a Convolutional Neural Network (CNN). A custom version of the AlexNet CNN (Krizhevsky et al., 2012) is trained on very high spatial resolution aerial images of Prades-Le-Lez (southern France). The validation is conducted on the town of Gigean (southern France) using a sliding window and AlexNet classifier.

This paper is organized as follows. Section 2 is a short literature review on manhole cover detection. Section 3 describes the proposed method. Section 4 reports and discusses the experimental results. Finally, Section 5 concludes the paper and presents new development perspectives.

2. RELATED WORK

Several methods using various types of images have been published in the literature to detect manhole covers.

The earliest were applied on digital photographic imagery. Tanaka and Mouri (Tanaka and Mouri, 2000) presented a detection method of round-shaped manhole covers based on morphological techniques. First, a black top-hat operation with a disk-shaped structure element is applied to extract circular shape components. Then, a masking operation is used to highlight the circular shaped components with a threshold. Finally, circular shapes that match the

*Corresponding author carole.delenne@umontpellier.fr

classical size of manhole covers are extracted. Relying on separability filters, a robust detection method for circular objects is given in (Niigaki et al., 2012). Instead of analyzing the intensity difference between manhole covers and their surroundings, the separability and uniformity of the image intensity distributions are calculated using the Bhattacharyya coefficient (Bhattacharya, 1946). Three indicators are defined and used to achieve the detection of the round-shaped manhole covers: circular objects, oriented separability and uniformity indicators.

The cited methods are sensitive to the thresholding level. In addition, they are prone to errors due to frequent occlusions, changes in illumination conditions and texture variations. Hence these solutions alone are not efficient in real-world conditions.

To overcome these problems, a multi-view approach combining two and three dimensional techniques and using surface images acquired by a moving van, was presented in (Timofte et al., 2014). The detection process consists in suppressing the highest number of undesirable objects (cars, shadows, trees, etc) from the rebuilt image of the road and in detecting the shapes corresponding to manhole covers using an improved Hough transformation (Cheng et al., 2006). A similar idea, but using LIDAR technology, was presented in (Tooke et al., 2011). First, road surface points are segmented and rasterized into 2D georeferenced intensity images. Then circular-shaped and rectangular-shaped manhole covers are detected using a Markov Chain. This approach is also used with rasterized 2D intensity images in (Guan et al., 2014, Yu et al., 2015). As mentioned in the introduction, many MLS based methods have performed well in detecting manhole covers. However, acquiring such LIDAR data at large scale remains a costly operation.

Breakthroughs in remote sensing technology have made a large amount of very high resolution aerial images readily available. Compared to manned vehicles, aerial surveillance has the advantage of higher mobility, larger cover scope (Lin et al., 2011) and cost-effectiveness.

Recently a circular geometrical filter based method was proposed in (Bartoli et al., 2015) to detect round-shaped manhole covers in very high resolution aerial images. After the segmentation of the image to retain only the road network, colorimetric indices are used to eliminate the vegetation and shadows. Finally, a circular filter is used to locate the manhole covers. In (Pasquet et al., 2016) a framework is proposed to automatically detect manhole covers in high resolution aerial images by combining the method based on the geometrical filter with a machine-learning SVM based approach. Results are encouraging, combination of the circular filter with deep-learning CNN instead of SVM can be envisaged to obtained better results than only use RBG channels.

3. MATERIALS AND METHODS

Inspired by the previous works, we implemented a new automatic detection system of manhole covers using aerial images and deep learning. The three-step methodology consists in: *i)* training AlexNet (Krizhevsky et al., 2012) on a ground truth dataset and *ii)* applying a sliding window on the images for the detection purpose. In order to improve detection results, step *iii)* consists in performing several iterations of boosting.

3.1 Study area and data

Two towns of Southern France have been chosen to develop and validate the approach: Prades-Le-Lez and Gigean. Very high res-

olution aerial images (5 cm resolution) were acquired in natural colours (RGB) in July 2014 on Gigean and June 2016 on Prades-Le-Lez (see Figures 1 and 2).

The Prades-Le-Lez database consists of 6 20000×20000 pixel georeferenced images. The exact position of the 605 manhole covers of the wastewater network was provided by Montpellier Métropole, the local operator. The Gigean dataset consists of one image of 17749×18361 pixels, and the position of 101 mahnole covers provided by CCBT, the local operator.

