

Article

A Few-Shot Defect Detection Method for Transmission Lines Based on Meta-Attention and Feature Reconstruction

Yundong Shi ¹, Huimin Wang ², Chao Jing ^{2,3} and Xingzhong Zhang ^{1,2,*}

¹ College of Software, Taiyuan University of Technology, Taiyuan 030024, China

² Shanxi Energy Internet Research Institute, Taiyuan 030024, China

³ College of Software, Shanxi Agricultural University, Jinzhong 030801, China

* Correspondence: zhangxingzhong@tyut.edu.cn

Abstract: In tasks of transmission line defect detection, traditional object detection algorithms are ineffective, with few training samples of defective components. Meta-learning uses multi-task learning as well as fine-tuning to learn common features in different tasks, which has the ability to adapt to new tasks quickly, shows good performance in few-shot object detection, and has good generalization in new tasks. For this reason, we proposed a few-shot defect detection method (Meta PowerNet) with a Meta-attention RPN and Feature Reconstruction Module for transmission lines based on meta-learning. First, in the stage of region proposal, a new region proposal network (Meta-Attention Region Proposal Network, MA-RPN) is designed to fuse the support set features and the query set features to filter the noise in anchor boxes. In addition, it has the ability to focus on the subtle texture features of smaller-sized objects by fusing low-level features from the query set. Second, in the meta-feature construction stage, we designed a meta-learner with the defect feature reconstruction module as the core to capture and focus on the defect-related feature channels. The experimental results show that under the condition, there are only 30 training objects for various types of component defects. The method achieves 72.5% detection accuracy for component defects, which is a significant improvement compared with other mainstream few-shot object detection. Meanwhile, the MA-RPN designed in this paper can be used in other meta-learning object detection models universally.

Keywords: few-shot learning; meta-attention; transmission lines; defects detection; feature regeneration



Citation: Shi, Y.; Wang, H.; Jing, C.; Zhang, X. A Few-Shot Defect Detection Method for Transmission Lines Based on Meta-Attention and Feature Reconstruction. *Appl. Sci.* **2023**, *13*, 5896. <https://doi.org/10.3390/app13105896>

Academic Editor: João M. F. Rodrigues

Received: 6 April 2023

Revised: 6 May 2023

Accepted: 9 May 2023

Published: 10 May 2023



Copyright: © 2023 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/>).

1. Introduction

As a necessary means to maintain the long-term stable transmission of electric power, the UAV (Unmanned Aerial Vehicle) inspection technology equipped with deep learning object detection has been effective in improving the automation of transmission line inspection, which is based on a large number of inspection images [1–3]. However, due to various factors such as temperature, weather, and human fallibility, there are many types of power defects, especially the scarcity of defect samples of key components such as insulators, vibration damper, and pins, resulting in low detection precision of these kinds of defects by object detection algorithms based on deep learning, which poses great challenges for the pilotless inspection of transmission lines [4].

Few-shot learning is a method to solve a problem with few training samples in the target domain, which can alleviate problems such as overfitting and the inability to converge deep-learning models in the case of few samples [5] and mainly contains metric learning, data augmentation, fine-tuning, meta-learning, and other methods [6]. Among them, data augmentation is the most commonly used technology for dealing with few-shot problems; for example, Bayraktar et al. [7] proposed a method to optimize the route of the target ground vehicle benefit from some geometric transformations, effectively expanding the sample and enhancing the segmentation ability of the model. In addition, [8]

presents a method to improve the performance of surface defect detection by combining GAN (Generative Adversarial Networks) [9] with classical methods. The synthetic samples generated by the GANs are then used to train a deep convolutional neural network (CNN) to perform image segmentation and detect defects. In terms of metric learning, Hsieh et al. [10] proposed a metric-based method with margin-ranking loss, which calculates the similarity between the features of region proposal and the features of ground truth, assisting the model in classifying few-shot objects. Li et al. [11] proposed the Deep Nearest Neighbor Neural Network (DN4) that replaces image-level feature measurements with local descriptors from images to classes in the final layer. When mapping features for query samples, DN4 calculates a similarity for each spatial feature and sums up all spatial similarities to obtain the similarity of the query sample. In terms of fine-tuning, Wang et al. [12] improved Faster R-CNN [13] by first training with base-class data, and in the fine-tuning stage, they froze the parameter weights of the early network and fine-tuned the top-level classifier and regressor with a balanced subset composed of base and novel classes. Sun et al. [14] trained the base-class data first, then fine-tuned it with a small amount of mixed data from base and novel classes to achieve few-shot object detection. In terms of meta-learning, Ye et al. [15] combined the adaptive algorithm Transformer [16] with the Set-to-Set function to map instances to the space of the corresponding task and construct different internal correlations within the set, enabling the model to perform specific learning for the object. Zhang et al. [17] proposed a meta-learning-based search and decoding method and introduced the Meta Navigator framework to find effective parameter adaptation strategies and apply them to different stages of the model to achieve few-shot classification.

At present, there are few studies on few-shot object detection for transmission lines, and the method of data augmentation is usually chosen to solve the problem. For example, Xu et al. [18] used GAN to expand the samples of small fittings, and Cui et al. [19] supplemented insulator defect samples by improving Cyclic-Consistent Generative Adversarial Networks (CycleGAN) [20]. However, such methods based on virtual sample construction have limited improvement in the detection precision of the classes with few samples and require high computing power and data quality in the source domain, as well as poor generalization, and can only perform data enhancement for a certain type of component defects, which requires repeated data expansion and model training operations when new few-shot detection tasks are added in subsequent inspection work. Choosing metric learning to measure different components is easy to operate, but the global features of defective components and normal components are more similar, resulting in metric learning that cannot solve the core problem of defect detection. Fine-tuning is suitable for detection tasks in which the base class is similar to the novel class. Zhang et al. [21] conducted pre-training on the YOLOv3 [22] model using the publicly available ImageNet dataset. Subsequently, they fine-tuned the model using a small number of vibration damper and wire clamp samples to achieve few-shot detection. On the other hand, Zhai et al. [23] pre-trained Faster R-CNN on artificial samples and fine-tuned it using insulator defects. Ultimately, they achieved a detection accuracy of 62.7% mAP for insulator defects with a real sample size of 184. However, the generalization of fine-tuning is poor and will lead to a decrease in detection precision when the type of components to be detected is changed. In contrast, by dividing the training data into several different small tasks to train the model, meta-learning not only improves the generalizability of the model under few sample conditions but also the training strategy of using fine-tuning for novel classes is applicable to defect detection tasks in which normal components are similar to defective components.

