AUTOMATED DETECTION OF MANHOLE COVERS IN MLS POINT CLOUDS USING A DEEP LEARNING APPROACH

Liyuan Qing[1], Ke Yang[2], Weikai Tan[1], Jonathan Li[1,2], Senior Member, IEEE*

[1]Department of Geography and Environmental Management, University of Waterloo
[2]Department of Systems Design Engineering, University of Waterloo
200 University Avenue West Waterloo, ON, Canada N2L 3G1
[*]Corresponding author: junli@uwaterloo.ca

IGARSS 2020 - 2020 IEEE International Geoscience and Remote Sensing Symposium | 978-1-7281-6374-1/20/$31.00 ©2020 IEEE | DOI: 10.1109/IGARSS39084.2020.9324137

ABSTRACT

Road manhole cover works as an important part of road construction. Timely detection can make a great progress in the development of road management. This paper proposes a rapid road manhole detection method using mobile LiDAR with state-of-the-art computer vision and deep learning techniques. Firstly, the road surface data is extracted from mobile laser scanning system(MLS). Then, the 2D geographic reference feature(GRF) images are formed from 3D point cloud. Finally, the object detector using deep learning technology was applied to locate and annotate the road manholes. Also, we adjusted the training model to present the better result with high confidence over 0.90. Compared with the previous method, the proposed method can correctly detect the manhole cover with higher rate of precision and F1-feature at 0.952 and 0.975 respectively, especially in the complex road situation.

Index Terms— Road manhole cover, road management, mobile laser scanning, deep learning.

1. INTRODUCTION

Road Manhole is regarded as one of the most non-negligible part in the road infrastructure. Manhole cover subsidence and the damages around manholes can cause the serious safety issues [1]. Therefore, regular and timely monitoring for road condition is very necessary for many transportation activities and it is very crucial for the citizen's safety and city's development [2].

In Ontario, Canada, there are about 40,000 km of highway pavement need to be inspected and maintained regularly [3]. It would be a huge workload for the traditional surveying that requires workers to walk along the road and use measuring instruments to collect the road information, which has low efficiency and high human-labor cost with potential safety issues [4]. Therefore, it is necessary to invite the novel technology into traffic data collection, so as to improve the efficiency.

In order to address the forehead problems, Mobile Laser Scanning(MLS) system is regarded as a reliable and cost-effective alternative for carrying out rapid road inspection and it has also been used in manhole cover detection [5]. It doesn't affect normal traffic on the road, which reduces the economic loss caused by traffic accident or closure and also releases the labor intensity and danger of manual measurement.

In the past few years, the computer vision was applied on the object detection [6]. Traditional object detection such us the Cascade classifier combined with Histograms of Oriented Gradient (HOG) descriptors or deformable part models (DPM) and Haar-like features selection or Support vector machines (SVM). However, when it comes to detect different objects,it is time-consuming. The Region Proposal + CNN(R-CNN) was proposed [7]to consider feature information from similar specific classes in the image so as to ensure a high recall rate with fewer windows, but it is inefficient to uniform all candidate boxes.Faster R-CNN [8] improved the accuracy by adding the region proposal network (RPN) and anchor box instead of the selective search method.However, the Faster R-CNN cannot meet the real-time requirement in terms of speed.

Nowadays, a novel neural network, YOLO (You Only Look Once [9] has made improvement for real time detection. It considers the object detection as a regression problem and inputs the whole graph to the network structure. YOLO can greatly raise the detecting speed and decrease the false positive ratio. The latest method YOLO v3 [10]improves the accuracy of small size of objects and speeds up the whole process comparing with the previous one.

In this paper, we trained the target images transformed from point cloud in MLS. The proposed method outperforms in varying road conditions. The main contributions include: 1) applying a latest deep learning technique is on road manhole cover detection; 2) generating and optimizing the training model for specific manhole cover detection.

2. METHOD

2.1. Data Preprocessing

In this research, we collected the road data from the Mobile Laser Scanning (MLS)system (RIEGL VMX-450), which efficiently provides the high quality of road data.

Since the manhole cover is supposed to be a part of the road surface infrastructure, the manhole cover detection can be considered as a kind of road surface surveying. Considering the further image processing and reduce the complexity of the point cloud data preprocessing, we applied the image segmentation by using the certain width curb to extract the road surface data from the original raw point cloud [11]. And then we formed 2D geographic reference features (GRF) from raw 3D point clouds. [12].

2.2. Manhole cover detection

In this paper, we proposed a supervised learning strategy to train a YOLOv3 model from a set of manually labeled training samples.

