

Multi-Objective Performance Evaluation of the Detection of Catenary Support Components Using DCNNs

Liu, Wenqiang; Liu, Zhigang; Nunez, Alfredo; Wang, Liyou; Liu, Kai; Lyu, Yang; Wang, Hongrui

DOI

[10.1016/j.ifacol.2018.07.017](https://doi.org/10.1016/j.ifacol.2018.07.017)

Publication date

2018

Document Version

Final published version

Published in

Proceedings of 15th IFAC Symposium on Control in Transportation Systems (CTS 2018)

Citation (APA)

Liu, W., Liu, Z., Nunez, A., Wang, L., Liu, K., Lyu, Y., & Wang, H. (2018). Multi-Objective Performance Evaluation of the Detection of Catenary Support Components Using DCNNs. In B. De Schutter, & A. Ferrara (Eds.), *Proceedings of 15th IFAC Symposium on Control in Transportation Systems (CTS 2018): Savona, Italy, June 6-8, 2018* (9 ed., Vol. 51, pp. 98-105). (IFAC-PapersOnLine; Vol. 51, No. 9). <https://doi.org/10.1016/j.ifacol.2018.07.017>

Important note

To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

Multi-Objective Performance Evaluation of the Detection of Catenary Support Components Using DCNNs

Wenqiang Liu*, Zhigang Liu *, Alfredo Núñez**, Liyou Wang*, Kai Liu*, Yang Lyu*, Hongrui Wang*

* School of Electrical Engineering, Southwest Jiaotong University
China (e-mail: Liuwq_2009@126.com).

** Section of Railway Engineering, Delft University of Technology
Netherlands (e-mail: A.A.NunezVicencio@tudelft.nl)

Abstract: The goal of this paper is to evaluate from a multi-objective perspective the performance on the detection of catenary support components when using state-of-the-art deep convolutional neural networks (DCNNs). The detection of components is the first step towards a complete automatized monitoring system that will provide actual information about defects in the catenary support devices. A series of experiments in an unified test environment for detection of components are performed using Faster-CNN, R-FCN, SSD, and YOLOv2. Through the comparison of different assessment indicators, such as precision, recall, average precision and mean average precision, the detection performance of the different DCNNs methods for the components of the catenary support devices is analyzed, discussed and evaluated. The experiment results show that among all considered methods, R-FCN is the more suitable for the detection of catenary support components.

© 2018, IFAC (International Federation of Automatic Control) Hosting by Elsevier Ltd. All rights reserved.

Keywords: Catenary, Railway Systems, Multi-Objective Performance Evaluation, Deep convolutional neural networks (DCNNs)

1. INTRODUCTION

With the rapid development of high-speed railway, a large number of new railway infrastructures will be constructed all over the world. The larger the infrastructure, the more challenges are faced when deciding infrastructure maintenance tasks. In this sense, automated monitoring systems capable to detect defects in the whole infrastructure will assure the safe operation of the complete railway system.

Two of the most important infrastructure components in railway system are the catenary and the track systems, as shown in Fig. 1. Potential failures of the catenary or track will severely threaten the railway traffic safety. For ensuring the safe operation of high-speed railway, a large number of monitoring and detection technology have been investigated, and the corresponding detection equipment has been developed and applied. In this paper the focus is on catenary support devices relying on image processing technology.

In (Cho and Ko, 2015), the scale-invariant feature transform (SIFT) is employed to track and locate the pantograph, and assessing the reliability of railway overhead power by measuring the stagger between the pantograph and contact wire. The speeded-up robust features (SURF) is proposed in (Yang *et al.*, 2013), to extract the features and detect the insulator. Then, the state of the insulator is evaluated according to the vertical grayscale statistic distribution. In (Zhang *et al.*, 2016), the contourlet transform (CT) along with the Chan-Vese (CV) model is proposed to detect and diagnose the insulator. A method for failure detection of the ear pieces is described in (Han *et al.*, 2017), which used the

