MULTI-CENTER PROTOTYPE ENHANCED MULTIMODAL LOW-RANK PROMPTING FOR PHASE-RESOLVED PARTIAL DISCHARGE RECOGNITION

Anonymous authors
Paper under double-blind review

ABSTRACT

Phase-Resolved Partial Discharge (PRPD) recognition aims to identify insulation defects using partial-discharge patterns of power equipment. Existing methods suffer from significant intra-class morphological diversity, inter-class feature overlap, as well as the efficiency-accuracy trade-off. To tackle these issues, we propose MCP-MMLoP, a dual-module framework consisting of a MMLoP module and a multi-center prototype (MCP) module. Specifically, the MMLoP module realizes parameter-efficient model adaptation via multimodal low-rank prompting to mitigate the efficiency-accuracy trade-off. Additionally, the MCP module equips each discharge class with multiple sub-prototypes to explicitly model diverse intra-class morphological patterns and provide complementary discriminative cues, alleviating recognition difficulties stemming from intra-class morphological diversity and inter-class feature similarity. Experiments on our self-constructed PRPD dataset covering six discharge classes show that MCP-MMLoP adds merely 0.0092M additional trainable parameters over the MMLoP baseline and achieves state-of-the-art performance, which validates that our framework enhances discriminative capacity for complex PRPD samples while retaining the parameter-efficiency benefits of low-rank prompt learning.

1 INTRODUCTION

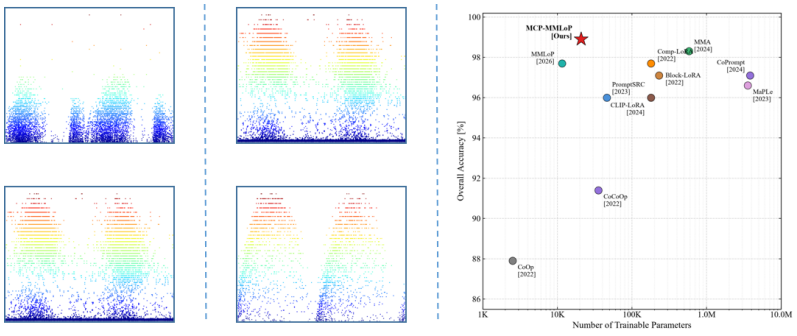
PRPD patterns carry rich discharge types information and are widely adopted for insulation condition assessment and partial discharge (PD) fault-type diagnosis of power equipment (Shahsavarian et al., 2021; Florkowski, 2020; Stone, 2005). The key issue is to learn generalizable partial discharge representations from PRPD diagrams captured under variable practical measurement conditions, as illustrated in Figure 1.

Existing PRPD recognition methods fall into two main categories: unimodal PRPD image recognition (Abubakar & Zachariades, 2024; Jamil et al., 2026)and multimodal PRPD recognition (Chouaibi et al., 2025; Jia et al., 2026a). Unimodal PRPD image recognition takes PRPD patterns as the primary input and carries out identification using handcrafted features or convolutional networks, which tend to lack physical interpretability and have limited modeling capability when handling complex samples (Wu et al., 2023; Yan et al., 2025). Multimodal PRPD recognition fuses multi-source information such as patterns and text, and leverages vision-language models (VLMs) to incorporate domain priors, which enhances modeling capacity for complex discharge samples and represents the mainstream research direction.

Multimodal methods can be further divided into traditional fusion methods and prompt learning-based methods. Traditional fusion methods integrate discharge waveforms, statistical features and pattern information through feature concatenation and weighted fusion. However, they rely on manually designed fusion rules, resulting in limited generalization capability and poor adaptability to few-shot scenarios. Recent advances in VLMs, especially CLIP (Radford et al., 2021), have provided a promising paradigm for few-shot recognition. Prompt learning freezes the backbone of pretrained VLMs and optimizes only a small set of task-specific parameters to achieve efficient model adaptation (Jia et al., 2022). Representative methods including CoOp (Zhou et al., 2022b), CoCoOp (Zhou et al., 2022a), MaPLe (Khattak et al., 2023a), and PromptSRC (Khattak et al., 2023b) have

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

significantly improved the flexibility and generalization performance of prompt-based adaptation. Therefore, we adopts prompt learning-based vision-language adaptation for PRPD recognition.



