

Article

Insulator Defect Detection Based on YOLOv8s-SwinT

Zhendong He ¹, Wenbin Yang ¹, Yanjie Liu ², Anping Zheng ¹, Jie Liu ^{1,*}, Taishan Lou ¹ and Jie Zhang ¹

¹ College of Electrical and Information Engineering, Zhengzhou University of Light Industry, Zhengzhou 450002, China; hezhendong@email.zzuli.edu.cn (Z.H.); 332101050066@zzuli.edu.cn (W.Y.); 1985001@zzuli.edu.cn (A.Z.); loutaishan@zzuli.edu.cn (T.L.); 2018007@zzuli.edu.cn (J.Z.)

² Sinohydro Bureau 3 Co., Ltd., Xi'an 710024, China; lyj_sd3@powerchina.cn

* Correspondence: 2006035@zzuli.edu.cn; Tel.: +86-186-3920-6027

Abstract: Ensuring the safety of transmission lines necessitates effective insulator defect detection. Traditional methods often need more efficiency and accuracy, particularly for tiny defects. This paper proposes an innovative insulator defect recognition method leveraging YOLOv8s-SwinT. Combining Swin Transformer and Convolutional Neural Network (CNN) enhances the model's understanding of multi-scale global semantic information through cross-layer interactions. The improved BiFPN structure in the neck achieves bidirectional cross-scale connections and weighted feature fusion during feature extraction. Additionally, a new small-target detection layer enhances the capability to detect tiny defects. The experimental results showcase outstanding performance, with precision, recall, and mAP reaching 95.6%, 95.3%, and 97.7%, respectively. This boosts detection efficiency and ensures high accuracy, providing robust support for real-time detection of tiny insulator defects.

Keywords: insulators; YOLOv8s; Swin Transformer; defect detection; BiFPN

1. Introduction

In the power system, insulators play a crucial role as they support and secure transmission lines, preventing current leakage and ensuring the regular operation of the circuits. However, insulators with defects may lead to arc discharges and leakage, thereby increasing the risk of fires, explosions, or electrical accidents in the power system. This poses adverse effects on the stability and reliability of the power system. Therefore, the regular inspection of insulators is of utmost importance for promptly identifying potential safety hazards and taking necessary maintenance and repair measures to ensure the secure operation of the power system. Insulator defect detection is vital for safeguarding the power system's safety, reliability, and efficient operation [1–4] and helps to reduce the risk of accidents and maintenance costs. Despite significant progress in insulator defect detection technology [5–7] over the past few decades, challenges persist in detecting defects in insulator images due to complex backgrounds and the minute nature of the defects.

There are several traditional detection methods for insulator defects. Yin et al. [8] proposed an innovative ultra-wideband microwave fault diagnostic system using the multimode transfer theory of dielectric waveguides. This system detects internal defects by leveraging the level difference of mode transfer. Mei et al. [9] introduced a microwave-based method for identifying internal defects in composite insulators and an automated detection system for full-size composite insulators. In experiments, they artificially created air gaps, carbonization, and conductive defects, and the developed system accurately located these defects. Jiang et al. [10] presented a terahertz imaging method utilizing edge detection algorithms. This approach employs the canny operator to extract defect edges and calculate the time interval of defect characteristic pulses to determine defect depth, resulting in three-dimensional defect imaging. Despite the effectiveness of these traditional methods, their processing speeds could be faster, which limits their suitability for real-time detection.



Citation: He, Z.; Yang, W.; Liu, Y.; Zheng, A.; Liu, J.; Lou, T.; Zhang, J. Insulator Defect Detection Based on YOLOv8s-SwinT. *Information* **2024**, *15*, 206. <https://doi.org/10.3390/info15040206>

Academic Editor: Katsuhide Fujita

Received: 6 March 2024

Revised: 28 March 2024

Accepted: 1 April 2024

Published: 6 April 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

2. Related Works

The continuous advancement of drone technology and deep learning techniques has made these methods' integration in the field of defect detection a trend. Regarding utilizing drones for state monitoring, Waqas et al. [11] proposed a novel and unique real-time OAM (Obstacle Avoidance Module) to prevent severe accidents during autonomous drone flight, coupled with the YOLOv3 algorithm for obstacle detection. Ali et al. [12] presented an autonomous UAV system that utilizes an improved Faster R-CNN for detecting diverse structural damages, subsequently mapping these detections onto GPS-denied environments. In recent years, deep-learning-based object detection algorithms have shown significant advancements in insulator detection for power transmission lines [13–16], including detecting multiple types of damage [17–19], and have outperformed traditional methods. Yang et al. [20] extended the YOLOv3 network, introducing a convolutional neural network that maintains detection speed while enhancing efficiency. Zhao et al. [21] presented a model combining an improved Faster R-CNN network for insulator localization with image segmentation using an adaptive threshold algorithm in the HSV color space. Insulator fault detection involves techniques like line detection, image rotation, and vertical projection. Hu et al. [22] introduced an AC-YOLO network to improve detection performance. This was achieved by introducing adaptive weight distribution multi-head self-attention modules and adaptive memory fusion and by integrating CBAM attention mechanisms. Zhang et al. [23] proposed DefGAN, a network for insulator defect detection, utilizing concave potential representations to improve classifier reliability. Defect scores are determined by anomaly probability and denoising autoencoder reconstruction error. Zhong et al. [24] developed the TOL-Framework, featuring a novel localization network for insulator target localization. An adversarial reconstruction model, trained exclusively with standard samples, assesses defect states, achieving high localization accuracy. Zhang et al. [25] introduced IL-YOLO, an insulator defect detection algorithm with three enhanced modules: IL-GAM, IL-C3, and IL-SPPFCSPC. Experimental results suggest that IL-YOLO offers notable advantages in handling complex backgrounds and multi-object challenges. However, these models may need more accuracy for detecting small-sized defects, which is critical for ensuring the reliability of overhead power transmission line operation and maintenance.

Fu et al. [26] proposed the I2D-Net which incorporates innovative modules to enhance defect localization in the presence of interference factors to address the challenge of detecting small-sized insulator defects. Li et al. [27] significantly boosted the small-target detection capabilities of the YOLOv7 network by leveraging spatial pyramid pooling and the innovative cross-stage partial channel (SPPCSPC) module, along with integrating the CA module and dynamic convolution module. Shuang et al. [28] introduced a Detail Feature Enhancement module. They integrated an auxiliary classification module into their insulator detection algorithm based on Faster R-CNN, significantly improving recognition accuracy for micro-defects. To address the impact of real-world factors such as lighting and shadows on image processing techniques, Cha et al. [29] employed deep architectures of CNNs to detect damage defects without the need to compute defect features. Liu et al. [30] presented a high-voltage power transmission line target and a defect detection method using deep learning object detection networks. They addressed imbalanced target category quantities by constructing a sizeable standardized dataset and optimized the network to improve feature extraction for small targets. Yu et al. [31] adopted a new lightweight network as the backbone of their designed network, combining denoising with target detection. They employed a novel lightweight network as the backbone for their work, integrating denoising with object detection. They introduced a new loss function, Focal EIOU, effectively addressing the sample imbalance for small objects and improving detection accuracy by suppressing background positives. Zheng et al. [32] enhanced the YOLOv7 model by integrating the Coordinate Attention [33] module and the HorBlock module, using the SIOU Focus Loss function to expedite model convergence and achieve precise detection of tiny target defects on power transmission lines. While deep learning methods

have improved detection speed and accuracy, the accumulation of convolutional layers may result in a gradual loss of information for small targets, which poses a challenge for effective feature extraction by detectors.

