Journal of Electronic Research and Application, 2025, Volume 9, Issue 3
http://ojs.bbwpublisher.com/index.php/JERA
ISSN Online: 2208-3510
ISSN Print: 2208-3502

Fault Detection for Split Pins of Power Transmission Fittings in UAV Inspections via Automatic Image Cropping-based Super-Resolution Reconstruction and Enhanced YOLOv8

Shihao Cui[1], Zhengyu Hu[1,2], Fangcheng Qiu[2], Qinglong Wang[1]*

[1]Department of Mechanical Engineering, North China Electric Power University, Baoding 071003, China
[2]Electric Power Research Institute of Yunnan Power Grid CO., LTD., Kunming 650217, China

**Author to whom correspondence should be addressed.*

Abstract: In modern industrial applications, ensuring the reliability of mechanical fittings is critical for maintaining operational safety and efficiency, particularly in power grid systems where split pins serve a pivotal role despite being susceptible to environmental degradation and failure. Existing UAV-based inspection systems are hampered by a low representation of split pin elements and complex backgrounds, leading to challenges in accurate fault detection and timely maintenance. To address this pressing issue, our study proposes an innovative fault detection method for split pins. The approach employs a three-step process: first, cropping operations are used to accurately isolate the fittings containing split pins; second, super-resolution reconstruction is applied to enhance image clarity and detail; and finally, an improved YOLOv8 network, augmented with inner-shape IoU and local window attention mechanisms, is utilized to refine local feature extraction and annotation accuracy. Experimental evaluations on a split pin defect dataset demonstrate robust performance, achieving an accuracy rate of 72.1% and a mean average precision (mAP) of 67.7%, thereby validating the method's effectiveness under challenging conditions. The proposed approach contributes to the field by specifically targeting the challenges associated with split pin detection in UAV-based inspections, offering a practically applicable and reliably precise method.

Keywords: Split pins; Fault detection; Power transmission fittings; YOLO; Deep learning

Online publication: 29 May, 2025

1. Introduction

Split pins play a crucial role as fasteners in transmission lines by being inserted into the bolts on connectors between power towers and transmission lines, thereby preventing bolts from loosening and detaching. However, long-term exposure to the natural environment renders these pins highly vulnerable to climatic conditions. For instance, persistent rainfall and strong winds can accelerate erosion, leading to abnormal wear or even complete shedding of split pins [1]. If such defects go undetected and unaddressed promptly, bolts may gradually loosen and eventually detach, triggering a chain reaction that fails other transmission line components [2]. This cascade of failures can lead to widespread outages and poses a significant threat to the safety and stability of the power grid. To mitigate these risks and ensure the secure and reliable operation of the power system, regular inspection and maintenance of transmission lines are imperative.

Traditional bolt inspection relies primarily on manual methods, which are both time-consuming and labor-intensive. The small size, large quantity, and wide distribution of bolts further complicate the process [3,4]. Recently, advances in unmanned aerial vehicle (UAV) technology and neural deep learning have spurred interest in automating bolt defect detection using UAV imagery [5]. Recent studies have proposed various deep learning-based methods for bolt and split pin defect detection. Jiao *et al*. introduced a Deformable-DETR network that incorporates multiple attention mechanisms and feature alignment to address limitations in modeling pixel-level contour-environment relationships, achieving promising detection performance [6]. Zhao *et al*. developed the AVSCNet network, which employs automatic visual shape clustering, extended RoI feature extraction, and bilinear interpolation-based feature enhancement and fusion to mitigate interference from inconsistent two-dimensional visual structures of bolts [7]. Sun *et al*. applied a grid-splitting approach to the original image, followed by recognition using an R-FCN network enhanced with class activation mapping and a large margin Softmax loss function [8]. Zhong *et al.* proposed a three-stage automatic defect detection system for split pins using PVANET++ with a novel anchor mechanism; split pins are first localized via PVANET++, the Hough transform, and the Chan–Vese model, and then defects are detected according to a predefined standard [9]. Similarly, Wang *et al.* designed a three-stage method for detecting missing split pins: YOLOv4 detects insulators and localizes connecting accessories by expanding prediction borders; Faster R-CNN with ResNet-101 then detects and classifies bolts; and finally, DenseNet121 identifies whether split pins are intact or missing [10].

While these detection methods enable general split pin fault detection, real-world UAV inspection images present additional challenges. Split pins often occupy only a small fraction of the image, and their high numbers further complicate detection. Moreover, complex backgrounds, variable lighting, and potential focus issues increase both the difficulty and uncertainty of split pin fault detection, necessitating higher accuracy and adaptability from detection methods. To address these challenges, we propose a practical detection method that sequentially applies image cropping, super-resolution reconstruction, and split pin fault detection. Our main contributions are as follows:

(1) By leveraging weights obtained from training on a large-scale transmission line dataset, we crop images of fittings containing split pins from the target dataset, effectively addressing the issue of a low proportion of split pin imagery.

(2) We apply super-resolution reconstruction to the cropped fitting images, thereby enhancing clarity and detail so that the reconstructed images more accurately display the characteristics and conditions of split pins, ultimately improving detection accuracy and reliability.