Our goal is meta-learning-based object detection for improving the efficiency of transmission line defect detection tasks under insufficient sample size conditions by combining meta-learning with defect detection. Specifically, we propose a meta-learning object detection algorithm (Meta PowerNet) for few-shot defect detection of transmission lines. First, a meta-attention region proposal network is designed to use a support set with prior

knowledge and low-level features with detailed information of the query set to improve the quality of the anchor boxes generated by the region's proposal network. Second, for the characteristics of transmission line defect detection, a defect feature reconstruction module is designed to reconstruct the support set meta-features according to normal and defective components, which not only reduces the computation of fusion meta-features of the support set and ROI features of the query set but also improves the fine-tuning effect of the detection head in the stage of meta-test. Third, this method can detect seven classes, including an insulator, vibration damper, and pin and their defects. Finally, the Meta-Attention Region Proposal Network can be embedded in other meta-learning object detection frameworks to effectively improve the detection precision.

The rest of this study is organized as follows: Section 2 discusses the related works on meta-learning and meta-learning object detection. Section 3 explains the details for each module. In Section 4, the implementation is explained, including the dataset; experimental parameters; evaluation indicators; and the results of ablation, comparison, and expanding the experiment. Finally, Section 5 concludes the paper and puts forward prospects for future work.

2. Related Works

2.1. Meta-Learning

Meta-learning, also known as learning to learn, is a method to learn “meta-knowledge” between different tasks through multi-task and cross-task training, which is a kind of a priori knowledge applicable to different tasks in order to improve the learning ability for new tasks in the future. This is one of the common approaches to solve the problem of having few samples (few-shot learning) [24]. The essence of meta-learning is to increase the generalization ability of the learner in multi-tasking, to train iteratively on a task-by-task basis, and to learn a set of initialization parameters that could perform well in different tasks so that only a few iterations are needed for a specific task to obtain better performance in a new task. The structure is shown in Figure 1.

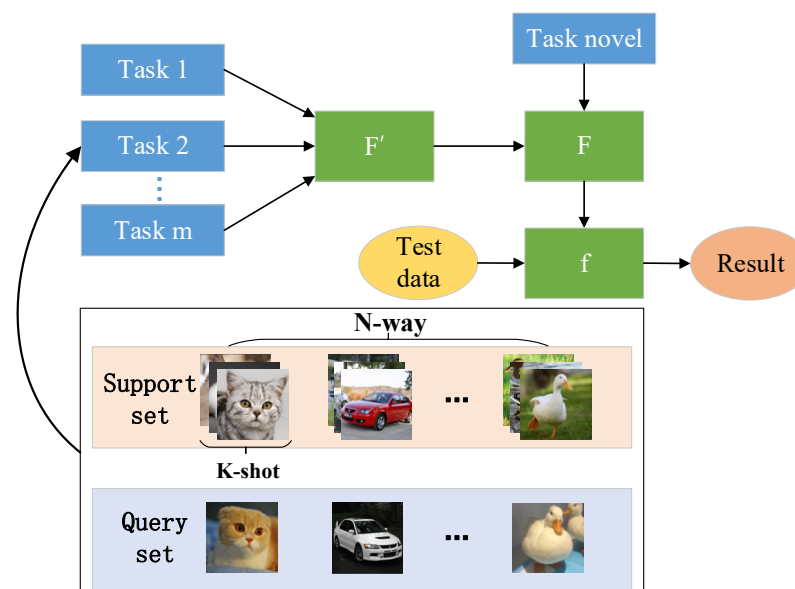


Figure 1. Architecture of meta-learning.

Meta-learning uses a sufficient amount of base class data and a small amount of novel class data as training samples. Training is divided into two stages, where the first stage (meta-training) is training on the base class data to learn a model with prior knowledge. In the second stage (meta-test), the base class data is mixed with the novel class data for further training, based on the model learned in meta-training, to learn the final model for the novel class. This not only speeds up the convergence of novel class learning but

also alleviates the catastrophic forgetting of the base class by the model. The two stages of training are divided into multiple tasks, each of which divides the data into support set data and query set data, with the support set containing N classes and each class containing K samples, referred to as the N -Way- K -Shot task. The model is tested on the query set data by learning the features of the support set data to verify the learning effect.

2.2. Meta Learning Object Detection

Meta-learning object detection builds on top of meta-learning classification and adds the task of object localization, needed to learn more effectively the rich prior knowledge from base classes. Based on meta-learning, few-shot object detection methods design meta-learners to train models with sufficient base-class samples during meta-training, learning universal meta-knowledge across tasks. During meta-testing, learned meta-knowledge is used to predict parameters based on class-related parameters of small-sample images.

Kang et al. [25] proposed a one-stage meta-learning object detection method based on YOLOv2 [26]. They converted the support set image of each class into an n -dimensional vector using convolution and used this to perform channel-wise reweighting of the feature of the query set, which was then predicted by the predictor head. In one-stage meta-learning object detection, the anchor box coordinates in the regression process were generated directly from the query set image features, but due to the small number of training samples, the regression process did not fully converge, and foreground and background objects were easily confused. Yan et al. [27] proposed Meta R-CNN with a meta-learner called PRN (Predictor-head Remodeling Network) based on Faster R-CNN, which constructed a class attention vector and fused it with ROI features of the query set to perform object detection. Xiao et al. [28] improved Meta R-CNN by introducing feature subtraction to enhance the fusion of support set class attention vectors and query features. Zhang et al. [29] combined the transformer with meta-learning to propose Meta-DETR, which learns class-agnostic meta-knowledge using an encoder–decoder architecture inspired by Deformable DETR [30]. Zhang et al. [31] proposed Meta Faster R-CNN, which utilizes the original Faster R-CNN for base class detection and designs Meta-RPN, a coarse-grained prototype matching network, and Meta-Classifier, a fine-grained prototype matching network for new class detection, generating anchors for new class objects and performing classification and regression.

Currently, meta-learning-based few-shot object detection performs well on public datasets dominated by natural scenes, but there is still a lot of room for improvement in defect detection. For example, in defect detection, the global similarity between normal and defective parts is high, which can prevent generic meta-learning methods from fully utilizing defect features for detection. In this paper, we propose a meta-learning-based object detection method specifically for defect detection to improve the detection precision of defective components in power transmission lines.

