Research on the Application of Deep Learning YOLOv3 in Aerial Patrol Inspection of Optical Cable Lines

Mingjiang Zhang*, Weihu Zhao, Hongwei Li, Feng Wang and Shuai Zhang

Institute of Information and Communication, National University of Defense Technology, Xi’an, Shaanxi, 710106, China

*Corresponding author’s e-mail: 422784285@qq.com

Abstract. Since the construction of engineering vehicles such as excavators in the vicinity of optical cable lines is the most important cause of the fault of optical cable lines, it is proposed to apply the deep learning YOLOv3 (You Only Look Once version 3) target detection algorithm to the engineering vehicle detection and warning of the aerial inspection image. Based on the Darknet deep learning framework, the rapid target detection of engineering vehicles is realized by making engineering vehicle datasets. The detection speed is 60ms per sheet, and the target recognition AP (Average Precision) value is 0.869, which exceeds the algorithms such as deep learning Faster R-CNN (Faster Region-based Convolutional Neural Network) algorithm and machine learning HOG+LBP+SVM (Histogram of Oriented Gradient + Local Binary Pattern + Support Vector Machine) algorithm. The research results can provide some reference for aerial patrol and inspection and early warning for optical cable lines.

1. Introduction

The optical cable network is an important basic communication facility in China. Its barrier-free and smooth transmission is the prerequisite for the stable operation of the national communication network. Optical cable lines inspection, is an important means of maintenance of fiber optic cable line, according to the cable line laying mode and geographical environment, the traditional patrol mode is generally divided into walking inspection or motorcycle inspection. In recent years, with the rapid development of unmanned industry and artificial intelligence field, the drone aerial photography inspection mode of optical cable lines is being accepted by more and more communication security and management departments. It has the advantages of high inspection efficiency and safety, and has become a research hotspot in the field of optical cable inspection[1]. After the aerial photos along the optical cable lines are collected, how to accurately and quickly detect the hidden dangers along the lines from the mass inspection images has become the key technical problem to realize the aerial inspection. In literature[2], the hidden danger of external invasion of engineering vehicles near the transmission line was detected and studied, and a target fusion detection scheme was proposed. First, the suspected target area was extracted by PBAS algorithm, and then VGG16 convolutional neural network was used for classification and identification. Although target detection of engineering vehicles was achieved, the process was relatively tedious. In the literature[3], aiming at the hidden danger of excavator construction near the optical cable lines, based on the deep-learning Faster R-CNN target detection algorithm, the target detection of engineering vehicles such as excavators in the UAV aerial photography environment was achieved. Although the detection effect was good, the detection accuracy still needed to be further improved. Based on this, this article through deep learning YOLOv3 target detection algorithm[4], with engineering vehicles such as excavators along the cable as a hidden target, and through simulation...
training, implements the target detection of hidden trouble along the cable. Compared with other traditional deep learning and machine learning target detection algorithms, the detection AP value has been greatly improved.

2. Engineering Vehicle Detection in Aerial Patrol Inspection of Optical Cable Lines

According to statistics, as shown in figure 1, the optical cable damage caused by engineering vehicles such as excavators accounts for more than half of all the events that cause transmission obstacles of optical cable lines[5]. Therefore, it is an important part of maintenance and management of the optical cable lines to conduct inspections on the engineering vehicles along the optical cable lines and then investigate the hidden dangers of construction.

In order to solve this problem, the engineering vehicles such as excavators along the route can be quickly detected in the collected mass photos of the route by means of drone aerial photography and deep learning target detection technology. Combined with the GPS coordinate information of the inspection photos, the inspection personnel can be assisted to quickly and accurately find the hidden danger of construction on the route, so as to realize the early warning of optical cable line faulty and improve the work efficiency of inspection. The aerial patrol inspection process based on deep learning target detection is shown in figure 2.

In the image processing link of aerial patrol inspection of optical cable lines, this paper will use deep learning YOLOv3 target detection algorithm to simulate the automatic target detection of engineering vehicles such as excavators.

3. Deep Learning YOLOv3 Target Detection Method

In 2006, professor Geoffrey Hinton[6] made deep learning a hot topic in academic circles. Up to now, deep learning target detection algorithms can be divided into two categories: one is region-based convolutional neural network, such as Fast R-CNN[7] and Faster R-CNN[8]. The other is regression-based convolutional neural network, which is typically represented by SSD(Single Shot MultiBox Detector)[9], YOLO[10] and YOLOv2[11] target detection algorithms.