(3) We incorporate local window attention into the connection between the backbone and neck networks of the YOLOv8 architecture to enhance local feature extraction. Additionally, we replace CIoU with inner-shape IoU to mitigate sample imbalance and introduce auxiliary bounding boxes to further optimize detection performance.

2. Dataset and PIN recognition method

2.1. Dataset

The dataset used in this study was provided by a company in Southwest China and consists of drone-captured images from actual transmission line inspections. It comprises 3369 images with resolutions ranging from (1000–8000)×(600–6000) pixels. The dataset includes 678 labels for abnormal pins and 3,433 labels for missing pins. Abnormal pins refer to pins that are misaligned, displaced to positions where detachment is likely, or partially detached. Missing pins indicate that a pin is absent from its designated position. **Figure 1** shows typical examples of normal, abnormal, and missing pins (from left to right, respectively).

The dataset is randomly divided into training and test sets in a 9:1 ratio, as shown in **Table 1**. The training set comprises 3,032 images for model learning and parameter tuning, while the test set consists of 337 images for evaluating the model's generalization ability on unseen data.

Figure 1. Schematic diagram of pin defects

Table 1. Training and test set statistics

Set	Images	Abnormal pin labels	Missing pin labels
Train	3032	602	3140
Test	337	76	293

2.2. Pin recognition method

Split pin defects in aerial images of transmission lines are difficult to identify (**Figure 2**). To address this, we propose a fault detection method tailored for UAV inspections. We propose a fault detection method tailored for UAV inspections. First, trained models perform segmentation on the input aerial images to distinguish those containing split pins from those without. This step increases the pixel proportion of split pins while reducing interference from complex backgrounds. Next, super-resolution reconstruction is applied to the segmented images to capture the distinctive shape and texture features of split pins more accurately. Following this, deep learning techniques are employed using a large dataset of annotated samples to develop a reliable detection model. To further improve accuracy, a local window attention mechanism is introduced, allowing the model to focus on critical areas and better extract split pin features. Finally, a coordinate transformation maps the detection results back to the original images, enabling manual verification by maintenance personnel and facilitating timely repairs. This method significantly reduces manual workload compared to traditional inspection techniques while ensuring rapid and precise identification of split pin faults to safeguard the safe and stable operation of transmission lines.