(a) (b) (c)

Figure 1: PRPD-recognition challenges: (a) Patterns within the same discharge category exhibit high morphological variability. (b) Features from distinct discharge types overlap with one another. (c) Models struggle to balance recognition accuracy and computational cost.

Existing prompt learning-based methods have demonstrated promising performance. For instance, MMA (Yang et al., 2024) realizes multimodal fusion via feature stacking and projection and obtains a concise architecture. MMLoP (Ghiasvand et al., 2026) adopts low-rank prompt parameterization and a shared cross-modal up-projection matrix to constrain prompts within a low-dimensional subspace, drastically reducing parameter counts while retaining cross-modal interaction modeling. Although these two methods mitigate the high-parameter issue, they yield relatively low recognition accuracy and fail to address intra-class diversity and inter-class similarity. This constitutes the first limitation. By contrast, MaPLe (Khattak et al., 2023a) inserts learnable prompts into both vision and text encoders to achieve precise cross-modal alignment, yet it suffers from heavy parameter and computational costs and thus cannot resolve the trade-off between recognition accuracy and computational overhead. This constitutes the other limitation.

To overcome the aforementioned limitations, this work proposes the Multi-Center Prototype Enhanced Multimodal Low-Rank Prompting (MCP-MMLoP) framework, which consists of a multi-center prototype (MCP) module and an MMLoP module (Ghiasvand et al., 2026). First, we introduce the MCP module, which maintains multiple sub-prototypes for each discharge class to explicitly characterize diverse intra-class morphologies and provide complementary discriminative cues. We further employ entropy-weighted exponential moving average (EMA) and prototype-space regularization to stabilize prototype updates and enlarge inter-class prototype separation, thereby mitigating the recognition challenges caused by intra-class diversity and inter-class similarity. Second, the MMLoP module adopts multimodal low-rank prompting to achieve parameter-efficient model adaptation, effectively alleviating the trade-off between recognition accuracy and computational overhead. The outputs of the two modules are weighted and fused for final recognition. What's more, we construct a PRPD dataset containing 1694 patterns spanning six discharge categories. Our main contributions are summarized as follows:

(1) We propose the MCP module, which assigns multiple sub-prototypes to each discharge class and characterizes diverse pattern morphologies within the same discharge category via maximum-similarity aggregation, mitigating the problems of intra-class variations and inter-class similarity in PRPD recognition.

(2) To resolve the efficiency-accuracy trade-off, we introduce the MMLoP module. This module constructs low-rank parameterized prompts and a shared cross-modal up-projection structure. By constraining prompt parameter distribution within a low-dimensional subspace, it drastically reduces parameter overhead while stably preserving cross-modal interaction modeling capacity.

(3) We build a PRPD dataset covering six typical discharge morphologies, including Suspension Discharge, Tip Discharge, Surface Discharge, Air Gap discharge, Metal Discharge, and noisy normal samples. Extensive experiments demonstrate that MCP-MMLoP achieves state-of-the-art performance on this dataset.

2 Related Work

2.1 PRPD Recognition

PRPD analysis is widely used for insulation fault diagnosis in power equipment. Early studies relied on statistical, fractal, and conventional machine-learning features (Sahoo et al., 2005; Basharan, 2018), while subsequent works adopted CNN-based models for PRPD or discharge-signal recognition (Peng et al., 2019; Liu et al., 2021; Sun et al., 2021). More recent studies explore image-based matching, data augmentation, and few-shot metric learning (Jia et al., 2026b; Li et al., 2025).

2.2 Prompt Learning

VLMs provide a promising foundation for few-shot recognition by transferring knowledge from large-scale image-text pretraining. CLIP (Radford et al., 2021) enables zero-shot recognition through aligned visual and textual representations. Building upon CLIP, CoOp (Zhou et al., 2022b) and CoCoOp (Zhou et al., 2022a) introduce learnable and instance-conditioned prompts, respectively, while MaPLe (Khattak et al., 2023a), PromptSRC (Khattak et al., 2023b), MMA (Yang et al., 2024), and CoPrompt (Roy & Etemad, 2024) further improve multimodal prompt adaptation. However, multimodal deep prompting can introduce considerable trainable parameters, motivating parameter-efficient alternatives including visual prompt tuning, adapters, and low-rank adaptation (Jia et al., 2022; Houlsby et al., 2019; Hu et al., 2022; Zanella & Ben Ayed, 2024). MMLoP (Ghiasvand et al., 2026) addresses this issue through low-rank prompt parameterization with a shared up-projection matrix, reducing trainable parameters to approximately 11.5 K while maintaining competitive performance.