histogram of oriented gradients (HOG) feature to express the rotary double-ear, and combined with the support vector machine (SVM) to recognize it. Then, the Gabor wavelet is utilized to detect the failure of the component. Authors in (Zhang *et al.*, 2017) combined the difference histograms of oriented gradients (DHOG) and AdaBoost algorithm to detect the auxiliary catenary wire, and judged the fault of the auxiliary catenary wire through the circular arc detection and segment clustering. In (Han *et al.*, 2016), segment clustering is proposed to first divide the independent area, and utilize the deformable part models (DPM) and latent SVM to detect the rod-insulator. The local normalization (LN) method to achieve the contrast enhancement of the rail image is proposed in (Li and Ren, 2012), and the defect localization based on projection profile (DLBP) is used to detect defects. A probabilistic model to differentiate fasteners to track and judge the fault of fasteners based on the likelihood probability was established in (Feng *et al.*, 2014). In many aspects, these traditional feature extraction and machine learning methods for object recognition have obtained positive results. However, their performance has been difficult to get improved. In recent years, with the rapid development of deep learning, there are methods applied in the industry to detect the fault of equipment. In (Faghih-Roohi *et al.*, 2016), a deep convolutional neural network is proposed for the analysis of image data for the detection of rail surface defects. In (Gibert *et al.*, 2017), a detection method was proposed which is able to improve the detection performance through combining multiple detectors within a multi-task learning framework. In (Chen *et al.*, 2017), the steady arm base was recognized and located with region-

based convolutional neural networks (RCNN) first, and then on the basis of the installation relationship, the steady arm was detected with Hough transform, and the slope of the angle of steady arm was calculated. Although the progress in recent research efforts in the literature (Jamshidi et al., 2017; Liu et al., 2017; Chen et al., 2018; Psuj, 2018), there is still a gap towards the application of DCNN techniques in the railway industry.

In the literature, different state-of-the-art deep convolutional neural networks (DCNNs) structures are available. However, for the specific task of detection in railway systems environments, not all the neural network structures will perform in the same way. Therefore, this paper proposes a simple assessment methodology based on multi-objective performance evaluation. The most advanced and the most representative structures are used for the detection of components in the catenary support, and the intention is to provide a systematic approach to support on their evaluation.

The rest of the paper is organized as follows. First, the inspection system is shown in Section 2. Four of the state-of-the-art DCNNs are introduced in Section 3. Next, the experiment results are analyzed and discussed in Section 4. Last, some conclusions are summarized in Section 5.



Fig. 1. Catenary and track devices

2. INSPECTION SYSTEM

XLN4C is a comprehensive inspection vehicle developed by the China Railway Inc., shown in Fig. 2. It is equipped with six detection and monitoring systems including comprehensive pantograph and catenary monitor system (CPCM-1C), catenary-checking video monitor system (CCVM-2C), catenary-checking on-line monitor system (CCLM-3C), high-precision catenary-checking monitor system (CCHM-4C), catenary and pantograph video monitor system (CPVM-5C), and ground monitor system for catenary and power supply equipment (CCGM-6C), and called as the 6C system. In particular, the 4C system is mainly used to detect and diagnose the condition of catenary by analyzing the collected 2D images as the one shown in Fig. 3. A large amount of catenary 2D images are acquired from this system and state of the art image processing technology used for the analysis.

Catenary support device has a complex structure with many components. In addition, its defect type and defect level are various, as shown in Appendix A. It is difficult to detect

quickly and accurately using the traditional image detection technology all those types of defects. In this paper, catenary support device images are used to train DCNNs, and evaluate the detection performance of different DCNNs.



Fig. 2. XLN4C inspection vehicle

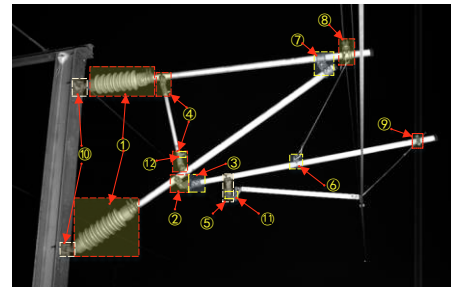


Fig. 3. Catenary support device

3. DCNN STRUCTURES OVERVIEW

Currently, DCNNs-based deep learning method has been the mainstream for solving object detection problems. There are a lot of the state-of-the-art DCNNs structures developed, proposed and used in the various fields. Among all, region-based and regression-based DCNNs structures are two main research directions, for which the representative algorithms are Faster RCNN, R-FCN, SSD, and YOLOv2. The core idea and network structure of these algorithms are described below.