In the research on target detection of engineering vehicles in the aerial photography environment, the Faster R-CNN target detection algorithm has shown good detection effect based on small-scale data sets, but its detection accuracy needs to be further improved[3]. In 2018, the YOLOv3 algorithm proposed by Redmon J[4] shows better detection effect in terms of detection speed and accuracy. As shown in figure 3, YOLOv3 uses the new darknet-53 feature extraction network structure[12] and adds a mixed method of residual network to darknet-19. Continuous 3×3 and 1×1 convolutional layer are used to expand the network structure to layer 53. In terms of prediction methods, 3 boundary boxes are predicted for each grid in YOLOv3. Although there are fewer boundary boxes than the 5 boundary boxes predicted by YOLO v2, the number of final boundary boxes is much larger than before due to the use of multiple scale feature graphs for prediction. In the network output layer, unlike the softmax classifier, YOLOv3 uses an independent logical classifier to use a multi-label classification for each prediction box to predict the classes that the boundary box might contain.
To sum up, YOLOv3 improves the previous YOLO series target detection algorithms' poor detection effect for small targets, making them achieve the optimal comprehensive performance in detection accuracy and speed. This provides a possibility to improve the quality and efficiency of the target detection of engineering vehicles in the aerial photography inspection environment. This paper will simulate the target detection of engineering vehicles in the aerial photography images of optical cable lines based on YOLOv3.

4. Engineering Vehicle Target Detection in Aerial Images based on YOLOv3

4.1. Experimental environment
- Hardware: GPU is NVIDIA GeForce GTX1050; Video memory is 4G; The hard disk is 1T+128SSD.
- Software: deep learning of Darknet framework, Windows10 operating system, compiled by VS2015, based on CUDA8.5, cudnn6.5, Opencv3.1.0 and Python3.5 software.

4.2. Dataset making
In order to obtain the picture data set of engineering vehicles in the aerial photography environment, a total of about 500 aerial photos including excavators, bulldozers and other engineering vehicles are selected from VEDAI[13] aerial photography data set. These vehicles have a small target in the image, and at the same time, the vehicle is parked randomly with shadow parts near the body, which is relatively close to the aerial photography reality. In order to expand the number of data sets, the number of aerial photos is expanded to 1000 by means of mirror image inversion. Second, the data set needs to be eventually modified to conform to the YOLOv3 training format. The idea is to first read the annotation file of VEDAI dataset, programmatically convert the dataset into a format consistent with VOC2007 dataset, and then generate corresponding XML files for all images in the Annotations folder, and put the generated training and test TXT files containing image names into the ImageSets\Main folder. Finally, the corresponding coordinates for each image are generated and the normalized location label TXT file aput into the Labels folder. The label file format is shown in figure 4, providing the category number of the engineering vehicle target in the image and the normalized location coordinate information. In this
paper, the engineering vehicles in the data set are classified into one category for testing. The category name is “ENG-Vehicle”.

![Figure 4. Label format for YOLOv3 dataset.](image)

### 4.3. Target detection network training process

In the training process of the engineering vehicle target detection network, the pre-training weight file darknet53.conv.74 based on the large data set ImageNet is loaded. Since only one type of target is detected, the three classes and filters in CFG configuration file are set as 1 and 18 respectively. In addition, the main parameter settings before training are shown in table 1.

| Parameter name | Value |
|----------------|-------|
| Batch          | 64    |
| Max_batches    | 2000  |
| Stepsize       | 1600,1800 |
| Momentum       | 0.9   |

| Parameter name       | Value |
|----------------------|-------|
| Subdivisions         | 64    |
| Policy               | steps |
| Learning_rate        | 0.001 |
| Decay                | 0.0005 |

According to the algorithm characteristics of YOLOv3, set Max_batches to 2000 and Learning_rate to 0.001. Combined with the convergent change characteristics of the loss function in the process of network training, step training mode is adopted, that is, when the iteration number reaches the value specified by Stepsize, the learning rate is multiplied by 0.1. Set Momentum to 0.9 and Decay to 0.0005.

In order to analyze the network training process and evaluate the training effect, the curve of average loss (Avg Loss) changing with the number of iterations (Batches) is drawn, as shown in figure 5. As shown in the figure, with the increment of Batches, Avg Loss finally reaches a smaller convergence state, indicating that the training effect of network is good.

![Figure 5. Average loss curve in network training.](image)

**Figure 5. Average loss curve in network training.**

![Figure 6. P-R curve of engineering vehicle target detection.](image)

**Figure 6. P-R curve of engineering vehicle target detection.**

### 4.4. Analysis of experimental results

Based on the trained target detection network, the test results of some engineering vehicles are shown in figure 6. The excavator and other engineering vehicles in the picture are successfully identified and selected by the rectangular box. The “ENG-Vehicle” marked on the box indicates that the identified target is the engineering vehicle.
As can be seen from the figure, this algorithm can identify most of the engineering vehicles such as excavators correctly on the test dataset, and successfully detect engineering vehicles even when there is interference from cars, trucks and other vehicles in the scene, showing good robustness. In order to evaluate the detection performance of the model on the engineering vehicle, a P-R (Precision-Recall) curve is drawn, as shown in figure 7. The Recall of horizontal axis and the Precision of vertical axis represent the missed detection and false detection of the target detection network respectively. The closer the area under the curve is to 1, the better the detection effect is. It can be seen from the figure that the P-R curve is relatively full, and the AP (Average Precision) can reach 0.869. Based on the experimental platform, the average detection time for each image is 60ms, and the detection speed is nearly real-time.

