LANDFILL DETECTION FROM RS IMAGES BY DEEP LEARNING

Hao Zhang, Yixuan Wu, Yating Zhou, Qian Ying, Zhonghao Zhang, Yongsheng Zhou

College of Information Science and Technology, Beijing University of Chemical Technology,
No.15 North Third Ring East Road, Chaoyang District, Beijing 100029, China
Email: 2017040261@mail.buct.edu.cn; 2017040412@mail.buct.edu.cn;
2017040413@mail.buct.edu.cn;zhyosh@mail.buct.edu.cn;

KEY WORDS: landfill, remote sensing, deep learning, detection

ABSTRACT: With the continuous acceleration of urbanization, the amount of municipal waste is increasing rapidly, and the problem of landfill is becoming more and more serious. Remote sensing (RS) technology has the characteristics of large-scale and high-frequency observation, which has significant advantages for the landfill detection. The traditional methods for detecting landfill are based on manual visual inspection and/or semi-manual interpretation. Those methods are usually costly, labor intensive, or low in efficiency and accuracy. For the above problems, this paper presents a detection method based on the deep learning technology. Firstly, 355 landfill images are searched and downloaded from Google Earth to build landfill datasets. Secondly, considering the difficulty of application and detection efficiency, YOLOv3 and YOLOv5 are selected for training. The model parameters are properly adjusted and tested in order to accurately identify the characteristics of the landfill. Finally, experiments are performed and the results verify that the method can achieve accurate detection of landfill.

1. INTRODUCTION

With the significant improvement of economic level, people's material consumption level is constantly improving, and the total amount of garbage is also growing rapidly. Illegal landfill has caused severe pollution to the surroundings (Liu, 2012). In recent years, incidents such as illegal landfill pollutes groundwater have emerged one after another (Wang, 2002). It is urgent to monitor the status and changes of illegal landfill continuously. However, field monitoring method currently has some problems such as large error and discontinuity (Qing, 2016).

Remote sensing (RS) technology is increasingly mature with the continuous development of space technology and payload technology. The resolution and image quality of remote sensing image is gradually increasing. Remote sensing image has been widely used in military, agriculture, resource survey and other fields. Remote sensing technology has the advantages of wide viewing angles and far sight distance which can be better used in the monitoring of illegal landfill site. It has been gradually applied to the monitoring of illegal stacking of garbage recent years (Wang, 2016).

In addition, it can be seen that traditional methods that based on satellite remote sensing image for detecting illegal landfill are mainly based on manual visual inspection and/or semi-manual interpretation. They usually have great shortcomings in automation and precision (Liu, 2009).

Deep learning is the developing trend in the field of artificial intelligence having deep network structure. Hinton proposed Greedy Layer-Wise Unsupervised Pretraining based on Deep Belief Network (DBN), and then proposed a deep structure of the multilayer automatic encoder several years ago (Sun, 2012). Moreover, it is widely used in animal husbandry industry (Su, 2020), fault diagnosis (Ren, 2017), human behavior recognition (Zhu, 2016) and so on.

In this paper we combine deep learning technology with remote sensing technology in order to detect landfills effectively in remote sensing images. YOLOv3 and YOLOv5 are selected. We demonstrate that by combining remote sensing and deep learning detection technology, the detection accuracy is accurate.

The rest of this paper is organized as follows. Section 2 mainly introduces the landfill datasets. Session 3 mainly shows methodology, including model and evaluation index. Section 4 shows the experiment and result discussion. The conclusion is discussed in Section 5.

2. BUILD LANDFILL DATASETS

In this paper, a total of 355 images are searched and downloaded from Google Earth. Firstly we change sizes of samples to make them the same which are different in the original images. Secondly, we use the image labeling tool LabelImg to label the remote sensing images manually according to the VOC2007 format. This is because we should obtain the .xml file containing information of corresponding images. Images obtained from Google Earth are in a limited scene. The data augmentation is a commonly used technique for increasing both the size and the diversity of labeled training sets by leveraging input transformations that preserve output labels (Buslaev, 2020). Thus, we implement basic image transformations, concluding flip, rotate, crop and so on.

