2025 International Conference on Intelligent Computing
July 26-29, Ningbo, China
https://www.ic-icc.cn/2025/index.php

FDML: An Improved Few-Shot Fault Detection Method for Transmission Lines Based on Meta-Learning

Yanwei Zhang[1], Bocheng Huang[2], Guocheng An, and Xiaolong Wang[3,4]

[1] The Artificial Intelligence Research Institute of Shanghai Huaxun Network System Co., LTD., Chengdu, Sichuan 610074, China
[2] Shanghai Key Lab. of Beidou Navigation and Location Services, College of Electronic Information and Electrical Engineering, Shanghai Jiao Tong Univ., Shanghai 200240, China
[3] School of Computer Science and Technology, Hangzhou Dianzi University, Hangzhou,310018, Zhejiang, China
[4] The 32nd Research Institute of China Electronics Technology Group Corporation, Shanghai, 201808, China

Abstract. In the task of fault detection in power transmission lines, certain fault categories suffer from the problem of insufficient samples, leading to inefficiencies in traditional object detection algorithms. Meta-learning, which employs multi-task learning and fine-tuning to extract common features across different tasks, performs well in few-shot object detection and demonstrates excellent generalization capabilities for new tasks. For this reason, an improved few-shot fault detection method based on meta-learning (FDML) is proposed in this paper. Firstly, to address the problem of distribution difference in data domains, a two-stage meta-learning training method is introduced to achieve model migration through meta fine-tuning. Secondly, we propose a support and query feature matching module (SQFM) to make the utmost of support features to assist detection, in which prototype features of the support class are accurately extracted in the first three stages of the backbone and then assigned to the query set features to highlight the class-specific representative features. To further integrate high-level feature before model prediction, a high-level semantic feature fusion module (HSFF) is designed to fuse RoI features and prototype features via combining the four feature fusion ways. Experimental results show that FDML effectively improves the few-shot object detection accuracy on the public dataset PASCALVOC and the fault dataset InsPLAD-fault, compared to the classic few-shot algorithms. Under the conditions of K= {5, 10, 20} shot in the fault dataset InsPLAD-fault, the mAP_{50} values are respectively 5.7%, 7.2% and 4.2% higher than the baseline network, which provides a solution for few-shot transmission line fault detection.

Keywords: transmission lines; fault detection; few-shot; meta learning; SQFM; HSFF

1 Introduction

With the development of contemporary society, the growing demand for electricity to sustain normal operations has driven the expansion of the national grid. As the main part of the grid, the stability of ultra-high voltage transmission lines plays a crucial role in ensuring the reliable operation of the entire system. However, due to prolonged exposure to the field (i.e., suffer from natural violations and external forces), the components on the lines are susceptible to damage, leading to various faults. Once faults occur, it will inevitably produce safety accidents without timely maintenance. Therefore, it is crucial to identify and repair faults in transmission lines promptly to prevent greater harm [1-3].

In recent years, UAV aerial photography and deep learning have been widely applied in transmission line inspections [4]. That is to say, the inspection method combining UAV aerial images and object detection algorithms has largely replaced the traditional manual inspection methods. Through the analysis of aerial images, faults and defects can be quickly identified and located for immediate repair. For instance, Zhou *et al.* proposed an automatic insulator fault detection method via leveraging the unique characteristics of insulators and optimizing YOLOv5 from four different aspects [5]. In order to overcome the challenge of real-time detection, Niu *et al.* proposed a lightweight algorithm named Comprehensive-YOLOv5, focusing on rapid localization and accurate identification of three common defects [6]. Similarly, reference [7] designed an improved insulator defect detection model based on YOLOv4 (ID-YOLO), aimed at detecting the relatively small insulator fault areas (e.g., bunch-drop) in aerial images. For addressing the problems of low accuracy and high computation cost in existing fault detection models, Yi *et al.* proposed the insulator and defect detection model YOLO-S, which provides an in-depth analysis of insulator localization and defect diagnosis [8]. In [9], Wang *et al.* enhanced insulator defect detection by constructing the ML-YOLOv5 model using knowledge distillation, with YOLOv5m as the teacher model. Via constructing a prior knowledge transfer model based on visual saliency, Hao *et al.* proposed a prior knowledge transfer and attention mechanism network (PKAMNet) to detect faulty insulator parallel gaps, which can ameliorate the phenomenon of missed and erroneous detection [10].

Although above deep learning-based methods have achieved lots of success in transmission line fault detection, there are still a difficult issue to be addressed. In practice, the majority of transmission line components are in normal working condition. So, due to various factors such as temperature, weather, and human fallibility, there are many fault types with the scarcity of samples. At the same time, it is essential to collect sufficient samples for training while using object detection algorithms based on deep learning. That is to say, the limited availability of samples often restricts the effectiveness of these methods, resulting in low detection accuracy. Consequently, how to effectively perform fault detection in scenarios with few samples has become one challenging task.

Few-shot learning is a hot-spot and promising approach to address the problem of limited training samples in the field of target detection, which can alleviate the phenomenons in deep-learning models such as over-fitting, difficult convergence and poor generalization [11]. This approach mainly includes transfer learning, meta learning, metric learning, data augmentation and so on. For example, Zhang *et al.* designed Meta-DETR, which operates entirely at image level without any region proposals to circumvent the constraint of inaccurate proposals in prevalent few-shot methods [12]. In [13], Zha *et al.* developed a novel two-stage weakly-supervised framework to address the few-shot fine-grained recognition task, in which well-designed modules are instantiated to achieve the background suppression and foreground alignment. To solve the issue of category confusion, Li *et al.* propose a two-stage fine-tuning approach via classification score calibration for remote sensing images, which follows the flowchart of base training and few-shot fine-tuning to train the detector [14]. At present, there are few studies on few-shot fault detection for transmission lines and this problem should be given wider attention. For instance, reference [15] designed a few-shot learning based two-stage insulator defect detection algorithm InsDef consisting of two stages: the insulator extraction stage and the defect recognition stage. Via constructing a multi-scale feature re-weighting network, Wang *et al.* proposed a two-stage method for insulator anomaly detection [16]. Shi *et al.* 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 [17].