The Transformer [34], initially designed for natural language processing tasks like machine translation, revolutionized sequence modelling through self-attention mechanisms, overcoming limitations seen in methods like LSTM [35]. Notably, the Vision Transformer (ViT) [36] bridged the Transformer and CNN for image processing by segmenting input images into patches, treating each patch as a token in the Transformer input sequence. Dian et al. [37] proposed the Faster R-Transformer which combined CNNs with self-attention mechanisms for aviation insulator detection and exhibited robustness across different conditions. Dai et al. [38] introduced the YOLO-Former, integrating ViT, the Convolutional Block Attention Module (CBAM), and the stem module for efficient foreign object detection. While ViT excels at capturing global information, it may struggle with local patterns within patches, which prompted the development of the Swin Transformer [39]. Liu et al. [40] introduced the YOLO-CSM model, which integrated the Swin Transformer and CBAM attention mechanisms alongside an extra detection layer. This model adeptly captures global information and essential visual features, boosting the ability to recognize tiny defects. Zhao et al. [41] addressed the small-sized, low signal-to-noise ratio and texture-detail-scarce targets by proposing a Res-SwinTransformer with a Local Contrastive Attention Network (RSLCANet). Experimental results showcase low false detection rates, high accuracy, and fast detection speed.

Simple defect detection can reveal the potential problems of insulators but cannot provide us with detailed information about the specific degree of the problems. Given this, we must upgrade the detection method to quantitative defect detection to gain a more comprehensive understanding of the health condition of insulators. Choi et al. [42] introduced a deep-learning-based network, SDDNet, tailored explicitly for segmenting concrete cracks in images. Composed of various enhanced modules, the model demonstrated its effectiveness in effectively eliminating complex backgrounds and features resembling cracks. Kang et al. [43] designed an innovative STRNet for pixel-level real-time crack segmentation in complex scenes. This network integrates modules and loss functions, simplifying its design while maintaining fast processing. They also proposed a new method for assessing image scene complexity. Ali et al. [44] developed an IDSNet for segmenting internal damage in concrete components at the pixel level using active thermography. Additionally, they created an AGAN to generate synthetic images for training this network.

To enhance the model's capability to detect minute insulator defects and to improve the extraction of semantic information from different feature layers, we have developed a novel object detection network, the YOLOv8s-SwinT, building upon the foundation of YOLOv8s. The main contributions of this paper are as follows:

- (1) We introduced a Swin Transformer-based Multi-Head Self-Attention (MSA) detection module into the YOLOv8s C2f module, enhancing global modelling during feature extraction.
- (2) The neck of YOLOv8s underwent modification, replacing the FPN + PAN [45,46] structure with the more efficient BiFPN [47] for improved bidirectional multiscale feature fusion, enhancing spatial and semantic information communication.
- (3) A tiny target detection layer was added to the detection head to enhance the model's capability to detect minute defects. By incorporating global and local information, YOLOv8s-SwinT demonstrates increased effectiveness in detecting tiny defects.

These modifications collectively contribute to developing YOLOv8s-SwinT, a robust object detection network tailored to address the challenges of detecting tiny insulator defects.

3. Swin Transformer

The ViT architecture uses MLP layers to capture spatial and two-dimensional features, offering unique benefits in tackling various tasks. Its self-attention mechanism functions

globally on features, distinguishing it from CNNs, and it lacks some inductive biases. Figure 1 showcases ViT’s design, which involves dividing input images into small blocks, converting them to vectors, and then processing them through a Transformer. The encoder follows a traditional Transformer approach, incorporating embedded patches, layer normalization, and MSA. The MSA combines results from V and Q, enhances them through MLPs, and returns them to their original structure. Regarding image classification, ViT incorporates a unique token into the input, with its output as the final prediction. This adaptability allows ViT to excel in image classification tasks and leverage its sequential processing capabilities.

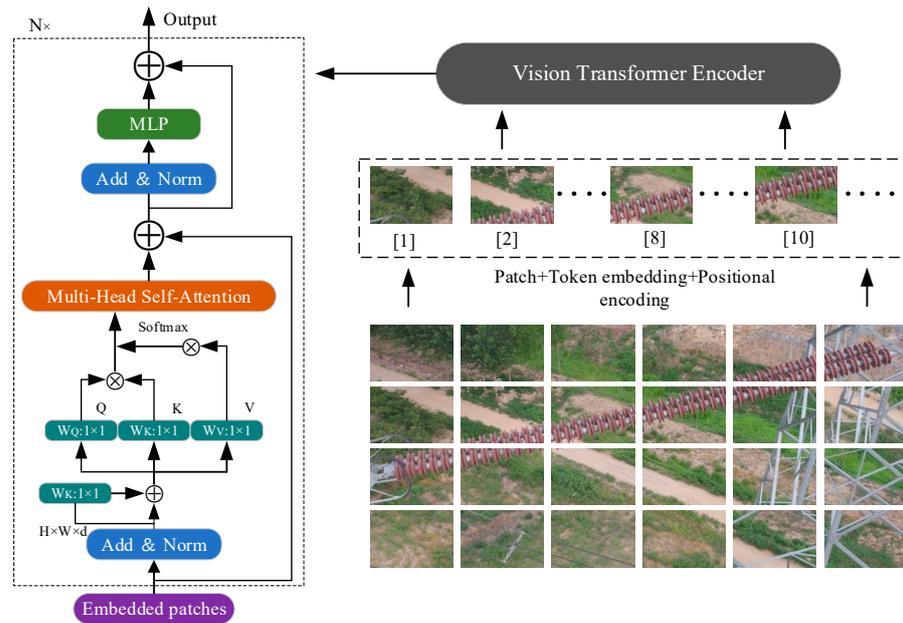


Figure 1. Diagram of the Vision Transformer network architecture.

While ViT achieves global self-attention modelling by dividing images into 16×16 blocks, it has certain limitations in extracting multiscale information. As the network depth increases, the number of blocks in ViT remains constant, restricting its capability for multiscale feature extraction. In contrast, the Swin Transformer adopts a hierarchical building approach, gradually reducing the number of blocks and expanding the receptive field of each block as the network depth increases, enabling better adaptation to multiscale information in visual tasks.

The Swin Transformer diverges from traditional convolutional methods, eliminating redundant information generation and addressing the quadratic complexity issue associated with global self-attention calculations in ViT models. It maintains linear computational complexity with increasing network layers, reflecting hierarchical feature mappings and enhancing performance in multiscale features and dense prediction tasks. The Swin Transformer further enhances the information interaction capability of neighborhood feature blocks through multi-level integrated embeddings. In summary, the illustrated model structure of the Swin Transformer in Figure 2 exhibits a hierarchical design achieved through window sliding calculations. By constraining self-attention within specific, non-overlapping local windows, it comprehensively models the features of each block’s neighborhood, resulting in more substantial global modelling outcomes.