3. Method

3.1. Overall Architecture

The key to applying meta-learning to defect detection is how to use a small number of defect samples to make the model learn the differences between normal components and defective components. Therefore, our method focuses on learning the features of defective components and the differences between them and normal components. Considering the regression problem in meta-learning object detection, a two-stage object detection method is adopted in our method, detecting the targets from coarse to fine granularity.

The framework of the Meta PowerNet model is shown in Figure 2, which mainly consists of four parts: feature extraction network, region proposal module, meta-learner, and predictor head. The feature extraction network is responsible for extracting features from the support set and the query set. The region proposal module, with MA-RPN (Meta-Attention Region Proposal Network) as its core, generates proposal regions of query images with fusing support set features, which could reduce noise in proposal regions. The meta-learner is used to learn the meta-features of each class in the support set as prior

information to assist the model in detecting the target in the query set. In addition, a defect feature reconstruction module is designed to enhance the feature representation of defects in meta-features. The predictor head is responsible for classifying and regressing the proposal region features of the query set to generate the final detection results. In addition, the normal components with sufficient samples are selected as base classes, while the defect components with few samples are treated as novel classes.

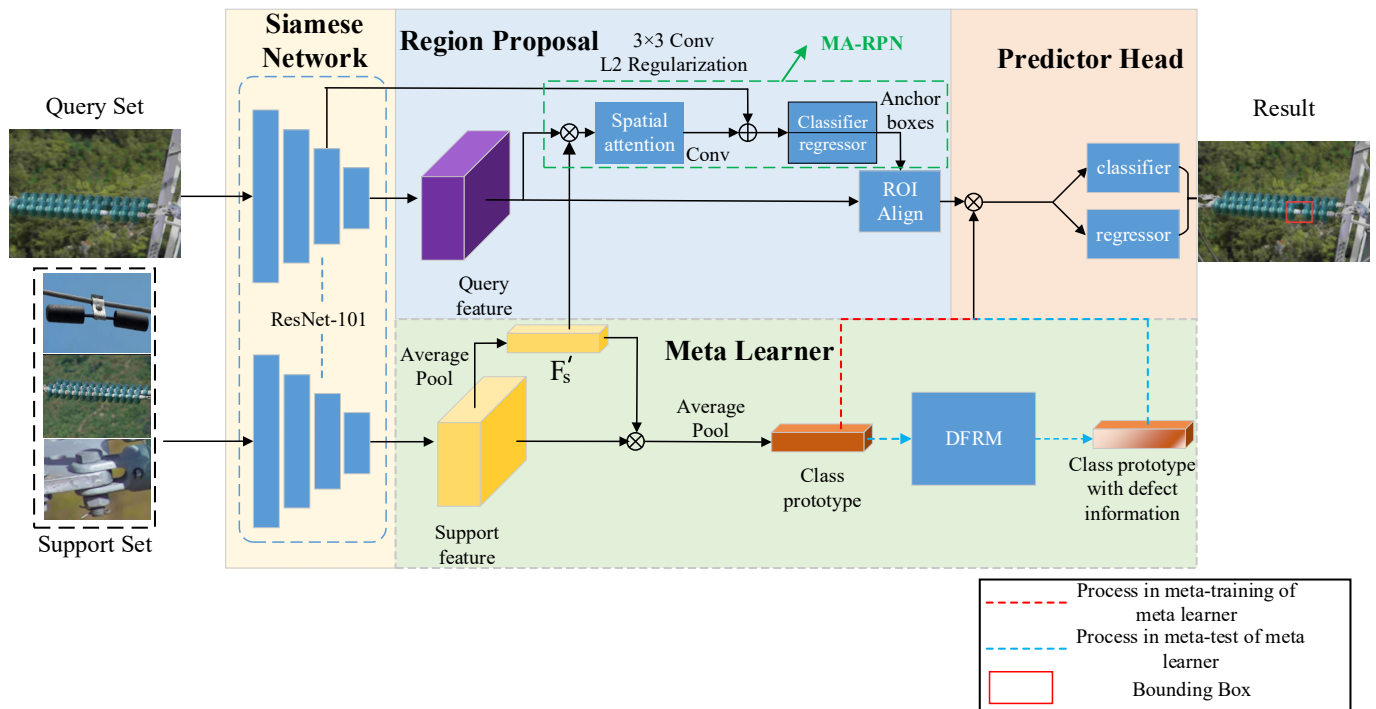


Figure 2. Network architecture of Meta PowerNet.

3.2. Feature Extraction Network

In this paper, we adopt the Siamese Neural Network [32] as the backbone to extract features from the query and support sets. This method is a commonly used approach for meta-learning-based object detection [25,27,31]. The input image of the query set is the original image that has been resized to a uniform size. After extracting the features using the ResNet101 backbone network, the image features at the res4 stage are used as the output.

As another branch of the Siamese neural network, the backbone network for extracting support set features also uses ResNet101 and shares weights with the backbone network for the query set. Moreover, since the background information in the support set images is useless for the model training, we crop the target object in support image with 16-pixel image context, zero-padded, and then resized to a square image of 320×320 .

3.3. Region Proposal

Region Proposal Network (RPN) is mainly used to generate anchor boxes for the predictor head and to perform foreground–background classification on these boxes. However, in the meta-test of meta-learning object detection, the training mode of RPN is fine-tuned based on the parameters learned in meta-training. Due to the small number of samples, RPN has limited regression ability for the new class targets in this way, resulting in low-quality anchor boxes for new classes and affecting the detection of new class targets by the predictor head.

We propose the Meta-Attention Region Proposal Network (MA-RPN), which utilizes support set features to enhance the expression ability of target classes in query set features and increases the model's attention to foreground regions through spatial attention mecha-

nism. Finally, by combining image details, the module could generate higher quality and more effective anchor boxes.