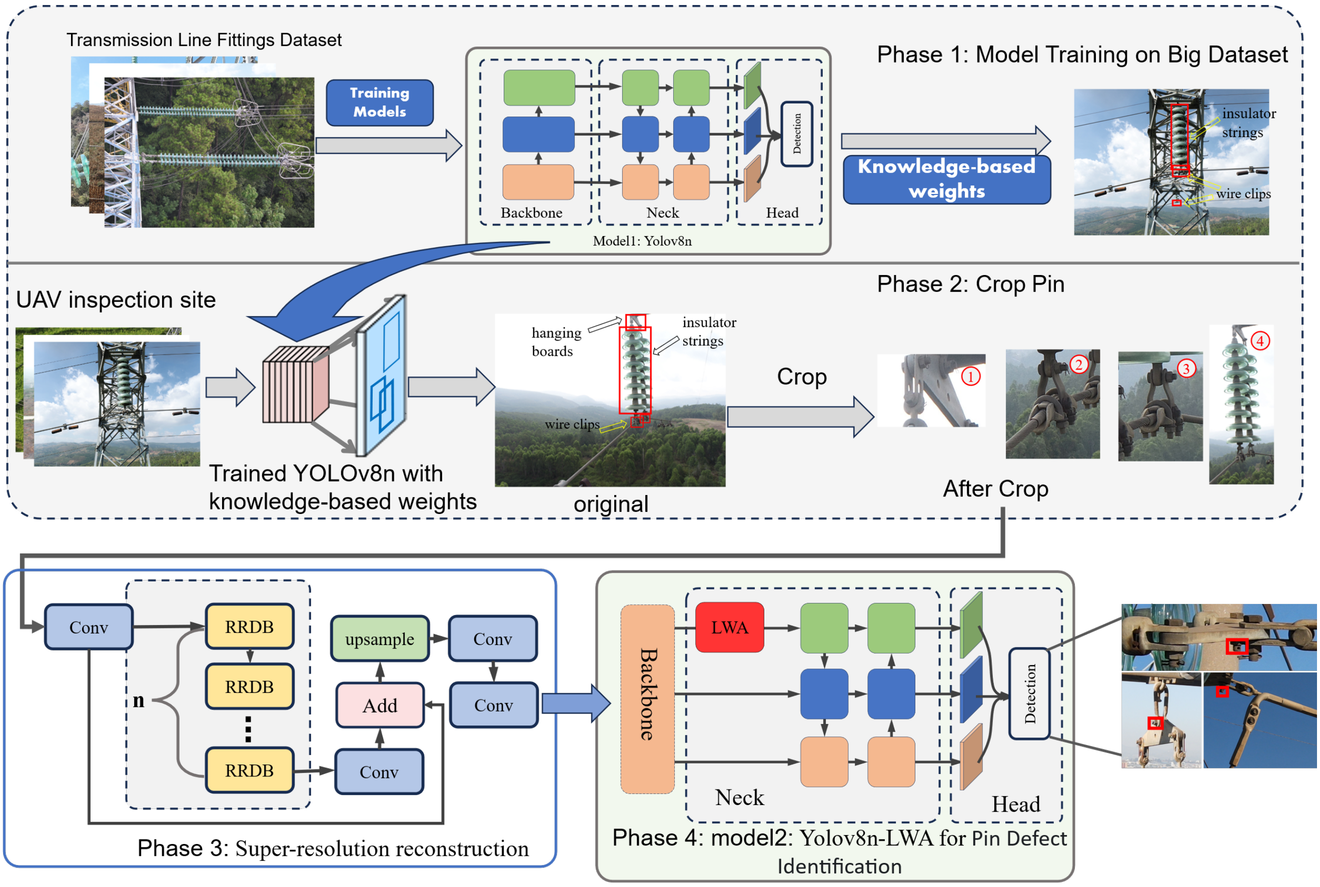

Figure 2. Pin recognition method

2.2.1. Positioning figures and tables

In practical production and UAV-based aerial inspections, the target for pin detection is much smaller compared to those in other public datasets or common transmission line fault detection tasks. This small target detection presents numerous challenges [11]. As illustrated in **Figure 3**, the label sizes in our drone dataset are primarily concentrated around 0.005×0.005 of the image area, which is only one-fourth—or even smaller—than those in datasets such as CPLID (used for insulation string defect detection) and the safety belt detection dataset from the

Guangdong Power Grid Intelligent Fieldwork Challenge [12,13]. Such small label sizes, combined with complex backgrounds, significantly hinder accurate detection. For instance, bolts at the connection points of angle steel on transmission towers, which are not designed with pins, may be mistakenly identified as pin defects by the model. To overcome these challenges, we observed that split pins are typically affixed to fittings, such as wire clips, hanging rings, connecting boards, hanging boards, and adjusting plates, which occupy a larger portion of the image and exhibit unique shapes. By detecting and cropping these fittings separately, we effectively increase the proportion of pin pixels and reduce background complexity. Additionally, since many fittings are connected to insulator strings, which are larger and easier to recognize, we also crop insulator strings to enhance the robustness of our method.

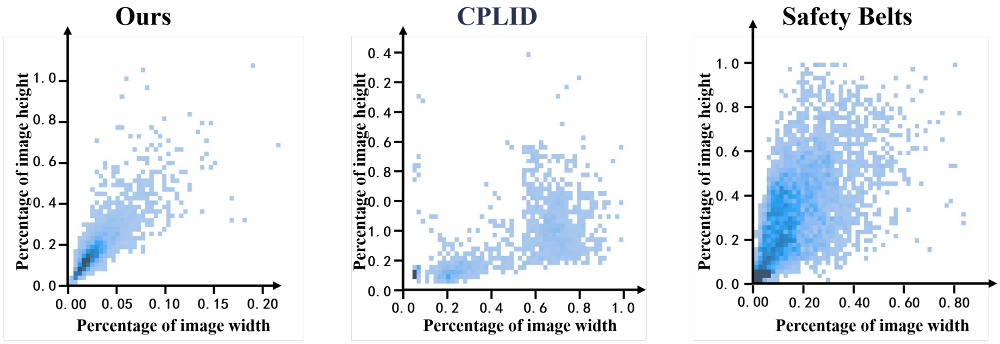

Figure 3. Comparison of Anchor box size

As shown in Phase 1 and Phase 2 of **Figure 2**, we adopt transfer learning by training a YOLOv8 model on a dataset of 7292 transmission line inspection images to detect wire clips, hanging rings, connecting boards, hanging boards, hanging boards, adjusting plates, and insulator strings [14]. The resulting weights are applied to the target pin defect dataset. To ensure complete coverage, the detected bounding boxes are expanded by 1.5 times before cropping. The pin defect labels are then transformed to correspond accurately to the cropped images. **Figure 2** presents a schematic of the image cropping process. **Table 2** summarizes the results: from 337 validation images, 638 cropped images were obtained, containing 101 abnormal pin labels and 274 missing pin labels. Although 7 abnormal and 34 missing pin labels were lost during cropping, the preservation rates reached 90.8% and 88.4%, respectively, demonstrating the efficacy of our cropping method.

Table 2. Label statistics after image cropping

Label statistics	Images	Abnormal pin labels	Missing pin labels
Validation	337	76	293
Cropped	638	101	274
Lost labels	—	7	34

2.2.2. Super-resolution reconstruction

Super-resolution reconstruction processes low-resolution images using known image information to restore detailed high-resolution features [15]. This technique enlarges small images, significantly enhances clarity, and makes textures, edges, and other fine features more distinguishable. Most cropped images are between 200 and 500 pixels in size, which is relatively small for deep learning-based analysis and target detection. The limited pixel count may hinder the detection model's ability to capture detailed features of pins, affecting accurate defect recognition. Moreover, small images restrict the model's performance, as high-definition inputs are required to fully exploit its capabilities.

In our work, super-resolution reconstruction is implemented using generative adversarial networks (GANs) [16]. Specifically, we adopt an approach inspired by ESRGAN [17], which improves upon SRGAN by incorporating residual dense blocks (RRDB) to enhance image sharpness and edge detail [18]. Moreover, we reference the Real-ESRGAN model [19], which employs advanced techniques such as a high-order degradation model and a U-Net discriminator with spectral normalization. This approach not only enhances image details but also effectively reduces artifacts. In our experiments, the resolution of the cropped images is doubled, resulting in significantly improved texture clarity and edge definition of the pins, as shown in **Figure 4**.