2.3 Prototype Learning

Prototypical Networks (Snell et al., 2017) represent each class with a single prototype, yet a single center may be insufficient for classes where patterns within the same discharge category exhibit high morphological variability. This issue is closely related to few-shot diagnosis under limited or ambiguous supervision (Feng et al., 2020; Li et al., 2023; Wang et al., 2023). Subsequent methods have explored mixture-based or MCP to capture heterogeneous class structures. Infinite Mixture Prototypes (Allen et al., 2019) model classes with multiple mixture components, while MCP employs multiple sub-prototypes with EMA, intra-class compactness, and inter-class separation to improve discrimination under limited supervision.

3 Method

3.1 Problem Definition and Overall Framework

Let the training set be $\mathcal{D} = \{(x_n, y_n)\}_{n=1}^{N}$, where x_n denotes a PRPD pattern and $y_n \in 1, \ldots, C$ denotes the discharge-type label ($C = 6$: Tip, Metal, Air Gap, Suspension, Surface, and Normal). MCP-MMLoP adopts a frozen CLIP dual-encoder as its backbone. Neither the image encoder $f(\cdot)$ nor the text encoder $g(\cdot)$ is updated, and only two types of lightweight parameters are optimized: the low-rank prompt factors $\{\mathbf{U}^{(l)}, \mathbf{V}_v^{(l)}, \mathbf{V}_t^{(l)}\}_{l=1}^{L}$ and the MCP $\{\mathbf{P}_c\}_{c=1}^{C}$.

The overall framework is illustrated in Figure 2. Given an input pattern x, the image encoder with visual prompts produces the prompted feature $\tilde{\mathbf{f}}_p$. The text encoder with textual prompts produces the prompted text feature $\tilde{\mathbf{g}}_k^p$, which is subsequently corrected by Uniform Drift Correction (UDC) to obtain $\hat{\mathbf{g}}_k$. The two paths compute class scores in parallel: the prompt-path score is $s_c = \mathrm{sim}(\tilde{\mathbf{f}}_p, \hat{\mathbf{g}}_c)$, while the prototype-path score is $g_c = \max_k \mathrm{sim}(\tilde{\mathbf{f}}_p, \mathbf{p}_{c,k})$. The final score $\ell_c = s_c + \beta g_c$ is converted into class probabilities through temperature-scaled softmax.

3.2 Low-Rank Multimodal Prompting

Following MMLoP, each layer-wise prompt matrix is parameterized as the product of two low-rank factors. Let the visual and textual prompts at the l-th layer be $\mathbf{P}_v^{(l)} \in \mathbb{R}^{V \times d_v}$ and $\mathbf{P}_t^{(l)} \in \mathbb{R}^{T \times d_t}$,

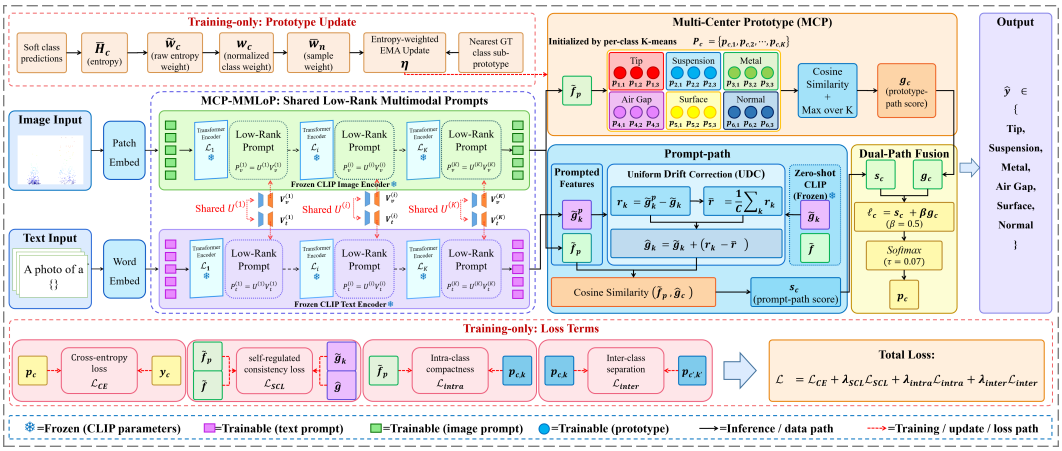


Figure 2: Overall framework of the proposed MCP-MMLoP.