As shown in Figure 3, in order to activate the important feature representation of the support set classes in query features, the global feature of the support set, which is obtained through global average pooling with support feature f_s^i , is fused with the query feature F_4 by MA-RPN. The specific method to obtain the global feature of support set will be introduced in the next section. More specifically, we use depth-wise cross-correlation [33] as the fusing method, the global feature of support set is used as a convolutional kernel to reweight the query set features F_4 , and the reweighted query feature F_5 can guide MA-RPN to produce anchor boxes relevant to classes of support set. Fusing process is shown in Equation (1).

$$F_5 = \text{GAP}(f_s^i) \otimes F_4 \quad (1)$$

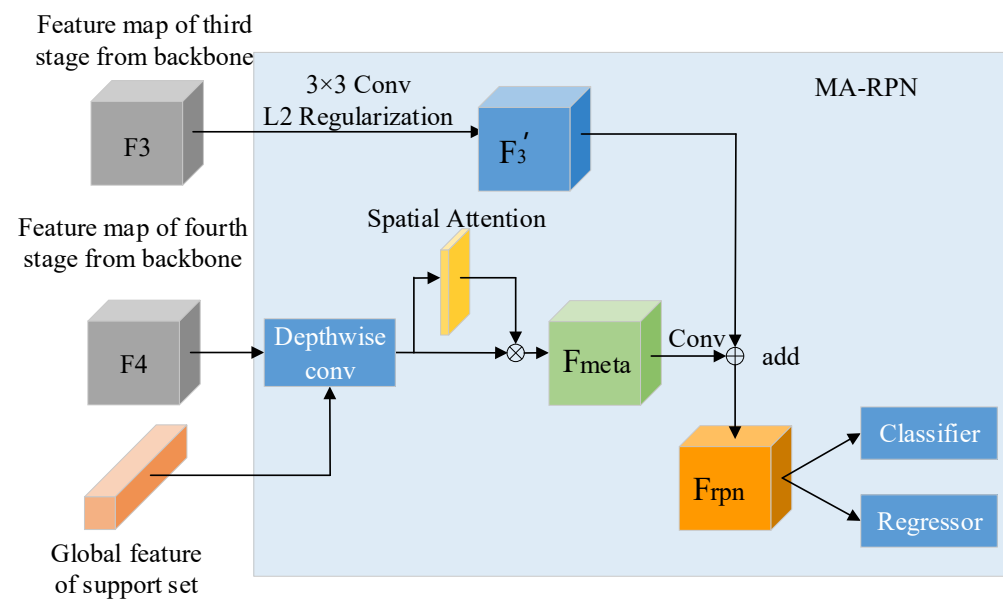


Figure 3. Architecture of MA-RPN.

Subsequently, MA-RPN trains a spatial weight matrix to weight the reweighted query features through spatial attention mechanism, highlighting the position representation of foreground objects in the features; we mark it as F_5' . Unlike traditional spatial attention, the input features here are the reweighted query features, which not only enhance the region proposal ability in meta-training but also carry more effective new-class features to participate in the spatial attention in meta-test, which can effectively improve the efficiency of fine-tuning and generate more accurate anchor boxes for new-class objects.

In addition, in order to reduce the loss of detailed information in the feature extraction network and improve the accuracy in predicting positions of anchor boxes, we also involve the low-level feature in the generation process of anchor boxes. First, we feed the query set feature F_3 , which carries more detailed information from the third level of backbone to a convolutional layer, which has a convolution kernel with a size of 3×3 , to adjust its dimension and size. After the L_2 regularization process, the new feature F_3' is output. Then, the feature F_5 is fused with it, so that the model can generate more effective candidate boxes by combining high-level semantic information with low-level detail information. The fusion method used here is inspired by the Feature Pyramid Network, where features F_5 and F_3' are added layer by layer to obtain feature F_{RPN} , which is used as the feature layer for anchor boxes extraction. The formula is shown in Equation (2). Subsequently, anchor boxes are classified as foreground or background, and their bounding boxes are regressed

through the same operations as in the original RPN. Finally, a specified number and size of anchor boxes are output and mapped to feature F_4 via ROI Align.

$$F_{RPN} = \text{conv}(F_5) + L_2(F_3) \quad (2)$$

3.4. Meta-Learner

As the most important module in the meta-learning-based object detection model, the meta-learner generates class prototypes as meta-features for each target class by extracting sufficient support set features to assist the model in classifying ROI features of query set. However, the overall features of the defective components of the power transmission line are similar to those of the normal components. The overall features of normal components have already been learned in meta-training. However, treating novel classes as completely new classes unrelated to base class in meta-test not only results in learning insufficiently the features of defective parts but also adds unnecessary redundancy to the predictor head.

We designed a meta-learner with the Defect Feature Reconstruction Module (DFRM) as its core according to the characteristics of defect detection in power transmission lines. By learning the channel differences between normal and defective components, the meta-learner reconstructs the meta-features to improve the attention of the predictor head to the channel relevant to defect class in the query set ROI features.

As shown in Figure 4, we perform global average pooling on the support set features to obtain a one-dimensional vector f_s^i , which is used as channel-level weights to reweight the original support set feature and enhance the representation of potential important features. Then, the output features are fed into the global average pooling layer to be compressed into one-dimensional features as meta-features, which will conduct channel-wise dot product with the query set ROI features in order to achieve channel feature selection. The one-dimensional vector f_s^i is the support set one-dimensional vector mentioned in the previous section, and the reason for choosing global average pooling to squeeze features is that this operation can retain more class-related information. To prevent confusion between meta-features of different classes, which would lead to ambiguity in classification task of predictor head, we adopted the Meta Loss from Meta R-CNN. The meta-features fed into a fully connected network for classification and adjusted the weights of the meta-learner through the feedback from loss function to enhance the distinguishability between different meta-features.

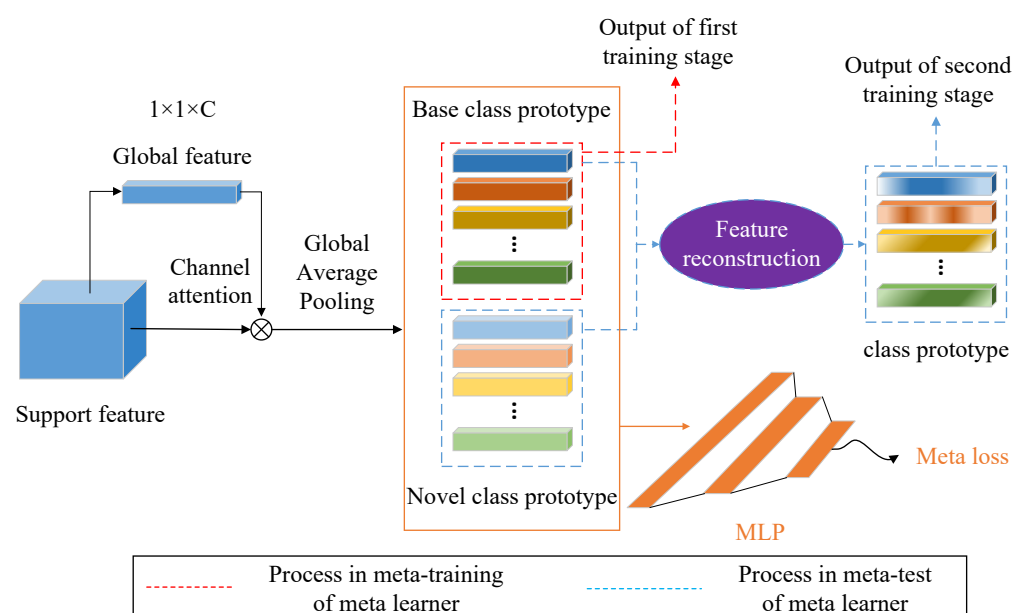


Figure 4. Architecture of Meta-Learner.