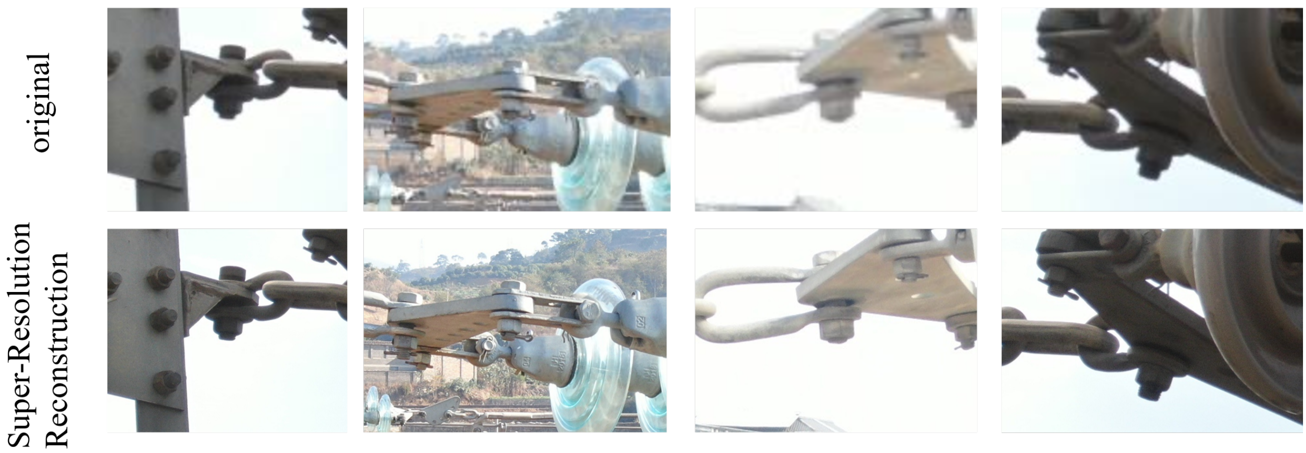

Figure 4. Images before and after super-resolution reconstruction

2.2.3. Detection model improvement

YOLOv8 has demonstrated outstanding performance in target detection and is thus chosen as the final model for identifying pin defects [20]. However, to further improve detection accuracy, we enhance YOLOv8 by integrating an Inner-shape-IoU loss to refine CIoU and by incorporating local window attention to boost feature extraction for pin defects.

The CIoU method in YOLOv8 evaluates the geometric relationship between the ground truth (GT) box and the predicted box by considering their relative position and shape. However, CIoU overlooks inherent attributes such as the box's shape and scale [21]. The Shape-IoU method addresses this limitation by explicitly incorporating these factors into the loss computation (**Figure 5**) [22]. Its formulation is as follows:

$$IoU = \frac{|B \cap B^{gt}|}{|B \cup B^{gt}|}, ww = \frac{2 \times (w^{gt})^{scale}}{(w^{gt})^{scale} + (h^{gt})^{scale}}, hh = \frac{2 \times (h^{gt})^{scale}}{(w^{gt})^{scale} + (h^{gt})^{scale}} \tag{1}$$

$$distance^{shape} = hh \times \frac{(x_c - x_c^{gt})^2}{c^2} + ww \times \frac{(y_c - y_c^{gt})^2}{c^2} \tag{2}$$

$$\Omega^{shape} = \sum_{t=w,h} (1 - e^{\omega t})^4 \tag{3}$$

$$\omega_w = hh \times \frac{|w - w^{gt}|}{max\,(w, w^{gt})}, \omega_h = ww \times \frac{|h - h^{gt}|}{max\,(h, h^{gt})} \tag{4}$$

Here, the scale factor adjusts to the target's size, while *ww* and *hh* serve as weight coefficients in the horizontal and vertical directions based on the ground truth (GT) box dimensions. Building on Shape-IoU and inspired by Inner-IoU [23], we introduce an auxiliary bounding box to compute an extended IoU loss. As shown in **Figure 5**, employing a larger auxiliary bounding box enlarges the effective regression range, which improves localization for small samples. The inner-IoU is computed as follows:

$$b_l^{gt} = x_c^{gt} - \frac{\omega^{gt} \times ratio}{2}, b_r^{gt} = x_c^{gt} + \frac{\omega^{gt} \times ratio}{2} \tag{5}$$

$$b_t^{gt} = y_c^{gt} - \frac{h^{gt} \times ratio}{2}, b_b^{gt} = y_c^{gt} + \frac{h^{gt} \times ratio}{2} \tag{6}$$

$$b_l = x_c - \frac{\omega \times ratio}{2}, b_r = x_c + \frac{\omega \times ratio}{2} \tag{7}$$

$$b_l = x_c - \frac{\omega \times ratio}{2}, b_r = x_c + \frac{\omega \times ratio}{2} \tag{8}$$