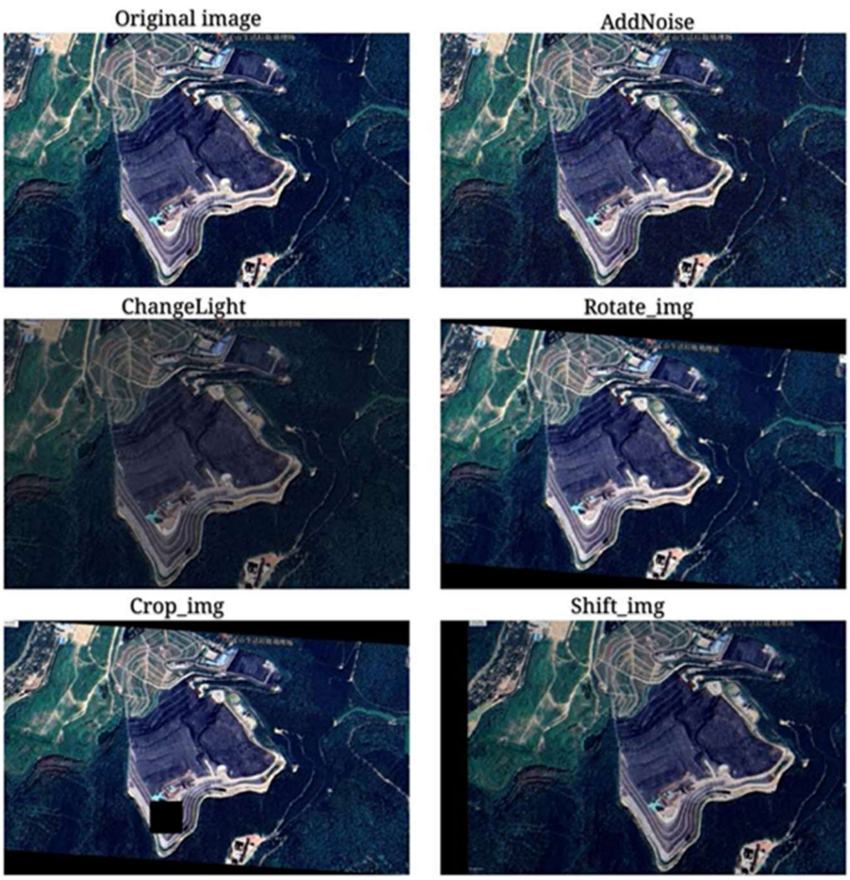

Figure 1. Illustration of the data augmentation.

Then, according to the ratio of 0.7:0.2:0.1, datasets are divided into three parts, i.e., train.txt (training set), test.txt (test set), and val.txt (validation set).

3. METHODOLOGY

3.1 Architecture of Target Detection Methods

Currently, the anchor-based target detection methods can be generally divided into types, i.e., two-stage target detection algorithm and one-stage target detection algorithm. The two-stage method takes a series of candidate frames obtained by the algorithm as samples. It is a two-stage cascaded network. The first stage proposes the candidate frame, and the second stage judges the candidate frame.

The one-stage method does not generate candidate frames. It directly converts the positioning problem of the target border into a regression problem, and the whole process of generating candidate boxes is integrated.

Two-stage method is better in detection accuracy and positioning accuracy, while one-stage method is faster than two-stage method and suitable for mobile devices and other platforms.

3.2 YOLO Algorithm

YOLO (You Only Look Once) algorithms have been widely used in the field of target detection. YOLO algorithms are better at detecting objects in different locations than that in two-stage algorithm (Duan, 2012). YOLO algorithm is to extract features from the input image through the feature extraction network and obtain a feature map of certain size. The YOLO algorithms divide an image into $S*S$ grid cells. Each grid cell predicts B objects and predicts 5 statistics, including x, y, w, h and a confidence level. Meanwhile, each grid cell which is classified as C should predict one category of information. The predicted output feature map has two dimensions, $S*S$ and $B*(5+C)$. The algorithms use the idea of regression to transform the target detection task into a regression problem which greatly accelerates speed of detection. The YOLO algorithms can learn the general information of the target and improve the accuracy of detection as well.