3.1 Faster RCNN

Faster RCNN (Ren et al., 2015) is a series of RCNN (Girshick et al., 2014). It is an integration of region proposal networks (RPNs) and Fast R-CNN (Girshick, 2015), which uses RPNs to extract the detection areas and then uses conventional neural networks (CNNs) to extract image features. Lastly, SoftMax function is used to achieve the classification and bounding boxes regression to locate the position of objects. The structure diagram is shown in Fig. 4.

3.2 Region-based fully convolutional networks (R-FCN)

R-FCN (Dai et al., 2016) is seen as an improved version of the Faster RCNN. It moved several full connection layers behind the region of interest (RoI) pooling layer to the front, which greatly improved the detection speed. Meanwhile, the

feature extraction network was replaced with ResNet. The structure diagram is shown in Fig. 5.

3.3 Single shot multibox detector (SSD)

SSD (Liu et al., 2016) can be simply seen as a combination of YOLO(v1) (Redmon et al., 2016) and anchor boxes idea, which is based on bounding boxes regression. By using a small convolution kernel on different feature maps to predict a series of box offsets of bounding boxes (Bboxes), the goal of the object detection can be achieved. The structure diagram is shown in Fig. 6.

3.4 YOLOv2

YOLOv2 (Redmon and Farhadi, 2016) is an improved version of YOLO(v1), which is also based on the bounding boxes regression. YOLOv2 draws on the anchor ideas in the Faster R-CNN, which samples on the convolution feature map with sliding window. Then, each centre predicts nine different sizes and proportions of the proposed box. Since there is no need for reshaping the convolution layer, the spatial information is kept, solving the shortcomings of YOLO (v1). The structure diagram is shown in Fig. 7.

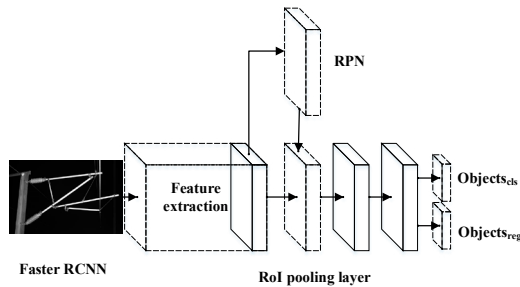


Fig. 4. Structure diagram of Faster RCNN

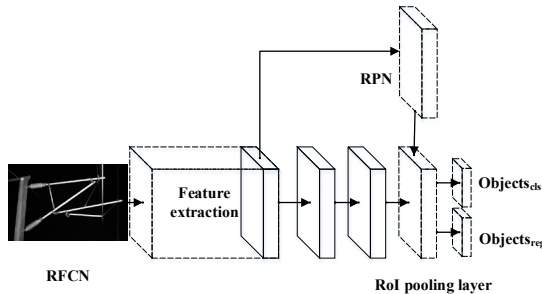


Fig. 5. Structure diagram of R-FCN

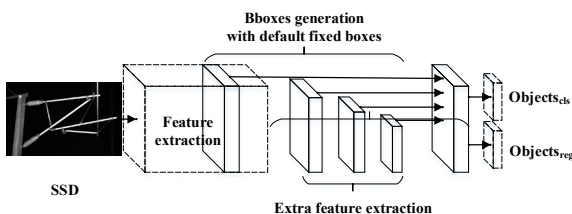


Fig. 6. Structure diagram of SSD

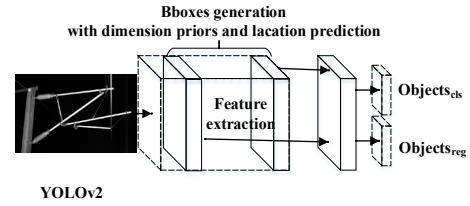


Fig. 7. Structure diagram of YOLOv2

4. EXPERIMENT AND RESULTS

In order to evaluate the detection performance of the four different DCNNs methods presented in the previous section, different indexes are chosen to assess their performance. The experiment environment is as follows. Deep learning open source framework Caffe (Jia et al., 2014), Ubuntu 14.04, 32GB RAM, CPU clocked at 3.2 GHz, and GTX 1080 GPU with 8GB memory.

4.1 Dataset and Parameter settings

The catenary dataset is made by the tool called “labelImg” provided from the website¹. The total image amount of dataset is 5022, among which the training dataset is 2417, the validation dataset is 1036 and the test dataset is 1569. The experiment parameter settings are as follows. Momentum and weight decay are 0.9 and 0.0005, and learning rate is 0.001, iterations are 15, 000.