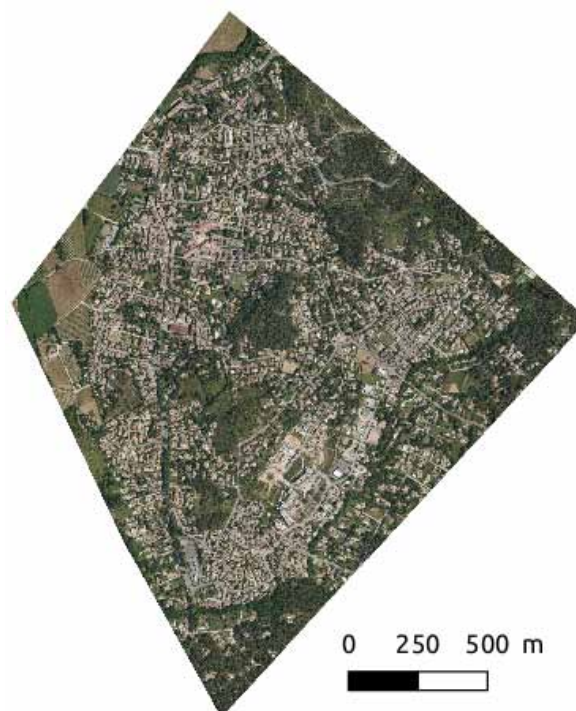Figure 1. Image of Gigean acquired at 5cm spatial resolution in July 2014

Figure 2. Image of Prades-le-Lez acquired at 5cm spatial resolution in June 2016.

To train the convolutional neural network, "ground truth" thumbnails with a size of 40×40 pixels, are extracted from the Prades-Le-Lez images. Two categories are considered: "manhole covers" and "other objects" which groups all the thumbnails which do not contain all or part of a manhole cover. Fig. 3 presents some thumbnails of the two categories.

Figure 3. Examples of ground truth RGB images: manhole covers (up) and other objects (bottom).

Different experiments were run using various image sizes ranging from 15×15 pixels to 80×80 pixels and the best results were obtained using a patch size of 40×40 pixels. Indeed, Figure 3 shows that a patch size of 40×40 is sufficient to take into account the necessary context around the manhole covers, as their typical size in France is 80 cm, *i.e* 16 pixels.

3.2 The Deep Convolutional Neural Network

Deep learning models have demonstrated high performance when used in various technical fields, such as computer vision and speech recognition, because of their capabilities in learning hierarchical deep features from large amounts of data. Namely, the deep convolutional neural network has shown exceptional performance in computer vision tasks such as classification (Szegedy et al., 2015). The well-known AlexNet architecture (Krizhevsky et al., 2012) is used in the present study. More sophisticated networks like GoogleNet (Szegedy et al., 2015) have been envisaged. Tests have been carried out with this network, but the results obtained were not conclusive and very time consuming. As GoogleNet requires 256×256 images, it is necessary to resize all the 40×40 thumbnails of the training database. This reduces database quality by blurring the images. Thus the results obtained with GoogleNet are not included in this work .

Figure 4 presents the customized AlexNet CNN that we have used. The process involves two main steps: feature extraction and classification.

3.2.1 Feature extraction. The deep CNN structure contains five *convolution layers* which will extract features from the input thumbnails. The spatial relationship between pixels is conserved by convolution. The four following AlexNet parameters (see (Krizhevsky et al., 2012) for more details) have been customized to fit with our needs:

- numbers of outputs of the convolution layer (*e.g.* 96 or 256);

- kernel size (*e.g.* 7×7, 5×5 or 3×3);

- stride s for the kernel displacement, reduced to 1 or 2 in order to extract more accurate features;

- padding $p = 0$, meaning that we do not compensate for the pixel losses due to the kernel size.

After the convolution, the Rectified Linear Unit function (ReLU) is used to eliminate the least informative thumbnails and the remaining ones are normalized.

Then the max pooling method, which is the most popular one used in the literature, is applied to output the maximum value in every subregion of the input data. The best results were obtained for following parameters: subregion size 3×3, stride=1 and padding=0.

3.2.2 Classification. The classification process is carried out by means of two fully connected layers, stacked at the end of the feature extraction step:

- "inner product" that merges all the outputs of the previous layers.

- "softmax" that computes the probability distribution over the two possible outcomes: "manhole covers" or "other objects".