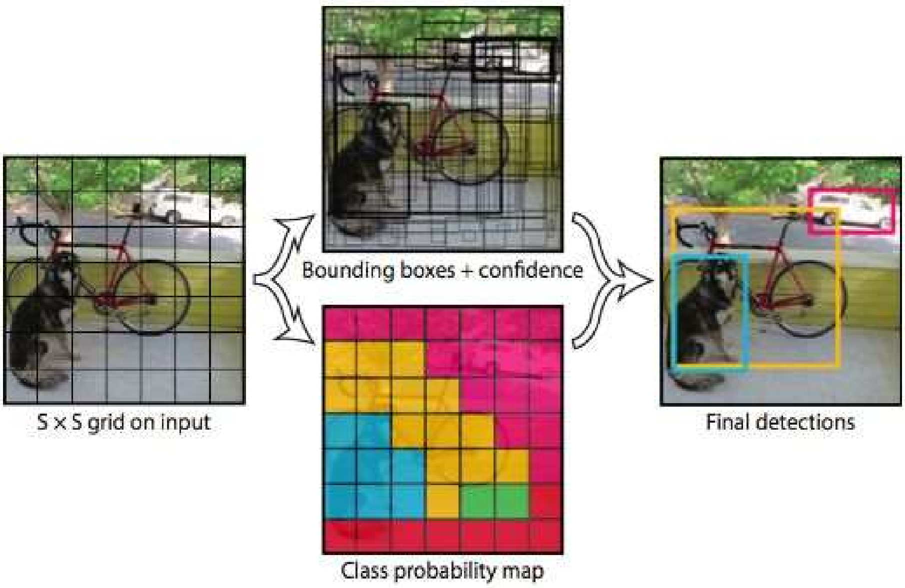

Figure 2. Flows of YOLO algorithms (Redmon, 2016).

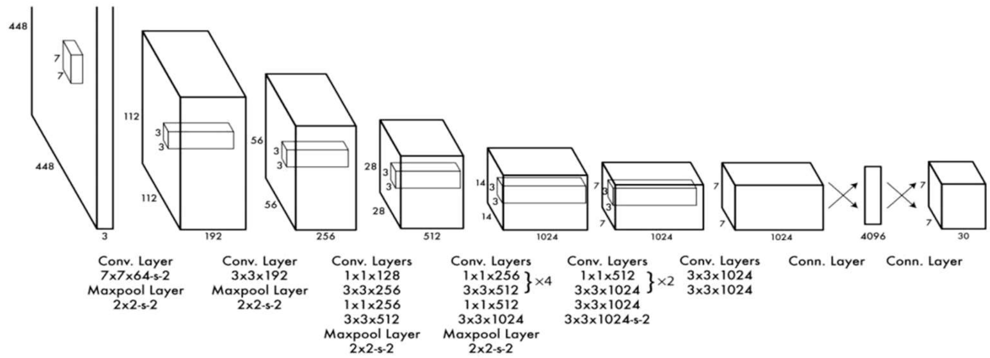

Figure 3. The model of YOLO convolutional neural network (Redmon, 2016).

YOLOv3 is optimized for classifier network. It uses Darknet53 network and it is combined with FPN network structure to derive predictions from convolution networks. It further improves the detection speed and reduces the background error detection rate.

YOLOv5 has a lighter and more flexible model, comparing with YOLOv3. Model designs of different sizes and complexity can be achieved through depth_multiple (scaling factor of channel) and width_multiple (layer scaling factor of BottleneckCSP module) coefficients. YOLOv5 takes full use of the latest version of PyTorch to mix precision. While YOLOv5 performs better than YOLOv3 in comprehensive performance through some built-in hyperparameter optimization strategies. The speed of small target recognition of YOLOv5 is greatly improved as well, reaching at 140FPS.

Table 1. Comparison of YOLOv5s, YOLOv5m and YOLOv5l.

	YOLOv5s	YOLOv5m	YOLOv5l
depth_multiple	0.33	0.67	1.0
width_multiple	0.50	0.75	1.0

3.3 Model Performance Evaluation Index

In the target detection algorithm, samples are usually divided into four cases, TP (true positive), FP (false positive), TN (true negative), FN (false negative) for the binary classification problem according to the combination of their real category and learner prediction category.