Despite these advancements, existing methods struggle with extremely scarce samples. Besides, there are low differentiation and confusing feature between different fault types of the same component. So, it is also a great challenge to achieve fine-grained identification with few samples. In response to the aforementioned analysis, hereinafter, we propose an improved few-shot fault detection method based on meta-learning named FDML. The main contribution are as follows.

(1) For alleviating the problem of distribution difference in data domains, a two-stage meta-learning training method is introduced to achieve model migration through meta fine-tuning.

(2) To make the utmost of support features to assist detection, a support and query feature matching module (SQFM) is designed to highlight the class-specific representative features, in which prototype features of the support class are accurately extracted and assigned to the query set features.

(3) A high-level semantic feature fusion module (HSFF) is designed to integrate RoI features and prototype features, which can effectively support the final model prediction.

(4) Via fusing the modules proposed above together, FDML is finally put forward. The experiments conducted on the datasets PASCALVOC and InsPLAD-fault demonstrate its superiority over other few-shot learning methods.

The rest of this study is organized as follows: The data preparation and related works on meta-learning are discussed and reviewed in Section 2. Section 3 describes the proposed algorithm for fault detection in detail. In Section 4, experimental results as well as performance evaluation and analysis are presented. In the end, we briefly summarize this article and provide the prospect for future research in Section 5.

2 Materials and Related Works

2.1 Data preparation and processing

Recently, a new dataset named InsPLAD [18] is proposed for inspecting power line assets to address multiple research gaps in the field. In the dataset, there are five different types of defects annotated on an image level (339 defect samples in total, cropped from the UAV images), allowing the evaluation of image classification and unsupervised anomaly detection methods. Table 1 shows the five fault types and the corresponding sample numbers. At the same time, InsPLAD-fault assets are shown in Figure 1. The first row represents the normal conditions, and the second row represents the fault samples.

Table 1. Description of InsPALD-fault.

Asset category	Fault type	Sample number
Yoke Suspension	rusty	49
Glass Insulator	missing cap	90
Polymer Insulator Upper Shackle	rusty	102
Lightning Rod Suspension	rusty	50
Vari-grip	rusty	48

Fig. 1. Sample images of InsPLAD, including normal and fault conditions.

Due to the low number of faults in transmission lines, in the article [18], InsPLAD-fault was adapted into two approaches: supervised fault classification and unsupervised anomaly detection. However, the proposed method aims to solve this problem via adopting few-shot learning methods as well as detect the precise location of the fault. For this goal, the Labelme image annotation tool was used to label the fault targets with rectangular boxes, as shown in Figure 2. Moreover, the fault dataset constructed in this study adopted the PASCAL VOC format, which consisted of three folders, including images in jpg format, annotation files in xml format, and image lists.

Fig. 2. Example of annotation, in which different color of rectangular boxes represent different fault types.

2.2 Few-Shot Training Method Based on Meta-Learning

2.2.1. Meta-Learning

Meta-learning, also known as learning to learn, is a promising method for few-shot object detection. Different from the traditional deep learning, meta-learning paradigm takes tasks as input units for iteration, and each task consists of support set images and query set images, which goal is to find the target belonging to the category of support set in query set images [31]. In other words, the model is tested on the query set data via learning the features of the support set data to verify the effect. The core idea of meta-learning is to increase the generalization ability in multi-tasking. Specifically, a set of initialization parameters can be obtained through meta-training, which can be effectively applied in different tasks and then, performing a few fine-tuning iteration for specific few-shot tasks can obtain satisfactory results. The architecture of meta-learning is shown correspondingly in Figure 3.

Fig. 3. The architecture of meta-learning.

2.2.2. Training Process

At present, meta-learning method adopts the standard setup of FSOD, given two disjoint sets of classes named the base class and the novel class. In the base-class dataset, there are sufficient annotated samples with bounding box annotations for each category, while each class has only K-shot annotated targets in the novel-class dataset. Generally, due to the scarcity of samples in the new task, the trained meta-learning model is directly used to test, which may result in poor meta-migration by excessive difference between data domains.

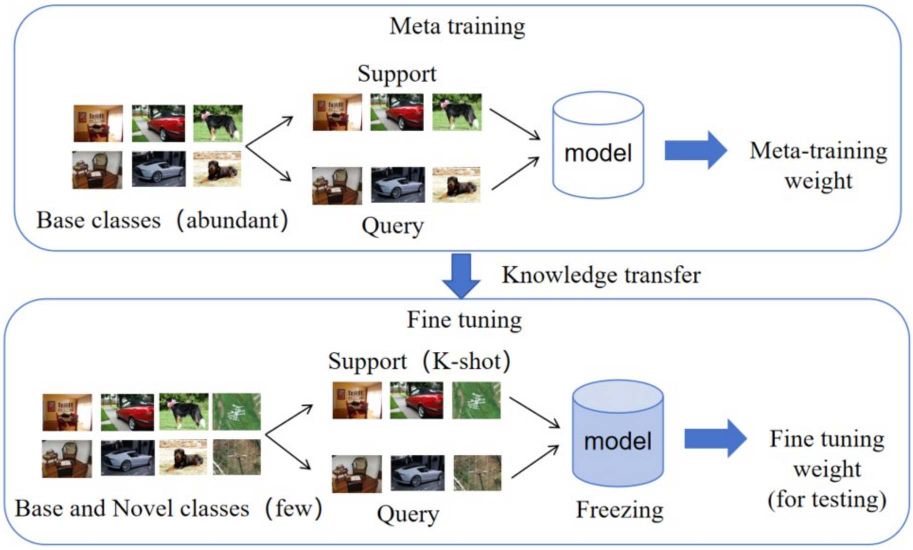

Fig. 4. The training process of meta-learning.