respectively. Then,

$$\mathbf{P}_v^{(l)} = \mathbf{U}^{(l)}\mathbf{V}_v^{(l)}, \qquad \mathbf{P}_t^{(l)} = \mathbf{U}^{(l)}\mathbf{V}_t^{(l)}, \tag{1}$$

where $\mathbf{U}^{(l)} \in \mathbb{R}^{T\times r}$ is the shared up-projection matrix shared across the visual and textual modalities, with T set to the visual prompt length; the visual branch reuses it after dimensional adaptation. $\mathbf{V}_v^{(l)} \in \mathbb{R}^{r\times d_v}$ and $\mathbf{V}_t^{(l)} \in \mathbb{R}^{r\times d_t}$ are the down-projection matrices, and the rank satisfies $r \ll \min(d_v, d_t)$. The inputs to the prompted Transformer layers are given by

$$\tilde{\mathbf{X}}_p^{(l)} = \{\mathbf{P}_v^{(l)},\ \mathbf{e}_{\text{cls}}^{(l)},\ \mathbf{e}_1^{(l)}, \ldots, \mathbf{e}_M^{(l)}\}, \qquad \tilde{\mathbf{Y}}_p^{(l)} = \{\mathbf{t}_{\text{SOS}}^{(l)},\ \mathbf{P}_t^{(l)},\ \mathbf{t}_1^{(l)}, \ldots, \mathbf{t}_N^{(l)},\ \mathbf{c}_k^{(l)},\ \mathbf{t}_{\text{EOS}}^{(l)}\}. \tag{2}$$

During training, the CLIP backbone θ_{CLIP} is completely frozen, and only the low-rank factors are optimized.

3.3 Classification Output of the Prompt Path and UDC

The prompted image feature is denoted as $\tilde{\mathbf{f}}_p$, the prompted text feature of class k as $\tilde{\mathbf{g}}_k^p$, and the zero-shot features of the frozen CLIP as $\tilde{\mathbf{f}}$ and $\tilde{\mathbf{g}}_k$. Prompt tuning introduces systematic drift across all class embeddings, i.e., a "base-class bias," thereby degrading generalization. To address this issue, UDC is adopted. The prompted text feature is decomposed into a zero-shot anchor and a residual, $\mathbf{r}_k = \tilde{\mathbf{g}}_k^p - \tilde{\mathbf{g}}_k$, and the mean residual shared across classes, $\bar{\mathbf{r}} = \frac{1}{C}\sum_k \mathbf{r}_k$, is subtracted:

$$\hat{\mathbf{g}}_k = \tilde{\mathbf{g}}_k + (\mathbf{r}_k - \bar{\mathbf{r}}), \qquad \hat{\mathbf{g}}_k \leftarrow \hat{\mathbf{g}}_k / \|\hat{\mathbf{g}}_k\|. \tag{3}$$

The corrected features retain only the class-discriminative residuals. The score of the prompt path for class c is computed using cosine similarity:

$$s_c = \text{sim}(\tilde{\mathbf{f}}_p, \hat{\mathbf{g}}_c) = \frac{\tilde{\mathbf{f}}_p \cdot \hat{\mathbf{g}}_c}{\|\tilde{\mathbf{f}}_p\|\ \|\hat{\mathbf{g}}_c\|}. \tag{4}$$

3.4 MCP Module

Accordingly, we introduce the concept of sub-prototypes to characterize a specific morphology within each discharge class.

$$\mathbf{P}_c = \{\mathbf{p}_{c,1}, \mathbf{p}_{c,2}, \ldots, \mathbf{p}_{c,K}\}, \qquad \mathbf{p}_{c,k} \in \mathbb{R}^D, \tag{5}$$

Where D denotes the feature dimension, consistent with $\tilde{\mathbf{f}}_p$. Sub-prototypes are initialized with the K-means cluster centers of per-class training features, so they correspond to distinct pattern

morphologies from the start of training. The cosine similarity between sample $\tilde{\mathbf{f}}_p$ and sub-prototype $\mathbf{p}_{c,k}$ is defined as

$$\mathrm{sim}(\tilde{\mathbf{f}}_p, \mathbf{p}_{c,k}) = \frac{\tilde{\mathbf{f}}_p \cdot \mathbf{p}_{c,k}}{\|\tilde{\mathbf{f}}_p\| \, \|\mathbf{p}_{c,k}\|}. \tag{6}$$