4.2 Evaluation Indexes

Some indexes are chosen including the precision and recall, average precision (AP), mean average precision (mAP) and frames per second (FPS). Some curves and charts are drawn including precision and recall curve (P-R curve) and loss curve.

$$\text{precision} = \frac{TP}{TP + FP} \times 100\% \quad (1)$$

$$\text{recall} = \frac{TP}{TP + FN} \times 100\% \quad (2)$$

$$AP = \int_0^1 p(r) dr \quad (3)$$

$$mAP = \frac{\sum_{q=1}^Q AP(q)}{Q} \quad (4)$$

where TP is true positive, FP is false positive, TN is true negative, Q is the number of the component class.

4.3 Experiment Results

1) Examples of component detection effects

In order to show the detection effect of the four different structures under the simple environment and complex environment, two sets of examples are shown in Fig. 8. One

¹ <https://github.com/tzutalin/labelImg>

only includes one set of catenary support device, and the

other contains two sets of catenary support device.

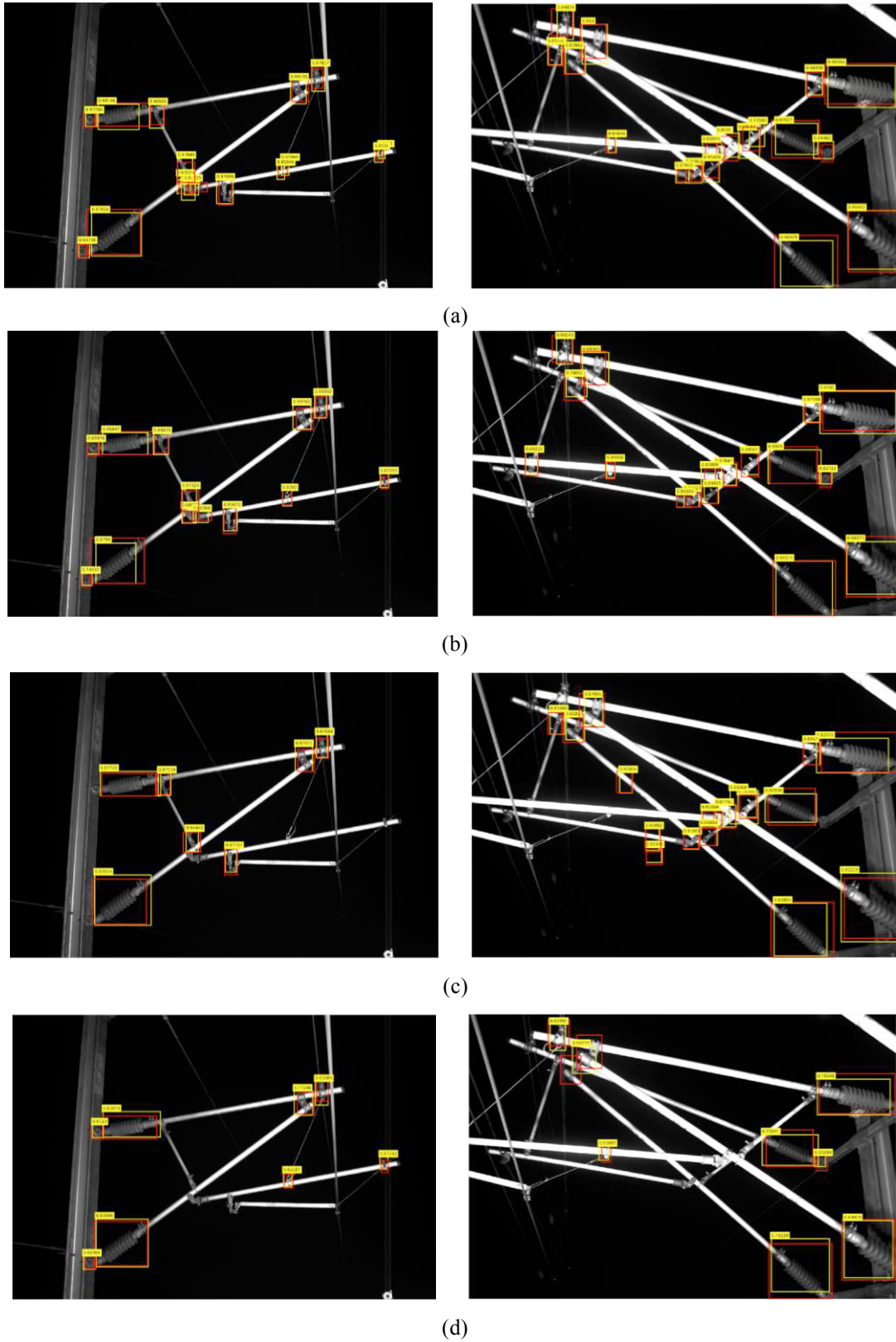


Figure. 8. Two examples of component detection effects for the four different DCNNs architecture. (a) row is the detection with Faster RCNN, (b) row is the detection with R-FCN, (c) row is the detection with SSD, and (d) row is the detection with YOLOv2.