The network is trained using a classical backpropagation scheme (not represented in Fig 4) which, at the end of classification step, modifies the inner weights of the convolution layers to improve the classification results. Once the network is fully trained, it can be used for the classification of any input image without further modification of its parameters.
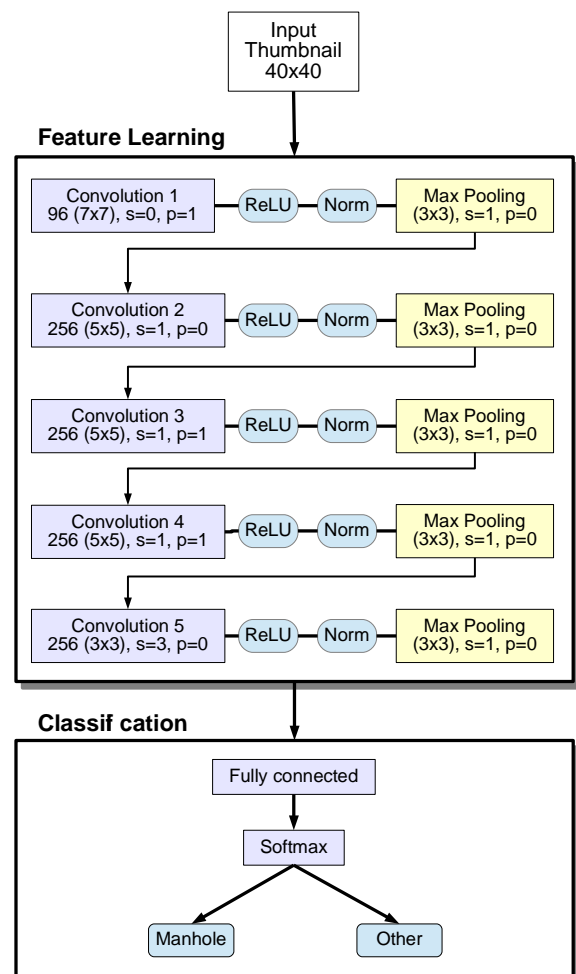

Figure 4. Customized Alex convolution neural network.

3.3 Sliding-Window Object Detection

We use the classical sliding window procedure to detect the manhole covers in the image. The window has the same size as the thumbnails used in the training of the CNN and is with a step of one pixel to scan the entire image. Note that a unique sliding window size can be used here as all manhole covers have the same dimensions and the two images used have the same resolutions. For each pixel position, the classifier gives its probability that the surrounding thumbnail belongs to each category. Hence, the problem of object detection is reduced to a set of local classification decisions, for which we will use an AlexNet customize classifier (Krizhevsky et al., 2012). In the rest of this article, we consider as manhole covers only the thumbnails with a likelihood greater than 90%, all the remaining thumbnails are classified as "other objects".

3.4 Boosting database

In order to improve the performance of the CNN, we have tested boosting, a greedy technique for ensemble learning (Li et al., 2005). The idea is to train a new network that learns to fix the errors of the previous one. Once a network is trained on the Prades-Le-Lez dataset, it is applied to the Prades-Le-Lez entire image and all the false positives *i.e.* objects incorrectly detected as manhole covers are added to the "other objects" category in a newly created training database. A new network is then trained with this dataset. In this paper, we employ boosting in a straightforward manner, working iteratively with the same network. Several iterations of boosting have been carried out and have improved the results, as shown in section 4.

3.5 Training and Test protocol

Deep learning algorithms require very large datasets for the training phase. To overcome the lack of images (only 605 thumbnails in the training dataset) and improve performance, a number of data augmentation techniques can be used to enlarge the size of the dataset. In this work, we employ the Keras library (Chollet, 2015) to increase the training database size by combining several data augmentation methods such as rotation, translation, horizontal flip, vertical and horizontal shift. From the combination of all transformations, we can obtain fifty images for the dataset from a single ground-truth image. Concerning the "other objects" category, we first choose, ten times more randomly extracted thumbnails from the dataset image and make sure that the thumbnails do not contain manholes. Context of each manhole cover is also added in this second category. As we do several boosting iterations, all false positives will be added to the "other objects" category. Finally, number of "manhole covers" thumbnails in the dataset is 18 405 and the "other objects" dataset size is 458 915.