$$b_t = y_c - \frac{h \times ratio}{2}, b_b = y_c + \frac{h \times ratio}{2} \tag{9}$$

$$inter = (min\,(b_r^{gt}, b_{rb}) - max\,(b_l^{gt}, b_l)) \times (min\,(b_b^{gt}, b_b) - max\,(b_t^{gt}, b_t)) \tag{10}$$

$$union = \omega^{gt} \times h^{gt} \times ratio^2 + \omega \times h \times ratio^2 \tag{11}$$

$$IoU^{inner} = \frac{inter}{union} \tag{12}$$

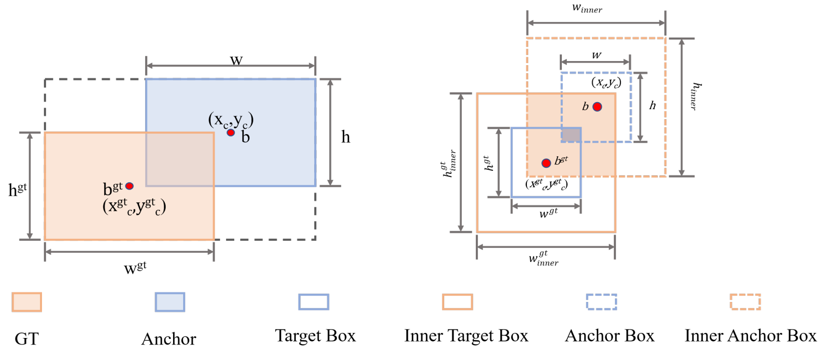

Figure 5. Calculation schematic of shape-IoU and inner IoU loss

In our study, the scaling coefficient (ratio) is set to 1.2. Integrating the auxiliary bounding box with the Shape-IoU method yields the Inner-shape-IoU loss:

$$L_{inner-shape-IoU} = L_{shape-IoU} + IoU - IoU^{inner} \tag{13}$$

2.2.4. Local window attention

Since the introduction of the self-attention mechanism in Transformer and the subsequent development of Vision Transformer [24,25], self-attention has become widely adopted in image analysis. EfficientViT further advanced this area with the Cascaded Group Attention mechanism, which partitions input features into groups for independent self-attention processing, reducing computational cost and enhancing efficiency [26].

Inspired by this, we propose local window attention. As illustrated in **Figure 6**, the input image is divided into multiple local windows, each processed independently with self-attention. The output from one window is then added to the input of the subsequent window, and cascading multiple attention modules enables the model to capture multi-level features effectively. **Figure 7** demonstrates that incorporating LWA after the SPPF module in YOLOv8n refines multi-scale fused features. This refinement allows the model to focus on critical features, thereby improving target localization and overall detection accuracy.

