Journal of Information Hiding and Multimedia Signal Processing ©2025 ISSN 2073-4212
Ubiquitous International Volume **16**, Number **4**, December **2025**

Attention-Aware Prompting for Learning New Faults on Insulators with Limited Data

Anh Trinh Hien

Institute of Information Technology
Vietnam Academy of Science and Technology
Hanoi City, Vietnam
hienanh@ioit.ac.vn

Tao Ngo Quoc*

Institute of Information Technology
Vietnam Academy of Science and Technology
Hanoi City, Vietnam
nqtao@ioit.ac.vn

Thanh-Tan Nguyen Thi

Electric Power Univiersity
Hanoi City, Vietnam
tanntt@epu.edu.vn

*Corresponding author: Tao Ngo Quoc

Received July 3, 2025, revised October 26, 2025, accepted October 28, 2025.

Abstract. *Monitoring and maintaining insulators is a critical task to ensure the safe and stable operation of electrical grid systems. However, traditional deep learning-based fault recognition methods typically struggle when faced with learning new types of faults with limited samples without degrading performance on previously learned classes—a challenge known as catastrophic forgetting. This phenomenon is further exacerbated in few-shot scenarios, where models are particularly prone to overfitting. To address these challenges, we propose framing the insulator fault recognition problem as a Few-Shot Class-Incremental Learning (FSCIL) task. In this paper, we introduce a novel framework, Attention-aware Adaptive Prompting for Insulator Diagnosis (AAPID), inspired by the Attention-aware Self-adaptive Prompt (ASP) method. Our approach utilizes a pre-trained Vision Transformer as a visual encoder and keeps its parameters fixed to avoid overfitting, thus leveraging its strong generalization capabilities. The core of AAPID is an adaptive prompting system, consisting of: (1) task-invariant prompts designed to capture general, shareable knowledge about insulator characteristics, and (2) task-specific adaptive prompts to encode discriminative information for each specific fault type. This facilitates effective knowledge transfer from known classes to newly introduced ones, even when training data is extremely limited. Experimental results on our constructed insulator image dataset demonstrate that our method effectively learns new fault classes with only a few training samples while maintaining high performance on previously learned fault classes, significantly outperforming state-of-the-art FSCIL and CIL methods.*
Keywords: Few-shot learning, Class-Incremental Learning, Insulators identification

---

1. **Introduction.** Insulators are essential components in electrical transmission and distribution systems [1],[21], serving the critical role of isolating and supporting power lines. Damage to insulators, such as cracks, breaks, or surface contamination, can cause electrical discharge incidents, disrupting power supply and threatening the safety of the entire grid. Therefore, early detection and monitoring of these faults

is paramount. Traditional manual inspection methods are time-consuming, costly, and pose risks to operational staff, leading to a promising shift toward automated inspection systems using computer vision and deep learning.

However, applying traditional deep learning models to this problem encounters two significant challenges. First, data regarding fault types is typically imbalanced and scarce. In practice, assembling a large and diverse dataset covering all possible faults on various insulator types is difficult. Second, the real world constantly changes; new types of insulators might be installed, or previously unseen fault patterns could emerge over time. An ideal model needs the ability to continuously learn new fault classes without retraining from scratch on the entire dataset—a task known as Class-Incremental Learning (CIL) [2]. The main challenge in CIL is catastrophic forgetting, wherein the model tends to lose knowledge of previously learned classes when trained on new ones.

Combining these two challenges creates an even more complex scenario, known as Few-Shot Class-Incremental Learning (FSCIL) [3]. In FSCIL, the model must learn new classes with only a few samples while still retaining knowledge of previously learned classes. This precisely reflects the problem encountered in insulator monitoring. Existing FSCIL methods typically fine-tune all neural network parameters, leading to overfitting on limited new samples and reducing the model's generalization capabilities on base classes. Conversely, recent prompt-based CIL methods, effective in preventing forgetting, require substantial amounts of data in incremental steps to train new task-specific prompts. This requirement makes them unsuitable for the FSCIL setting, where data for new tasks is extremely limited.

To address these limitations, we propose a new framework called Attention-aware Adaptive Prompting for Insulator Diagnosis (AAPID), strongly inspired by the Attention-aware Self-adaptive Prompt (ASP) method. Similar to ASP, our approach keeps the parameters of a pre-trained Vision Transformer (ViT) fixed to leverage its strong representational capabilities and avoid overfitting. The core of AAPID is an adaptive prompt system, divided into two components: (1) Task-Invariant Prompts (TIPs): These prompts are designed to learn general knowledge shared across all classes, such as basic shape, color, and structural features of insulators. (2) Self-adaptive Task-Specific Prompts (TSPs): These prompts are generated by a prompt encoder to capture distinctive identification information for each fault type. Critically, this encoder facilitates knowledge transfer from known classes to new classes, eliminating the need for substantial data to learn prompts for new tasks.