Table 2. Binary classification confusion matrix.

Actual Class	Predicted Class	
	Positive	Negative
Positive	TP (True Positive)	FN (False Negative)
Negative	FP (False Positive)	TN (True Negative)

The accuracy of the detection is determined by Precision (TP/(TP+FP)) and Recall (TP/(TP-FN)). In general, the higher the Recall is, the lower the accuracy is. However, it is not enough to use accuracy and Recall to measure the quality of a deep learning network model. To better evaluate model performance, mAP values are used. AP value is the area of two axes, Precision and Recall respectively. mAP is the average of AP values which are in various categories. The range of mAP is from 0 to 100. As the indicators of measuring detection accuracy in target detection, when mAP value is close to 100, the detection accuracy of the model is very high.

4. EXPERIMENTAL AND DISCUSSION

4.1 Experimental Process

Based on the above two datasets, YOLOv3 algorithm is used to train datasets in the first step. After using the official code to convert datasets of the initial VOC2007 format into the format which suits YOLOv3 training, anchors values based on the above datasets are obtained by using *K*-means clustering algorithm. The model parameters are adjusted then the test results are obtained finally.

YOLOv5 algorithm is used to train the same two datasets in the second step. In the paper we use the PyTorch Deep Learning Framework to build the YOLOv5 network in the following environments: graphics GTX 1050, memory 16G, operating system Windows 10, CUDA10.2, PyTorch 1.6, some necessary dependency libraries and the development environment is Python 3.8.5. The initial datasets of VOC2007 format are then converted into YOLO format which is suitable for training in YOLOv5 by using official code. In addition, we configure the datasets in the .yaml file. Then record locations of the training set, the validation set and the test set and record the number of detection classes. After entering from the training entrance train.py file and adjusting the parameters to satisfy the overall needs, we begin to train. During the training process of YOLOv5, indicators of training loss and performance are automatically saved in the tensorboard and results.txt files. Statistics in results.txt are drawn as results.png when the process is complete at last.

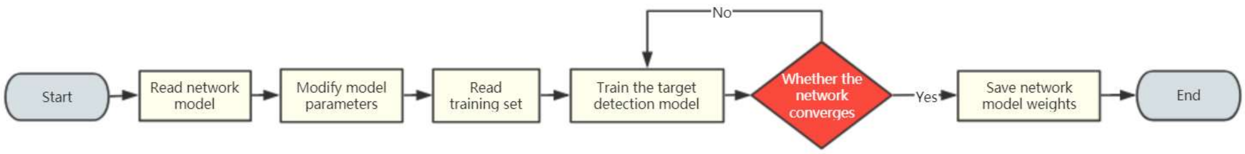

Figure 4. The target detection algorithm model.

4.2 Results and Analysis

YOLOv3 and YOLOv5 series models are trained under Dataset 1 and Dataset 2, and the results are shown in Table 3.

Table 3. The results of each model under different datasets.

	Module	Precision	Recall	mAP.5	mAP.95
Dataset1	YOLOv3	0.976	0.904	0.937	
	YOLOv5s	0.949	0.872	0.936	0.65
	YOLOv5m	0.953	0.942	0.948	0.686
	YOLOv5l	0.952	0.947	0.957	0.693
Dataset2	YOLOv3	**0.992**	0.983	0.985	
	YOLOv5s	0.984	**0.994**	**0.994**	0.777
	YOLOv5m	0.986	0.992	**0.994**	0.812
	YOLOv5l	0.98	0.992	**0.994**	**0.826**

It can be concluded from Table 3 that in the detection of landfills, both YOLOv3 and YOLOv5 series models can achieve high detection accuracy. The mAP value of YOLOv5 series models is slightly higher than that of YOLOv3.

The detection accuracy of each model under Dataset 1 and Dataset 2 is compared. The detection accuracy of each model trained under Dataset 2 is higher than that of each model trained under Dataset 1. It can be seen that in the case of the same model, through data augmentation, the number of samples is increased, and the diversity of samples is enriched. As a result, the accuracy of the model is improved within the same training time.