To address this issue, we adopt the two-stage training paradigm commonly used in transfer learning methods, as shown in Figure 4. In the meta-training stage, the model is trained on abundant samples of base classes. After, in the fine-tuning phase, the meta-learning model is fine-tuned via using a balanced dataset consisting of both base and novel classes, which contains N classes and only K-shot samples as the supporting set, referred to as the N-way-K-shot task. By this means, the proposed method helps the model better adapt to the new data distribution, and alleviate the distribution differences of data domains. Specially, public datasets with rich data, such as PASCALVOC, COCO, etc. are usually selected as base classes while adopting specific few-shot dataset as novel classes for meta-learning object detection.

3 Methods

3.1 Overview of the Proposed Method

In this paper, the few-shot transmission-line fault detection method based on meta-learning is improved to deal with sample scarcity. The overall structure of improved method is shown in Figure 5, which mainly consists of feature extraction backbone, RPN, detecting head and two proposed modules, i.e., support and query feature matching module (SQFM) and high-level semantic feature fusion module (HSFF). Specifically, the model is built as a few-shot detector based on the two-stage algorithm Faster R-CNN [19], following the setting of Meta R-CNN framework [20]. For the feature extraction network, FDML establish the dual branch connected network, in which Resnet-101 [21] pre-trained on Imagenet is adopted as the backbone. At the same time, it should be noted that the weight parameters are shared in the feature extraction process between supports branch and query branch. Moreover, through the proposed SQFM, the class prototype of support set is built via extracting the middle-level features in the first three stages of the backbone and assigned to the query branch for feature reconstruction. Subsequently, the query branch is RoI aligned through the RPN network, where the query characteristics after interacting with the supporting branch can be provided. After alignment, we employ the last stage of the backbone to extract the high-level semantic feature of two branches, namely, the RoI features generated by the query branch and the class-level prototype features of the support branch. At last, these features are input into the HSFF for more efficient feature fusion, and then output to the detection head for final prediction. It is worth noting that the support set images are resized to 224*224 as the support branch input after clipping the instance, and the single scale feature map is adopted for object detection without FPN structure [22].

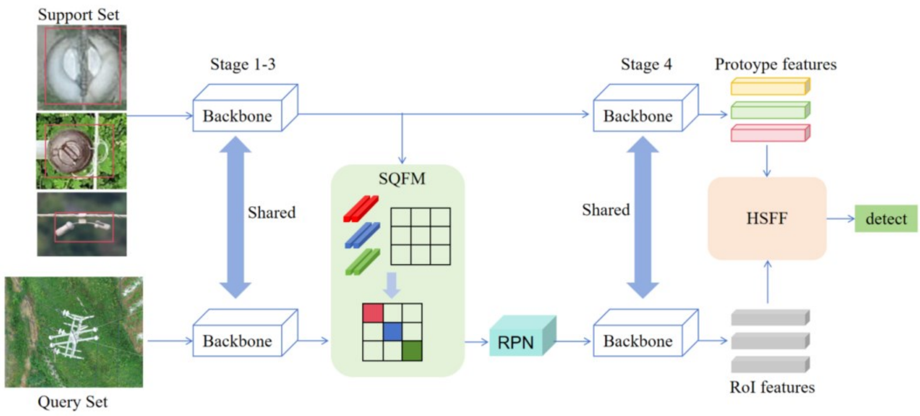

Fig. 5. Overview of the proposed method (FDML).

3.2 Support and Query Feature Matching Module (SQFM)

As mentioned before, the interaction between support features and query features is very critical in meta-learning methods. At present, the query features is usually reconstructed by the support features which are pooled. Generally, the mid-level features with more detailed information in the early stages of the network are not fully utilized. Therefore, inspired by the fine-grained feature aggregation module (FFA) [23], we hereinafter propose a novel module SQFM (seeing Figure 6), which can be divided into prototype extraction stage and prototype assignment stage.

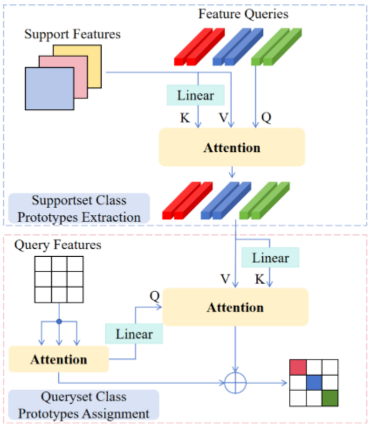

Fig. 6. The illustration of SQFM composed of prototypes extraction stage and prototypes assignment stage.

Broadly speaking, support features are extracted into the class-level fine-grained prototype via modeling the relationship between classes and then assigned to the query features, which helps the model extract and learn key information.