In summary, our contributions include:

- Being the first to formulate the real-world insulator fault recognition problem as an FSCIL task and analyze the limitations of existing methods in this application context.
- Proposing AAPID, a novel prompt-based framework designed to learn new fault classes with limited samples while minimizing catastrophic forgetting.
- Demonstrating through experiments on our constructed insulator image dataset that AAPID significantly outperforms state-of-the-art CIL and FSCIL methods in both learning new classes and maintaining performance on previously learned classes.

2. **Related Works.** Early automated methods for fault detection [4] in insulators relied primarily on traditional image processing techniques for manual feature extraction; however, these approaches were often sensitive to environmental variations. Recently, deep learning models, particularly Convolutional Neural Networks (CNNs) [5],[20], have demonstrated superior effectiveness. Nevertheless, these models inherently have limitations: they are designed for a fixed set of fault classes and require large, static datasets for training, lacking adaptability to new fault types emerging in real-world scenarios.

To overcome these limitations associated with static models, the field of Class-Incremental Learning (CIL) [6] has emerged, aiming to enable models to sequentially learn new classes without forgetting previously acquired knowledge. Major CIL methodologies include rehearsal-based approaches that store previous samples, regularization techniques designed to preserve important model parameters, and knowledge distillation methods [7] intended to transfer knowledge from older models. However, a common weakness across these approaches is the necessity for substantial data during each incremental learning task. In industrial monitoring practice, data on newly emerging faults is often scarce, creating an even more challenging scenario known as Few-Shot Class-Incremental Learning (FSCIL), which requires models to learn from only a few samples.

Current FSCIL [8] methods typically emphasize developing feature extractors with strong generalization capabilities or devising learning strategies to avoid overfitting on limited data. Nonetheless, most existing methods usually fine-tune all network parameters during base class training, potentially leading to overfitting on these classes and impeding the model's ability to transfer knowledge [9] to new classes. In contrast, our proposed Attention-aware Adaptive Prompting for Insulator Diagnosis (AAPID) method keeps the pretrained encoder parameters fixed to mitigate this risk. Thus, the primary differentiator of our work from these FSCIL methods is our use of a fixed encoder and an adaptive prompt system, which specifically avoids the overfitting and knowledge degradation common in full-network fine-tuning.

Recently, prompt-based approaches have become a promising direction within CIL, maintaining the parameters of large-scale models such as Vision Transformers (ViTs) [10] fixed while only fine-tuning a small set of new parameters known as "prompts." Methods such as L2P and DualP have effectively demonstrated the use of prompts to store task-specific knowledge. However, a significant limitation of these approaches in FSCIL contexts is their requirement for substantial data in incremental tasks to effectively train new prompts. With very limited data, training these prompts typically results in severe overfitting, rendering them unsuitable for FSCIL scenarios. Our proposed AAPID method addresses precisely this limitation by adaptively generating prompts without requiring extensive amounts of new data. Specifically, while methods like L2P and DualP require significant data to train new task-specific prompts, AAPID's key innovation is its ability to generate these prompts from the input itself via an encoder, making it uniquely suited for the data-scarce FSCIL scenario.

3. Proposed Method.

3.1. Problem Definition. The Few-Shot Class-Incremental Learning (FSCIL) problem in the context of insulator fault recognition can be formally defined as follows:

The model sequentially learns from a series of sessions $t \in \{0, 1, \ldots, T\}$, each associated with datasets $D_0, D_1, \ldots, D_T$. At learning session t, the model does not have access to data from previous sessions $\{0, \ldots, t-1\}$.

Each session t has a dataset:

$$D_t = \{(x_{t,i}, y_{t,i})\}_{i=1}^{N_t}, \tag{1}$$

where $x_{t,i}$ is an insulator image, and $y_{t,i} \in Y_t$ is the corresponding fault class label. The label sets of different sessions are disjoint:

$$Y_t \cap Y_{t'} = \emptyset \quad \text{for all} \quad t \neq t'. \tag{2}$$

Thus, each incremental learning session introduces completely new fault classes.

The initial learning session ($t = 0$) is called the base task. In this session, the model is trained on a large dataset D_0, consisting of abundant samples of common fault types (e.g., "Normal," "Minor cracks"). Subsequent sessions ($t > 0$) represent incremental tasks, formulated as an N-way K-shot problem. For instance, a 5-way 5-shot session indicates that the model must learn 5 new fault classes, each represented by only 5 images.