To improve the detection of differences between normal and defective components in the same category, the differences mainly exist in a few channels of image features. In the meta-testing phase, we proposed the Defect Feature Reconstruction Module (DFRM) to calculate the difference values between the corresponding meta-features of normal and defective components on each channel. The DFRM multiplies the vector of difference values with the meta-feature of normal components to enhance the values corresponding to channels with different values that are large. After fusing with the ROI features of query set, the predictor head can pay more attention to the channels related to defects.

First, the element-wise subtraction between normal and defective meta-features is computed, followed by taking absolute value to obtain the initial difference values. Then, through standardization and *ReLU* as activation function, the small difference values are suppressed to zero as noise, which could reduce their impact on distinguishing the defect-related features in predictor head. The formula is shown in Equation (3), where F_B represents the meta-feature of base class, F_N refers to the meta-feature of novel class, and F_R represents the reconstructed meta-feature.

$$F_R = \text{ReLU}(\text{Stand}(|F_B - F_N|)) \times F_B + F_B \quad (3)$$

3.5. Predictor Head

We adopt the most commonly used predictor head in current two-stage object detection algorithms, which is the predictor head in Faster R-CNN. It consists of a classification branch and a regression branch, and both branches process ROI features through fully connected layer with activation functions to output the classification scores and bounding box coordinates of the object regions in query set image. The loss function for the classification is cross-entropy loss, and Smooth L1 loss is the loss function of regression branches. Finally, the total loss function of our model consists of classification loss $Loss_{R\ cls}$ and regression loss $Loss_{R\ reg}$ of RPN, classification loss $Loss_{P\ cls}$ and regression loss $Loss_{P\ reg}$ of predictor head, and meta-loss $Loss_{meta}$, as shown in Equation (4).

$$oss = Loss_{R\ cls} + Loss_{R\ reg} + Loss_{P\ cls} + Loss_{P\ reg} + Loss_{meta} \quad (4)$$

Due to the high similarity between the normal components and the defect components, meta-test phase takes the training strategy of that fine-tuning the predictor head based on the pre-trained model in meta-training phase. This approach reduces computation while maintaining detection accuracy for the base class and avoids catastrophic forgetting caused by the introduction of novel class data.

4. Experiment

4.1. Dataset

We validate the model's detection effect using both a self-built few-shot defect dataset on power transmission lines and a public dataset. The self-built dataset includes 3845 images of three important components in power transmission lines, namely, an insulator, vibration damper, and pin, as well as 571 corresponding images of defective components covering four types of defects, including insulator self-explosion, insulator damage, damper missing, and pin missing. The training set, validation set, and test set are divided in a ratio of 7:2:1. In addition, according to the few-shot object detection dataset standard, we selected 30 objects with more obvious features for each class in the above training dataset to construct a few-shot dataset. Specifically, seven few-shot annotation files corresponding to each class are provided to generate support set data for the meta-learner. Since we plan to experiment with 30 shots of each class, every few-shot annotation file contains several image names, which include 30 annotations of the corresponding class. As multiple objects may be selected in one image, the few-shot training dataset consists of 141 images. An example of the power transmission line dataset is shown in Figure 5.

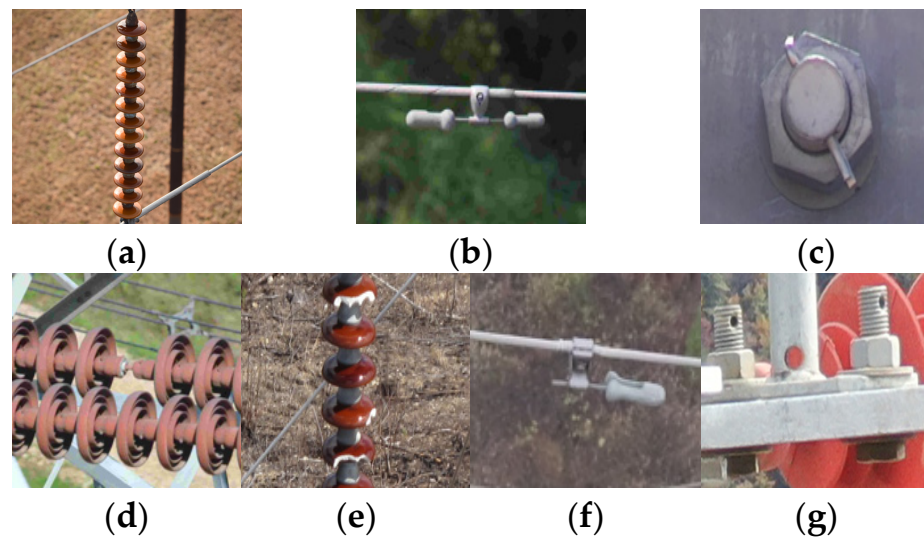


Figure 5. Examples of the self-built dataset. (a) Insulator; (b) vibration damper; (c) pin; (d) insulator self-explosion; (e) insulator damage; (f) damper missing; (g) pin missing.

For the public dataset, we used the PASCAL VOC2007 few-shot detection dataset, which contains 20 common classes in natural scenes, such as a person, bicycle, bus, bird, and sofa. Fifteen of these classes were selected as base classes, while the other five were selected as novel classes. This dataset is the same as the original PASCAL VOC2007, except it includes 20 few-shot annotation files that correspond to each class.

4.2. Implementation

The experimental hardware configuration in this study consisted of 16 GB memory, an Intel i7-6800K CPU, and an NVIDIA GeForce RTX 3090 GPU. The software configuration included an Ubuntu 18.04 operating system, CUDA 11.3, PyTorch 1.10.2, as well as the object detection training frameworks mmdetection and mmfewshot for few-shot object detection training.

