Corrosion Detection in Transformers Based on Hierarchical Annotation

Yong Cao[1(✉)], Yinian Zhou[2], Zhao Zhang[3,4], Wenjun Wu[1], Xihai Chen[3,4], Sha Yang[3,4], and Baili Zhang[1]

[1] School of Computer Science and Engineering, Southeast University, Nanjing, China
[2] School of Computer and Software, Nanjing University of Information Science and Technology, Nanjing, China
[3] State Key Laboratory of Smart Grid Protection and Control, Nanjing, China
[4] Nari Group Corporation, Nanjing, China

Abstract. Transformers are prone to metal corrosion due to long-term exposure in the external environment. It is necessary to detect corrosion in time to prolong their service life-time. Detecting metal corrosion based on deep learning is becoming a feasible method at present. However, because of the irregularity and detachability of metal corrosion, traditional annotation approach results in such problems as ambiguity and uncertainty in the annotating process. Thus, a novel hierarchical annotation approach is proposed in this paper. The rationale for this initiative is as follows: traditional annotation approach has been used to annotate a large area covering the range of corrosion, as long as the area is visually continuous and adjacent to the corrosion that cannot be clearly divided. The annotation result is recorded as training set A; secondly, in the annotating boxes from the first step, the areas with obvious and relatively independent features are re-annotated to form the second level of nested annotation. The annotation result is recorded as training set B. Finally, Faster R-CNN and YOLOv5 models were trained by A, B training sets respectively in the experiment. It can be concluded that the detection performance with hierarchical annotation approach proved better than traditional annotation approach for Faster R-CNN and YOLOv5 models.

Keywords: Deep learning · Faster R-CNN · YOLOv5 · Object detection · Hierarchical annotation

1 Introduction

Corrosion is considered as a destructive phenomenon that affects most transformers. It is the result of deterioration caused by metallic materials' surface and internal microstructure reacting with corrosive environments [1, 2]. Due to long-term exposure in the external environment, corrosion has become a common defect to transformers. Corrosion can continuously impair the transformers' service life and cause economic losses before it is tackled promptly. Hence, it is necessary to regularly monitor their running condition.

Currently, non-destructive methods are usually applied in corrosion detection, such as X-ray [3], local wavenumber [4], infrared thermography [5, 6], magneto-optic imaging

[7] and camera. Among them, X-ray detection technology has been proved inconvenient in practice, and it's also harmful to the surrounding environment. The detection technology that applies local wavenumbers with different frequencies is more complex in data collection. Using infrared thermography to collect and detect corrosion demands higher requirements for the collecting environment. Dudziak utilizes a neural-like algorithm to detect metal corrosion by magneto-optic imaging based upon Faraday rotation of polarized light, which is more technically demanding.

Compared with the above-mentioned detection methods, using Convolutional Neural Networks (CNN) [8] to perform corrosion detection with the RGB images collected by camera which can bring about such advantages as lower collecting cost, simpler operation and lower professional requirements. Following AlexNet [9], with the emergence of VGG [10], GoogleNet [11], ResNet [12] and other CNN models, object detection has achieved rapid development. One solution is recommended in literature [13] by utilizing sliding window to intercept regions, then to detect corrosion by CNN. However, this method costs more time-consuming. Fast R-CNN [14], Faster R-CNN [15], YOLO [16–18] and other object detection models based on region proposal can greatly shorten the detection time and ensure precision and recall remain unchanged. However, due to the irregular shape and detachability of corrosion, these object detection models cannot directly achieve satisfying results. Therefore, this paper proposes a novel hierarchical annotation approach: Firstly, traditional annotation approach has been applied to annotate a large area covering the range of corrosion, as long as the area is visually continuous and adjacent to the corrosion that cannot be clearly divided; then, in the annotating boxes from the first step, the areas with obvious and relatively independent features are re-annotated to form the second level of nested annotation.

In comparison experiment, two annotation approaches are respectively applied to annotate 1180 same pictures and be recorded as two different training sets. Besides, 206 pictures are annotated by traditional annotation approach as the test set; then are compared the detection results of Faster R-CNN and YOLOv5 trained by different training sets. The experimental findings indicate that the detection results of models trained by hierarchical annotation approach are better than those of models trained by traditional annotation approach. Besides, precision and recall of Faster R-CNN are better than those of YOLOv5.