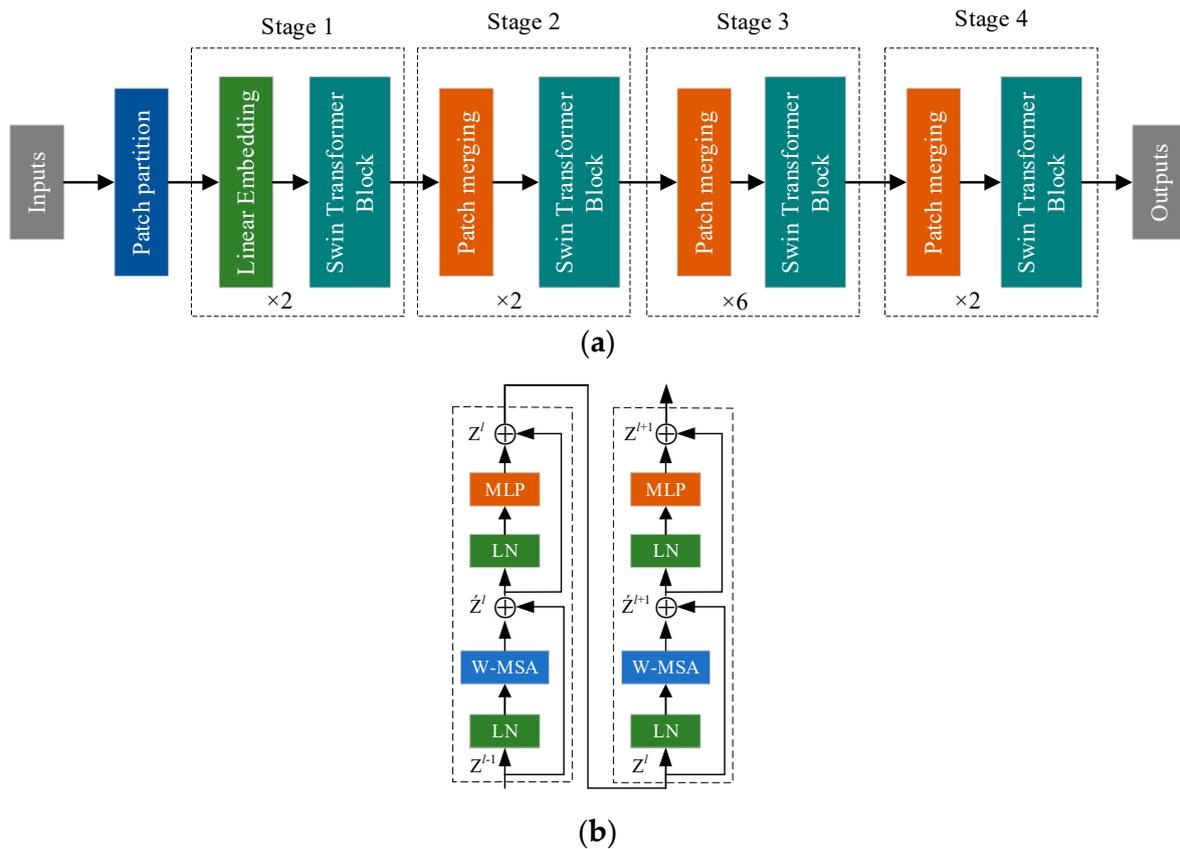


Figure 2. Swin Transformer network architecture. (a) Diagram of the Swin Transformer model structure and (b) Swin Transformer block.

Figure 2a outlines the Swin Transformer model with four distinct stages. It partitions the input image into smaller 4×4 patches, forming patch blocks. In the initial stage, these patches undergo transformation and embedding before being processed by the Swin Transformer block. The subsequent stages merge patches, gradually enriching the feature dimensions.

Figure 2b illustrates the Swin Transformer block, which comprises two Transformer blocks. The first block incorporates normalization, MSA, and MLP. However, MSA’s global self-attention approach poses computational complexities. The second block introduces a sliding window mechanism, facilitating information exchange between adjacent windows. This approach maintains computational efficiency while enhancing performance in multiscale tasks, allowing for flexible information processing for improved visual detection accuracy.

This paper successfully integrates the outstanding characteristics of the Swin Transformer with the C2f module in YOLOv8s, creating a novel network architecture. This innovation enables the network to efficiently capture global and local information on feature maps, establishing a complementary relationship between feature extraction and information fusion. As a result, it dramatically enhances the network’s detection accuracy.

4. Improved YOLOv8s Algorithm

The original YOLOv8s object detection network focuses on capturing local information in feature maps to enhance the model’s receptive field. However, for highly similar data, it may struggle to extract fundamental entity-specific features [48]. To address this issue, we propose a new object detector for detecting insulator damage defects, termed YOLOv8s-SwinT (as shown in Figure 3). This model integrates Swin Transformer modules into the backbone and neck, facilitating global self-attention modelling during feature extraction. The traditional feature pyramid structure in YOLOv8s is replaced by an improved BiFPN

structure, dynamically balancing features with different weight information across different scales. A small object detection layer is introduced to enhance the network’s ability to detect insulator damage defects.

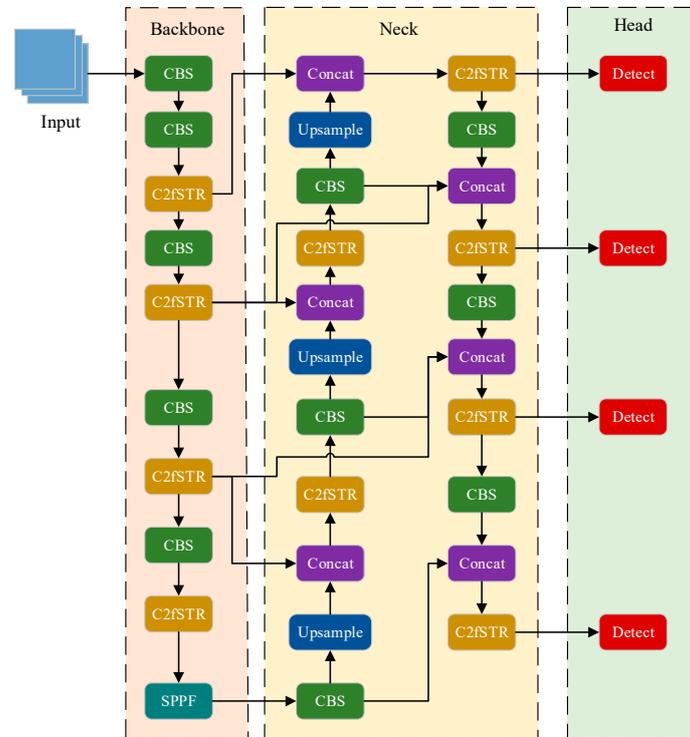


Figure 3. YOLOv8s-SwinT network architecture.

4.1. C2fSTR Module

Currently, deep-learning-based object detection methods primarily employ CNNs as backbone networks. However, convolutional structures are constrained by the size of convolutional kernels, limiting their focus to local regions of the feature map and rendering them less sensitive to global information. This limitation is particularly problematic for detecting small defective targets characterized by small volumes and limited feature information, which often results in feature loss and an increased risk of false negatives. This paper introduces a C2fSTR module designed to overcome the limitations of convolutional structures, enabling the model to better capture global gradient flow information while remaining lightweight. Its structure is shown in Figure 4.

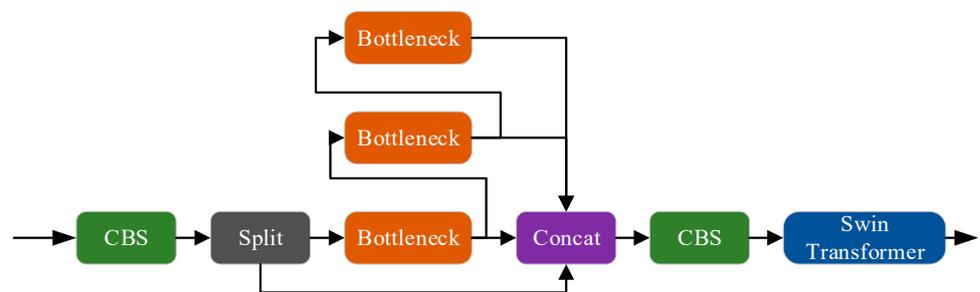


Figure 4. C2fSTR module.