The training process adopts the Stochastic Gradient Descent (SGD) algorithm to optimize the model parameters. The input image size is uniformly adjusted to 1200×800 pixels, and the batch size is set to four. In the meta-training phase, normal components are used as the base class for training, with an initial learning rate of 0.02, weight decay rate of 0.002, momentum of 0.9, and 15,000 iterations of training for each batch. For each base class (insulator, vibration damper, and pin), 200 objects are randomly selected as training samples from the training set and only one object is taken from each image. In the meta-test phase, normal components (base class, including insulator, vibration damper, and pin) and defect components (new class including insulator self-explosion, insulator damage, damper missing, and pin missing) samples are mixed to feed into the model for training. A total of 30 objects are selected as support set training samples for each class, and the number of iterations is set to 1000. Other parameters are the same as in the meta-training phase. In addition, the validation set and test set divided in Section 3.1 are used for both the meta-training phase and meta-test phase. To ensure the accuracy of the experiment, only the annotated boxes related to the training classes are retained for evaluation metric calculation during validation and testing. For example, only the annotated boxes of the base class are treated as ground truth in the evaluation of the meta-training phase, while in the meta-test phase, annotated boxes of both the base class and novel class participate in the calculation of evaluation.

4.3. Evaluation Indicators

We evaluate the detection performance using commonly used evaluation metrics in object detection and few-shot object detection, namely AP_{50} and mAP (Mean Average

Precision). AP_{50} represents the average precision calculated at an IOU threshold of 50%, as shown in Equation (7).

$$Precision = \frac{Num(TP)}{Num(TP) + Num(FP)} \quad (5)$$

$$AP = \frac{\sum Precision}{Num(TotalObjects)} \quad (6)$$

$$mAP = \frac{\sum AP}{Num(Classes)} \quad (7)$$

In addition, following the widely used standard in few-shot object detection, we set the support set sample size to 1, 5, and 10 in each training task to evaluate the performance of the MA-RPN with the public dataset under different training conditions.

4.4. Performance in Difference Hyperparameters

In order to fully exploit the performance of the model and verify the effectiveness of our method, we investigated the performance of the model under different optimization methods and learning rates at 30-shot. The experimental results are shown in Table 1.

Table 1. Performance in difference optimization method and learning rate.

Optimization Method	Learning Rate	mAP (Base and Novel)
SGD	0.01	72.4
	0.015	72.6
	0.02	72.8
	0.025	72.7
	0.03	72.5
Meta-SGD [34]	—	73.0

We found that using Meta-SGD [34], which can change the update direction and learning rate, optimizes the model and performs better than using SGD with a fixed learning rate. In the comparison of the results of using SGD with different learning rates, the learning rate of 0.02 is optimum. In order to effectively compare with other few-shot learning algorithms, in the subsequent experiments, the model will use the optimal learning rate of 0.02 during training. However, if the best detection performance is desired, the Meta-SGD method should be used to optimize the model.

4.5. Ablation

We conducted ablation experiments on MA-RPN and the feature reconstruction module DFRM. For the ablation experiment on MA-RPN, the traditional RPN structure was used to generate proposal regions instead of MA-RPN. For the ablation experiment on the DFRM, the original meta-features and ROI features from the query set were directly fused during the meta-test phase. The results are shown in Table 2. It can be seen that adding MA-RPN alone increased the mAP of defect components by 1.1%, and the defect detection performance of small-sized components such as pins was significantly improved, demonstrating that MA-RPN can not only effectively reduce noise in anchor boxes but can also generate more accurate candidate regions for small target objects by introducing low-level details. After adding the DFRM alone, the detection mAP increased by 2.0%, which was a significant improvement. This suggests that reconstructing meta-features based on the differences between defect and normal component features helps the model focus on defect-related features during classification in predictor head. Finally, the best detection performance was achieved with the model that combined both modules, verifying that the two modules do not conflict with each other and do not affect detection accuracy.

Table 2. Results of ablation experimental.

MA-RPN	DFRM	Novel Classes				
		Insulator Self-Explosion	Insulator Damage	Damper Missing	Pin Missing	Mean
✓	✓	78.5	72.6	80.6	46.2	69.4
		79.3	73.1	82.2	47.7	70.5
		80.9	75.0	82.2	47.5	71.4
✓	✓	81.9	76.2	82.4	49.8	72.5

4.6. Comparison Experiment

We compared the proposed detection algorithm with other mainstream few-shot object detection algorithms with a self-built few-shot defect dataset on a power transmission line; the results are presented in Table 1 for comparison. The training parameters of each model refer to the optimized experimental parameters in the mmfewshot algorithm framework, ensuring that the models achieve optimal performance. The detection results are shown in Figure 6.

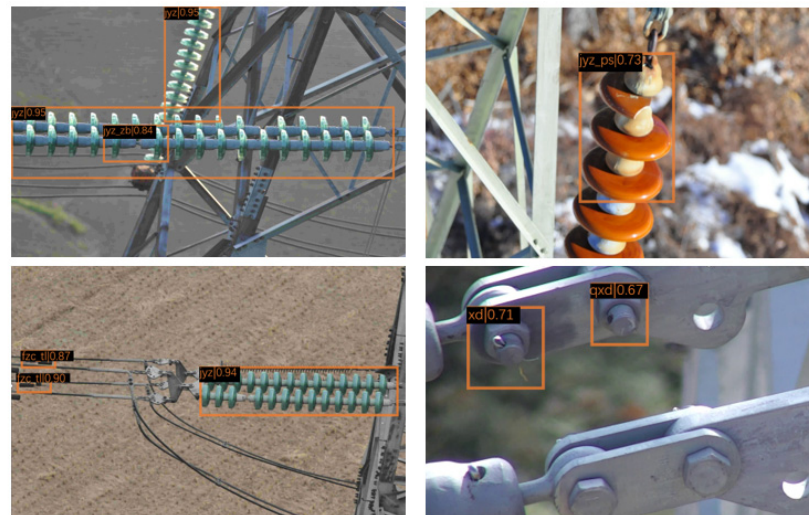
**Figure 6.** Detection result.

Table 3 presents and compares the detection performance of the proposed detection algorithm with other mainstream few-shot object detection algorithms based on the fully trained models. The following conclusions can be drawn: (1) Our proposed method achieved the highest overall mAP among mainstream methods, with significant performance improvements, indicating that Meta PowerNet is more suitable for defect detection tasks of transmission lines under small sample conditions. Specifically, the detection accuracy of normal components such as insulators and vibration dampers is slightly improved compared to other methods, but the detection performance of defect components such as insulator self-explosions and missing vibration dampers is significantly better than other methods, indicating that the main advantage of our method compared to other methods is the high detection accuracy of defect parts. (2) Meta PowerNet also performed well in detecting pins and pins missing, indicating that the model has better detection performance for small target objects and can be used for the defect detection of components with small sizes on transmission lines. (3) When detecting different types of defects of the same component, such as insulator self-explosion and insulator damage, our proposed method also demonstrated good discrimination.