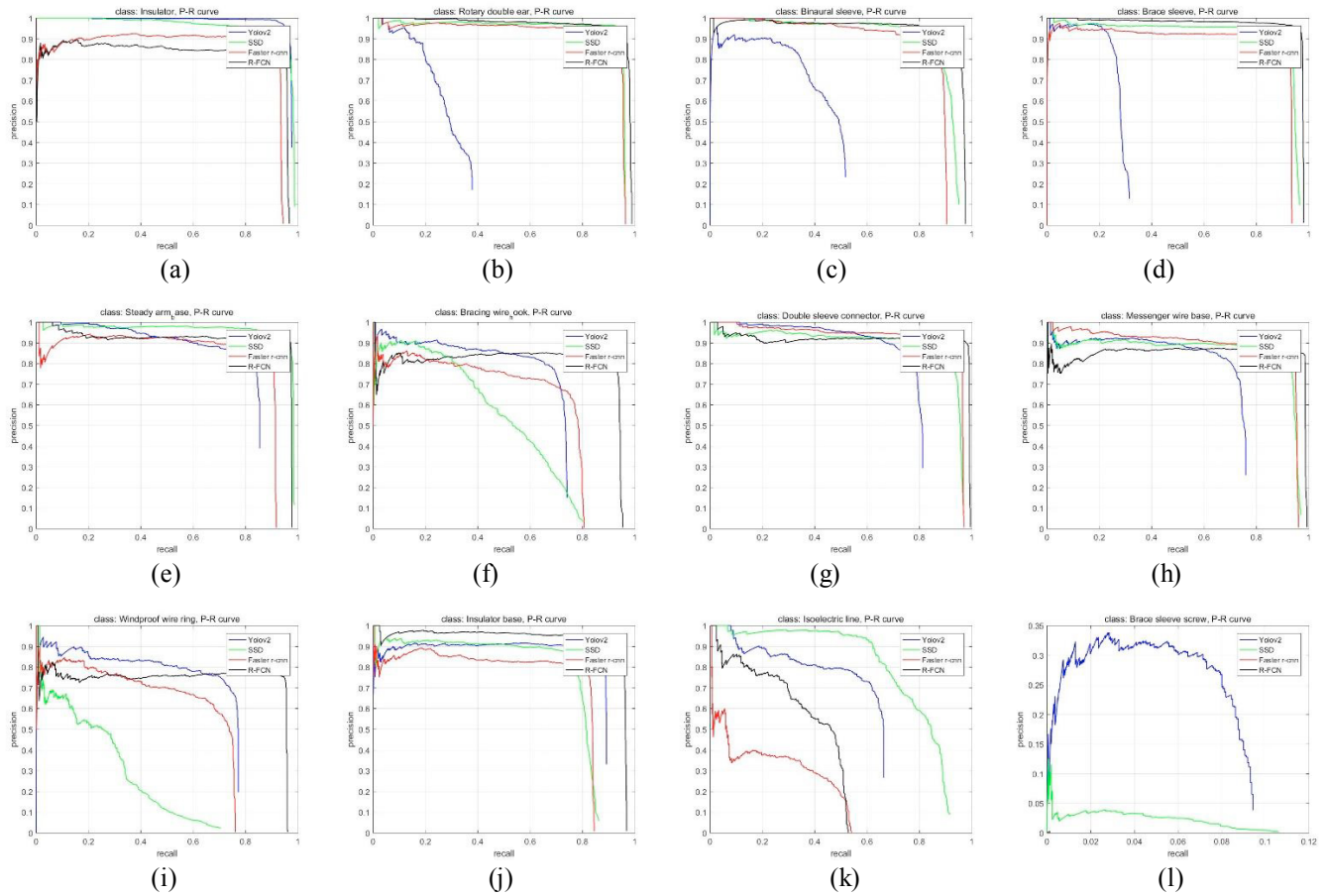


Figure. 9. P-R Curve for the detection of all the parts of the catenary support device (a) Insulator, (b) Rotary double-ear, (c) Binaural sleeve, (d) Brace sleeve, (e) Steady arm base, (f) Bracing wire hook, (g) Double sleeve connector, (h) Messenger wire base, (i) Windproof wire ring, (j) Insulator base, (k) Isoelectric line, (l) Brace sleeve screw, among, (a)~(j) are the large objects, (k) and (l) are the small objects.

Table 1. Detection average precision on catenary dataset for the four different DCNNs architecture

Catenary Dataset	Insulator	Rotary double-ear	Binaural sleeve	Brace sleeve	Steady arm base	Bracing wire hook	Double sleeve connector	Messenger wire base	Windproof wire ring	Insulator base	Isoelectric line	Brace sleeve screw	mAP	FPS
Faster RCNN	0.783	0.842	0.785	0.796	0.767	0.563	0.843	0.804	0.508	0.65	0.181	0	0.627	1.46
R-FCN	0.757	0.88	0.861	0.88	0.835	0.732	0.846	0.787	0.682	0.846	0.334	0	0.703	2.02