The three models of YOLOv5 are also compared. YOLOv5s has the smallest depth and the smallest width of the feature map among the three models of YOLOv5. YOLOv5m and YOLOv5l are constantly deepening and widening on this basis. It can be seen that the mAP.95 value of YOLOv5l is higher than the mAP.95 value of YOLOv5m, and the mAP.95 value of YOLOv5m is higher than the mAP.95 value of YOLOv5s. This is due to that YOLOv5s has the smallest network, and the detection accuracy is the lowest. On this basis, YOLOv5m network and YOLOv5l network have continuously deepened and widened the network, and the detection accuracy has been continuously improved.

4.1.1 The Detection Results of YOLOv3

Figure 5 and Figure 6 show the detection results of YOLOv3 in Dataset1 and Dataset2.

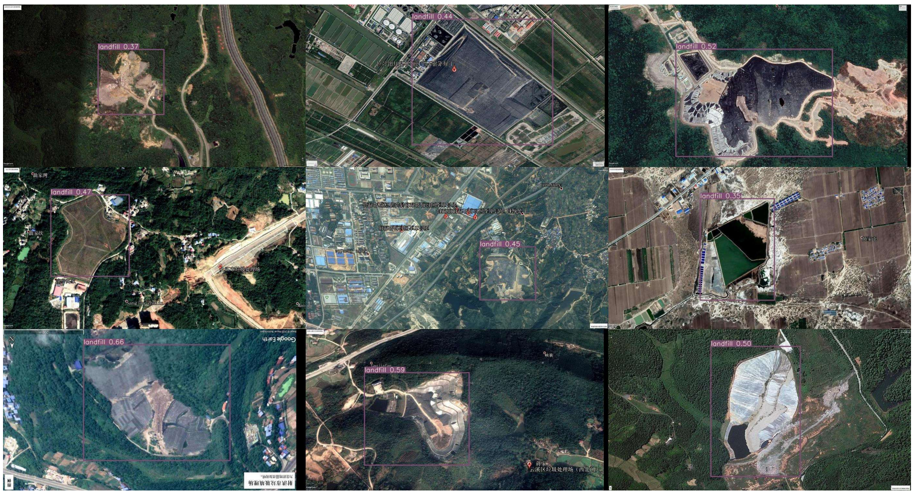

Figure 5. Detection results of YOLOv3 in Dataset1.

Figure 6. Detection results of YOLOV3 in Dataset2.

From Figure 5 and Figure 6, it can be seen that YOLOv3 can initially identify landfills and can basically satisfy the needs of landfill detection. By comparing Figure 5 and Figure 6, it can be concluded that the detection accuracy of YOLOv3 under Dataset 2 is higher than the detection accuracy of YOLOv3 under Dataset 1. The augmentation of the data set samples increases the detection accuracy.

4.1.2 The Detection Results of YOLOv5

Figure 7 and 8 show the detection results of YOLOv5s in Dataset1 and Dataset2. It can be seen that YOLOv5s has high detection accuracy and can achieve accurate detection on landfill, which can basically meet the needs of landfill detection in practical applications. According to Figure 7 and Figure 8, it can be concluded that the detection accuracy under Dataset 2 is greater than that under Dataset 1.

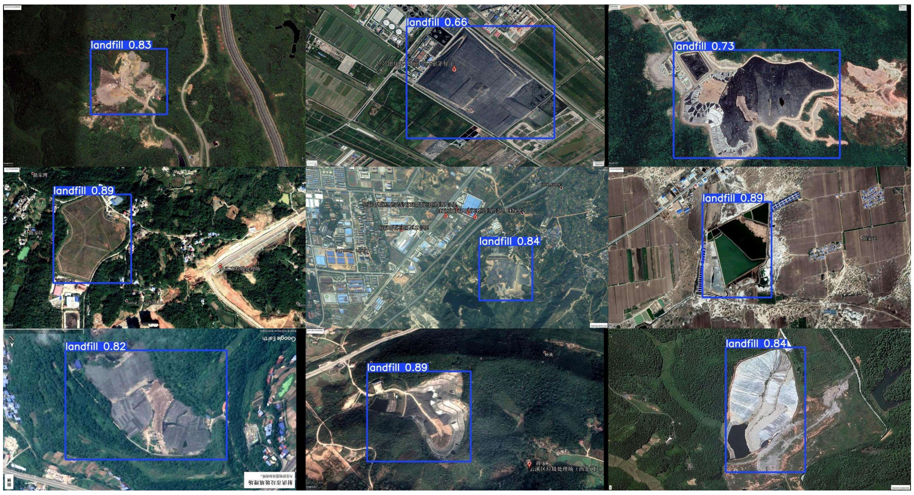

Figure 7. Detection results of YOLOv5s in Dataset1.