In summary, our contributions in this work include:

1. A novel data annotation approach, hierarchical annotation approach, is proposed.
2. Minimum bounding box algorithm is applied for merging the intersecting boxes.
3. A novel formulation is proposed for recalculating precision in view of the features of corrosion.

2   Related Works

The shape and size of corrosion is irregular, which is caused by the size of equipment and the spread of corrosion. Simultaneously, the detachability of corrosion makes some components of transformers can be regarded as a whole piece of corrosion or as several individual corrosion.

Fig. 1. The results annotated by traditional approach

Figure 1 shows the results annotated by traditional approach. However, this approach has several drawbacks. Due to the detachability of corrosion, the screws as the parts of corrosion in Fig. 1(left), while other screws in another picture, Fig. 1(right), are annotated as independent corrosion; besides, in Fig. 1(left), the corrosive screw in the upper right corner and the non-corrosive screw in the lower right corner are both not annotated, which means the screws with different types have the same annotation results. These drawbacks indicate that using traditional annotation approach to annotate corrosion will generate ambiguity and uncertainty.
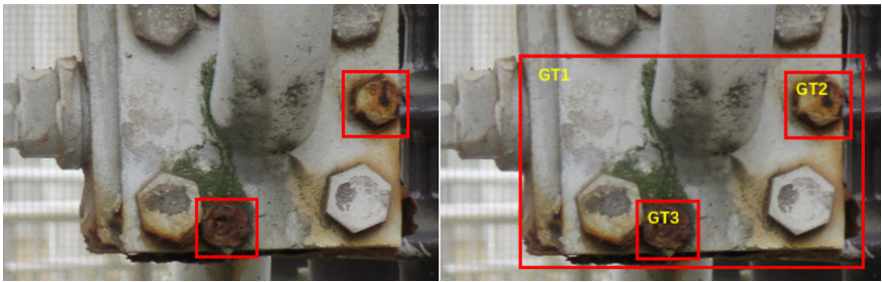

Fig. 2. The results annotated by hierarchical approach

Hence, to prevent the ambiguity and uncertainty in the process of annotation, one naive solution is presented: try to use small boxes instead of large boxes, as shown in Fig. 2(left). Therefore, after adopting this solution, the number of boxes with smaller size will increase, while the number of boxes with larger size will decrease to some extent. But this approach cannot solve this issue. In Fig. 2(left), the surrounding corrosion around boxes is omitted. Definitely, we can continue to refine the sizes of boxes to make it closer to corrosion. However, this operation will increase the workload of data annotation and the difficulty of problem solved.

Considering that Faster R-CNN and YOLO models based on region proposal utilize the predefined anchors with different sizes and shapes to detect objects with different sizes and shapes. Therefore, corrosion is detected mainly by the anchors with similar sizes. For the above reasons, this paper proposes a novel hierarchical annotation app-roach: Firstly, traditional annotation approach is used to annotate a large area within the range of corrosion, as long as it is visually continuous, adjacent to the corrosion that

cannot be clearly divided. For example, in Fig. 2(right), we apply a large box, ground truth (GT)1, to annotate corrosion; Next, in the annotating box of the first step, the corrosion with obvious and relatively independent features is re-annotated to form the second level of nested annotation. As shown in Fig. 2(right), we apply the boxes, GT2 and GT3, to annotate the corrosion with obvious and relatively independent features.

Ambiguous problem generated by applying traditional approach to annotate corrosion can be solved by adopting hierarchical annotation approach, while being compatible with traditional annotation approach. Simultaneously, it obviously increases the number of GT and further achieves the effect of data enhancement.
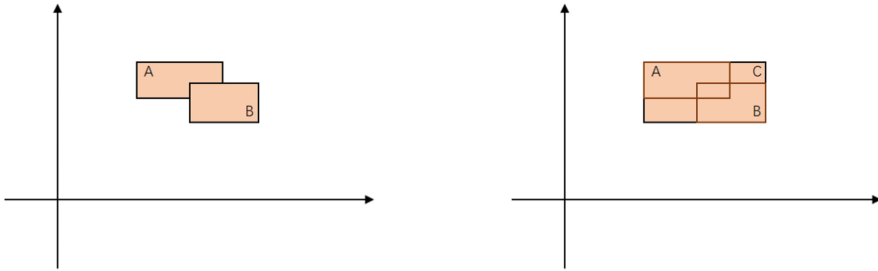

Fig. 3. Minimum bounding box algorithm

However, using hierarchical annotation approach results in lots of overlaps and nesting in the annotation results. Therefore, when using a model trained by hierarchical annotation approach to detect corrosion in transformers, the boxes in detecting result also have lots of overlaps and nesting. For this reason, this paper uses minimum bounding box algorithm to merge multiple intersecting boxes to one box, and the box is recorded as the final box.