SSD	0.877	0.856	0.818	0.834	0.876	0.459	0.812	0.782	0.223	0.683	0.716	0.003	0.662	2.30
YOLOv2	0.886	0.261	0.381	0.225	0.73	0.587	0.703	0.615	0.586	0.743	0.511	0.027	0.521	3.70

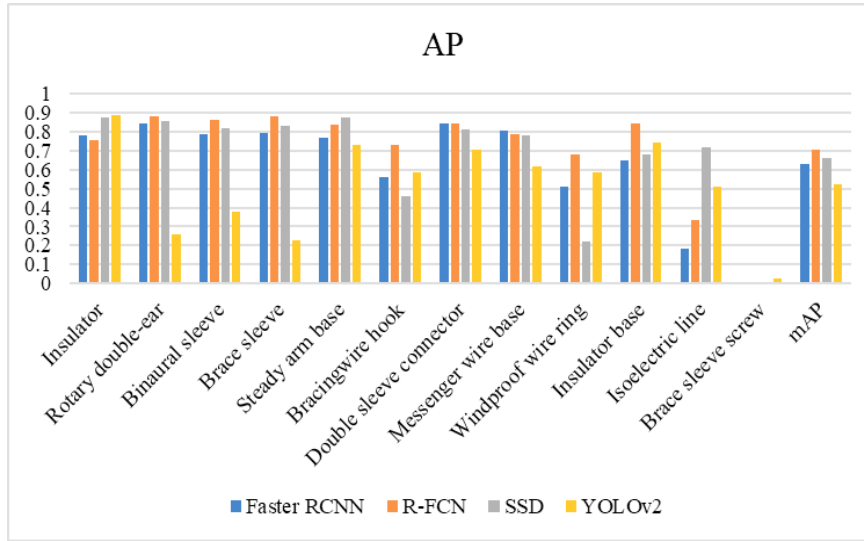


Figure. 10. Detection average precision on catenary dataset for the four different DCNNs architecture

2) P-R curve

For analyzing the relationship between the false detection rate and missed detection rate of each component for four different DCNNs method, the curve of the precision and recall (P-R curve) of every part is drawn up, as shown in Fig. 9.

3) AP and mAP

For comparing the detection accuracy and detection efficiency of different models, the mathematical statistics for the test dataset are carried out through AP and mAP as well as FPS, as show in Table 1 and Fig. 10.

4) Training loss

To measure the robustness of the training model, the training loss curve is plotted as the number of iterations increases, as shown in Fig. 11.