An FSCIL model typically consists of two main components: a feature extractor f_θ parameterized by θ, and a classifier h_ψ parameterized by ψ. The ultimate goal is that after each session t, for any input image x belonging to any class observed from session 0 to session t, the model accurately predicts the fault class label:

$$y = h_\psi(f_\theta(x)). \tag{3}$$

3.2. Model Architecture. To address the Few-Shot Class-Incremental Learning (FSCIL) problem in insulator fault recognition, we propose a novel framework called Attention-aware Adaptive Prompting for Insulator Diagnosis (AAPID). An overview of the AAPID architecture is depicted in Figure 1. Instead of fine-tuning the entire neural network, our method keeps the parameters of a pre-trained Vision Transformer (ViT) fixed and only learns a small set of parameters called "prompts." This approach leverages the robust generalization capabilities of the ViT model and prevents overfitting when learning from limited data.

The core of AAPID is an adaptive prompting system, separated into two primary components designed to learn distinct types of knowledge: task-invariant prompts and task-specific prompts. Task-invariant prompts, labeled as "Prompt init" in Figure 1, are crafted to capture general, shareable knowledge across all classes. In our problem setting, these represent common insulator characteristics unrelated to specific fault types, such as circular shape, ceramic material, and basic color. Inspired by the ASP paper, to ensure these prompts remain truly invariant, all tokens within the prompt are initialized with the same values. This encourages the attention mechanism of the ViT to consistently process these prompts across different tasks, thereby encapsulating only non-specific fault information.

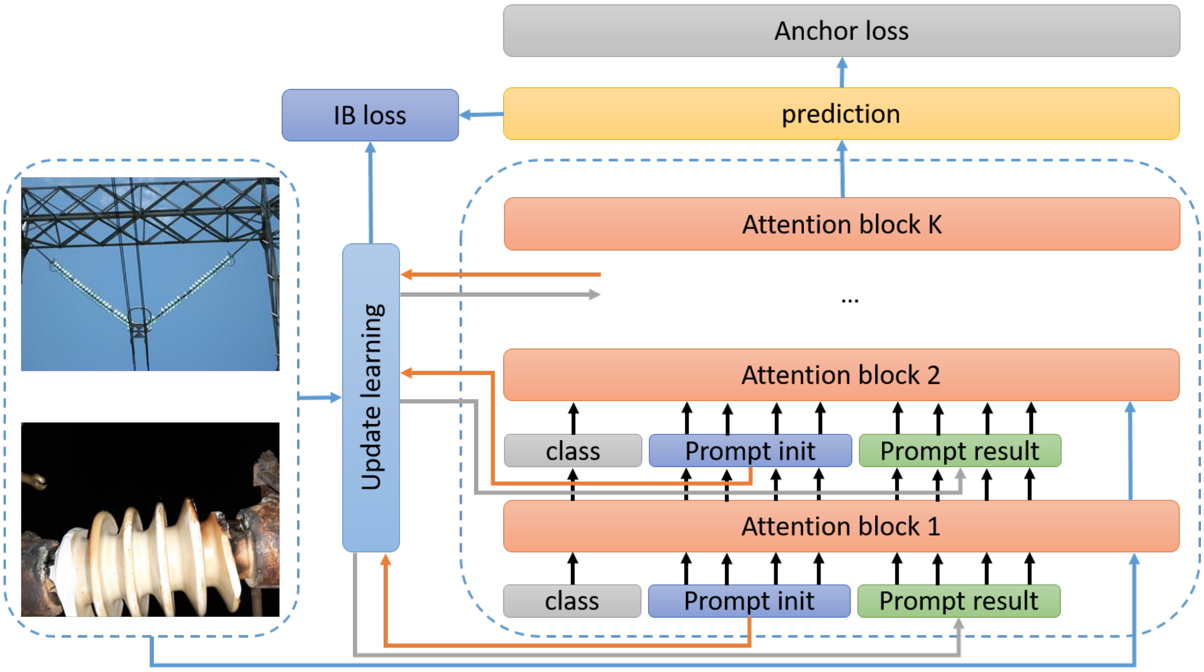

FIGURE 1. Overview diagram of the architecture of the proposed AAPID method. The model consists of attention blocks from ViT, a prompt learning module ("Update learning"), and is optimized by two loss functions, IB loss and Anchor loss.

To distinguish various fault types, the model must grasp task-specific information. Rather than training new prompts for each fault type with extensive data, we employ an "Update learning" module (prompt encoder) to automatically generate task-specific prompts ("Prompt result" in Figure 1) from an input image.