For example, the boxes, A and B in Fig. 3(left), are generated by detection models; then, the orange box C in Fig. 3(right) is generated by minimum bounding box algorithm. Simultaneously, the box C is recorded as the final box instead of the boxes, A and B.

3   Experiment

In this Sects. 3.1 and 3.2 introduce the experimental environment and evaluating indicators; Sect. 3.3 uses traditional annotation approach to annotate the training set and test set, trains several object detection models, and calculates their precision and recall; Sect. 3.4 uses hierarchical annotation approach to re-annotate the training set while the test set remains unchanged, and repeats the above experimental process; Sect. 3.5 compares and analyzes experimental results.

3.1   Experiment Environment

The experimental software and hardware environment are shown in Table 1.

Table 1. Experimental software and hardware environment

CPU	I7-9800X/8@4.4 GHz
RAM	32 GB
GPU	Nvidia RTX2080Ti@11G
OS	Centos7
Deep learning framework	PyTorch
Object detection model	Faster R-CNN, YOLOv5
The volumes of training set, test set	1180, 206
Evaluating indicators	Precision, Recall, IoU

3.2 Evaluating Indicators

In this experiment, this paper uses model's precision and recall as evaluating indicators. The classification results can be divided into four categories according to the forecast results and the actual results. The confusion matrix for binary classification is shown in Table 2.

Table 2. Confusion matrix for binary classification

		Forecast result	
		Positive	Negative
Actual result	Positive	TP	FN
	Negative	FP	TN

We use GTs to represent the number of GT and PRs to represent the number of predicted boxes. Therefore, for one picture, the calculation formulations of precision (P) and recall (R) as follows:

$$GTs \neq 0, PRs \neq 0. P = \frac{TP}{TP + FP}, R = \frac{TP}{TP + FN} \tag{1}$$

$$GTs \neq 0, PRs = 0. P = 0, R = 0 \tag{2}$$

$$GTs = 0, PRs \neq 0. P = 0, R = 0 \tag{3}$$

$$GTs = 0, PRs = 0. P = 1, R = 1 \tag{4}$$

For multiple pictures:

$$P = \frac{p_1 + p_2 + \cdots + p_n}{N}, R = \frac{r_1 + r_2 + \cdots + r_n}{N} \tag{5}$$

Eventually, Intersection over Union (IoU) between the predicted box and the ground truth is applied to judge whether a predicted box is correct. IoU is an evaluation metric, which is used to measure how close an annotation or test output lines up with a ground truth. In this paper, the value of IoU greater than 0.5 indicates that one predicted box is correct. The calculation formulation is shown in Fig. 4.
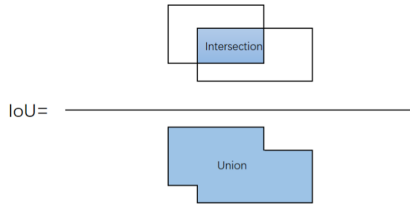

Fig. 4. IoU calculation formulation

In addition, the detachability of corrosion may result in one predicted box only is a part of GT. Although the box has correctness to some extent, it is judged as error because its IoU is less than 0.5. For example, the boxes in Fig. 5(left), PR1 and PR2, are the parts of GT with correctness to some extent, but their IoU are both less than 0.5, thus they are judged as errors.
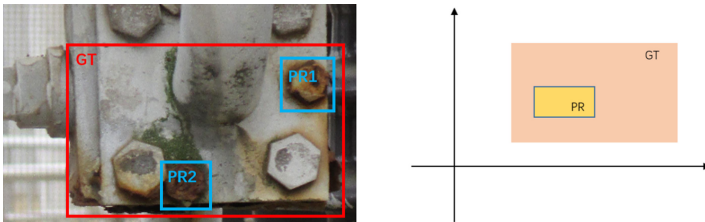

Fig. 5. The predicted boxes as the parts of GT

Thus, in order to weaken the impact of such conditions. According to the characteristics of corrosion, a novel calculation formulation is proposed. If one predicted box satisfies this condition, IoU $\leq 0.5 \& \frac{GT \cap FR}{FR} > 0.98$, which is regarded as no forecast. After omitting such predicted boxes, the novel precision obtained is recorded as valid-P. Since those predicted boxes regarded as errors are omitted, the value of valid-P will be greater than the value of P.