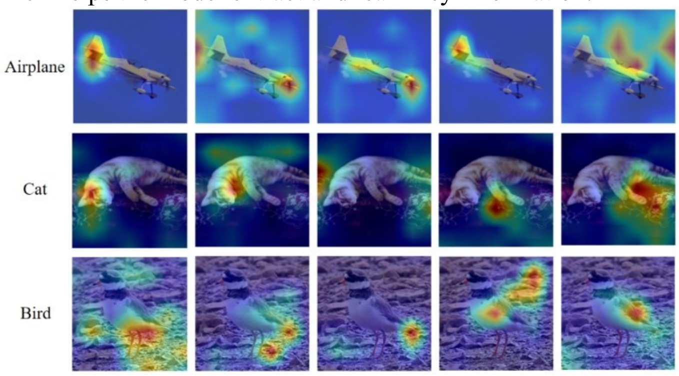

Fig. 7. Attention heat-maps of feature queries. The attention mask is generated from the attention mechanism of supportset class prototypes extraction in SQFM.

In the supportset class prototype extraction stage, inspired by the idea of incorporating object queries to encode location information in DETR [24], we introduce a new component named feature queries as learnable embedding to extract class-level prototypes and refine them into a representative set of features as a guide. As shown in Figure 7, different details of class-related feature can be successfully captured in these feature queries. Specifically, given a set of features queries Q and the support set feature F_s of the feature extraction network, we employ the linear layer to adjust the dimension of F_s and then apply the cross-attention mechanism to realize class prototypes extraction. The specific formula is as follows.

$$P = Attention(Q, F_s W_0, F_s) = soft\max(\frac{Q(F_s W_0)^T}{d})F_s \tag{1}$$

In the equation, d represents the feature dimension of feature queries. In addition, W_0 is the linear matrix that projects F_s into the dimension space of Q, while P denotes the extracted class-level fine-grained prototype. Notably, each category has its unique feature query, and here Q represents the collection of feature query for all categories.

After prototypes extraction, it is crucial to match fine-grained prototype to the queryset feature map. To this end, we also employ two attention mechanisms to achieve the prototypes assignment. Specifically, in order to highlight the significant features of queryset, the self-attention mechanism is adopted to process the query features firstly. By this means, the enhanced query features are conducive to the subsequent feature matching. Further, given the enhanced queryset feature F_q and the extracted class-level prototype P, we adjust their feature dimension through the fully connected layer as input, and the cross-attention mechanism is used for prototypes assignment as well. The specific formulas for the above process are as follows.

$$F_q = Selfattention(q, k, v) \tag{2}$$

$$F_a = Attention(F_q W_1, PW_1, P) = soft\max(\frac{(F_q W_1)(PW_1)^T}{\sqrt{d}})P \tag{3}$$

where q, k, and v are obtained by linear transformation of the original query features, while W_1 represents the linear matrix shared by F_q and P, aiming to project them into the same feature dimension. According to the method, F_a which denotes the feature obtained by prototypes assignment are generated.

Finally, the residual connecting method is adopted to aggregate the prototype assignment feature F_a and the original query features F_q to generate final features F_{out} after complete feature matching. The computation can be written as follows.

$$F_{out} = F_q + F_a \tag{4}$$

In the meta-training stage, feature queries of base classes are randomly initialized and trained extensively. In the fine-tuning phase, it is challenging to train feature queries from scratch since there are only limited examples of novel classes. So, to address this issue, the most compatible feature queries from base classes are copied to apply in novel classes detection [23]. In a word, SQFM not only effectively makes a distinction between targets and background via learning the expression of inter-class similarity and difference between different objects, but also focuses on the representative class-specific features while filtering out the features unrelated to the support classes which is more suitable for shallow fine-grained information extraction. In addition, the residual connection retains more original feature information, which is conducive to the subsequent high-level feature fusion.

3.3 High-Level Semantic Feature Fusion Module (HSFF)

While adopting the meta-learning method for FSOD, it is vital to perform feature fusion for high-level semantic information alignment before final detection. That is, feature fusion between the RoI features generated by query images and class-level prototype features is an indispensable step in meta-learning method [25]. Up to now, many previous meta-learning methods employ element-wise multiplication to achieve the fusion process, i.e., highlighting related features. However, in this way, the model can only learn the similarities in the same class, ignoring the differences between classes. For this goal, we propose a novel high-level semantic feature fusion module (HSFF) via combining element-wise multiplication, addition, subtraction and channel cascade methods to measure the intra-class similarity and inter-class difference between features for mutual complementation, as shown in Figure 8.

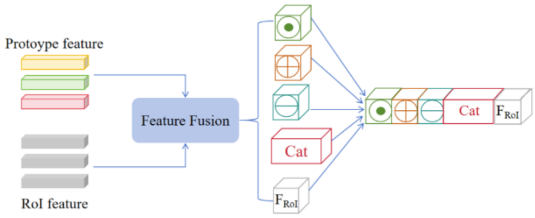

Fig. 8. The illustration of HSFF composed of four fusion ways.

Specifically, the common features can be highlighted through element-wise multiplication, while performing element-wise addition operations enable the network to focus on specific features being equivalent to utilize attention weights. Via element-wise subtraction, the distance between RoI features and prototype features can be directly measured, i.e., the difference between features can be obtained. At the same time, applying channel concat operation for feature fusion is a self-learning operation method that automatically learn how to combine features. Subsequently, after refining the relationship of features through the aforementioned four feature fusion methods, we connect them with the original features in order to retain the original RoI information. The specific formula is as follows.

$$F = concat(f_1(F_{RoI} \odot F_p), f_2(F_{RoI} \oplus F_p), f_3(F_{RoI} - F_p), f_4(F_{RoI}, F_p), F_{RoI}) \quad (5)$$

Where f_1 to f_4 are independent fully connected layer to adjust dimension. F_{RoI} represents the RoI feature of the query branch, while F_P denotes the class-level prototype feature generated by the support branch. Finally, the fused features are cascaded through channels and adjusted to the output feature F for prediction.