3.6 Object Detection evaluation

To evaluate our automatic object detection system, we have followed the procedure adopted for the Pascal VOC challenge (Everingham et al., 2010). A detection is said to be correct if the overlapping area a_0 (Eq. 1) between the predicted bounding box B_p and the ground truth bounding box B_{gt} is greater than 50%, using this formula:

$$a_0 = \frac{area(B_p \cap B_{gt})}{area(B_p \cup B_{gt})}, \qquad (1)$$

where $(B_p \cap B_{gt})$ denotes the intersection of the predicted and ground truth bounding boxes and $(B_p \cup B_{gt})$ their union.

Note that, similarly to the procedure adopted for the Pascal VOC challenge (Everingham et al., 2010), when there is a multiple detection of the same object in an image with bounding boxes that have an intersection area greater than 30%, the boxes are merged with the non maximum suppression technique (Neubeck and Gool, 2006). If the intersection area is less than 30%, only one is counted as a correct detection; the others are all counted as false positives.

Our system is then evaluated through the computation of *precision*, *recall* and *F-measure* as expressed below:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \qquad (2)$$

where TP is the number of correctly classified manhole covers and FP the thumbnails wrongly classified as manhole covers (hence TP+FP the total amount of thumbnails classified as manhole covers).

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \qquad (3)$$

where FN is the number of manholes which has not been recognized. So TP+FN is the total amount of existing manhole covers.

$$\text{F} - \text{measure} = \frac{2(\text{Precision} * \text{Recall})}{\text{Precision} + \text{Recall}} \qquad (4)$$

4. EXPERIMENTAL RESULTS AND DISCUSSION

4.1 Results and discussion

Four experimental networks have been applied on the aerial images: 1) the original AlexNet network with the default parameters presented in (Krizhevsky et al., 2012), 2) the boosted version of the network, 3) the boosted version with the cleaned database and 4) the customized version of AlexNet with the cleaned database i.e. with the parameters presented in Figure 4.

We trained the original AlexNet network on a database which contains the thumbnails of all the manhole covers identified in the operator's database. This network yields a huge number of detections with 1575 thumbnails classified as manhole covers and only 49% of real manhole covers detected. Most of the detections being false positives, the precision is hardly better than 3%.

The second experiment consists in boosting this network, as explained before, by adding all the false positive detections obtained on Prades-Le-Lez with the original network to the "other objects" category, and training it again. Five boosting iterations have been done. At the fifth iteration, the number of detections has greatly reduced, with 136 thumbnails classified as "manhole

Figure 5. Examples of detected manholes covers, false positives and non detected manhole covers by the customized network. Green boxes correspond to detected manholes covers, red boxes to false positives.

Network	Precision	Recall	F-Measure
Original	3.11%	49%	5.84
Boosted	31.62%	43%	36.44
Boosted V2	32.48%	41%	38.93
Customized	68.49%	50%	57.80

Table 1. Precision, recall and F-Measure of the four networks for a likelihood threshold of 90%.

covers". The recall also slightly decreased to 43% while precision reached 31.6%. The analysis of these results shows that several real manhole covers whose locations were reported in the operator's database were hardly visible on the image, as they were covered by cars, shadows or vegetation.

Consequently, a third test was carried out with a cleaned training database from which 296 thumbnails corresponding to manhole covers that were not entirely visible on the image were retrieved. To avoid the detection of other patterns classically encountered near manhole covers such as pedestrian crossings or pavements, we also added the immediate surroundings of the manhole covers to the "other objects" category. This yielded better results as the number of correctly detected manhole covers increased up to 51%. The total number of detected objects increased also as it reached 162 compared to 136 with the previous network. However this increase did not deteriorate the precision which reached 32.48% because most of the added objects were true positives, *i.e.* manhole covers.

Several tests were performed to customize the network by modifying some parameters. Best results were obtained with the values given in Figure 4. Having noted a better recall with the cleaned database, the customized network was trained using only the cleaned database. The results obtained are significantly better than in the previous two cases. Indeed, for a recall of 50% the precision is of 68.49%. In addition, when using this network the number of false positives becomes lower than the number of manhole covers that are detected.

Table 1 presents the precision, recall and F-Measure of all the networks for a likelihood threshold of 90%. It can be seen that for all three criteria, the customized network has the highest scores while the original AlexNet network has the lowest scores. However, these results are specific to a given likelihood threshold and in order to evaluate the network's efficiency, it is necessary to analyse its results over a larger range of likelihood values.