The design inspiration for the C2fSTR module is drawn from the C3 module and ELAN, incorporating a specialized convolutional module that effectively fuses feature maps of different scales, enhancing the model’s receptive field and detection accuracy. The ELAN (Effective Long-Range Aggregation Network) module is a neural network

component primarily intended for capturing remote dependencies. The ELAN module proficiently extracts global contextual data during processing, enhancing the model's accuracy and efficiency. In YOLOv8s, the C2f module is applied to both the backbone and neck networks to merge feature maps of different scales, making it a critical component of the network structure. This paper introduces the Swin Transformer module after the C2f module to further improve the model's performance. The Swin Transformer extracts global features through self-attention mechanisms and sliding windows, reinforcing global information interaction in the feature maps. This combination effectively addresses the shortcomings of traditional convolution in global information acquisition, particularly enhancing the robustness and accuracy of detecting small defective targets. In summary, the proposed C2fSTR module ingeniously integrates the strengths of the C2f and the Swin Transformer, enabling the model to capture global information better while maintaining lightweight characteristics. This results in improved performance in the task of detecting small defective targets.

4.2. Small Object Detector

In traditional object detection networks like the original YOLOv8s, feature fusion typically begins from the third layer of features. We introduce a Small Object Detection Layer on top of the original YOLOv8s algorithm to enhance the network's capability to detect small objects. The role of this layer is to incorporate the second layer of features into the feature fusion network, preserving more shallow semantic information. Specifically, we introduce additional information about small objects by adding an initially unmerged 160×160 feature map in the feature extraction network. To effectively process this new feature map, we perform an upsampling operation in the feature fusion network followed by a downsampling operation. These operations increase detection layers to four, enhancing the network's perceptual capabilities and sensitivity to small objects.

This refined architecture enables the network to comprehensively capture semantic information in the images, mainly when dealing with small objects. By making these adjustments to the original YOLOv8s algorithm, we aim to achieve superior performance in object detection tasks, especially in addressing the challenges posed by small objects.

4.3. Improved BiFPN

Traditional FPN structures in object detection networks use a top-down unidirectional information flow for feature fusion. PANet, as shown in Figure 5a, improves this process by introducing an additional bottom-up path. This enhancement aids in more effective information transmission and the retention of shallow-level features. Building upon PANet, BiFPN, illustrated in Figure 5b, further refines the feature fusion process. The original BiFPN network performs fusion on layers 3 to 7 out of seven feature layers, employing a specific strategy to consider nodes with only one input edge contributing less to the network. BiFPN eliminates feature fusion nodes from layers 3 and 7 to reduce computational complexity. Additionally, a cross-scale connection method is introduced for improved feature representation. This method adds an extra edge to directly fuse features from the feature extraction network with features of relative sizes from the bottom-up path.

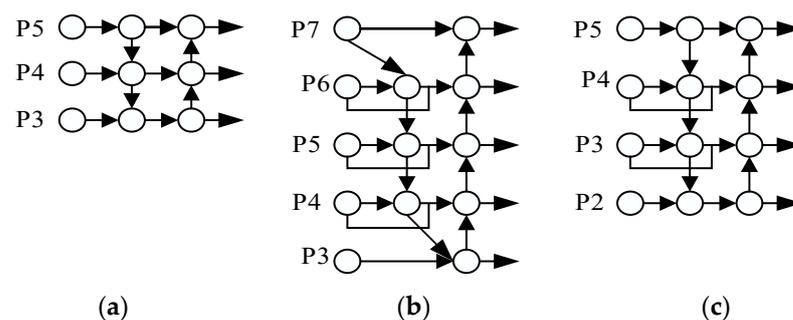


Figure 5. Feature fusion network design. (a) PANet, (b) BiFPN, and (c) Improved BiFPN.

This study applied enhancements to BiFPN by introducing cross-scale connections for increased feature fusion without substantially increasing computational costs. The modified network architecture, depicted in Figure 5c, exhibits slightly higher computational complexity but demonstrates excellent performance in small-object detection tasks. The essential advantage of this structure is its ability to preserve more shallow semantic information while maintaining relatively deep semantic information, thereby improving the network's overall perceptual capabilities.

5. Experimental Results and Analysis

5.1. Image and Label Databases

This paper adopts the Chinese Power Line Insulator Dataset (CPLID) [49] as the foundational data, which consists of 848 images of insulators. The dataset includes both images of insulators in normal conditions and those with defects. The original resolution of the images is 1728×1296 pixels. Due to the relatively sparse data volume, the limitations of the original data may restrict the accuracy and robustness of deep learning methods. Therefore, this study augmented the dataset to 2500 images by using data augmentation techniques such as geometric transformations, adding Gaussian noise, blackening arbitrary rectangular regions in the images, and adjusting brightness and contrast. Some examples of data augmentation are shown in Figure 6.

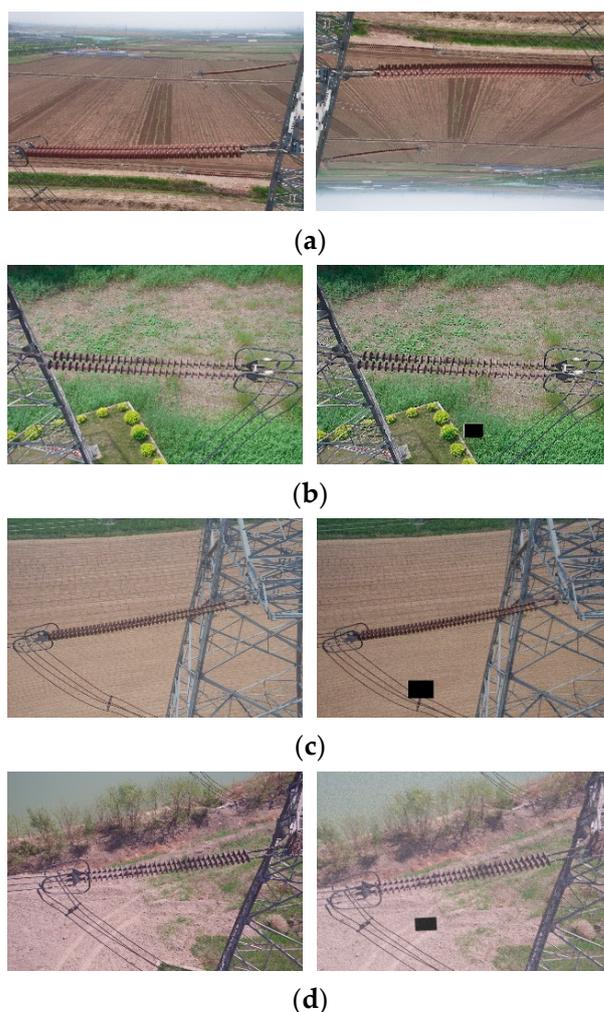


Figure 6. Some examples of data augmentation. (a) Geometric transformation, (b) blackening arbitrary rectangular regions, (c) blackening arbitrary rectangular regions and adjusting brightness, and (d) blackening out arbitrary rectangles and adding Gaussian noise.