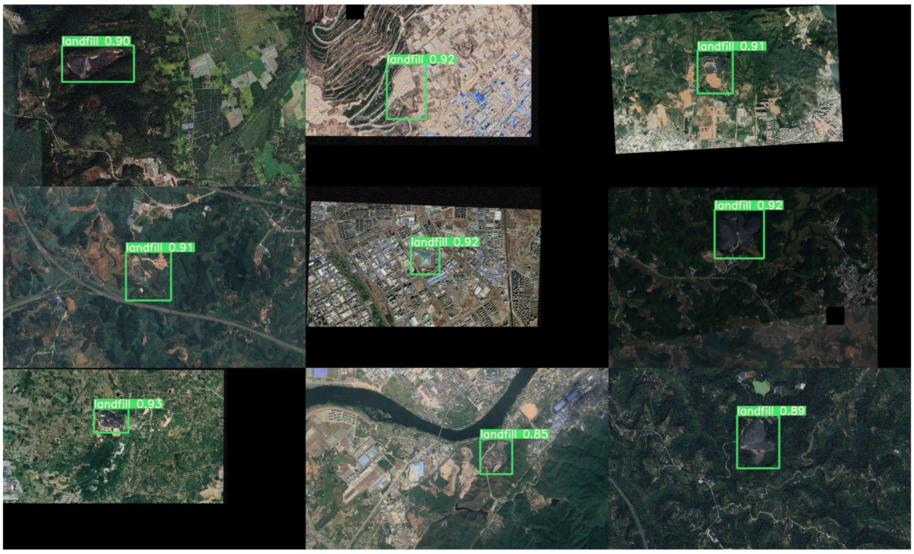

Figure 8. Detection results of YOLOv5s in Dataset2.

Figure 9 and 10 show the detection results of YOLOv5m in Dataset1 and Dataset2. It can be seen that the detection accuracy of YOLOv5m is excellent, and it can achieve accurate detection of landfills. Furthermore, both the detection accuracy of Dataset 1 and Dataset 2 are higher than that of YOLOv5s.

Figure 9. Detection results of YOLOv5m in Dataset1.

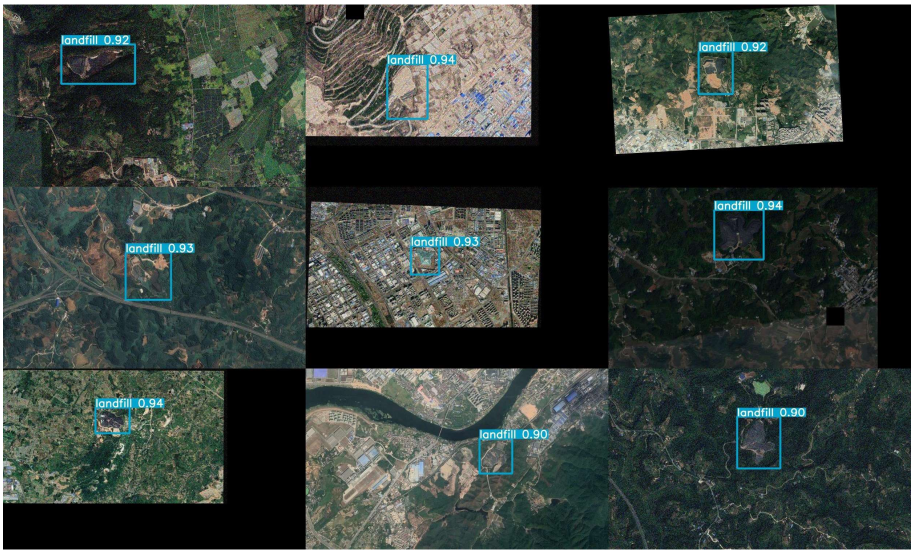

Figure 10. Detection results of YOLOv5m in Dataset2.