The YOLOv3 is the updated vision of YOLO. It had the 3 outperformed features: a) Using a pyramid network structure. b) Using logistic regression classifier instead of the softmax logistics [13]. c) Using the Darknet-53 which has the largest points operation per second among other backbones(such as ResNet-101, Darknet-19), its mean average precision also has well-improved [10]. This network utilizes na, nb, nw, nh as coordinates for each of bounding boxes. In a image, when one cell moves (la , lb)from top left of the region, considering the bounding box prior has width as pw and height as ph , then the predictions correspond to:

$$p_a = \sigma\left(n_a\right) + l_a \tag{1}$$

$$p_b = \sigma\left(n_b\right) + l_b \tag{2}$$

$$p_w = p_w e^{n_w} \tag{3}$$

$$p_h = p_h e^{n_h} \tag{4}$$

This network only assigns one boundary box to each ground truth object and each layer has 255 outputs: 85 outputs per anchor [4 box coordinates + 1 object confidence + 80 class confidences], times 3 anchors. Since we use 2 classes in this detection, we reduce the number of filters=[4 + 1 + n] * 3, where n is the class count as 2. This modification should be made to the layer preceding each of the 3 layers. Also write classes=2 in each layer. The class 1 represents the manhole cover, and class 0 represents the road background. In the purpose of training the specific model for detecting manhole covers, we calibrated the training model and analyzed the training result. Some parameters can be helpful to analyze the performance of the training model and optimize the calibration for the model.

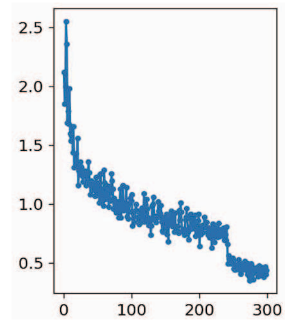


Fig. 1. $\mathcal{L}_{GIoU}$ when epochs =300

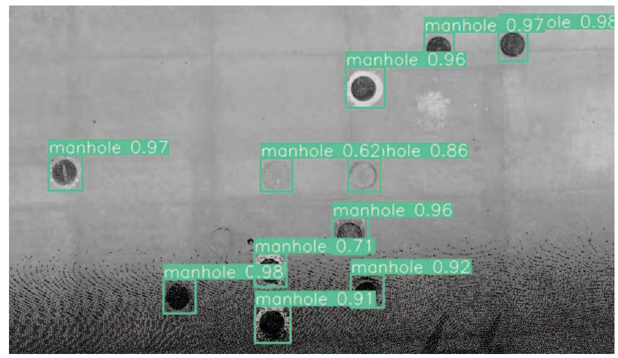


Fig. 2. Manhole cover label (epochs=200, batch-size=1)

Generalized Intersection over Union(GIoU) is a new metric that can be used as a 2D bounding box regression loss. It could test the measurement's performance in image detection [14]. Using GIoU loss $\mathcal{L}_{GIoU}$in training YOLOv3 can significantly improve the model's performance.

$$\mathcal{L}_{GIoU} = 1 - GIoU \tag{5}$$

which means the less $\mathcal{L}_{GIoU}$, the larger $GIoU$.

And the precision and the recall equations are also used to measure the performance of the model. Precision rate and recall rate are two important evaluation indicators in the target classification. In the field of information retrieval, precision is related to the predicted results. Recall rate is the proportion of relevant document in the returned result to all relevant documents.

$$Precision = \frac{TP}{TP + FP} \qquad Recall = \frac{TP}{TP + FN} \tag{6}$$

where TP, FP and FN are the number of True positive, False positive and False negative samples.

3. RESULTS AND DISCUSSIONS

The 2D images were collected by using the approach mentioned in [12].We trained the manhole covers which are mainly made of cast ion and in regular graphic figure. We put

Table 1. PARAMETER SETTING OF TRAINING MODEL

Epochs	100	200	300
Batch size	1	1	1
Batch size	2	2	2

36 manhole covers images with the corresponding labels into the training model together. The setting of input image size is 416*416. Meanwhile, the parameters of training model were adjusted to be applied to manhole cover detection. We alternated the model with different epochs=100, 200, 300 and different batch sizes(in Table 1). The performance of training models and output results of detection were discussed. Finally, we got the detection results (in Fig.2) by using the optimized training model. The comparison was also made with previous studies in detecting manhole covers in complex road manhole cover situation. All the training and detecting processes are running on GPU (Nvidia GTX 1080).