We used the Labeling tool to annotate the actual boxes in the images, and the annotated objects were categorized as ordinary insulators and damaged insulators. Subsequently, the annotated insulator dataset was divided into training, validation, and test sets in a ratio of 7:2:1. The number and categories of labels in the dataset are shown in Figure 7.

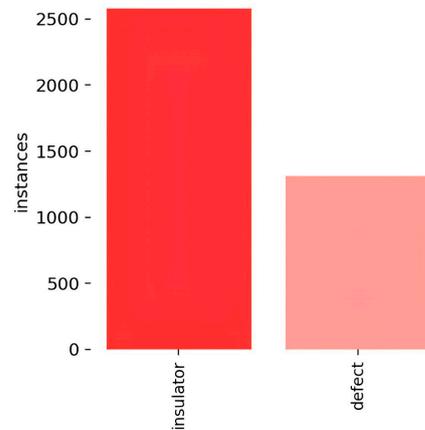


Figure 7. Labels and label distribution.

The sizes and numbers of images in the training set, validation set, and test set are shown in Table 1.

Table 1. Dataset partitioning.

Data	Image Size	Image Numbers
training set	1728 × 1296	1750
validation set	1728 × 1296	500
test set	1728 × 1296	250

5.2. Experimental Environment and Parameters

The software environment and hardware parameters used in the experimental process are presented in Table 2.

Table 2. Experimental environment configuration.

Parameter	Configuration
cpu	i7-11800H
gpu	GeForce RTX 3080 Laptop
python	3.8
pytorch	1.10
optimizer	SGD
momentum	0.937
Weight decay	0.0005
Batch size	16
initial learning rate	0.01
epochs	200

5.3. Experimental Results

Ablation experiments were conducted to validate the positive impact of the proposed improvement strategies on the network, and they were trained on the insulator dataset. The results are summarized in Table 3, where “√” indicates the adoption of the corresponding improvement method, and “×” indicates its absence. SOD stands for Small Object Detection layer, and IBiFPN represents the Improved BiFPN.

Table 3. Ablation experiment results.

Model	C2fSTR	SOD	IBiFPN	Precision/%	Recall/%	mAP@0.5/%
YOLOv8s	×	×	×	92.1	91.6	94.3
	✓	×	×	94.6	94.4	96.4
	×	×	✓	93.8	92.7	95.8
	×	✓	×	92.3	92.0	94.9
	✓	✓	✓	95.6	95.3	97.7

As shown in Table 3, we first used the YOLOv8s model as the base network and tested the effect of adding the C2fstr module. The results indicate an improvement of 2.5% in precision, a 2.8% in recall, and a 2.1% increase in mAP@0.5. This suggests that the C2fSTR module directs the network’s attention to local information in feature maps and enhances the model’s detection capability by adding the Swin Transformer module to increase the global information focus. In the second experiment, we compared the impact of adding the improved BiFPN. The results show that, compared to the base network, the precision improved by 1.7% and the recall by 1.1%, and mAP@0.5 increased by 1.5%. This indicates that the new feature fusion network can integrate more feature layers effectively. Moreover, in the feature fusion process, the network emphasizes input features with substantial contributions, enhancing its learning capability. The impact of adding the Small Object Detection Layer was compared in the third experiment. The results indicate that precision and recall increased by 0.3% and 0.4%, respectively, as compared to the base network, and mAP@0.5 improved by 0.6%. This suggests that the Small Object Detection Layer addition enhances the network’s capability to detect tiny defects, lowering the false-negative rate. Finally, experiments combining all three improvement methods were conducted. The results highlight that the combined improved algorithm performs the best, achieving a precision of 95.6%, a recall of 95.3%, and mAP@0.5 reaching 97.6%, meeting the accuracy requirements for insulator defect detection.

To affirm the superiority of our method, we compared it with YOLOv5s, Faster R-CNN, and YOLOv7, as depicted in Figure 8. The results reveal that our improved algorithm outperforms the other models in precision, recall, and mAP@0.5.

Table 4 compares various evaluation metrics for the different models, all based on the dataset constructed in this paper. As shown in Table 4, our proposed model demonstrates a 3.6% increase in F1 score and a 3.4% increase in mAP compared to YOLOv8s. Compared to the two-stage Faster R-CNN network, the F1 score increased by 2.7%, and mAP increased by 5.4%. Against YOLOv5s, the proposed network achieved a 1% increase in the F1 score and a 1.9% increase in mAP. In contrast to YOLOv7, the F1 score increased by 2.8%, and mAP increased by 4.2%. Regarding computational complexity, YOLOv5s leads with a performance of 105 FPS and has shorter training and inference times. Although our model has a lower FPS (88) and a longer training time (3 min 34 s), it achieves the highest accuracy of 97.7% in mAP@0.5, surpassing all other models.

Table 4. Comparison experiment of the different models.

Model	F1Score/%	mAP@0.5/%	Train Time/Epoch	Inference Time	GFLOPs	FPS
Faster R-CNN	92.7	92.3	5'48"	62.5 ms	370.21	16
YOLOv5s	94.4	95.8	2'09"	9.52 ms	16.0	105
YOLOv7	92.6	93.5	2'52"	10.2 ms	26.7	98
YOLOv8s	91.8	94.3	3'16"	10.86 ms	28.4	92
Ours	95.4	97.7	3'34"	11.36 ms	29.3	88