Figure 11 and 12 show the detection results of YOLOv5l in Dataset1 and Dataset2. It can be seen that the detection accuracy of YOLOv5l is particularly excellent, which can realize accurate detection of landfills. And the detection accuracy is the best among the 3 models of YOLOv5.

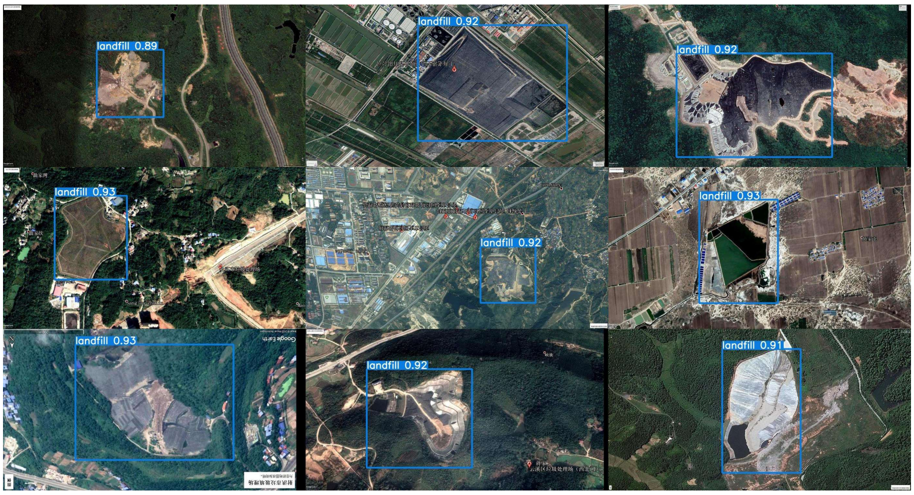

Figure 11. Detection results of YOLOv5l in Dataset1.

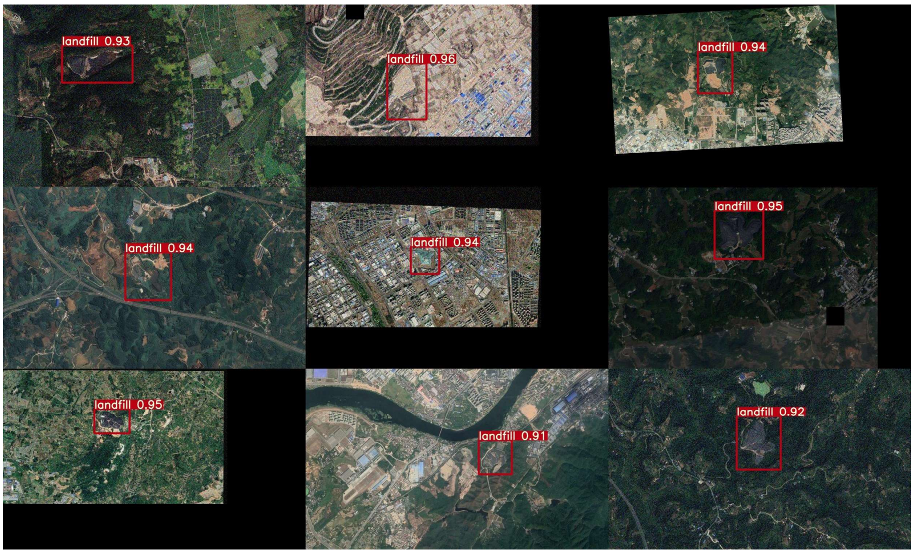

Figure 12. Detection results of YOLOv5l in Dataset2.

5. CONCLUSION

The traditional landfill detection methods based on remote sensing images usually have low detection accuracy and low level of automation. This paper presents a method based on deep learning to identify landfills. Excellent detection results have been achieved. However, landfill detection from remote sensing images by deep learning involves many theories, methods and technologies and this method is not mature at present, and there are many issues needed to be addressed and solved. In the future, we will perform optimization from the following aspects.