The prototype-path score for class c is defined as the maximum similarity among its K sub-prototypes (soft assignment, where a sample only needs to match its most similar morphology):

$$g_c(\tilde{\mathbf{f}}_p) = \max_{k \in \{1,\dots,K\}} \mathrm{sim}(\tilde{\mathbf{f}}_p, \mathbf{p}_{c,k}). \tag{7}$$

Max aggregation enables each sub-prototype to independently characterize one pattern morphology, avoiding the averaging of different morphologies into a single center. This constitutes the fundamental distinction between multi-center and single-center prototypes.

3.5 EMA Prototype Update

Prototypes need stable evolution during training. Direct prototype update via batch gradients is vulnerable to perturbations from confusing samples such as Tip-Metal boundary samples. We thus adopt the EMA update. The batch-level class marginal distribution and information entropy are computed first. Let $\boldsymbol{q}_n$ denote the soft label of sample n. The marginal probability of class c in a batch is:

$$\bar{p}_c = \frac{1}{B} \sum_{n=1}^{B} q_{n,c}, \tag{8}$$

and its information entropy $\bar{H}_c = -\bar{p}_c \log \bar{p}_c$ measures the dispersion of the current predictions for this class. A larger entropy indicates greater uncertainty in class discrimination, and thus its contribution to prototype updates should be reduced:

$$\tilde{w}_c = \frac{1}{\bar{H}_c}, \qquad w_c = \frac{\tilde{w}_c}{\frac{1}{C}\sum_{j=1}^{C} \tilde{w}_j}, \tag{9}$$

where the normalization ensures $\frac{1}{C}\sum_c w_c = 1$, preventing systematic drift in the update magnitude. The sample-level weight is obtained by aggregating the class weights according to the soft label: $\bar{w}_n = \sum_{c=1}^{C} q_{n,c} \, w_c$.

For sample n, its most similar sub-prototype in the ground-truth class is $k_n^* = \arg\max_k \mathrm{sim}(\tilde{\mathbf{f}}_{p,n}, \mathbf{p}_{y_n,k})$. The corresponding sub-prototype is updated via EMA with weight $\bar{w}_n$:

$$\mathbf{p}_{y_n,k_n^*}^{t+1} = \eta \, \mathbf{p}_{y_n,k_n^*}^{t} + (1-\eta) \, \bar{w}_n \, \tilde{\mathbf{f}}_{p,n}, \tag{10}$$

where the momentum $\eta \in [0, 1)$ is set to 0.99 in the final model, following the sensitivity analysis in Table 3. The contributions of samples from high-entropy (uncertain) classes are automatically attenuated, preventing confused samples from significantly perturbing the prototype locations.

3.6 Prototype Space Regularization

To mitigate intra-class dispersion and inter-class overlap, we adopt two lightweight regularizers:

Intra-class compactness: samples are encouraged to move toward the most similar sub-prototype of their corresponding class,

$$\mathcal{L}_{\mathrm{intra}} = \frac{1}{B} \sum_{n=1}^{B} \left(1 - g_{y_n}(\tilde{\mathbf{f}}_{p,n})\right). \tag{11}$$

This loss only acts on the nearest sub-prototype, allowing different morphologies within the same class to form their own compact clusters rather than forcing the entire class into a single cluster.

Inter-class separation: a cosine hinge constraint is imposed on arbitrary pairs of sub-prototypes from different classes, penalizing their similarity when it exceeds the margin m:

$$\mathcal{L}_{\text{inter}} = \frac{1}{|\mathcal{S}|} \sum_{(c,k)\neq(c',k')} \max\left(0,\ \text{sim}(\mathbf{p}_{c,k}, \mathbf{p}_{c',k'}) - m\right), \tag{12}$$

where $\mathcal{S}$ denotes the set of all cross-class sub-prototype pairs, and $m = 0.2$ is the margin. The hinge formulation drives the loss to zero once the required margin is satisfied, thereby avoiding unnecessary over-separation.