Table 3. Performance comparison of this method and mainstream few-shot object detection.

Methods	Base Classes				Novel Classes					mAP
	Ins ¹	Damper ²	pin	Mean	Ins-Exp ³	Ins-Dam ⁴	Dam-Mis ⁵	Pin-Mis ⁶	Mean	
FSRW [25]	80.7	79.5	46.3	68.8	76.0	72.3	76.8	42.7	66.9	67.7
MPSR [35]	82.9	81.4	50.1	71.4	78.6	72.5	79.5	45.3	68.9	70.0
Meta R-CNN [27]	82.6	80.9	49.5	71.0	77.8	70.2	79.2	45.1	68.0	69.3
Attention RPN [36]	83.1	81.7	50.6	71.8	79.8	73.8	80.9	46.5	70.2	70.9
FSCE [14]	84.5	82.6	51.1	72.7	81.2	75.3	81.7	48.0	71.8	72.2
Meta PowerNet	84.7	83.1	51.7	73.1	81.9	76.2	82.4	49.8	72.5	72.8

¹ ins: insulator. ² damper: vibration damper. ³ ins-exp: insulator self-explosion. ⁴ ins-dam: insulator damage. ⁵ dam-mis: damper missing. ⁶ pin-mis: pin missing.

4.7. Expand

This experiment extends MA-RPN to Meta R-CNN and Attention RPN models, both of which are also meta-learning-based object detection algorithms. The PASCAL VOC2007 split1 few-shot dataset was used for training and testing. The experimental results are compared in Table 4.

Table 4. Comparison of expansion effects of MA-RPN.

Method \ Shot	1	5	10
Meta R-CNN [27]	54.1	63.8	65.0
Meta R-CNN [27] + MA-RPN	54.7	65.2	65.5
Attention RP [36]	37.8	56.5	58.6
Attention RPN [36] + MA-RPN	38.1	58.4	59.0

We compared Meta R-CNN and Attention RPN models with and without the MA-RPN module under the conditions with 1, 5, and 10 support samples per class. The results showed that MA-RPN improved the detection precision of both models by 0.6% and 0.3% for 1-shot, 1.4% and 0.9% for 5-shot, and 0.5% and 0.4% for 10-shot. The results indicating that the improvement was weak under 1-shot but most significant under 5-shot. As for the 10-shot condition, the performance gain began to decrease due to the insufficient feature extraction of the meta-learner and the fine-tuning effect of the RPN. Overall, the results demonstrated the general applicability of the MA-RPN module for the two models under different support sample conditions.

5. Conclusions and Future Work

We propose a meta-learning-based method for few-shot defect detection in transmission lines, which aims to address the problem of insufficient samples in the task. First, a region proposal network that combines support set features to generate anchor boxes is proposed to improve the quality of anchor boxes. Second, a meta-learner based on the defect feature reconstruction module is designed to improve the detection accuracy of defect components by considering the data characteristics of normal and defect components in transmission lines. Experimental results show the effectiveness of the Meta PowerNet compared with other mainstream few-shot object detection methods. Moreover, the proposed MA-RPN can be extended to other meta-learning-based object detection methods to enhance detection performance.

As for future work, we plan to divide the detection process into two steps. First, traditional object detection methods will be employed to locate the key components to be detected. Subsequently, the region of interest will be enlarged based on the detected results, and the enlarged image will be input into a meta-learning image classification network for defect recognition. This approach can not only avoid the problem of local feature loss but can also improve the detection accuracy of the base class.

Author Contributions: Conceptualization, X.Z. and Y.S.; methodology, Y.S. and H.W.; software, Y.S. and H.W.; validation, Y.S., H.W. and C.J.; writing—original draft preparation, Y.S.; writing—review and editing, C.J. and X.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Key R&D Program of Shanxi Province, grant number 2022ZDYF100.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The PASCAL VOC 2007 dataset with the few-shot annotations is available at <https://github.com/open-mmlab/mmfewshot/blob/main/tools/data/detection/voc/README.md> (accessed on 5 May 2023). For privacy reasons, the self-built dataset cannot be made fully public. Readers can contact the corresponding author for details.

Acknowledgments: Additionally, the authors would like to thank all the reviewers who participated in the review.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Ma, Y.P.; Li, Q.W.; Chu, L.L.; Zhou, Y.Q.; Xu, C. Real-Time Detection and Spatial Localization of Insulators for UAV Inspection Based on Binocular Stereo Vision. *Remote Sens.* **2021**, *13*, 230. [CrossRef]
2. Zhao, J.L.; Zhang, X.Z.; Dong, H.Y. Defect detection in transmission line based on scale-invariant feature pyramid networks. *Comput. Eng. Appl.* **2022**, *58*, 289–296.
3. Chen, S.Y.; Fu, Z.J. Multi-scale transmission line component detection incorporating efficient attention. *Comput. Eng. Appl.* **2023**, *1*, 1–11. Available online: <http://kns.cnki.net/kcms/detail/11.2127.TP.20230103.1221.002.html> (accessed on 6 April 2023).
4. Liu, C.Y.; Wu, Y.Q. Research progress of vision detection methods based on deep learning for transmission lines. *Proc. CSEE* **2023**, *29*, 1–24. Available online: <http://kns.cnki.net/kcms/detail/11.2107.TM.20220831.1053.002.html> (accessed on 6 April 2023).
5. Zhao, K.L.; Jin, X.L.; Wang, Y.Z. Survey on Few-shot Learning. *J. Softw.* **2021**, *32*, 349–369.
6. Li, A.; Luo, T.; Lu, Z.W.; Xiang, T.; Wang, L.W. Large-Scale Few-Shot Learning: Knowledge Transfer with Class Hierarchy. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 7205–7213.
7. Ertugrul, B.; Burla, N.K.; Aras, U.E.; Numan, C. Traffic congestion-aware graph-based vehicle rerouting framework from aerial imagery. *Eng. Appl. Artif. Intell.* **2023**, *119*, 105769.
8. Bayraktar, E.; Tosun, B.; Altintas, B.; Celebi, N. Combined GANs and Classical Methods for Surface Defect Detection. In Proceedings of the 2022 30th Signal Processing and Communications Applications Conference (SIU), Safranbolu, Turkey, 15–18 May 2022.
9. Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative Adversarial Networks. *Commun. ACM* **2020**, *63*, 139–144. [CrossRef]
10. Hsieh, T.I.; Lo, Y.C.; Chen, H.T.; Liu, T.L. One-shot object detection with co-attention and co-excitation. In Proceedings of the 33rd International Conference on Neural Information Processing Systems, Vancouver, BC, Canada, 8–14 December 2019; Curran Associates Inc.: Red Hook, NY, USA, 2019; Volume 245, pp. 2725–2734.
11. Li, W.B.; Wang, L.; Xu, J.L.; Huo, J.; Gao, Y.; Luo, J.B. Revisiting Local Descriptor based Image-to-Class Measure for Few-shot Learning. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 7253–7260.
12. Wang, X.; Thomas, E.H.; Trevor, D.; Joseph, E.G.; Yu, F. Frustratingly Simple Few-Shot Object Detection. *arXiv* **2020**, arXiv:2003.06957.
13. Xiao, Y.; Lepetit, V.; Marlet, R. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 1137–1149.
14. Sun, B.; Li, B.H.; Cai, S.C.; Yuan, Y.; Zhang, C. FSCE: Few-Shot Object Detection via Contrastive Proposal Encoding. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 19–25 June 2019; pp. 7348–7358.
15. Ye, H.J.; Hu, H.; Zhan, D.C.; Sha, F. Few-shot learning via embedding adaptation with set-to-set functions. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 14–19 June 2020; pp. 8805–8814.
16. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, L.; Polosukhin, I. Attention Is All You Need. *arXiv* **2017**, arXiv:1706.03762.
17. Zhang, C.; Ding, H.H.; Li, G.S.; Li, R.B.; Wang, C.H.; Shen, C.H. Meta Navigator: Search for a Good Adaptation Policy for Few-shot Learning. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 19–25 June 2021; pp. 9415–9424.