According to the verification, the AP obtained by YOLOv3 in this paper is 0.428 higher than the mAP obtained by the traditional HOG+LBP+SVM machine learning algorithm in [13] which is also based on the VEDAI dataset. In addition, based on the same experimental environment and engineering vehicle dataset, the deep-learning target detection algorithm Faster R-CNN is adopted in this paper for comparison. The experimental results show that the AP based on this algorithm is 0.659 and the average detection time for each image is 109ms. The experimental results are shown in table 2. It is obvious that YOLOv3 is more suitable for the target detection of engineering vehicles under the aerial photography environment.

| Target detection method | Target detection object | mAP  | Detection speed (f/s) |
|-------------------------|-------------------------|------|----------------------|
| HOG+LBP+SVM             | Various vehicles (including engineering vehicles) | 0.441 | /                    |
| Faster R-CNN            | Engineering vehicles    | 0.659 | 9.2                  |
| YOLOv3                  | Engineering vehicles    | 0.869 | 16.6                 |

In addition, in the second picture of figure 6, there is a rectangular box of overlapping detection. Through analysis, this is mainly because the number of aerial photography data sets of engineering vehicles is limited, which leads to the insufficient learning of advanced semantic feature information of engineering vehicles by target detection network in the model training stage.

5. Conclusion
By making aerial photography dataset of engineering vehicles and based on deep learning target detection YOLOv3 algorithm, this paper simulates and realizes the target detection of engineering vehicles such as excavators under the aerial photography environment on the computer. Experimental results show that YOLOv3 algorithm is higher than HOG+LBP+SVM and Faster R-CNN target detection algorithms.
detection algorithm in both detection accuracy and detection speed, with good detection effect, which can realize aerial inspection of engineering vehicles and early warning of hidden dangers along the optical cable lines. It has certain reference significance for the application and popularization of aerial patrol inspection mode in the maintenance and management of optical cable lines in the future. In order to achieve better detection effect, the next research should optimize the network structure with a more appropriate anchor scale and collect more datasets to improve the detection ability of the trained model.

References

[1] Guo Xuanhong, Liu Ziliang, Zheng Chengcheng, et al. Application of UAV to oilfield cable maintenance [J]. China High-tech Zone, 2018, (01): 47.

[2] Yan Chunjiang, Wang Wei, Fang Hualin, et al. Intrusion detection for engineering vehicles under the electric transmission line based on deep learning[J]. Information Technology, 2018, 42(07):28-33+38.

[3] Zhang Mingjiang, Li Hongwei, Zhao Weihu, et al. Application of Deep Learning in the Patrol and Inspection of Military Optical Cable Lines by UAV[J]. Study on Optical Communications, 2018(06): 57-61.

[4] Redmon J, Farhadi A. YOLOv3: An incremental improvement[J]. arXiv, 2018: 1804.02767.

[5] Wang Pengfei, Yao Qi, Xing Jianfeng. Discussion on the Method of Obstacle Repair of Optical Cable Lines[J]. Information & Communications, 2013, (01):165-166.

[6] Hinton G, Osindero S, Teh Y W A. Fast Learning Algorithm for Deep Belief Nets[J]. Nature, 2006, 18(7):1527-1554.

[7] Girshick R. Fast R-CNN [C] // Proceedings of the IEEE International Conference on Computer Vision, 2015:1440-1448.

[8] Ren S, He K, Girshick R, et al. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 39(06):1137-1149.

[9] Liu W, Anguelov D, Erhan D, et al. SSD: Single Shot MultiBox Detector[C] // European Conference on Computer Vision. Springer International Publishing, 2016: 21-37.

[10] Redmon J, Divvala S, Girshick R, et al. You Only Look Once: Unified, Real-Time Object Detection [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 779-788.

[11] Redmon J, Farhadi A. YOLO9000: Better, faster, stronger[J]. arXiv, 2016: 1612.08242.

[12] Li Yunpeng, Hou Lingyan, Wang Chao. Moving objects detection in automatic driving based on YOLOv3[J]. Computer Engineering and Design, 2019, 40(04): 1139-1144.

[13] Razakarivony S, Jurie F. Vehicle Detection in Aerial Imagery: A small target detection benchmark[J]. Journal of Visual Communication and Image Representation, 2016, 34:187-203.