3.1. Model's optimization

We trained the model with different epochs, and found that the value of GIoU loss had a sharp decline when the epochs are equal to at around 240 (Fig.1). Meanwhile, the GIoU loss value was decreased to 0.50 and lower. When epochs are set at over 240, the accuracy of bounding box's prediction locating the target objects can be significantly improved.

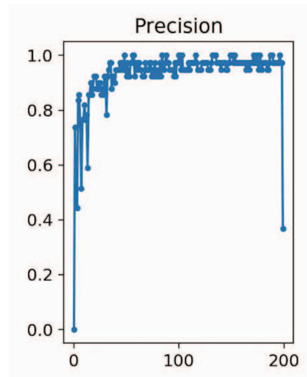
Comparing with different settings, the best result was found from the training model with the epochs=200 and batch size=1. The main precision closely to 0.973 and recall rises and stays at 0.995. From Fig.2, when the value of epoch is over 100, the precision rate stays at the same level, which explains that the epoch should be set above 100 times to get a better precision rate. Since the precision rate stays steadily and increase the size of epochs will need more time, there is no need to train the model with more epochs to increase the precision.

Although the training model with epochs=300 has less $\mathcal{L}_{GIoU}$ value than the training model with epochs=200, but from the manhole covers' annotation in output images, there is a 0.36 possibility shown in Fig.4. That appeared an over fitting phenomenon. We also found that batch size =1 can annotate more manhole covers, though increasing the batch size can reduce the training time, it also influenced the detection results.It shows that when the input images are in a limited number, it is better to minimize the batch size. Above all, we set epochs=200,batch size=1 for training model.

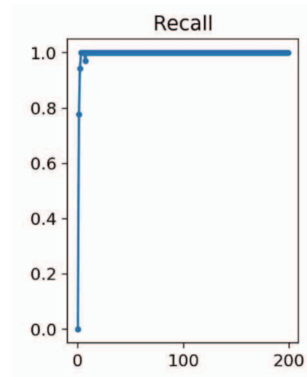
3.2. Comparative Study

In order to prove the performance of the YOLOv3 in detecting process, we detected the same objects so as to make a comparison with the previous method in [15]. However, due to the poor manhole cover's surrounding situation and condition, such as the covers printed with the road mark, thick dirt and mud, the previous method failed to detect the road manhole covers. The proposed method showed a better performance on detecting the manhole covers in the same tested images(Fig.5), even when the covers are not highly contrast to the road background.

We calculated the rate of precision, recall and F1-measure to evaluate the performances of those two models. The rate was calculated based on the tested images. Shown in Table 2, the proposed model outperforms the previous algorithms in successfully detecting the manhole covers. It can detect the manhole covers even those previous method failed to detect (shown in Fig.5). Above all, the proposed method can better reached the standard of manhole cover detection and it is a more appropriate alternative for manhole surveying.



(a)



(b)

Fig. 3. (a) Precision rate and (b) Recall rate when epochs =200, batch-size=1.

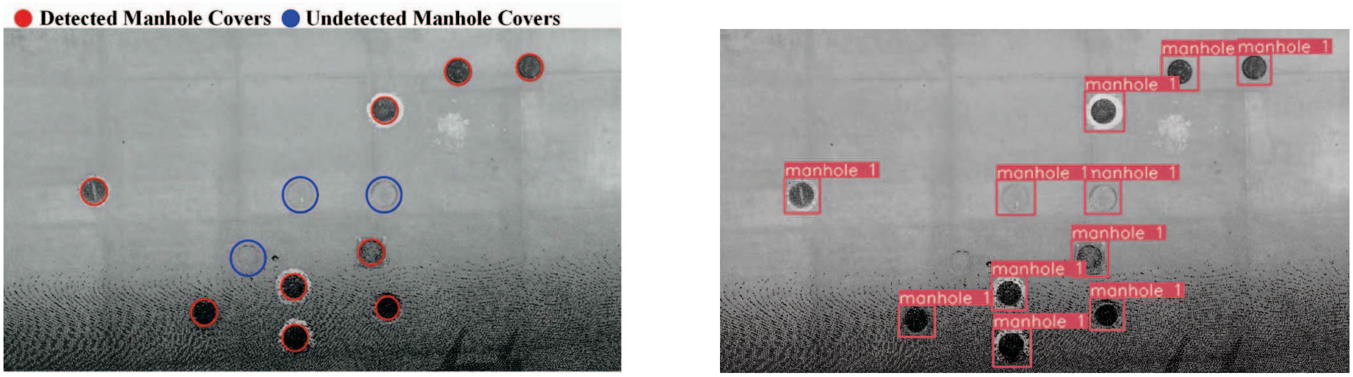
(a)

(b)

Fig. 4. Manhole cover annotation(a) epochs =300, batch-size=1 (b) epochs =200, batch-size=1.