18. Xu, H.Q.; Yu, J.B.; Liang, C.; Zhang, X.H. Detection Method for Small Metal Defects of Improved RPN Transmission Line Based on GAN. *Chin. J. Electron Devices* **2021**, *44*, 1409–1416.
19. Cui, K.B.; Pan, F. A CycleGAN small sample library amplification method for faulty insulator detection. *Comput. Eng. Sci.* **2022**, *44*, 509–515.
20. Zhu, J.Y.; Park, T.; Isola, P.; Efros, A.A. Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 2242–2251.
21. Zhang, Y.X.; Wu, G.P.; Liu, Y.Z.; Yang, S.; Xu, W.Z. Transfer learning of transmission line damper and clamp detection based on YOLOv3 network. *J. Comput. Appl.* **2020**, *40*, 188–194.
22. Redmon, J.; Farhadi, A. YOLOv3: An Incremental Improvement. *arXiv* **2018**, arXiv:1804.02767.
23. Zhai, Y.J.; Yang, K.; Wang, Q.M.; Wang, Y.R. Disc Insulator Defect Detection Based on Mixed Sample Transfer Learning. *Proc. CSEE* **2020**, *7*, 2867–2877.
24. Li, F.C.; Liu, Y.; Wu, P.X.; Dong, F.; Cai, Q.; Wang, Z. A Survey on Recent Advances in Meta-Learning. *Chin. J. Comput.* **2019**, *44*, 422–446.
25. Kang, B.Y.; Liu, Z.; Wang, X.; Yu, F.; Feng, J.S.; Darrell, T. Few-shot Object Detection via Feature Reweighting. In Proceedings of the IEEE International Conference on Computer Vision, Seoul, Republic of Korea, 27 October–2 November 2019; pp. 8419–8428.
26. Redmon, J.; Farhadi, A. YOLO9000: Better, Faster, Stronger. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 6517–6525.
27. Yan, X.P.; Chen, Z.L.; Xu, A.; Wang, X.X.; Liang, X.D.; Lin, L. Meta R-CNN: Towards General Solver for Instance-level Low-shot Learning. In Proceedings of the IEEE International Conference on Computer Vision, Seoul, Republic of Korea, 27 October–2 November 2019; pp. 9576–9585.
28. Ren, S.Q.; He, K.M.; Girshick, R.; Sun, J. Few-shot Object Detection and Viewpoint Estimation for Objects in the Wild. *IEEE Trans. Pattern Anal. Mach. Intell.* **2022**, *45*, 3090–3106.
29. Zhang, G.J.; Luo, Z.P.; Cui, K.W.; Lu, S.J. Meta-DETR: Few-Shot Object Detection via Unified Image-Level Meta-Learning. *arXiv* **2021**, arXiv:2103.11731.
30. Zhu, X.Z.; Su, W.J.; Lu, L.W.; Li, B.; Wang, X.G.; Dai, J.F. Deformable DETR: Deformable Transformers for End-to-End Object Detection. *arXiv* **2020**, arXiv:2010.04159.
31. Li, B.; Wu, W.; Wang, Q.; Zhang, F.Y.; Xing, J.L.; Yan, J.J. Meta Faster R-CNN: Towards Accurate Few-Shot Object Detection with Attentive Feature Alignment. *arXiv* **2021**, arXiv:1812.11703.
32. Bromley, J.; Guyon, I.; LeCun, Y.; Säckinger, E.; Shah, R. Signature verification using a “siamese” time delay neural network. *Adv. Neural Inf. Process. Syst.* **1993**, *6*, 25–44. [[CrossRef](#)]
33. Han, G.X.; Huang, S.Y.; Ma, J.W.; He, Y.C.; Chang, S.F. SiamRPN++: Evolution of Siamese Visual Tracking with Very Deep Networks. *arXiv* **2018**, arXiv:2104.07719.
34. Li, Z.G.; Zhou, F.W.; Chen, F.; Li, H. Meta-SGD: Learning to Learn Quickly for Few Shot Learning. *arXiv* **2017**, arXiv:1707.09835.
35. Wu, J.X.; Liu, S.T.; Huang, D.; Wang, Y.H. Multi-Scale Positive Sample Refinement for Few-Shot Object Detection. *Lect. Notes Comput. Sci.* **2020**, *12361*, 456–472.
36. Fan, Q.; Zhuo, W.; Tang, C.K.; Tai, Y.W. Few-Shot Object Detection with Attention-RPN and Multi-Relation Detector. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 14–19 June 2020; pp. 4012–4021.

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.