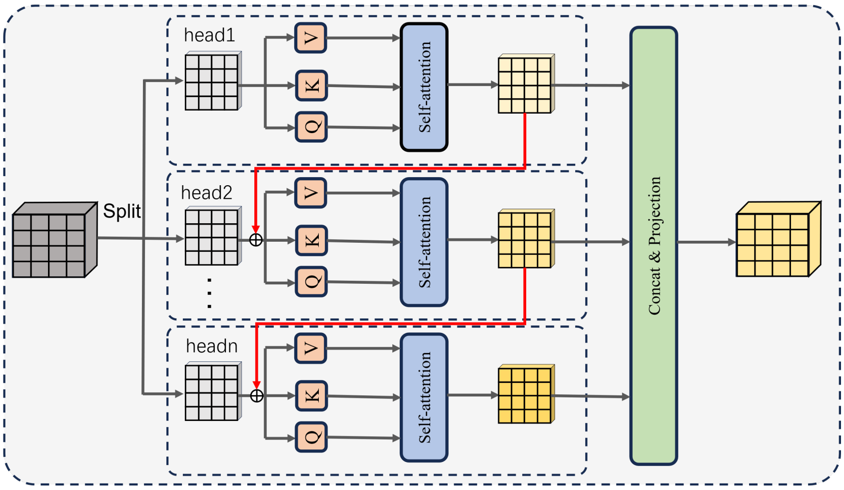

Figure 6. Local window attention

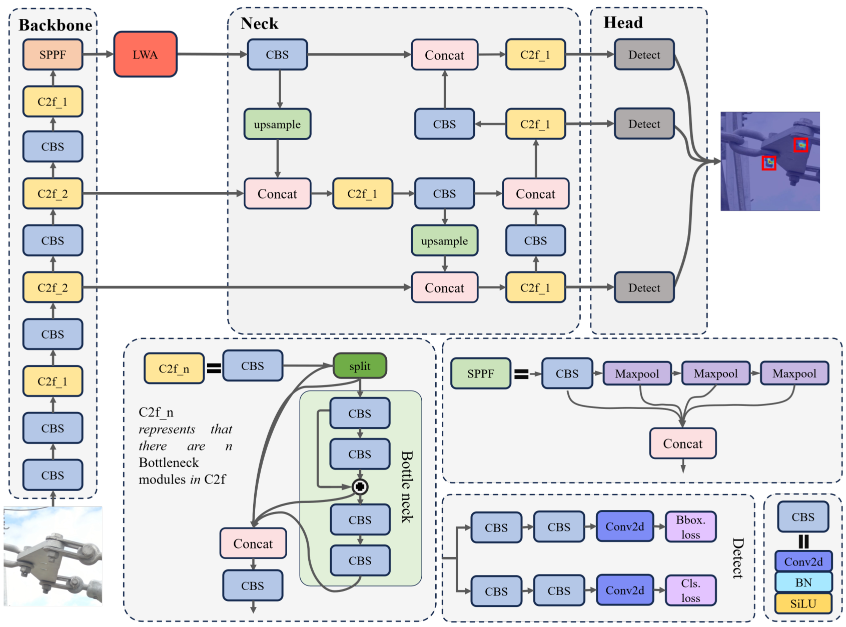

Figure 7. Yolov8n with local window attention

3. Experiment results and discussion

3.1. Experimental environment and evaluation metrics

Experiments were conducted using an RTX 4090D GPU (24 GB VRAM) paired with an Intel® Xeon® Platinum 8474C CPU on Ubuntu 20.04. We utilized Python 3.8, PyTorch 1.11.0, and CUDA 11.3. The model was trained for 300 epochs using SGD (initial learning rate: 0.01) with a batch size of 32 and 8 workers. Additionally, we evaluated performance using Precision, Recall, and Average Precision (*AP*). Precision is defined as the ratio of true positives (*TP*) to the sum of TP and false positives (*FP*). Recall is the ratio of *TP* to the sum of *TP* and false negatives (*FN*). *AP* represents the area under the precision-recall curve, comprehensively measuring performance.

These metrics are calculated as follows:

$$Precision = \frac{TP}{TP+FP}, Recall = \frac{TP}{TP+FP}, AP = \int_0^1 P(R)dR \quad (14)$$

Among them, TP represents the number of positive examples that are correctly classified. FP represents the number of negative examples that are wrongly classified as positive examples. FN represents the number of positive examples that are wrongly classified as negative examples. P represents Precision rate, and R represents Recall rate.

3.2. Experimental results

To evaluate the impact of cropping and super-resolution reconstruction on pin defect detection, we conducted five experiments using the YOLOv8n model on three image types: original images, cropped images, and cropped images with super-resolution. The image types are defined as follows: Type 1 comprises the original images; Type 2 includes cropped images achieved through shearing; Type 3 consists of cropped images with super-resolution applied solely to the test set; and Type 4 involves cropped images with super-resolution applied to the entire dataset.

Table 3 summarizes results at various resolutions. For original images at 1280×1280, the mean AP (mAP) was 25.8%; abnormal pin Recall was 14.5% with an AP of 18.9%, and missing pin AP was 32.7%, indicating insufficient baseline performance. After cropping the images and training at 640×640, both Precision and Recall exceeded 50%, with mAP rising to 59.1%. This improvement confirms that cropping effectively enhances pin defect recognition. However, training cropped images at 1280×1280 yielded only a 0.9% mAP increase, and the performance of missing pin detection declined. Applying super-resolution reconstruction solely on the cropped test set further reduced mAP to 58.6%, likely due to noise introduced during upsampling.

Table 3. Results comparison for different image types and resolutions

Image type	Image size	Metric	Mean (%)	Abnormal pin (%)	Missing pin (%)
1	1280×1280	Precision	43.9	48.0	33.9
		Recall	28.4	14.5	42.3
		AP	25.8	18.9	32.7
2	640×640	Precision	65.5	58.9	72.1
		Recall	53.0	43.9	62.0
		AP	59.1	51.0	67.3
3	1280×1280	Precision	73.6	69.9	77.3
		Recall	52.8	46.5	59.1
		AP	60.0	54.6	65.5
4	1280×1280	Precision	66.4	59.1	73.8
		Recall	63.1	59.4	66.8
		AP	64.2	58.6	69.8

Conversely, applying super-resolution reconstruction to all cropped images and training at 1280×1280 resulted in an average Recall of 63.1%—the highest across experiments—and a mAP of 64.2%, a 4%

improvement over non-super-resolved images. This demonstrates that super-resolution reconstruction enhances high-frequency details and improves feature extraction for pin defect detection. An ablation study was performed on cropped, super-resolved images at 1280×1280 to assess the contributions of inner-shape IoU and local window attention. **Table 4** shows that Inner-shape-IoU Alone: Increased overall AP by 1.3%, though Recall decreased slightly, possibly due to its focus on precise localization. LWA Alone: Improved the AP for missing pins from 69.8% to 72.2%, with a modest decline in average Recall. Combined (YOLOv8n-LWA): Achieved an AP of 61.0% for abnormal pins (up 2.5%) and 74.5% for missing pins (up 4.7%), with an overall Precision increase of 5.7% and a 4% mAP improvement. This suggests that LWA compensates for the localization emphasis of inner-shape IoU, effectively enhancing defect detection.