To ensure the prompt encoder generalizes effectively to unseen fault types, we adopt the Information Bottleneck (IB) principle. The encoder training is optimized by an IB loss function, aiming to: (1) maximize mutual information between the generated prompts and fault labels (embedding semantic fault information), and (2) minimize mutual information between the prompts and the input data (excluding irrelevant noise from the images). This directs prompts to focus on core fault characteristics, enhancing recognition performance for classes with limited samples.

Alongside the IB loss, we implement an additional similarity-based loss called Anchor Loss to enhance differentiation between fault classes. After passing through attention blocks, the model produces a feature vector for each input image. Anchor Loss aims to cluster feature vectors of images belonging to the same fault class closely together and near the class prototype in the feature space, while simultaneously pushing prototypes of different classes apart. This yields clear, distinct feature clusters for each fault class, facilitating more effective performance of the final classification layer ("prediction").

Ultimately, the overall loss function of the AAPID model is a weighted combination of the IB loss and Anchor Loss, enabling the model to effectively generate meaningful prompts and robustly discriminate between different fault classes.

3.3. **Loss function.** The training process of the AAPID model is guided by a composite loss function consisting of two main components: Information Bottleneck (IB) Loss to learn highly generalizable prompts and Anchor Loss to optimize the discriminative capability of the feature space.

3.3.1. *Information Bottleneck (IB) Loss.* To ensure that the task-specific prompts (TSP) generated by the prompt encoder contain useful information regarding fault classes while eliminating unnecessary noise from input images, we employ the Information Bottleneck (IB) principle. The objective of IB is to achieve an optimal balance between compressing input data and retaining important predictive information. The IB objective function can be represented as follows:

$$L_{IB} = I(P; X) - \gamma I(P; y) \tag{4}$$

where $I(P;X)$ is the mutual information between prompt P and input image X, and $I(P;y)$ is the mutual information between prompt P and fault label y. Minimizing this loss is equivalent to minimizing the information from input images retained by prompts while maximizing label-relevant information.

Since direct computation of mutual information is infeasible, we utilize a variational lower bound approximation. The practical IB loss we minimize is formulated as:

$$L_{IB} = E[-\log P(y|x)] + \beta_{KL} \cdot KL(P(p|x)||r(p)) \tag{5}$$

This loss function consists of two components:

- *Classification Loss:* The first component, $E[-\log P(y|x)]$, is the standard cross-entropy loss, ensuring accurate prediction of fault labels from input images.
- *KL Divergence Regularization:* The second component is the Kullback-Leibler (KL) divergence between the prompt distribution $P(p|x)$ generated by the encoder and a prior distribution $r(p)$ (usually Gaussian). This component acts as a regularization mechanism, preventing prompts from encoding excessive irrelevant information from specific samples, thereby enhancing generalization.

3.3.2. *Anchor Loss.* To further enhance the discriminative capability of the model, we employ an additional similarity-based loss termed Anchor Loss. The goal of this loss function is to minimize intra-class feature distances and maximize inter-class feature distances. Specifically, we aim to bring features of images belonging to the same fault class closer to their class prototype.

The Anchor Loss is defined based on cosine similarity as:

$$L_{anchor} = 1 - \frac{\hat{c}_k^T f_{\theta,p}(x_k)}{||\hat{c}_k|| \cdot ||f_{\theta,p}(x_k)||} \tag{6}$$

where $f_{\theta,p}(x_k)$ is the feature vector of image x_k belonging to fault class k. Instead of calculating the class centroid c_k using the entire dataset at each step, we employ a more efficient approach by selecting an anchor sample $\hat{x}_k$ at the beginning of each epoch. This anchor sample is the one most similar to the true class centroid, and its feature vector $\hat{c}_k = f_{\theta,p}(\hat{x}_k)$ serves as an estimated class centroid for loss computation.

3.3.3. *Overall Loss Function.* The final loss function for training the AAPID model is a weighted combination of the above two components:

$$L_{total} = L_{IB} + \lambda L_{anchor} \tag{7}$$

4. **Experiments.**

4.1. **Dataset.** To comprehensively evaluate the effectiveness of our proposed AAPID method, we conducted experiments on two datasets: a standard benchmark dataset, CIFAR100 [11], for comparison with state-of-the-art methods, and a practical application dataset constructed by our team, named VPI-Fault-2500, to validate real-world applicability.

The first dataset, CIFAR100, is widely used as a standard in image classification tasks, consisting of 60,000 color images sized 32×32 pixels across 100 distinct classes. Following the standard setup in FSCIL research, we divided these 100 classes into 60 base classes used for initial model training. The remaining 40 classes were allocated to 8 consecutive incremental learning sessions, each configured as a 5-way 5-shot task. This arrangement allows assessment of the model's capability in continuously learning a sequence of few-shot tasks and its resilience against catastrophic forgetting.