3.3 Traditional Annotation Approach

Firstly, 1180 specimens were annotated by traditional annotation approach for training Faster R-CNN and YOLOv5 models. And this approach was also used to annotate 206 specimens as the test set. Then, VGG16 and Res101 respectively as backbone networks, momentum SGD [19] as optimizer were applied to train Faster R-CNN models;

DarkNet53 [17] as backbone network, momentum SGD as optimizer were applied to train YOLOv5 model. Finally, the above trained models were used to test and calculate the values of P, R and valid-P. The values of P, R and valid-P in the three models are shown in Table 3.

Table 3. The values of P, R and valid-P with traditional annotation approach

Model	P	R	valid-P
YOLOv5	33.23%	27.03%	35.78%
Faster R-CNN + VGG16	57.94%	52.60%	64.26%
Faster R-CNN + Res101	69.85%	68.51%	76.87%

3.4 Hierarchical Annotation Approach

Next, 1180 specimens were re-annotated by hierarchical annotation approach for training Faster R-CNN and YOLOv5 models, while the test set remain unchanged. Then, VGG16 and Res101 respectively as backbone networks, momentum SGD as optimizer were applied to retrain Faster R-CNN models; DarkNet53 as backbone network, momentum SGD as optimizer were applied to retrain YOLOv5 model. Finally, the above trained models were used to test.

After using minimum bounding box algorithm to merge the intersecting boxes of each picture, we calculate the values of P, R and valid-P. The values of P, R and valid-P in the three models are shown in Table 4.

Table 4. The values of P, R and valid-P with hierarchical annotation approach

Model	P	R	valid-P
YOLOv5	36.59%	35.38%	41.01%
FasterR-CNN + VGG16	58.87%	60.25%	66.52%
FasterR-CNN + Res101	82.45%	83.13%	95.74%

3.5 Major Findings and Discussion

In this experiment, the comparison results of each index between traditional annotation approach and hierarchical annotation approach are shown in Table 5.

Faster R-CNN + Res101 model has the best detection result, whose values of P, R are both higher than 80%. What's more, the value of valid-P even exceeds 95%. Besides, we have other experimental findings:

Table 5. The improvement of each index after adopting hierarchical annotation

Model	P↑	R↑	valid-P↑
YOLOv5	3.36%	8.35%	5.23%
FasterR-CNN + VGG16	0.93%	7.65%	2.26%
FasterR-CNN + Res101	12.60%	14.62%	18.87%

- The values of P, R and valid-P of Faster R-CNN model are greater than those of YOLOv5 model.
- In Faster R-CNN model, using Res101 as backbone network has greater values of P, R and valid-P than using VGG16.
- After training with hierarchical annotation approach, YOLOv5 and Faster R-CNN + VGG16 have a slight improvement in the values of P and valid-P and have a great improvement in the value of R; Besides, the values of P, R and valid-P of Faster R-CNN + Res101 have great escalation.

  The reasons as follows:

- Faster R-CNN is two-stage, while YOLOv5 is single-stage. Faster R-CNN firstly filters out a large number of background regions through region proposal networks, so that subsequent classification can pay more attention to detecting corrosion, which contributes to the classification results. Therefore, the detection time in Faster R-CNN is longer, but the values of P, R and valid-P are greater than those of YOLOv5.
- The model structure of Res101 is more complicated than VGG16. Res101 has more convolution layers than VGG16, while its gradient can be better backpropagation by using batch normalization [20] and Rectified Linear Units [21]. Simultaneously, Res101 enables the model to be fully trained through the residual module [12]. Therefore, the values of P and R in Res101 are greater than those of VGG16 (Fig. 6).
- After using hierarchical annotation approach, the number of GT obviously increases in the training set, which is conducive to data enhancement; The object detection model detects objects with different sizes by predefined anchors with different sizes. Therefore, hierarchical annotation approach solves the ambiguity caused by traditional annotation while increasing the number of GT. In conclusion, the values of P, R and valid-P can be greatly improved.