Generally speaking, through the proposed HSFF, the model effectively focuses on class-specific features and fuse high-level semantic features via combining the information from the support and the query for final detection.

4 Results and Analysis

In this section, we realize a lot of experiments on FDML and compare the experimental results with other representative few-shot object detection algorithms. To confirm the effectiveness of each module, the contributions of different modules are discussed and the visual detection results are presented.

4.1 Dataset

To validate the effect of proposed method, we use both a public dateset and the few-shot fault dataset on transmission lines (i.e. InsPLAD-fault mentioned before). For the public dataset, the PASCAL VOC 2007 and 2012 few-shot dataset are adopted, containing 20 common classes which are divide into 15 base classes and 5 novel classes. It should be noted that there are three different class divisions for a more comprehensive evaluation which are shown in table 2, and experiments are performed under K= {1, 2, 3, 5, 10} shot settings. Following the evaluation standard widely used in FSOD, VOC2007 and VOC2012 training sets are used for training while VOC2007 test set is used for evaluation. For InsPLAD-fault, the training set and test set are divided in a ratio of 1:1 as well, and the base classes remain the 15 classes from the Pascal VOC dataset, while the novel classes are replaced with the 5 fault classes. Due to the challenge of fine-grained recognition in transmission line faults, this paper set the support set sample number to 5, 10, and 20 in each task (i.e., K= {5, 10, 20} shot).

Table 2. Class divisions of Pascal VOC datasets

Split set	Novel classes	Base classes
Novel Split 1	bird, bus, cow, motorbike, sofa	49
Novel Split 2	aeroplane, bottle, cow, horse, sofa	90
Novel Split 3	boat, cat, motorbike, sheep, sofa	102

4.2 Experimental Setup and Evaluation Metrics

In this paper, the experimental hardware configuration comprises 24GB memory, an Intel i7-14700K CPU and an NVIDIA GeForce RTX4090 GPU, while the model experiment environment includes Python 3.8, CUDA 11.8 and Pytorch 2.0. While model training, the stochastic gradient descent (SGD) optimizer is adopted to update the parameters with a batch size of 8 scenes. In the meta training stage, the model is trained 20k iterations on the base classes dataset, while the learning rate is set to 0.005, and decayed at 17k iteration by a factor of 0.1. In the fine-tuning stage, the learning rate is set to 0.001 and the model is fine-tuned to convergence via using the balanced dataset of base classes and novel classes under K-shot, in which the relevant Settings follow the Meta R-CNN framework.

In order to comprehensively evaluate the performance of the proposed method, the mAP_{50} is used as the evaluation metric, which represents the mean average precision of total class calculated at an IOU threshold of 0.5. The specific formulas are shown in Equation (6)-(9).

$$Precision = \frac{TP}{TP + FP} \tag{6}$$

$$Recall = \frac{TP}{TP + FN} \tag{7}$$

$$AP = \int_0^1 P(R)d(R) \tag{8}$$

$$mAP = \frac{\sum AP}{Num(Classes)} \tag{9}$$

In the four mentioned equations, *TP* represents the number of objects correctly located and recognized by the model, *FP* represents the number of false positive samples predicted mistakenly as positive by the model, and *FN* denotes the number of targets failed to correctly detect.

4.3 Comparative Experiment

In order to verify the effectiveness and generalization of FDML, we compare it with other mainstream few-shot object detection methods, including FSRW, Meta R-CNN, TFA, FSDetView, MPSR, FSCE, Meta DETR, and FPD [12, 23, 26-30], on three splits of public dataset PASCAL VOC. Table 3 lists quantitative results correspondingly.

Table 3. FSOD results (mAP_{50}) on the three splits of Pascal VOC dataset.

Set	Novel Split 1					Novel Split 2					Novel Split 3				
Method/shot	1	2	3	5	10	1	2	3	5	10	1	2	3	5	10
FSRW	14.8	15.5	26.7	33.9	47.2	15.7	15.3	22.7	30.1	40.5	21.3	25.6	28.4	42.8	45.9
Meta R-CNN	19.9	25.5	35.0	45.7	51.5	10.4	19.4	29.6	34.8	45.4	14.3	18.2	27.5	41.2	48.1
Meta DETR	40.6	51.4	58.0	59.2	63.6	**37.0**	36.6	43.7	49.1	**54.6**	41.6	45.9	52.7	58.9	60.6
TFA	39.8	36.1	44.7	55.7	56.0	23.5	26.9	34.1	35.1	39.1	30.8	34.8	42.8	49.5	49.8
FSCE	44.2	43.8	51.4	61.9	63.4	27.3	29.5	43.5	44.2	50.2	37.2	41.9	47.5	54.6	58.5
FPD	40.1	55.3	58.2	62.3	64.9	24.8	36.8	44.9	47.6	51.8	38.5	44.8	49.7	57.9	59.0
FDML(ours)	**45.1**	**56.4**	**61.6**	**63.2**	**65.3**	35.1	**40.1**	**47.9**	**50.4**	53.9	**42.5**	**49.9**	**55.8**	**61.1**	**62.3**

Obviously, the experimental results show that FDML achieves better performance than other methods, especially in terms of 2-shot and 3-shot with the most significant advantages. While adopting split 1, the values of mAP associated with FDML are the greatest under all shot settings. Specifically, FDML improves mAP by 3.4% compared with the best classical method FPD under 3-shot. Furthermore, the algorithm achieves the highest or sub-optimal detection accuracy in the split 2, in which it has a significant mAP improvement of 3.3%, 3.0%, 1.3% respectively than the best result of other methods, under K={2, 3, 5} shot. As for the split 3 condition, our approach demonstrates enhanced mAP by a margin of 2.2% in comparison to the optimal algorithm (i.e., Meta DETR) under 5-shot setting. At the same time, it can be concluded that the method maintain great performance while increasing the sample size gradually. However, when the sample size is extremely scarce, such as 1-shot, the propose method is not as effective as some methods like Meta-DETR. This is because the designed SQFM may be difficult to capture more class representative features and the feature fusion of HSFF is limited, which make detection challenging. Overall, these results demonstrate that FDML is more apt for few-shot object detection.