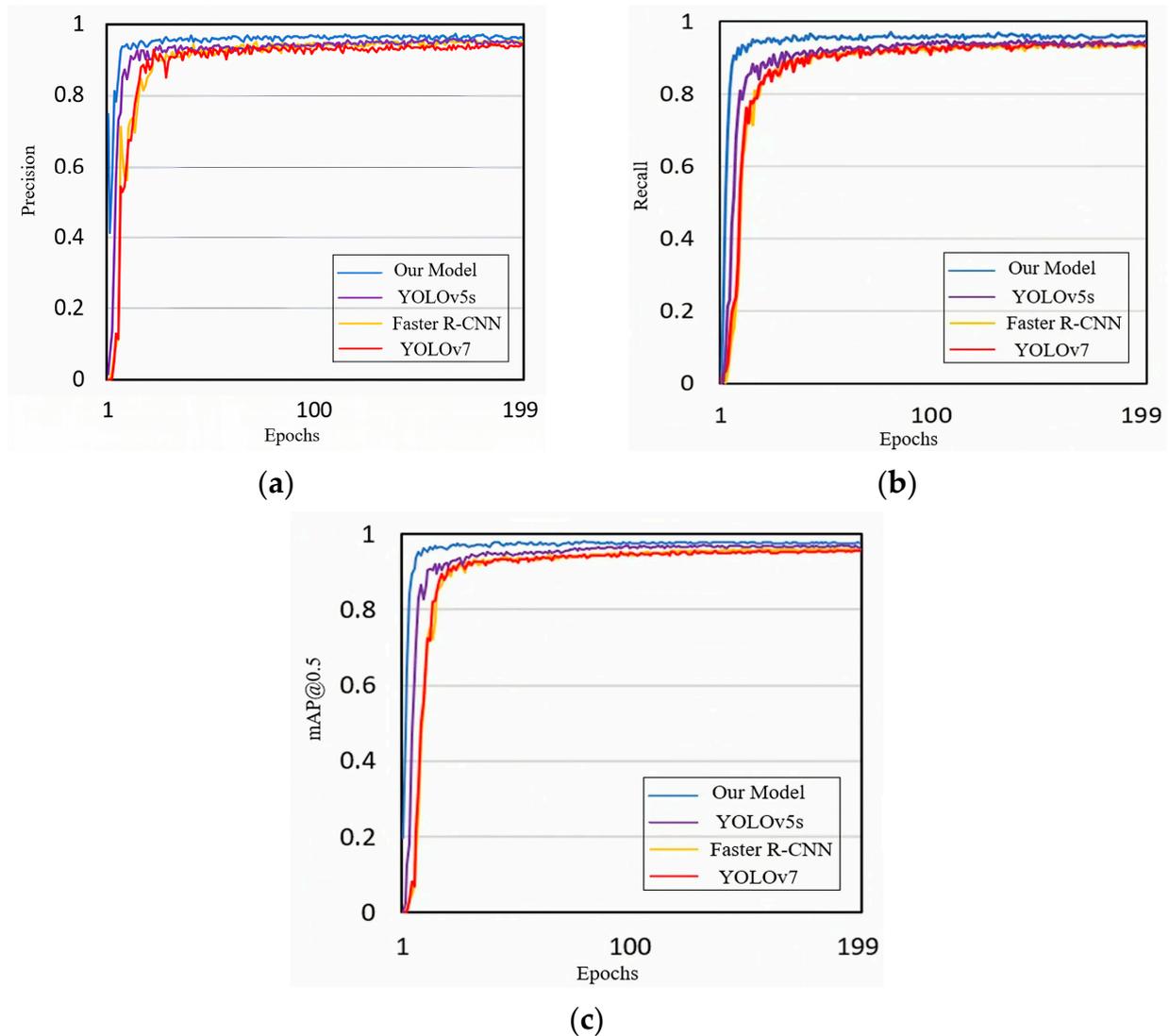


Figure 8. Comparison of training results for different models. (a) Precision, (b) recall, and (c) mAP@0.5.

In summary of the experiments above, the enhanced YOLOv8s algorithm proposed in this paper demonstrates a significant advantage in detection accuracy. The added C2fSTR module directs attention to local information in feature maps and emphasizes global information, enhancing the network's feature extraction capability. Including the Small Object Detection Layer and the improved feature fusion network enables the integration of more scale feature layers, improving small object detection and overall performance. Despite a slight decrease in speed, it still meets real-time requirements for practical engineering in insulator defect detection.

To validate the generalization ability and robustness, we specifically selected small objects and targets in complex environments for testing in the test set. The results, shown in Figure 9, highlight that the improved algorithm excels in identifying micro-defects in insulators and accurately recognizes insulator targets in challenging environments. The defect labels in the figure all indicate damaged defects.

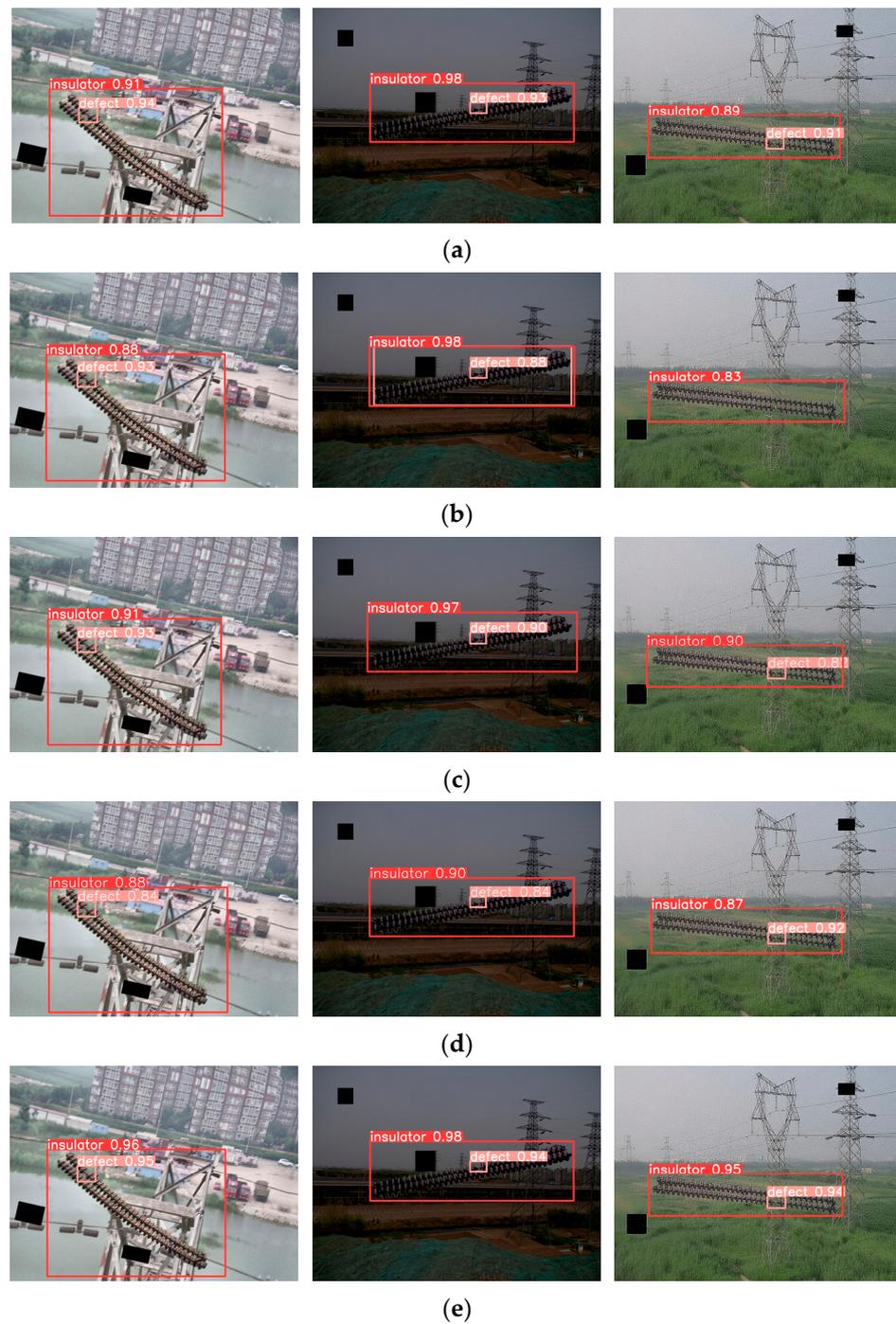


Figure 9. Detection results of different algorithms. (a) Faster R-CNN, (b) YOLOv5s, (c) YOLOv7, (d) YOLOv8s, and (e) ours.

6. Discussion

This paper proposes an insulator defect detection method using YOLOv8s-SwinT to accurately identify insulators and their defects in transmission line images. The experimental analysis shows that integrating the Swin Transformer module with the C2f module in YOLOv8s improves the effective utilization of local and global information in feature maps during feature extraction, resulting in richer semantic information at different levels and improved detection accuracy. Replacing the original FPN + PAN structure with the enhanced BiFPN structure and adding the Small Object Detection Layer strengthens the feature map representation and enhances micro-defect detection. Compared to the original

network, the improved model exhibits a 3.5% increase in precision, a 3.7% increase in recall, and a 3.4% improvement in mAP@0.5 with no significant decrease in detection speed, meeting real-time high-precision requirements. Compared to other detection models, this method shows significant advantages.