Fig. 6 presents all the results obtained with the four tested networks according to a variation of the chosen threshold. The original AlexNet network, red dots, has the lowest precision of the tested networks because it has the highest rate of false positives. The boosted network, green dots, produced better results for a similar recall. However, its maximum recall score is far lower than the first network. The precision is further improved with the boosted network trained using the cleaned database. In this case, the recall does not fall below 45% for an overall higher precision that still is lower that 50%. The customized network using the cleaned database yields the best results with a precision of 75% for a recall of 49%. Although better recall scores have been reached using the first network (66%), the corresponding precision is lower due to the large number of false positives. This renders the first network less interesting than the remaining three. In the case of the customized network, even if the recall does not exceed 61%, the precision can exceed 70%. For a given recall, the customized network produces better precision scores than other

networks. The customized network detects almost 50% of the manhole covers in the aerial images.

These results are better than those obtained by (Pasquet et al., 2016), using a machine-learning SVM approach jointly with a low level approach. In their case, for a precision of 66% the recall is only of 45% with a simpler database and numerous preprocessing steps, whereas for the same precision value we obtain a recall of 54% without any additional preprocessing.

An analysis of the false positives, shows that they mostly occur in heavily textures areas such as as shown in figure 5. To correct this problem, we are considering combining the CNN to a circular filter presented in (Bartoli et al., 2015). Instead of training the network with only the three RGB channels, it will be trained on four entries with the fourth being the result obtained with the circular filter. This will add a geometrical "shape" information to the RGB reflectance values. There are several ways of combining these inputs such as presented in (Park et al., 2016). A first method would consist in merging the inputs and training the network with four channels. It is also possible to chose the "location" where the combination will be carried out *i.e.* either in the start, middle or at the end of the network. Merging these data should constrain the network to recognize circular patterns and increase the recall by detecting manhole covers which are currently not detected, maybe because they do not stand out in the surrounding context *i.e.* the contrast between the manhole cover and the asphalt background is not strong enough for the identification based solely on RGB information. Adding an information on the shape would allow all circular objects on the ground to be included in the network and would add the likelihood score of the manhole covers which have low contrast.
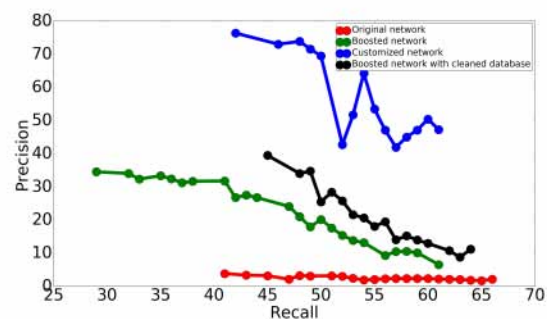

Figure 6. Precision *vs* Recall curve.

5. CONCLUSION

In this work, we have put forward a procedure that automatically detects manhole covers using aerial RGB images. We have trained it on the Prades-le-Lez database and tested it on the Gigean dataset. Four different network configurations were tested and compared in terms of recall, precision and F-measure. The preliminary results are encouraging and indicate that the model can efficiently detect these small objects in very high resolution aerial images with a recall that is higher than 50% and an average precision of 60%. To improve its performance, we are currently working on the combination of an information on the circular shape of the objects. Adding information on the geometrical shape of the objects may help in reducing false detection for objects that have low contrast with their surroundings. This will further improve precision. In addition, using simple preprocessing techniques to restrict our detection to specific regions of the aerial images

where these objects are usually found, such as roads and pavements, may enhance the detection rate by eliminating false positives. It is also interesting to note that these results were obtained using only two classes of objects while it is common practice in image processing to use several classes (Huynh et al., 2016), mainly to help the network extract more class-specific features and improve its detection rate. Training the network to recognize a third class such as inlet grates, which are urban objects that have similar size and are also located along roads and pavements, may allow the network to optimize the learning phase and extract more precise features. This is the work we are currently undertaking.

6. ACKNOWLEDGMENT

REFERENCES

Bartoli, O., Chahinian, N., Allard, A., Bailly, J.-S., Chancibault, K., Rodriguez, F., Salles, C., Tournoud, M.-G. and Delenne, C., 2015. Manhole cover detection using a geometrical filter on very high resolution aerial and satellite images. In: *Join Urban Remote Sensing Event*, Lausanne, Switzerland.

Bhattacharya, A., 1946. On a measure of divergence between two multinomial populations. *The Indian Journal of Statistics* pp. 401–406.