In order to verify the applicability and superiority of proposed method in fault detection, our research conducts comparative experiments with the latest few-shot learning methods (especially meta-learning method), e.g., Meta R-CNN, FSDetView, FSCE and FPD [20, 23, 28, 30]. It should be noted that, via applying the idea of model migration, the model trained on the common dataset (i.e., base classes) is directly used for fine-tuning to be adapted to the fault domain (i.e., novel classes).

Table 4. Comparison experiment in InsPLAD-fault.

Method/shot	5	10	20
Meta R-CNN	29.5	45.8	53.1
FsDetView	27.3	34.9	40.6
FSCE	42.1	56.4	61.1
FPD	39.0	54.3	58.9
FDML (ours)	43.9	58.1	62.6

Without loss of generality, Table 4 lists the quantitative results of different methods on InsPLAD-fault, from which one can see that FDML exhibits the most superior performance among mainstream methods under all conditions of sample size. This indicate that FDML is more apt for fault detection tasks of transmission line. To be specific, FDML respectively achieves 43.9%, 58.1%, 62.6% average precision under K={5, 10, 20} shot, which outperforms the above methods by 1.8%-16.6%, 1.7%-23.2%, and 1.5%-20.0%, respectively. So, these experimental results demonstrate that FDML can not only extract useful features, but also thoroughly explore the relations between features for model predicting. At the same time, these significant improvements verifies that the improved method can be better adapted to specific scenes than other methods, i.e., has better generalization performance. Overall, FDML has strong engineering application value.

4.4 Ablation Study

To validate the effectiveness of two improved modules proposed in this paper, corresponding ablation experiments are conducted via enabling them progressively. Table 5 quantifies the effect of each module.

Table 5. Results of ablation experiment in PASCAL VOC.

Set	Novel Split 1					Novel Split 2					Novel Split 3				
Method/shot	1	2	3	5	10	1	2	3	5	10	1	2	3	5	10
Baseline	39.3	51.2	54.3	58.8	60.7	30.2	33.9	42.3	44.6	48.9	36.1	44.3	48.3	54.5	55.4
+SQFM	43.7	53.2	57.8	62.8	64.4	34.9	37.3	47.2	49.8	52.7	38.1	47.5	51.7	57.9	58.2
+HSFF	40.7	51.7	55.1	61.6	63.8	31.2	34.6	44.1	47.8	50.3	39.9	46.4	49.0	56.9	57.7
Both	**45.1**	**56.4**	**61.6**	**63.2**	**65.3**	**35.1**	**40.1**	**47.9**	**50.4**	**53.9**	**42.5**	**49.9**	**55.8**	**61.1**	**62.3**

It is clear that, all the metrics are significantly improved after introducing SQFM. Specifically, compared to the baseline, the values are respectively improved by 4.0%, 5.2% and 3.4% of three splits under 5-shot setting. As for 10-shot setting, the detection mAP increased by 3.7%, 3.8%, 2.8%, respectively. Also, the metrics of the rest shot settings are significantly improved compared with the baseline, demonstrating that SQFM can not only effectively extract prototype features, but also assign them to the query branch for highlighting class-specific features. To further illustrate the effectiveness of SQFM, the feature heat map visualization technology Grad-CAM is adopted to display the feature assignment and provide a more intuitive understanding, as shown in Figure 9. Comparing the second and third column, one can find that the representative feature can be enhanced through the class-specific prototype assignment. So, this directly proves that SQFM is beneficial to the model prediction.

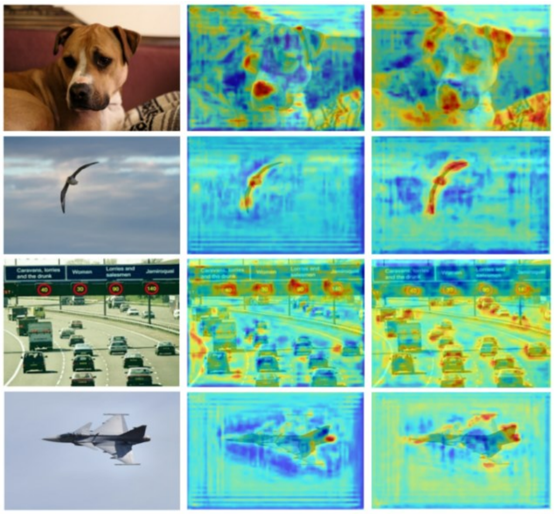

Fig. 9. The visualization of SQFM in which the first column is original image, the second column represents query feature, and the third column represents feature through SQFM.

After incorporating the HSFF alone, it can be observed that the larger the sample size, more significant the improvement of the proposed method. Specifically, the metrics are improved by 3.1%, 1.4% and 2.3% respectively, under 10-shot setting. This indicates that RoI features and prototype features can be fused effectively via combining multiple fusion methods through HSFF before the final prediction. Due to insufficient feature fusion caused by too few samples, the improvement under 1-shot and 2-shot conditions are very limited. This is because the randomness dominated by sample personality characteristics is too high.

Up until now, we have separately demonstrated the validity of two sub-modules. The last row in table 6 is the combination of two modules. Obviously, the detection performance is greater than that of any module alone which verifies that two modules do not conflict with each other. Compared to the baseline, the values are respectively improved by 4.6%, 5.0%, and 6.9% under 10-shot condition. In summary, the SQFM has a more obvious improvement on detection performance and we can further enhance the effect while adopting two modules together.