4. CONCLUSIONS

In this research paper, we focused on applying the novel deep learning detection method on the road manhole cover detection and discussed the performance of this model. The high quality road manhole cover data was collected from MLS sys-



(a) (b)

Fig. 5. Comparison between(a) Previous method and (b) Proposed method in complex road background .

Table 2. QUANTITATIVE EVALUATION

	Precision	Recall	F1-measure
Previous method	0.857	1.000	0.923
YOLOv3 method	0.952	1.000	0.975

tem and 2D georeference figures were generated from the 3D-LiDAR point cloud data.Then we got the labeled samples as input to train and test the model to get the best weights for the detection procedure. We also alternated the simple factors to find the better training model. Comparing with the previous method, the advantages of YOLOv3 can be clearly described with the higher precision and F1-measure of 0.952 and 0.975 respectively, and also can be applied in complex situation.

In conclusion, the proposed method can be better applied with advanced MLS for road real-time surveying and has a indispensable role in road planning. Furthermore, in order to get the better training model, we need to train variety of labeled images and adjust this model to detect more classes of items.

5. ACKNOWLEDGMENT

This work was supported in part by the National Natural Science Foundation of China under Grants 41871380 and the Natural Sciences and Engineering Research Council of Canada under Grant 50503-10284.

6. REFERENCES

[1] Q. Zhao, Z. Li, W. Hu, X. Meng, and H. Zhang, "Driving Comfort Evaluation for Manhole Covers and Pavement around Manholes," *Adv. Mater.Sci.Eng.*, vol. 2019, p. 1293619, Oct. 2019, doi: 10.1155/2019/1293619.

[2] M. Cui, D. Han, and J. Wang, "An efficient and safe road condition monitoring authentication scheme based on fog computing," *IEEE Internet Things J.*, vol. 6, no. 5, pp. 9076–9084, 2019.

[3] Office of the Auditor General of Ontario, "3.10 Ministry of Transportation—Road Infrastructure Construction Contract Awarding and Oversight," Tech. Rep., 2016. [Online]. Available: http://www.auditor.on.ca/en/content/annualreports/arreports/en16/v1_310en16.pdf

[4] Y. Yu, H. Guan, and Z. Ji, "Automated detection of urban road manhole covers using mobile laser scanning data," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 6, pp. 3258–3269, 2015.

[5] Z. Wei, M. Yang, L. Wang, H. Ma, X. Chen, and R. Zhong, "Customized mobile lidar system for manhole cover detection and identification," *Sensors*, vol. 19, no. 10, p. 2422, 2019.

[6] L. Ma, Y. Li, J. Li, C. Wang, R. Wang, and M. A. Chapman, "Mobile laser scanned point-clouds for road object detection and extraction: A review," *Remote Sens.*, vol. 10, no. 10, p. 1531, 2018.

[7] G. Gkioxari, R. Girshick, and J. Malik, "Contextual action recognition with r* cnn," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2015, pp. 1080–1088.

[8] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," in *Adv. Neural. Inf. Process Syst.*, 2015, pp. 91–99.

[9] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 779–788.

[10] J. Redmon and A. Farhadi, "Yolov3: An incremental improvement," *arXiv preprint arXiv:1804.02767*, 2018.

[11] H. Guan, J. Li, Y. Yu, C. Wang, M. Chapman, and B. Yang, "Using mobile laser scanning data for automated extraction of road markings," *ISPRS J. Photogramm. Remote Sens.*, vol. 87, pp. 93–107, 2014.

[12] Y. Yu, J. Li, H. Guan, C. Wang, and J. Yu, "Automated detection of road manhole and sewer well covers from mobile lidar point clouds," *IEEE Geosci. Remote. Sens. Lett.*, vol. 11, no. 9, pp. 1549–1553, 2014.

[13] J. Redmon and A. Farhadi, "Yolo9000: better, faster, stronger," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 7263–7271.

[14] H. Rezatofighi, N. Tsoi, J. Gwak, A. Sadeghian, I. Reid, and S. Savarese, "Generalized intersection over union: A metric and a loss for bounding box regression," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, 2019, pp. 658–666.

[15] Y. Yu, H. Guan, D. Li, C. Jin, C. Wang, and J. Li, "Road manhole cover delineation using mobile laser scanning point cloud data," *IEEE Geosci. Remote. Sens. Lett.*, vol. 17, no. 1, pp. 152–156, Jan 2020.