Cheng, Y., Xiong, Z. and Wang, Y., 2006. Improved classical Hough transform applied to the manhole cover's detection and location. *Optical Technique* 32(S1), pp. 504–508.

Chollet, F., 2015. Keras. `https://github.com/fchollet/keras`.

Everingham, M., Van Gool, L.and Williams, C. K. I., Winn, J. and Zisserman, A., 2010. The pascal visual object classes (voc) challenge. *International Journal of Computer Vision* 88(19), pp. 303–338.

Guan, H., Yu, Y., Li, C. and Wang, Q., 2014. Automated extraction of manhole covers using mobile LiDAR data. *Remote Sensing Letters* 5(12), pp. 1042–1050.

Huynh, C. K., Le, T. S. and Hamamoto, K., 2016. Convolutional neural network for motorbike detection in dense traffic. In: *2016 IEEE Sixth International Conference on Communications and Electronics (ICCE)*, pp. 369–374.

Krizhevsky, A., Sutskever, I. and Hinton, G. E., 2012. Imagenet classification with deep convolutional neural networks. *in Advances in neural information processing systems* pp. 1097–1105.

Li, X., Wang, L. and Sung, E., 2005. A study of adaboost with svm based weak learners. In: *Proceedings. 2005 IEEE International Joint Conference on Neural Networks, 2005.*, Vol. 1, pp. 196–201 vol. 1.

Lin, A. Y. M., Novo, A., Har-Noy, S., Ricking, N. and Stamatiou, K., 2011. Combining GeoEye-1 Satellite Remote Sensing, UAV Aerial Imaging, and Geophysical Surveys in Anomaly Detection Applied to Archaeology. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 4(4), pp. 870–876.

Neubeck, A. and Gool, L. V., 2006. Efficient non-maximum suppression. In: *18th International Conference on Pattern Recognition (ICPR'06)*, Vol. 3, pp. 850–855.

Niigaki, H., Shimamura, J. and Morimoto, M., 2012. Circular objet detection based on seperability and uniformity of feature distributions using bhattacharrya coefficient. In: *21st International conference on pattern regocnition (ICPR 2012)*, Tsukuba, Japan, pp. 2009–2012.

Park, E., Han, X., Berg, T. L. and Berg, A. C., 2016. Combining multiple sources of knowledge in deep cnns for action recognition. In: *2016 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pp. 1–8.

Pasquet, J., Desert, T., Bartoli, O., Chaumont, M., Delenne, C., Subsol, G., Derras, M. and Chahinian, N., 2016. Detection of manhole covers in high-resolution aerial images of urban areas by combining two methods. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 9(5), pp. 1802–1807.

Rajendra, S. and Chandraskaran, V., 2014. Beyond sliding windows: Monitoring urban expansion of Tiruchirapalli town of Tamil Nadu state, India using IRS-1C satellite data. In: *international federation of surveyors*.

Rogers, C., Hao, T., Costello, S., Burrow, M., Metje, N., Chapman, D., ... and Saul, A., 2012. Condition assessment of the buried utility service infrastructure: a proposal for integration. *Tunnelling and Underground Space Technology* 28, pp. 202–211.

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V. and Rabinovich, A., 2015. Going deeper with convolutions. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–9.

Tanaka, N. and Mouri, M., 2000. A detection method of cracks and structural objects of the road surface image. In: *Proc. IAPR Workshop Mach. Vis. Appl., Tokyo, Japan*, pp. 387–390.

Timofte, R., Zimmermann, K. and Van Gool, L., 2014. Multiview traffic sign detection, recognition, and 3d localisation. *Machine Vision and Applications* 25(3), pp. 633–647.

Tooke, T., van der Laan, M., Coops, N. and Christen, A., 2011. Classification of Residential Building Architectural Typologies using LiDAR. In: *International Joint Urban Remonte Sensign Event, Munich*, pp. 221–224.

Wang, L., Wang, S., Li, W. and Li, J., 2014. Monitoring urban expansion of the Greater Toronto area from 1985 to 2013 using Landsat images. In: *IEEE Geoscience and Remote Sensing Symposium, Quebec City*, pp. 4264–4267.

Yu, Y., Guan, H. and Ji, Z., 2015. Automated detection of urban road manhole covers using mobile laser scanning data. *IEEE Transactions on Intelligent Transportation Systems* 16(6), pp. 3258–3269.