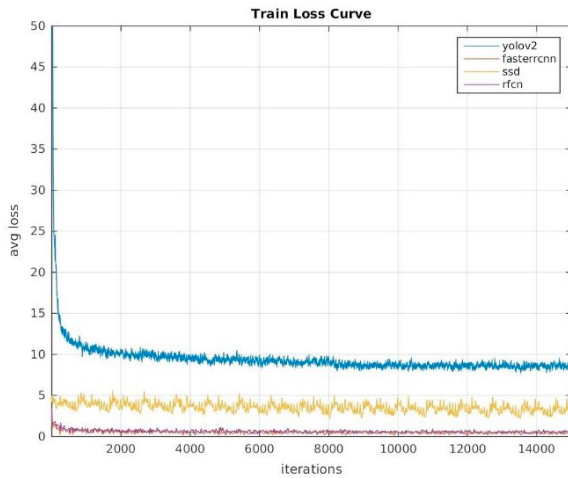


Fig. 11. Training loss for the four different DCNNs architecture

4.4 Results analysis

- (1) From Fig. 8, the detection results based on regression method, SSD and YOLOv2 have the result of missed detection even if it is under the simple background environment (as shown in Fig. 8(c)-left and Fig. 8(d)-left). Among the detection results with Faster RCNN and R-FCN, the latter performs better.
- (2) It can be found from Table 2 and Fig. 10 that the detection results with R-FCN are the best among these DCNNs, whether it is a single object AP or the whole mAP. For the smaller objects isoelectric line and brace sleeve screw, the regression-based SSD and YOLOv2 achieve some effects though they are not ideal. In addition, for the bigger objects, the SSD also performs relatively good and its speed is also very fast.
- (3) Fig. 9 shows that the P-R curve of the YOLOv2 is the worst. It is difficult to achieve a good balance between the precision and recall. In other words, if the model wants to improve its false detection rate, then it has to sacrifice its missed detection rate. In contrast, the P-R curve of R-FCN which have a better performance.
- (4) As seen from Fig. 11, the robustness of the training model of the region-based DCNNs, Faster RCNN and R-FCN, are the best. Their convergence speed is also faster. To the opposite, the loss curve of SSD fluctuated greatly and YOLOv2 has a low convergence speed.

5. CONCLUSION

Combined with the most advanced DCNNs learning method technology, the development of advanced detection methodologies is crucial for the railway infrastructure maintenance. This paper compared the state-of-the-art and the most representative DCNNs learning algorithms with different multi-objective evaluation indexes.

Through the above results analysis, for catenary image detection, the region-based R-FCN is more suitable for the detection of catenary support components. However, its shortcoming is a poor performance in detection of the smaller objects. A possible further research direction could be to incorporate the capabilities of SSD and YOLOv2 into an integrated scheme that make use the structure R-FCN achieve the desired results. SSD and YOLOv2 could be used as a RPNs, and the proposed regions would be sent into the R-FCN to improve the detection effects of the DCNN.

This paper does not study the characteristics of each network individually. It gives a general comparative result, which expects that the follow-up research work is able to focus on a core framework for optimization and improvement, and ultimately establishing a detection network that is more suitable for the maintenance of the railway infrastructure.