1. After acquiring the remote sensing images of the landfill, this method only performs simple processing on the remote sensing images to form a dataset for training. Next, we will conduct a deeper analysis of the remote sensing image information of the landfill and investigate how to use existing methods to improve the efficiency of remote sensing image preprocessing and feature extraction.

2. At present, although this method can detect landfills in different environments, there are still many shortcomings in terms of detection accuracy. We will continue to try different models and add tricks in model training and other methods to improve accuracy.

REFERENCES

Buslaev, A., Iglovikov, V. I., Khvedchenya, E., et al, 2020. Albumentations: fast and flexible image augmentations. Information, 11(2): 125.

Duan, Z., Li, S., Hu, J., Yang, J., Wang, Z., 2020. A review of deep learning target detection methods and their mainstream frameworks. Progress in laser and optoelectronics, 57 (12), pp. 59-74.

Fan, H., Huang, L., 2004. Monitoring and management of municipal solid waste landfill by remote sensing and geographic information system. Jiangxi chemical industry, (04), pp. 50-52.

He, K., Zhang, X., Ren, S., et al., 2015. Deep residual learning for image recognition. Computer Vision and Pattern Recognition.

Hinton, G., Osinderos, Teh, Y., 2006. A fast learning algorithm for deep belief nets. Neural Computation, 18 (7), pp. 1527-1554.

Liu, Y., Ren, Y., Wei, C., Wang, A., Zhou, H., Chi, Y., 2009. Application of Beijing-1 small

satellite to monitor irregular garbage dump. Journal of Remote Sensing, 13 (02), pp. 320-326.
Liu, Z., He, L., Dong, J., Meng, R., Song, B., Xi, B., 2012. Groundwater pollution risk assessment of simple landfill site. Environmental Science Research, 25(07), pp. 833-839.
Qin, H., 2016. Research on supervision method of urban domestic waste based on domestic high resolution remote sensing images. China Construction Informatization 2016(04), pp. 75-77.
Qin, H., Zhang, X., 2015. Research on monitoring and management of urban landfill based on domestic high resolution remote sensing images. China Construction Informatization, (21), pp. 76-78.
Redmon, Joseph, et al, 2016.You Only Look Once: Unified, Real-Time Object Detection. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 779–788.
Ren, H., Qu, J., Chai, Y., Tang, Q., Ye, X., 2017. In-depth learning in the field of troubleshooting research status and challenges. Control and decision-making, 32(08), pp. 1345-1358.
Su, H., Zheng, D., 2020. Applied research on the target detection algorithm of pig images based on deep learning technology. Jilin Agricultural University Journal, 2020-10-17, pp. 1-8.
Sun, Z., Xue, L., Xu, Y., Wang, Z., 2012. Review of deep learning research. Computer application research, 29 (08), pp. 2806-2810.
Erhan, D., Szegedy, C., Toshev, A., et al, 2014. Scalable object detection using deep neural networks. 2014 IEEE Conference on Computer Vision and Pattern Recognition.
Wang, C., Yin, S., Meng, B., Ma, W., Zhu, L., Wu, C., 2016. Remote sensing monitoring and analysis of irregular waste sites in Beijing Tianjin Hebei region. High tech communication, 26 (z1), pp. 799-807.
Wang, Y., Shen, X., Xie, J., 2006. Overview of remote sensing image classification methods. Remote Sensing Information, 2006 (05), pp. 67-71.
Wang, Y., Zhao, Y., 2002. Groundwater pollution by municipal solid waste landfill in Beitiantang District, Beijing. Hydrogeology and Engineering Geology, 2002 (06), pp. 45-47 + 63.
Zhao, J., Zang, K., Gong, H., Zhao, W., 2005. Application of remote sensing technology in garbage location and treatment in Beijing. Journal of Capital Normal University (Natural Science Edition), (03), pp. 109-113.
Zhu, Y., Zhao, J., Wang, Y., Zheng, B., 2016. Summary of human behavior Al recognition algorithms based on deep learning. Journal of Automation, 42 (06), pp. 848-857.