In order to further analyze the effectiveness of each module in transmission line fault detection, we conducted several ablation experiments on the dataset InsPLAD-fault. Four models are established, which are the baseline, the baseline with SQFM, the baseline with HSFF and the FDML (i.e., with both two modules). The ablation experimental results on the test set are correspondingly presented in Table 6.

Table 6. Results of ablation experiment in InsPLAD-fault.

Baseline	SQFM	HSFF	shot		
			5	10	20
√			38.2	50.9	58.4
√	√		39.5	52.4	60.5
√		√	40.2	56.9	61.2
√	√	√	43.9	58.1	62.6

Obviously, from the second line, one can be seen that the detection accuracy can be improved by 1.3%, 1.5% and 2.1% under K={5, 10, 20} shot, respectively. So, this directly demonstrates that the SQFM can effectively extract the prototypes of fault to highlight the representative features as well. Comparing the first and third rows in Table 6, the values are respectively improved by 2.0%, 6.0% and 2.8% after adopting HSFF. This is because HSFF can fuse high-level features for exploring the relations between them. The last row denotes that two modules are integrated. It is clear that, the metrics in it are all greater compared to the other rows. Specifically, FDML increases mAP_{50} by 5.7%, 7.2% and 4.2%, respectively. As a result, this verifies that the strategies in these modules are reasonable and our model improve the most under 10-shot in fault detection. In conclusion, in the field of transmission line fault detection, the SQFM can slightly improve the detection performance and the HSFF has a more significant improvement, which is more suitable for fault detection scenarios.

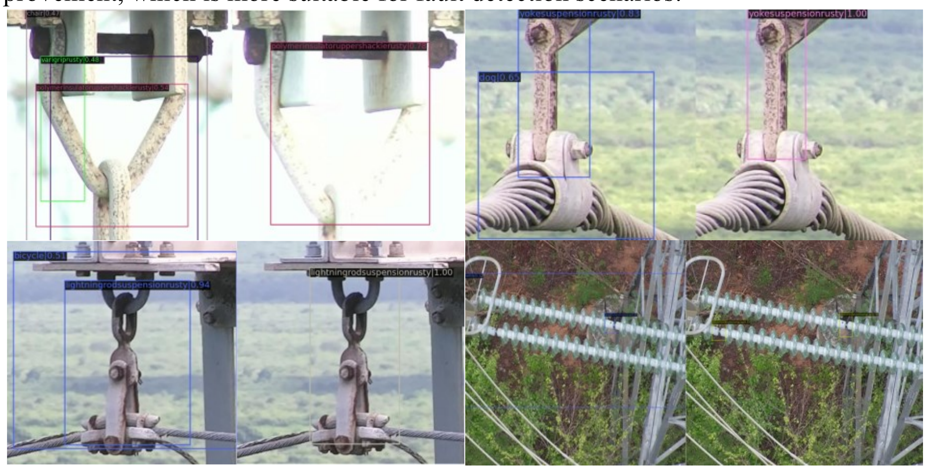

Fig. 10. Detection results.

4.5 Detection Result Visualization

To provide a more intuitive presentation of the effectiveness of the improved algorithm on fault detection, the detection results before and after improvement are visualized and compared, as shown in Figure 10. The left column represents the detection results of baseline model while the right column denotes the detection results of FDML.

As can be seen, there are a large number of false alarms in the left column, which proves that the it is difficult for the baseline algorithm to distinguish between the different fault classes. By contrast, FDML has the ability of accurate recognition. At the same time, it is evident that our model exhibits significantly improved confidence scores compared to the baseline for all types of faults. Of cause, we will focus on designing modules for glass insulator missing cap which is very different from other faults.

5 Conclusion

In this paper, we propose an improved few-shot fault detection method based on meta-learning. First, a two-stage meta-learning training method is introduced to address the problem of distribution difference in data domains. Second, a support and query feature matching module (SQFM) is proposed to optimize the use of support features to assist detection via highlighting class-specific representative features. Moreover, a high-level semantic feature fusion module (HSFF) is designed to further integrate high-level feature before model prediction via combining the four feature fusion ways along with identity. Experimental results demonstrate the effectiveness and superiority of FDML compared to other mainstream few-shot object detection (FSOD) methods. On average, the mAP_{50} values are respectively 5.7%, 7.2% and 4.2% higher than the baseline under K= {5, 10, 20} shot.

6 Limitation

As for future work, we will make the attempt to explore the seperation of localization and classification tasks while adopting UAV images which are large in size. At the same time, more open-source fault detection data from transmission lines will be introduced for training to develop a model with improved generalization and adaptability.

Acknowledgments. This research was funded by National key R&D Program of China, Project Number 2023YFC3006700, Topic Five Number: 2023YFC3006705.

References

1. Liu, C.; Wu, Y. Research Progress of Vision Detection Methods Based on Deep Learning for Transmission Lines. Proceedings of the CSEE. 2023, 43, 7423-7445.
2. Panahi, H.; Zamani, R.; Sanaye-Pasand, M.; Mehrjerdi, H. Advances in transmission network fault location in modern power systems: Review, outlook and future works. IEEE Access. 2021, 9, 599-615.
3. Kanwal, S.; Jiriwibhakorn, S. Artificial intelligence based faults identification, classification, and localization techniques in transmission lines-a review. IEEE Lat. Amer. Trans. 2023, 21, 1291-1305.