3.7 Dual-Path Fusion and Overall Objective

The scores from the two paths are weighted and fused:

$$\ell_c = s_c + \beta\, g_c(\tilde{\mathbf{f}}_p), \tag{13}$$

where β is the fusion coefficient, set to 0.5 by default. The class probabilities are obtained using temperature-scaled softmax:

$$p_c = \frac{\exp(\ell_c/\tau)}{\sum_{c'=1}^{C} \exp(\ell_{c'}/\tau)}, \tag{14}$$

where the temperature τ follows the CLIP convention and is set to 0.07. The prompt path provides cross-modal semantic discrimination, while the prototype path provides multiple morphological centers within each class; the two paths are complementary.

The overall training objective is defined as

$$\mathcal{L} = \mathcal{L}_{\text{CE}} + \lambda_{\text{SCL}}\mathcal{L}_{\text{SCL}} + \lambda_{\text{intra}}\mathcal{L}_{\text{intra}} + \lambda_{\text{inter}}\mathcal{L}_{\text{inter}}, \tag{15}$$

where $\mathcal{L}_{\text{CE}}$ denotes the cross-entropy loss computed from the fused probabilities, and $\mathcal{L}_{\text{SCL}}$ denotes the self-regulated consistency loss from MMLoP, which consists of a feature-level L1 constraint and a logit-level symmetric KL constraint:

$$\begin{aligned}\mathcal{L}_{\text{SCL}} = \lambda_1\|\tilde{\mathbf{g}}_p - \tilde{\mathbf{g}}\|_1 + \lambda_2\|\tilde{\mathbf{f}}_p - \tilde{\mathbf{f}}\|_1 \\ + \tfrac{1}{2}D_{\text{KL}}\big(\text{sim}(\tilde{\mathbf{f}}_p, \tilde{\mathbf{g}}_p)\,\|\,\text{sim}(\tilde{\mathbf{f}}, \tilde{\mathbf{g}})\big) \\ + \tfrac{1}{2}D_{\text{KL}}\big(\text{sim}(\tilde{\mathbf{f}}, \tilde{\mathbf{g}})\,\|\,\text{sim}(\tilde{\mathbf{f}}_p, \tilde{\mathbf{g}}_p)\big),\end{aligned} \tag{16}$$

4 Experiments

4.1 Dataset and Experimental Protocol

Dataset. We evaluate MCP-MMLoP on a laboratory-based PRPD dataset containing 1,694 patterns from six discharge classes: Tip, Suspension, Surface, Air Gap, Metal, and Normal. Each pattern is represented as a PRPD scatter plot with phase angle on the horizontal axis and discharge amplitude on the vertical axis. The dataset is divided into training, validation, and test sets with 1353, 167, and 174 samples, respectively, as summarized in Table 4 of Appendix A.

Experimental Protocol. The dataset is stratified by class and randomly split at a ratio of 8:1:1.

4.2 MAIN RESULT

Due to class imbalance in our dataset, we report overall accuracy(OA) and macro-F1. All methods use the same CLIP ViT-B/16 backbone, data splits, and data augmentation. We compare against CLIP zero-shot recognition, prompt-learning methods including CoOp (Zhou et al., 2022b), CoCoOp (Zhou et al., 2022a), MaPLe (Khattak et al., 2023a), PromptSRC (Khattak et al., 2023b), MMA, and CoPrompt, as well as parameter-efficient adaptation methods including CLIP-LoRA (Zanella & Ben Ayed, 2024), Block-LoRA, and Comp-LoRA (Hu et al., 2022). MMLoP (Ghiasvand et al., 2026) serves as the primary low-rank prompting baseline. For fair comparison, all methods use the text template "a photo of {}" when applicable. Trainable parameter counts are reported together with recognition performance in Table 1.

Table 1: Performance comparison between our method and other baseline methods (%).