Table 4. Ablation study on YOLOv8 enhancements

Model	Type	Precision (%)	Recal1 (%)	AP (%)
YOLOv8n	Mean	66.4	63.1	64.2
	Abnormal pin	59.1	59.4	58.6
	Missing pin	73.8	66.8	69.8
+Inner-shape-IoU	Mean	76.0	57.5	65.5
	Abnormal pin	69.1	51.5	59.6
	Missing pin	82.8	63.5	71.4
+LWA	Mean	74.0	61.4	66.2
	Abnormal pin	68.6	56.4	60.2
	Missing pin	79.4	66.3	72.2
Yolov8n-LWA	Mean	72.1	61.6	67.7
	Abnormal pin	66.3	56.5	61.0
	Missing pin	77.9	66.8	74.5

Figure 8 displays heat maps comparing YOLOv8n and YOLOv8n-LWA. The enhanced model reduces both false negatives and false positives, with higher confidence in pin defect detection.

yolov8n

Yolov8n-law

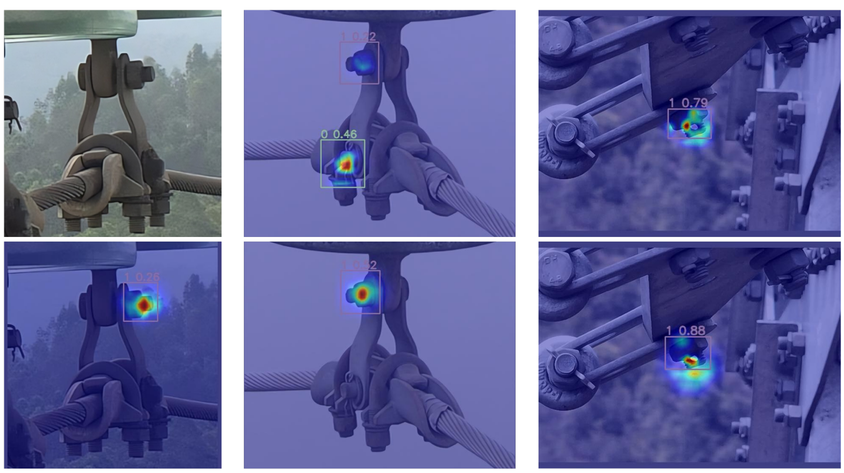

Figure 8. Heat maps comparing YOLOv8n and YOLOv8n-LWA (“0” and “1” denote the two defect types: abnormal pin and missing pin)

3.3. Robustness analysis

To evaluate the performance of the YOLOv8n-LWA model under challenging conditions, we applied a strong exposure enhancement to all test images to simulate sunlight irradiation. Subsequently, we analyzed the model's detection results on this enhanced dataset.

As shown in **Table 5**, at a resolution of 1280×1280, prolonged exposure of test images led to a decrease in model accuracy, with the mAP dropping by 3.2%. **Figure 9** illustrates that this degradation may be attributed to factors such as overexposure and noise, which impede the model's ability to accurately detect targets. Nevertheless, the overall performance remains satisfactory, with the mAP still reaching 64.5%. These findings indicate that the YOLOv8n-LWA model maintains robust detection capabilities under adverse conditions, preserving target recognition to a considerable extent in real-world environments.

Table 5. Experimental results of the model

Image condition	Metric	Mean (%)	Abnormal pin (%)	Missing pin (%)
Original	Precision	72.1	66.3	77.9
	Recall	61.6	56.5	66.8
	AP	67.7	61	74.5
Strong exposure	Precision	70.8	64.8	76.8
	Recall	59.5	56.5	62.4
	AP	64.5	58.5	70.6

Figure 9. Detection results of long-exposure images. ("0" and "1" denote the two defect types: abnormal pin and missing pin.)

4. Conclusion

This paper presents a novel detection method for identifying pin defects in aerial images of transmission lines. Our approach involves three key steps: cropping fittings that contain pins, applying super-resolution reconstruction to enhance image clarity, and utilizing the YOLOv8-LWA model for defect recognition. This integrated strategy significantly improves the detection of pin defects in aerial imagery.

The method leverages pre-trained weights from large transmission line datasets to precisely crop fittings containing pins, retaining 90% of defects and amplifying their representation through duplicate cropping. Super-resolution reconstruction enhances image clarity, enabling high-resolution training and improving defect discriminability. The YOLOv8-LWA architecture introduces two key innovations: (1) Inner-Shape-IoU, which refines bounding box localization by prioritizing target geometry and scale, and (2) Local window attention, which captures spatially contextual defect features through position-aware self-attention. Evaluated on cropped datasets, the model achieves a precision of 72.1%, a recall of 61.6%, and an mAP of 67.7%. Notably, it maintains robust performance (mAP=64.5%) on long-exposure aerial images, demonstrating strong generalization across challenging imaging conditions.

While effective, this study has limitations. The training dataset remains constrained in size and defect diversity, potentially limiting the model's adaptability to rare or region-specific pin anomalies. Furthermore, the cropping algorithm requires optimization to improve defect retention rates during preprocessing. Future work should prioritize expanding the dataset with aerial imagery from diverse geographical regions and environmental conditions (e.g., fog, rain, or snow) to enhance generalizability. Integrating adaptive cropping thresholds and exploring few-shot learning techniques could further strengthen the framework's practicality for large-scale infrastructure inspections.

Funding

Fundamental Research Funds for the Central Universities (2023MS134)

Disclosure statement

The authors declare no conflict of interest.