In addition to the standard dataset, we constructed the VPI-Fault-2500 dataset to evaluate the model's effectiveness in practical scenarios. This dataset contains 2500 insulator images classified into five labels: Normal, Surface Crack, Rim Chip, Body Breakage, and Severe Surface Contamination, with 500 images per label. We simulated the FSCIL scenario by dividing this dataset into three base classes (Normal, Surface Crack, Rim Chip) and a single incremental learning session. The incremental session involves the remaining two rare fault classes (Body Breakage and Severe Surface Contamination), configured as a 2-way 5-shot task, meaning the model learns from only 5 images per new fault class. This scenario realistically emulates situations where a system, initially trained on common faults, must rapidly adapt to recognizing new fault types with limited examples.

4.2. **Experiment setup.** In this section, we detail the research questions, evaluation metrics, comparative methods, and implementation settings used to assess the effectiveness of the proposed AAPID approach.

4.2.1. *Research Questions (RQs).* Our experiments are designed to address three key research questions:

- *RQ1:* How does the overall performance of our proposed AAPID compare against state-of-the-art CIL and FSCIL methods on both benchmark and practical application datasets?
- *RQ2:* To what extent does each component of the AAPID architecture (Task-Invariant Prompts, Self-adaptive Task-Specific Prompts, and Anchor Loss) contribute to the model's final performance?
- *RQ3:* Does the AAPID method effectively learn a feature space in which different fault classes (both base and few-shot classes) are distinctly separated?

4.2.2. *Evaluation Metrics.* We utilize standard metrics in FSCIL research for comprehensive performance evaluation:

- *Average Accuracy (A_{avg}):* The mean accuracy across all classes observed throughout all learning sessions ($t = 0, \ldots, T$). This metric assesses the model's overall performance.
- *Performance Dropping (PD):* The decline in performance, calculated as $A_0 - A_T$, where A_0 is the accuracy on base classes after the first session and A_T is the accuracy on the same classes after the final session. This metric quantifies catastrophic forgetting.
- *Harmonic Accuracy (HAcc):* The harmonic mean between the accuracy on base classes (A_o) and the average accuracy on new classes (A_n) after the final session. This metric evaluates the balance between learning new classes and retaining knowledge of old classes.

4.2.3. *Comparative Methods (Baselines).* We compare AAPID against three main categories of methods to provide a comprehensive view:

- *Classic CIL methods:* iCaRL [12], Foster [13].
- *Advanced FSCIL methods:* CEC [14], FACT [15], TEEN [16].
- *Prompt-based CIL methods:* L2P [17], DualP [18], CodaP [19]. Since these methods were initially not designed for few-shot scenarios, we also implemented enhanced versions (L2P+, DualP+, CodaP+) by replacing their linear classifiers with prototypical networks, a technique demonstrated to significantly improve performance.

4.2.4. *Implementation Details.* All experiments were conducted using PyTorch. We used the pre-trained ViT-B/16 model on ImageNet-1K as the backbone encoder. Input images were resized to 224×224 pixels. The model was trained for 20 epochs in the base learning session using SGD as the optimizer, with a learning rate set to 0.01 for both datasets. Batch size was set to 48. The prompt length was 3 for CIFAR100 and 5 for VPI-Fault-2500. The hyperparameter λ in the loss function was selected through cross-validation.

4.2.5. *Overall Performance Comparison (RQ1).* Tables 1 and 2 present detailed performance comparisons between AAPID and other methods on the CIFAR100 and VPI-Fault-2500 datasets.

On the CIFAR100 dataset (Table 1), AAPID demonstrates superior performance across all primary metrics. Specifically, the Average Accuracy (A_{avg}) of AAPID is the highest at 89.5%, indicating the best overall performance across all learning sessions. Critically, AAPID has the lowest Performance Dropping (PD) value of 4.8%, highlighting its exceptional capability to mitigate catastrophic forgetting. Moreover, the highest Harmonic Accuracy (HAcc) of 86.1% confirms an excellent balance between retaining old knowledge and acquiring new knowledge. Original prompt-based CIL methods (L2P, DualP, CodaP) exhibit significantly lower HAcc values, verifying their unsuitability for few-shot scenarios.

On the VPI-Fault-2500 dataset (Table 2), a similar trend is more pronounced. AAPID achieves an A_{avg} of 94.2% and an HAcc of 92.5%, significantly outperforming the second-best method, TEEN. This clearly indicates that AAPID is not only effective on benchmark datasets but also highly efficient for practical applications in insulator fault recognition, where the capability to learn from limited examples is critical.