Fig. 6. Outputs of Faster R-CNN + Res101 trained by hierarchical annotation approach

4   Conclusion

In this paper, Faster R-CNN and YOLOv5 models are applied to detect corrosion in transformers. Through preliminary experiments, it is found that precision and recall of models trained by traditional annotation approach are lower than expected. Hence, a novel hierarchical annotation approach is proposed by utilizing the characteristics of corrosion. Ultimately, according to experimental findings, the models' precision and recall have been greatly improved after adopting hierarchical annotation approach.

References

1. Roberge, P.R.: Handbook of Corrosion Engineering. McGraw-Hill Education, New York (2019)
2. Roberge, P.R.: Corrosion Engineering. McGraw-Hill Education, New York (2008)
3. Dunn, W.L., Yacout, A.M.: Corrosion detection in aircraft by X-ray backscatter methods. Appl. Radiat. Isot. **53**(4–5), 625–632 (2000)
4. Gao, T., Sun, H., Hong, Y., et al.: Hidden corrosion detection using laser ultrasonic guided waves with multi-frequency local wavenumber estimation. Ultrasonics **108**, 106182 (2020)
5. Doshvarpassand, S., Wu, C., Wang, X.: An overview of corrosion defect characterization using active infrared thermography. Infrared Phys. Technol. **96**, 366–389 (2019)
6. Wicker, M., Alduse, B.P., Jung, S.: Detection of hidden corrosion in metal roofing shingles utilizing infrared thermography. J. Build. Eng. **20**, 201–207 (2018)
7. Dudziak, M.J., Chervonenkis, A.Y., Chinarov, V.: Nondestructive evaluation for crack, corrosion, and stress detection for metal assemblies and structures. In: Nondestructive Evaluation of Aging Aircraft, Airports, and Aerospace Hardware III. International Society for Optics and Photonics, vol. 3586, pp. 20–31 (1999)
8. LeCun, Y., Boser, B., Denker, J.S., et al.: Backpropagation applied to handwritten zip code recognition. Neural Comput. **1**(4), 541–551 (1989)
9. Krizhevsky, A., Sutskever, I., Hinton, G.E.: ImageNet classification with deep convolutional neural networks. In: Advances in Neural Information Processing Systems, vol. 25, pp. 1097–1105 (2012)
10. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556 (2014)
11. Szegedy, C., Liu, W., Jia, Y., et al.: Going deeper with convolutions. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 1–9 (2015)
12. He, K., Zhang, X., Ren, S., et al.: Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778 (2016)
13. Yao, Y., Yang, Y., Wang, Y., et al.: Artificial intelligence-based hull structural plate corrosion damage detection and recognition using convolutional neural network. Appl. Ocean Res. **90**, 101823 (2019)
14. Girshick, R.: Fast R-CNN. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 1440–1448 (2015)
15. Ren, S., He, K., Girshick, R., et al.: Faster R-CNN: towards real-time object detection with region proposal networks. arXiv preprint arXiv:1506.01497 (2015)
16. Redmon, J., Divvala, S., Girshick, R., et al.: You only look once: unified, real-time object detection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 779–788 (2016)
17. Redmon, J., Farhadi, A.: Yolov3: an incremental improvement. arXiv preprint arXiv:1804.02767 (2018)
18. Jocher, G., Stoken, A., Borovec, J., et al.: ultralytics/yolov5: v3.1 - bug fixes and performance improvements (2020). https://doi.org/10.5281/zenodo.4154370
19. Sutskever, I., Martens, J., Dahl, G., et al.: On the importance of initialization and momentum in deep learning. In: International Conference on Machine Learning, pp. 1139–1147. PMLR (2013)
20. Ioffe, S., Szegedy, C.: Batch normalization: accelerating deep network training by reducing internal covariate shift. In: International Conference on Machine Learning, pp. 448–456. PMLR (2015)
21. Nair, V., Hinton, G.E.: Rectified linear units improve restricted Boltzmann machines. In: ICML (2010)