The algorithm proposed in this paper focuses solely on detecting damage defects in insulators. However, insulators may also suffer from other defects, such as cracks and contamination, which may limit the algorithm. Additionally, the deficiency of the algorithm lies in its inability to detect defects quantitatively, merely identifying damage without offering a detailed assessment of severity. This limitation hampers a comprehensive understanding of the health of insulation, thereby impeding accurate issue evaluation. Future research will integrate autonomous drone technology and expand the insulation image dataset to encompass various defects. Through this integration, we aim to conduct comprehensive insulation inspections, facilitating a more precise evaluation of its health status.

Author Contributions: Conceptualization, Z.H. and W.Y.; methodology, Y.L.; software, A.Z.; validation, J.L. and T.L.; formal analysis, J.Z.; investigation, Z.H. and W.Y.; resources, Z.H. and W.Y.; data curation, J.L., A.Z., and T.L.; writing—original draft preparation, W.Y.; writing—review and editing, Z.H.; visualization, J.L. and J.Z.; supervision, A.Z.; project administration, Z.H. and W.Y.; funding acquisition, Z.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Natural Science Foundation of China (No. 62293510, 62103076, 62003312) and the Key Science and Technology Program of Henan Province (No. 232102221032, 232102221011).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are contained within the article.

Conflicts of Interest: Author Yanjie Liu was employed by the company Sinohydro Bureau 3 Co., Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Palangar, M.F.; Mohseni, S.; Mirzaie, M.; Mahmoudi, A. Designing an automatic detector device to diagnose insulator state on overhead distribution lines. *IEEE Trans. Ind. Inform.* **2021**, *18*, 1072–1082. [\[CrossRef\]](#)
2. Zhang, X.; Zhang, Y.; Liu, J.; Zhang, C.; Xue, X.; Zhang, H.; Zhang, W. InsuDet: A fault detection method for insulators of overhead transmission lines using convolutional neural networks. *IEEE Trans. Instrum. Meas.* **2021**, *70*, 5018512. [\[CrossRef\]](#)
3. Tao, X.; Zhang, D.; Wang, Z.; Liu, X.; Zhang, H.; Xu, D. Detection of power line insulator defects using aerial images analyzed with convolutional neural networks. *IEEE Trans. Syst. Man Cybern. Syst.* **2018**, *50*, 1486–1498. [\[CrossRef\]](#)
4. She, L.; Fan, Y.; Xu, M.; Wang, J.; Xue, J.; Qu, J. Insulator breakage detection utilizing a convolutional neural network ensemble implemented with small sampled data augmentation and transfer learning. *IEEE Trans. Power Deliv.* **2021**, *37*, 2787–2796. [\[CrossRef\]](#)
5. Liu, J.; Hu, M.; Dong, J.; Lu, X. Summary of insulator defect detection based on deep learning. *Electr. Power Syst. Res.* **2023**, *224*, 109688. [\[CrossRef\]](#)
6. Wei, D.; Hu, B.; Shan, C.; Liu, H. Insulator defect detection based on improved Yolov5s. *Front. Earth Sci.* **2024**, *11*, 1337982. [\[CrossRef\]](#)
7. Singh, G.; Stefenon, S.F.; Yow, K.-C. Interpretable visual transmission lines inspections using pseudo-prototypical part network. *Mach. Vis. Appl.* **2023**, *34*, 41. [\[CrossRef\]](#)
8. Yin, J.; Wu, R.; Yang, J.; Xu, X.; Fan, C. An ultra-wideband microwave diagnostic method for detecting cracks in polymeric insulators. *IEEE Trans. Antennas Propag.* **2020**, *69*, 457–464. [\[CrossRef\]](#)
9. Mei, H.; Jiang, H.; Chen, J.; Yin, F.; Wang, L.; Farzaneh, M. Detection of internal defects of full-size composite insulators based on microwave technique. *IEEE Trans. Instrum. Meas.* **2021**, *70*, 6007710. [\[CrossRef\]](#)
10. Mei, H.; Jiang, H.; Yin, F.; Wang, L.; Farzaneh, M. Terahertz imaging method for composite insulator defects based on edge detection algorithm. *IEEE Trans. Instrum. Meas.* **2021**, *70*, 4504310. [\[CrossRef\]](#)
11. Waqas, A.; Cha, Y.J. Obstacle Avoidance Method for Autonomous UAV for Structural Health Monitoring. In *Society for Experimental Mechanics Annual Conference and Exposition*; Springer Nature Switzerland: Cham, Switzerland, 2023; pp. 149–154.
12. Ali, R.; Kang, D.; Suh, G.; Cha, Y.-J. Real-time multiple damage mapping using autonomous UAV and deep faster region-based neural networks for GPS-denied structures. *Autom. Constr.* **2021**, *130*, 103831. [\[CrossRef\]](#)