4.2.6. *Ablation Study (RQ2).* To clearly understand the role of each component, we conducted an ablation study by systematically removing individual modules from AAPID. The results are presented in Table 3. Removing any component resulted in decreased performance on the A_avg metric for both datasets, validating the rationale behind our design and indicating that the components mutually support each other. Notably, omitting the Task-result Prompts (w/o TRP) caused the most significant performance drop, emphasizing its crucial role in learning new classes.

TABLE 1. Performance comparison on CIFAR100 dataset (setup: 60+8x5-way 5-shot)

Method	A_{avg}(%)↑	PD(%)↓	HAcc(%)↑
iCaRL	78.4	27.2	57.5
Foster	73.9	35.9	11.0
CEC	81.0	19.0	64.1
FACT	78.6	21.7	55.5
TEEN	87.3	8.8	81.2
L2P	71.1	37.0	0.0
DualP	70.6	36.9	0.1
CodaP	72.0	37.4	0.0
DualP+	81.1	8.5	75.3
AAPID (Ours)	**89.5**	**4.8**	**86.1**

TABLE 2. Performance comparison on VPI-Fault-2500 dataset (setup: 3+2-way 5-shot)

Method	A_{avg}(%)↑	PD(%)↓	HAcc(%)↑
iCaRL	85.1	15.4	78.2
Foster	82.6	20.1	65.7
TEEN	91.5	6.2	88.3
DualP+	88.9	8.0	85.1
AAPID (Ours)	**94.2**	**3.1**	**92.5**

TABLE 3. Ablation Study Results on Average Accuracy (A_avg) (%)

Method	CIFAR100	VPI-Fault-2500
AAPID (Full model)	89.5	94.2
w/o Task-init Prompts (TIP)	88.9	93.6
w/o Task-result Prompts (TRP)	87.8	92.1
w/o Anchor Loss	88.7	93.4
w/o IB Loss	88.2	92.9

4.2.7. *Quality Study (RQ3).* To address the third research question (RQ3), "Does the AAPID method effectively learn a feature space in which different fault classes are distinctly separated?", we employed the t-SNE technique for dimensionality reduction and visualization of the features extracted from the test samples (Query Set) on the VPI-Fault-2500 dataset. The figure 2 above provides a visual comparison between the feature spaces learned by a baseline method and our AAPID method.

Baseline Feature Space (Left Column): The baseline method results in a chaotic and ineffective feature space. In the Query Set plot, data points from all five classes (represented by five different colors) are densely mixed and indistinguishable, forming a large, central cluster. There is no evident structure or separation among classes. This indicates that the baseline model completely failed to learn discriminative features, resulting in an inability to differentiate between base classes and new few-shot classes. This scenario vividly illustrates catastrophic forgetting and the baseline's incapability of learning effectively from limited data.

AAPID Feature Space (Right Column): In stark contrast, our AAPID method yields a highly structured and well-organized feature space. In the Query Set plot, the five distinct classes are clearly separated into five discrete and coherent clusters. Each cluster is dense, homogeneous in color, and maintains a clear margin from other clusters. This demonstrates that AAPID not only preserves knowledge of previously learned classes but also effectively captures unique features for new classes using only a few examples, subsequently organizing them systematically within the feature space.

This visual analysis strongly affirms RQ3, demonstrating that AAPID indeed learns a superior feature space, where classes are distinctly represented with high separability. The quality of this feature space fundamentally supports the impressive quantitative performance reported in previous result tables.