References

[1] Zhang Y, Yuen K, 2022, Bolt Damage Identification based on Orientation-Aware Center Point Estimation Network. Structural Health Monitoring, 21(2): 438–450. https://doi.org/10.1177/14759217211004243

[2] Yang L, Fan J, Liu Y, et al., 2020, A Review on State-of-the-Art Power Line Inspection Techniques. IEEE Transactions on Instrumentation and Measurement, 69(12): 9350–9365. https://doi.org/10.1109/TIM.2020.3031194

[3] Toth J, Gilpin-Jackson A, 2010, Smart View for a Smart Grid — Unmanned Aerial Vehicles for transmission Lines. 2010 1st International Conference on Applied Robotics for the Power Industry, Montreal, QC, Canada, 1–6. https://doi.org/10.1109/CARPI.2010.5624465

[4] Chen J, Liu Z, Wang H, et al., 2017, 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.

[5] Liang H, Zuo C, Wei W, 2020, Detection and Evaluation Method of Transmission Line Defects Based on Deep Learning. IEEE Access, 8: 38448–38458.

[6] Jiao R, Li K, Zhang X, et al., 2024, Bolt Defect Detection Method Based on Multiple Attention and Feature Alignment. Electric Power Information and Communication Technology, 22: 21–29.

[7] Zhao Z, Qi H, Qi Y, et al., 2020, Detection Method Based on Automatic Visual Shape Clustering for Pin-Missing Defect in Transmission Lines. IEEE Transactions on Instrumentation and Measurement, 69(9): 6080–6091.

[8] Sun J, Liu G, Liu H, 2023, Fine-Grained Detection of Pin Defects Based on Improved R-FCN and Class Activation Mapping. Guangdong Electric Power, 36: 50–57.

[9] Zhong J, Liu Z, Han Z, et al., 2019, A CNN-Based Defect Inspection Method for Catenary Split Pins in High-Speed Railway. IEEE Transactions on Instrumentation and Measurement, 68(8): 2849–2860.

[10] Wang H, Shao Y, Zou S, et al., 2021, Detection of Cotter Pins Missing of Connection Fittings on Transmission Lines of Power System. Proceedings of the 2021 40th Chinese Control Conference (CCC), 6873–6879.

[11] Xiao L, Wu B, Hu Y, 2021, Missing Small Fastener Detection Using Deep Learning. IEEE Transactions on Instrumentation and Measurement, 70: 1–9.

[12] Tao X, Zhang D, Wang Z, et al., 2020, Detection of Power Line Insulator Defects Using Aerial Images Analyzed With Convolutional Neural Networks. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 50(4): 1486–1498.

[13] Liu L, Huang K, Bai Y, et al., 2024, Real-time Detection Model of Electrical Work Safety Belt based on Lightweight Improved YOLOv5. Journal of Real-Time Image Processing, 21(4): 151.

[14] Weiss K, Khoshgoftaar TM, Wang D, 2016, A Survey of Transfer Learning. Journal of Big Data, 3(1): 9.

[15] Li K, Yang S, Dong R, et al., 2020, Survey of Single Image Super-Resolution Reconstruction. IET Image Processing, 14(11): 2273–2290.

[16] Goodfellow I, Pouget-Abadie J, Mirza M, et al., 2020, Generative Adversarial Networks. Commun. ACM, 63(11): 139–144.

[17] Wang X, Yu K, Wu S, et al., 2018, ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks. Proceedings of the European Conference on Computer Vision (ECCV) Workshops.

[18] Ledig C, Theis, L, Huszar-Ferenc C, et al., 2017, Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017.

[19] Wang X, Xie L, Dong C, et al., 2021, Real-ESRGAN: Training Real-World Blind Super-Resolution with Pure Synthetic Data. Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) Workshops, 1905–1914.

[20] Córdova-Esparza TJ, Romero-González DM, Alenjandro J, 2023, A Comprehensive Review of YOLO Architectures in Computer Vision: From YOLOv1 to YOLOv8 and YOLO-NAS. Machine Learning and Knowledge Extraction, 5(4): 1680–1716.

[21] Zheng Z, Wang P, Liu W, et al., 2020, Distance-IoU Loss: Faster and Better Learning for Bounding Box Regression. Proceedings of the AAAI Conference on Artificial Intelligence, 34: 12993–13000.

[22] Zhang H, Zhang S, 2023, Shape-IoU: More Accurate Metric Considering Bounding Box Shape and Scale. arXiv preprint. https://doi.org/10.48550/arXiv.2312.17663

[23] Zhang H, Xu C, Zhang S, 2023, Inner-IoU: More Effective Intersection over Union Loss with Auxiliary Bounding Box. arXiv preprint. https://doi.org/10.48550/arXiv.2311.02877

[24] Vaswani A, Shazeer N, Parmar N, et al., 2017, Attention is All You Need. Advances in Neural Information Processing Systems, 30.

[25] Dosovitskiy A, Beyer L, Kolesnikov A, et al., 2020, An Image is Worth 16x16 words: Transformers for Image Recognition at Scale. arXiv preprint. https://doi.org/10.48550/arXiv.2010.11929

[26] Liu X, Peng H, Zheng N, et al., 2023, EfficientViT: Memory Efficient Vision Transformer with Cascaded Group Attention. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 14420–14430.

Publisher's note

Bio-Byword Scientific Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.