13. Song, Z.; Huang, X.; Ji, C.; Zhang, Y. Deformable YOLOX: Detection and Rust Warning Method of Transmission Line Connection Fittings Based on Image Processing Technology. *IEEE Trans. Instrum. Meas.* **2023**, *72*, 2504321. [[CrossRef](#)]
14. Wei, N.; Li, X.; Jin, J.; Chen, P.; Sun, S. Detecting insulator strings as linked chain structure in smart grid inspection. *IEEE Trans. Ind. Inform.* **2022**, *19*, 9019–9027. [[CrossRef](#)]
15. Li, J.; Xu, Y.; Nie, K.; Cao, B.; Zuo, S.; Zhu, J. PEDNet: A Lightweight Detection Network of Power Equipment in Infrared Image Based on YOLOv4-Tiny. *IEEE Trans. Instrum. Meas.* **2023**, *72*, 5004312. [[CrossRef](#)]
16. Lu, X.; Quan, W.; Gao, S.; Zhao, H.; Zhang, G.; Lin, G.; Chen, J.X. An Outdoor Support Insulator Surface Defects Segmentation Approach via Image Adversarial Reconstruction in High-Speed Railway Traction Substation. *IEEE Trans. Instrum. Meas.* **2022**, *71*, 5023619. [[CrossRef](#)]
17. Cha, Y.J.; Choi, W.; Suh, G.; Mahmoudkhani, S.; Büyüköztürk, O. Autonomous structural visual inspection using region-based deep learning for detecting multiple damage types. *Comput.-Aided Civ. Infrastruct. Eng.* **2018**, *33*, 731–747. [[CrossRef](#)]
18. Wang, Y.; Wang, H.; Xin, Z. Efficient detection model of steel strip surface defects based on YOLO-V7. *IEEE Access* **2022**, *10*, 133936–133944. [[CrossRef](#)]
19. Tzelepakis, A.; Leontaris, L.; Dimitriou, N.; Koukidou, E.; Bollas, D.; Karamanidis, A.; Tzovaras, D. Automated Defect Detection In Battery Line Assembly Via Deep Learning Analysis. In Proceedings of the 10th ECCOMAS Thematic Conference on Smart Structures and Materials, Patras, Greece, 3–5 July 2023.
20. Yang, Z.; Xu, Z.; Wang, Y. Bidirection-fusion-YOLOv3: An improved method for insulator defect detection using UAV image. *IEEE Trans. Instrum. Meas.* **2022**, *71*, 3521408. [[CrossRef](#)]
21. Zhao, W.; Xu, M.; Cheng, X.; Zhao, Z. An Insulator in Transmission lines recognition and fault detection model based on improved faster RCNN. *IEEE Trans. Instrum. Meas.* **2021**, *70*, 5016408. [[CrossRef](#)]
22. Hu, Y.; Wen, B.; Ye, Y.; Yang, C. Multi-Defect Detection Network for High-Voltage Insulators Based on Adaptive Multi-Attention Fusion. *Appl. Sci.* **2023**, *13*, 13351. [[CrossRef](#)]
23. Zhang, D.; Gao, S.; Yu, L.; Kang, G.; Wei, X.; Zhan, D. DefGAN: Defect detection gans with latent space pitting for high-speed railway insulator. *IEEE Trans. Instrum. Meas.* **2020**, *70*, 5003810. [[CrossRef](#)]
24. Zhong, J.; Liu, Z.; Yang, C.; Wang, H.; Gao, S.; Nunez, A. Adversarial reconstruction based on tighter oriented localization for catenary insulator defect detection in high-speed railways. *IEEE Trans. Intell. Transp. Syst.* **2020**, *23*, 1109–1120. [[CrossRef](#)]
25. Zhang, Q.; Zhang, J.; Li, Y.; Zhu, C.; Wang, G. IL-YOLO: An Efficient Detection Algorithm for Insulator Defects in Complex Backgrounds of Transmission Lines. *IEEE Access* **2024**, *12*, 1. [[CrossRef](#)]
26. Fu, Q.; Liu, J.; Zhang, X.; Zhang, Y.; Ou, Y.; Jiao, R.; Li, C.; Mazzanti, G. A Small-Sized Defect Detection Method for Overhead Transmission Lines Based on Convolutional Neural Networks. *IEEE Trans. Instrum. Meas.* **2023**, *72*, 3524612. [[CrossRef](#)]
27. Li, K.; Wang, Y.; Hu, Z. Improved YOLOv7 for Small Object Detection Algorithm Based on Attention and Dynamic Convolution. *Appl. Sci.* **2023**, *13*, 9316. [[CrossRef](#)]
28. Shuang, F.; Wei, S.; Li, Y.; Gu, X.; Lu, Z. Detail R-CNN: Insulator Detection based on Detail Feature Enhancement and Metric Learning. *IEEE Trans. Instrum. Meas.* **2023**, *72*, 2524414. [[CrossRef](#)]
29. Cha, Y.J.; Choi, W.; Büyüköztürk, O. Deep learning-based crack damage detection using convolutional neural networks. *Comput. Aided Civ. Infrastruct. Eng.* **2017**, *32*, 361–378. [[CrossRef](#)]
30. Liu, Z.; Wu, G.; He, W.; Fan, F.; Ye, X. Key target and defect detection of high-voltage power transmission lines with deep learning. *Int. J. Electr. Power Energy Syst.* **2022**, *142*, 108277. [[CrossRef](#)]
31. Yu, Z.; Lei, Y.; Shen, F.; Zhou, S.; Yuan, Y. Research on Identification and Detection of Transmission Line Insulator Defects Based on a Lightweight YOLOv5 Network. *Remote Sens.* **2023**, *15*, 4552. [[CrossRef](#)]
32. Zheng, J.; Wu, H.; Zhang, H.; Wang, Z.; Xu, W. Insulator-defect detection algorithm based on improved YOLOv7. *Sensors* **2022**, *22*, 8801. [[CrossRef](#)]
33. Hou, Q.; Zhou, D.; Feng, J. Coordinate attention for efficient mobile network design. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Nashville, TN, USA, 20–25 June 2021; pp. 13713–13722.
34. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention is all you need. In Proceedings of the Annual Conference on Neural Information Processing Systems 2017, Long Beach, CA, USA, 4–9 December 2017.
35. Qin, C.J.; Wu, R.H.; Huang, G.Q.; Tao, J.F.; Liu, C.L. A novel LSTM-autoencoder and enhanced transformer-based detection method for shield machine cutterhead clogging. *Sci. China Technol. Sci.* **2023**, *66*, 512–527. [[CrossRef](#)]
36. Dosovitskiy, A.; Beyer, L.; Kolesnikov, A.; Weissenborn, D.; Zhai, X.; Unterthiner, T.; Dehghani, M.; Minderer, M.; Heigold, G.; Gelly, S.; et al. An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale. *arXiv* **2020**, arXiv:2010.11929.
37. Dian, S.; Zhong, X.; Zhong, Y. Faster R-Transformer: An efficient method for insulator detection in complex aerial environments. *Measurement* **2022**, *199*, 111238. [[CrossRef](#)]
38. Dai, Y.; Liu, W.; Wang, H.; Xie, W.; Long, K. YOLO-Former: Marrying YOLO and transformer for foreign object detection. *IEEE Trans. Instrum. Meas.* **2022**, *71*, 5026114. [[CrossRef](#)]
39. Liu, Z.; Lin, Y.; Cao, Y.; Hu, H.; Wei, Y.; Zhang, Z.; Lin, S.; Guo, B. Swin transformer: Hierarchical vision transformer using shifted windows. In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), Montreal, BC, Canada, 11–17 October 2021; pp. 10012–10022.

40. Liu, C.; Ma, L.; Sui, X.; Guo, N.; Yang, F.; Yang, X.; Huang, Y.; Wang, X. YOLO-CSM-Based Component Defect and Foreign Object Detection in Overhead Transmission Lines. *Electronics* **2023**, *13*, 123. [[CrossRef](#)]
41. Zhao, T.; Cao, J.; Hao, Q.; Bao, C.; Shi, M. Res-SwinTransformer with Local Contrast Attention for Infrared Small Target Detection. *Remote Sens.* **2023**, *15*, 4387. [[CrossRef](#)]
42. Choi, W.; Cha, Y.J. SDDNet: Real-time crack segmentation. *IEEE Trans. Ind. Electron.* **2019**, *67*, 8016–8025. [[CrossRef](#)]
43. Kang, D.H.; Cha, Y.J. Efficient attention-based deep encoder and decoder for automatic crack segmentation. *Struct. Health Monit.* **2022**, *21*, 2190–2205. [[CrossRef](#)] [[PubMed](#)]
44. Ali, R.; Cha, Y.J. Attention-based generative adversarial network with internal damage segmentation using thermography. *Autom. Constr.* **2022**, *141*, 104412. [[CrossRef](#)]
45. Wang, B.; Chen, S.; Wang, J.; Hu, X. Residual feature pyramid networks for salient object detection. *Vis. Comput.* **2020**, *36*, 1897–1908. [[CrossRef](#)]
46. Yang, J.; Fu, X.; Hu, Y.; Huang, Y.; Ding, X.; Paisley, J. PanNet: A deep network architecture for pan-sharpening. In Proceedings of the IEEE International Conference on Computer Vision, Vancouver, BC, Canada, 7–14 July 2001; pp. 5449–5457.
47. Tan, M.; Pang, R.; Le, Q.V. Efficientdet: Scalable and efficient object detection. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 18–24 June 2022; pp. 10781–10790.
48. Liu, Z.; Sun, L.; Zhang, Q. High Similarity Image Recognition and Classification Algorithm Based on Convolutional Neural Network. *Comput. Intell. Neurosci.* **2022**, *2022*, 2836486. [[CrossRef](#)] [[PubMed](#)]
49. Fawzi, A.; Samulowitz, H.; Turaga, D.; Frossard, P. Adaptive data augmentation for image classification. In Proceedings of the IEEE International Conference on Image Processing, Phoenix, AZ, USA, 25–28 September 2016.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.