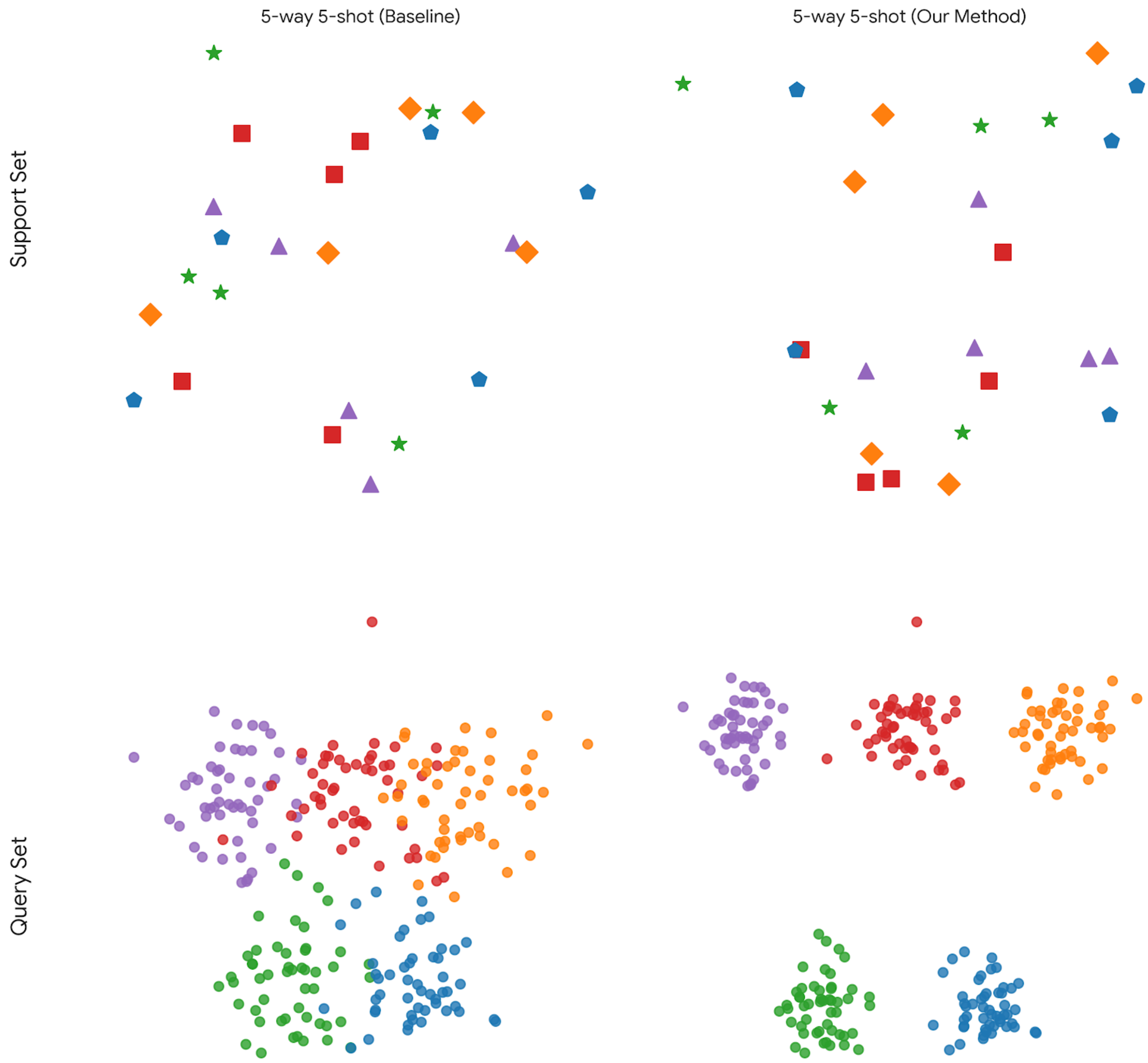

FIGURE 2. The t-SNE visualization of ablation study on VPI-Fault-2500.

5. **Conclusions.** In this work, we highlighted critical challenges inherent to practical insulator fault recognition: the scarcity of data for rare fault types and the limited adaptability of traditional deep learning models when learning new fault classes. We formalized this problem within the framework of Few-Shot Class-Incremental Learning (FSCIL) and analyzed the limitations of existing Class-Incremental Learning (CIL) and FSCIL methods in this specific context.

To address these challenges, we proposed Attention-aware Adaptive Prompting for Insulator Diagnosis (AAPID), a novel prompt-based incremental learning framework inspired by the ASP method. By keeping the parameters of a pre-trained Vision Transformer model fixed and only learning an adaptive prompting system—comprising task-invariant prompts to capture general knowledge and task-specific prompts generated by an encoder guided by the Information Bottleneck principle—our method effectively learns new classes with just a few samples without needing to retrain the entire network.

Experimental results conducted on both the benchmark dataset CIFAR100 and our constructed practical application dataset VPI-Fault-2500 demonstrated the superior performance of AAPID compared to various advanced CIL and FSCIL methods. Quantitative analysis through performance tables and qualitative analysis via t-SNE visualization of feature spaces consistently showed AAPID's exceptional capability to learn new fault classes while robustly retaining knowledge of previously learned classes. Our study not only provides an effective solution for insulator monitoring but also establishes a promising pathway for applying FSCIL to a broader range of industrial inspection problems where data on novel fault classes is inherently limited.