REFERENCES

- Cho, C.J. and Ko, H. (2015). Video-based dynamic stagger measurement of railway overhead power lines using rotation-invariant feature matching. *IEEE Transactions on Intelligent Transportation Systems*, 16(3), 1294-1304.
- Yang, H.M., Liu, Z.G., Han, Y., and Han, Z.W. (2013). Defective condition detection of insulators in electrified railway based on feature matching of speeded-up robust features. *Power System Technology*, 37(8), 2297-2302.
- Zhang, G.N., Liu, Z.G., and Han, Y. (2016). Automatic recognition for catenary insulators of high-speed railway based on contourlet transform and Chan-Vese model. *Optik - International Journal for Light and Electron Optics*, 127(1), 215-221.
- Han, Ye., Liu, Z.G., Geng, X., and Zhong, J.P. (2017). Fracture detection of ear pieces in catenary support devices of high-speed railway based on HOG features and two-dimensional Gabor transform. *Journal of the China Railway Society*, 39(2), 52-57.
- Zhang, G.N., Liu, Z.G., Han, Y., and Han, Z.W. (2017). Loss fault detection for auxiliary catenary wire of high-speed railway catenary wire holder. *Journal of the China Railway Society*, 39(5), 40-46.
- Han, Y., Liu, Z.G., Lee, D.J., Zhang, G.N., and Deng, M. (2016). High-speed railway rod-insulator detection using segment clustering and deformable part models. In *Image Processing (ICIP), 2016 IEEE International Conference on*, 3852-3856, IEEE.
- Li, Q.Y. and Ren, S.W. (2012). A real-time visual inspection system for discrete surface defects of rail heads. *IEEE Transactions on Instrumentation and Measurement*, 61(8), 2189-2199.
- Feng, H., Jiang, Z.G., Xie, F.Y., Yang, P., Shi, J., and Chen, L. (2014). Automatic fastener classification and defect detection in vision-based railway inspection systems. *IEEE Transactions on Instrumentation and Measurement*, 63(4), 877-888.
- Faghih-Roohi, S., Hajizadeh, S., Núñez, A., Babuska, R., and De Schutter, B. (2016, July). Deep convolutional neural networks for detection of rail surface defects. In *Neural Networks (IJCNN), 2016 International Joint Conference on*, 2584-2589, IEEE.
- Gibert, X., Patel, V., and Chellappa, R. (2017). Deep multitask learning for railway track inspection. *IEEE Transactions on Intelligent Transportation Systems*, 18(1), 153-164.
- Chen, D.J., Zhang, W.S., & Yang, Y. (2017). Detection and recognition of high-speed railway catenary locator based on Deep Learning. *Journal of University of Science and Technology of China*, 47(4), 320-327.
- Jamshidi, A., Faghih-Roohi, S., Hajizadeh, S., Núñez, A., Babuska, R., Dollevoet, R., Li, Z., and De Schutter, B. (2017). A big data analysis approach for rail failure risk assessment. *Risk Analysis*, 37(8), 1495-1507.
- Liu, Z.G., Wang, L.Y., Yang, C.J. and Han, Z.W. (2017). A high-precision loose strands diagnosis approach for isoelectric line in high-speed railway. *IEEE Transactions on Industrial Informatics*, DOI 10.1109/TII.2017.2774242.
- Chen, J.W., Liu, Z.G., Wang, H.R., Nunez, A. and Han, Z.W. (2018). Automatic defect detection of fasteners on the catenary support device using deep convolutional neural network. *IEEE Transactions on Instrumentation and Measurement*, 67(2), 257-269.
- Psuji, G. (2018). Multi-sensor data integration using deep learning for characterization of defects in steel elements. *Sensors*, 18(2), 292.
- Girshick, R., Donahue, J., Darrell, T., and Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 580-587.
- Girshick, R. (2015). Fast R-CNN. In *Proceedings of the IEEE international conference on computer vision*, 1440-1448.
- Ren, S.Q., He, K.M., Girshick, R., and Sun, J. (2015). Faster R-CNN: Towards real-time object detection with region proposal networks. In *Advances in neural information processing systems*, 91-99.
- Dai, J.F., Li, Y., He, K.M., and Sun, J. (2016). R-FCN: Object detection via region-based fully convolutional networks. In *Advances in neural information processing systems*, 379-387.
- Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.Y., and Berg, A.C. (2016, October). SSD: Single shot multibox detector. In *European conference on computer vision*, 21-37, Springer, Cham.
- Redmon, J., Divvala, S., Girshick, R., and Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 779-788.
- Redmon, J. and Farhadi, A. (2016). YOLO9000: better, faster, stronger. *arXiv preprint arXiv:1612.08242*.
- Jia, Y.Q., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R., Guadarrama, S., and Darrell, T. (2014, November). Caffe: Convolutional architecture for fast feature embedding. In *Proceedings of the 22nd ACM international conference on Multimedia*, 675-678, ACM.

Appendix A

Table 2. Defect type and level of the components of catenary support device

	Part name	Defect type	Defect level
①	Insulator	Cracking	B
		Flashover	B
②	Rotary double-ear	Cracking	A
		Cotter pin losing	A
③	Binaural sleeve	Cracking	A
④	Brace sleeve	Cracking	A
⑤	Steady arm base	Cracking	A
		Nut looseness	A
⑥	Bracing wire hook	Cracking	A
		Nut looseness	A
⑦	Double sleeve connector	Cracking	A
		Nut looseness	A
⑧	Messenger wire base	Cracking	A
		Opposite direction	C
		Balance line losing	C
⑨	Windproof wire ring	Cracking	A
		Nut looseness	B
⑩	Insulator base	Cotter pin losing	B
⑪	Isoelectric line	Line looseness	C
⑫	Brace sleeve screw	Nut looseness	B

Note: Isoelectric line is the joint area between steady arm base and registration arm, and brace sleeve screw is subarea of the brace sleeve.