4. 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. 2020, 50, 1486-1498.
5. Zhou, M.; Li, B.; Wang, J.; He, S. Fault detection method of glass insulator aerial image based on the improved YOLOv5. IEEE Trans. Instrum. Meas.. 2023, 72, 1-10.
6. Niu, S.; Zhou, X.; Zhou, D.; Yang, Z.; Liang, H.; Su, H. Fault Detection in Power Distribution Networks Based on Comprehensive- YOLOv5. Sensors 2023, 23, 6410.
7. Hao, K.; Chen, G.; Zhao, L.; Li, Z.; Liu, Y.; Wang, C. An insulator defect detection model in aerial images based on multiscale feature pyramid network. IEEE Trans. Instrum. Meas. 2022, 71, 1-12.
8. Yi, W.; Ma, S.; Li, R. Insulator and defect detection model based on improved yolo-s. IEEE Access 2023, 11, 215-226.
9. Wang, T.; Zhai, Y.; Li, Y.; Wang, W.; Ye, G.; Jin, S. Insulator Defect Detection Based on ML-YOLOv5 Algorithm. Sensors 2024, 24, 204.
10. Hao, S.; An, B.; Ma, X.; Sun, X.; He, T.; Sun,S. PkAMNet: A transmission line insulator parallel- gap fault detection network based on prior knowledge transfer and attention mechanism. IEEE Trans. Power. Deliv . 2023, 38, 3387-3397.
11. Liu, Y.; Zhang, H.; Zhang, W.; Lu, G.; Tian, Q.; Ling, N. Few-Shot Image Classification: Current Status and Research Trends. Electronics. 2022, 11, 1752.
12. Zhang, G.; Luo, Z.; Cui, K.; Lu, S.; Xing, E. Meta-DETR: Image-level few-shot detection with inter-class correlation exploitation. IEEE Trans. Pattern Anal. Mach. Intell . 2023, 45, 832-843.
13. Zha, Z.; Tang, H.; Sun, Y.; Tang, J. Boosting few-shot fine-grained recognition with background suppression and foreground alignment. IEEE Trans. Circuits Syst. Video Technol.. 2023, 33, 3947-3961.
14. Li, R.; Zeng, Y.; Wu, J.; Wang, Y.; Zhang, X. Few-shot object detection of remote sensing image via calibration. IEEE Geos. Remote Sens. Letters . 2022, 19, 1-5.
15. Lu, Z.; Li, Y.; Shuang, F.; Han, C. InsDef: Few-shot learning-based insulator defect detection algorithm with a dual-guide attention mechanism and multiple label consistency constraints. IEEE Trans. Power. Deliv . 2023, 38, 4166-4178.
16. Wang, Z.; Gao, Q.; Li, D.; Liu, J.; Wang, H.; Yu, X.; Wang, Y. Insulator anomaly detection method based on few-shot learning. IEEE Access. 2021, 9, 970-980.
17. 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.
18. Silva, A.; Felix, H.; Simoes, F.; Teichrieb, V.; Santos, M.; Santiago, H.; Sgotti, V.; Neto, H. InsPLAD: A dataset and benchmark for power line asset inspection in uav images.Internationnal. Journal. Remote Sens. 2023, 44, 7294-7320.
19. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards real-time object detection with region proposal networks. IEEE Trans. Pattern Anal. Mach. Intell . 2017, 39, 1137-1149.
20. Yan, X.; Chen, Z.; Xu, A.; Wang, X.; Liang, X.; Lin, L. Meta R-CNN: Towards general solver for instance-level low-shot learning. IEEE International Conference on Computer Vision. 2019, 9576-9582.
21. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. IEEE Conference on Computer Vision and Pattern Recognition. 2016, 770-778.
22. Lin, T.; Dollár, P.; Girshick, R.; He, K.; Hariharan, B.; Belongie, S. Feature pyramid networks for object detection. IEEE Conference on Computer Vision and Pattern Recognition. 2017, 936-944.

23. Wang, Z.; Yang, B.; Yue, H. Fine-Grained Prototypes Distillation for Few-Shot Object Detection. Proceedings of AAAI Conference on Artificial Intelligence. 2024.
24. Carion, N.; Massa, F.; Synnaeve, G. End-to-end object detection with transformers. European Conference on Computer Vision. 2020, 213-229.
25. Han, G.; Huang, S.; Ma, J. Meta Faster R-CNN: Towards Accurate Few-Shot Object Detection with Attentive Feature Alignment. Proceedings of AAAI Conference on Artificial Intelligence. 2024, 780-789.
26. Kang, B.; Liu, Z.; Wang, X.; Yu, F.; Feng, J.; Darrell, T. Few-shot object detection via feature reweighting. IEEE International Conference on Computer Vision. 2019, 8419-8428.
27. Wang, X.; Thomas, E.; Darrell, T. Frustratingly Simple Few-Shot Object Detection. International Conference on Machine Learning. 2020.
28. Xiao, Y.; Lepetit, V.; Marlet, R. Few-shot object detection and viewpoint estimation for objects in the wild. IEEE Trans. Pattern Anal. Mach. Intell . 2023, 45, 3090-3106.
29. Wu, J.; Liu, S.; Huang, D.; and Wang, Y. Multi-scale positive sample refinement for few-shot object detection. European Conference on Computer Vision. 2020, 456-472.
30. Sun, B.; Li, B.; Cai, S.; Yuan, Y.; Zhang, C. FSCE: Few-shot object detection via contrastive proposal encoding. IEEE Conference on Computer Vision and Pattern Recognition. 2021, 7352-7362.
31. Finn, C.; Abbeel, P.; Levine, Sergey. Model-agnostic Meta-Learning for Fast Adaptation of Deep Networks International Conference on Machine Learning. 2017, 2640-3498.