While our proposed AAPID method demonstrates strong performance, we acknowledge its limitations. First, the method's effectiveness is closely tied to the quality of the pre-trained ViT backbone. As the encoder parameters are kept fixed to mitigate catastrophic forgetting , a significant domain gap between the pre-training data (e.g., ImageNet) and the specific target domain of insulator faults could potentially limit the upper bound of performance. The adaptive prompts are designed to bridge this gap, but their effectiveness may be constrained by the quality of the frozen features. Second, from a practical deployment perspective, while AAPID is highly efficient during training, the inference phase still requires a full computational pass through the large-scale ViT model. This inference cost might be a consideration for real-time monitoring systems with constrained hardware resources. Future work could explore domain-specific pre-training for the encoder or investigate lightweight, efficient ViT architectures as an alternative backbone.

Acknowledgment. This work is partially supported by the project code NVCC02.06/25-25

REFERENCES

[1] Borghei, Moein, and Mona Ghassemi. "Insulation materials and systems for more-and all-electric aircraft: A review identifying challenges and future research needs." IEEE Transactions on Transportation Electrification 7.3 (2021): 1930-1953.
[2] Chen, Xiuwei, and Xiaobin Chang. "Dynamic residual classifier for class incremental learning." Proceedings of the IEEE/CVF International Conference on Computer Vision. 2023.
[3] Hersche, Michael, et al. "Constrained few-shot class-incremental learning." Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2022.
[4] Singh, Laxman, et al. "Design of thermal imaging-based health condition monitoring and early fault detection technique for porcelain insulators using Machine learning." Environmental Technology & Innovation 24 (2021): 102000.
[5] Zhang, Xingtuo, et al. "InsuDet: A fault detection method for insulators of overhead transmission lines using convolutional neural networks." IEEE Transactions on Instrumentation and Measurement 70 (2021): 1-12.
[6] Wang, Siyuan, et al. "Fault Recognition Method for Substation Equipment Based on Lifelong Learning." 2023 International Conference on Image Processing, Computer Vision and Machine Learning (ICICML). IEEE, 2023.
[7] Khan, Hikmat, Nidhal Carla Bouaynaya, and Ghulam Rasool. "Brain-inspired continual learning: Robust feature distillation and re-consolidation for class incremental learning." IEEE Access (2024).
[8] Zhang, Jinghua, et al. "Few-Shot Class-Incremental Learning for Classification and Object Detection: A Survey." IEEE Transactions on Pattern Analysis and Machine Intelligence (2025).
[9] Angriawan, Muhamad. "Transfer Learning Strategies for Fine-Tuning Pretrained Convolutional Neural Networks in Medical Imaging." Research Journal of Computer Systems and Engineering 4.2 (2023): 73-88.
[10] Thisanke, Hans, et al. "Semantic segmentation using Vision Transformers: A survey." Engineering Applications of Artificial Intelligence 126 (2023): 106669.
[11] Chi, Zhixiang, et al. "Metafscil: A meta-learning approach for few-shot class incremental learning." Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2022.
[12] Ye, Zejun, et al. "CtF: Mitigating Visual Confusion in Continual Learning Through a Coarse-To-Fine Screening." International Conference on Intelligent Computing. Singapore: Springer Nature Singapore, 2024.
[13] Wang, Fu-Yun, et al. "Foster: Feature boosting and compression for class-incremental learning." European conference on computer vision. Cham: Springer Nature Switzerland, 2022.
[14] Qin, Zhili, et al. "Rethinking few-shot class-incremental learning: A lazy learning baseline." Expert Systems with Applications 250 (2024): 123848.
[15] Ma, Kailang, et al. "Instance-wise batch label restoration via gradients in federated learning." The Eleventh International Conference on Learning Representations. 2022.
[16] Wang, Qi-Wei, et al. "Few-shot class-incremental learning via training-free prototype calibration." Advances in Neural Information Processing Systems 36 (2023): 15060-15076.
[17] Khan, Muhammad Gul Zain Ali, et al. "Supplementary: Introducing Language Guidance in Prompt-based Continual Learning."

[18] Ma'sum, Muhammad Anwar, et al. "PIP: Prototypes-Injected Prompt for Federated Class Incremental Learning." Proceedings of the 33rd ACM International Conference on Information and Knowledge Management. 2024.
[19] Liu, Chenxi, et al. "Few-shot class incremental learning with attention-aware self-adaptive prompt." European Conference on Computer Vision. Cham: Springer Nature Switzerland, 2024.
[20] Md. Fazle Rabbi, Md. Nahid Sultan, Mahmudul Hasan, Md. Zahidul Islam,"Tribal Dress Identi cation using Convolutional Neural Network." Journal of Information Hiding and Multimedia Signal Processing. Volume 14, Number 3, September 2023
[21] Pho Hai Dang,"VPZL: Visual prompt-guided zero-shot learning for insulator defect detection.",Journal of Information Hiding and Multimedia Signal Processing. Volume 16, Number 3, September 2025