Method	Params.	Tip	Suspension	Surface	Air Gap	Metal	Normal	OA	Macro-F1
Zero-shot CLIP (Radford et al., 2021)	0	89.7	0.0	0.0	0.0	0.0	0.0	20.1	5.7
CoOp (Zhou et al., 2022b)	2.5 K	97.4	88.9	84.6	80.0	79.5	96.3	87.9	88.4
CoCoOp (Zhou et al., 2022a)	35.4 K	100.0	83.3	80.8	92.0	89.7	96.3	91.4	90.6
MaPLe (Khattak et al., 2023a)	3.56 M	94.9	100.0	100.0	96.0	92.3	100.0	96.6	97.2
PromptSRC (Khattak et al., 2023b)	46.08 K	100.0	100.0	96.2	92.0	89.7	100.0	96.0	96.2
MMA (Yang et al., 2024)	0.58 M	94.9	100.0	100.0	96.0	100.0	100.0	98.3	98.6
CoPrompt (Roy & Etemad, 2024)	3.82 M	97.4	100.0	100.0	92.0	94.9	100.0	97.1	97.5
CLIP-LoRA (Zanella & Ben Ayed, 2024)	0.18 M	94.9	100.0	96.2	92.0	94.9	100.0	96.0	96.5
Block-LoRA (Hu et al., 2022)	0.23 M	94.9	100.0	96.2	96.0	97.4	100.0	97.1	97.5
Comp-LoRA (Hu et al., 2022)	0.18 M	94.9	100.0	96.2	100.0	97.4	100.0	97.7	98.1
MMLoP (Ghiasvand et al., 2026)	11.56 K	100.0	100.0	96.2	100.0	92.3	100.0	97.7	98.1
MCP-MMLoP(Ours)	20.8 K	100.0	100.0	96.2	100.0	97.4	100.0	**98.9**	**99.0**

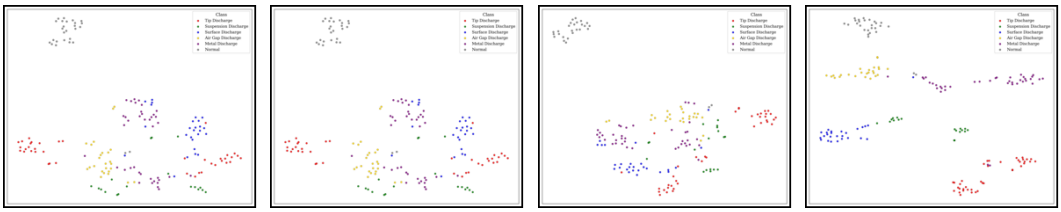
Table 1 summarizes the comparison results of different prompt learning methods on the PRPD recognition dataset. We highlight three main observations. First, prompt learning substantially improves zero-shot PRPD recognition performance. Vanilla zero-shot CLIP only achieves an OA of 20.1% and produces zero accuracy for five of the six discharge classes, whereas CoOp improves the OA to 87.9%, demonstrating the necessity of adapting pre-trained visual-language models to PRPD recognition tasks. Second, MMLoP yields prominent parameter efficiency via low-rank prompt design. With only 11.56 K trainable parameters, MMLoP achieves 97.7% OA and outperforms MaPLe (96.6%), while consuming merely 0.3% of MaPLe's 3.56 M parameters. This verifies that low-rank prompt adaptation is effective for achieving parameter-efficient PRPD recognition. Third, our proposed MCP-MMLoP achieves the overall best recognition performance. Equipped with the multi-center prototype module, MCP-MMLoP obtains 98.9% OA and 99.0% macro-F1 with 20.8 K trainable parameters, surpassing MMLoP by 1.2% and 0.9% in OA and macro-F1, respectively. In particular, the recall of Metal Discharge increases significantly from 92.3% to 97.4%, demonstrating that multi-center prototypes effectively handle intra-class morphological diversity and inter-class feature similarity, thereby refining discriminative feature learning for PRPD patterns. Overall, MCP-MMLoP achieves state-of-the-art performance on the constructed PRPD dataset.

4.3 VISUAL ANALYSIS

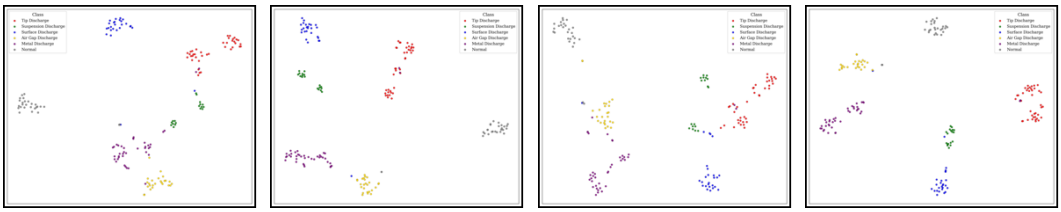
We use t-SNE to visualize test features from competing baselines and MCP-MMLoP, and calculate silhouette scores on original features to quantify intra-class compactness and inter-class separation.

From Figure 3, early prompt-learning methods show severe cluster overlap. Other baselines partially improve feature separation but remain troubled by mixed similar PRPD samples; MMLoP still has partial cross-category overlap. In comparison, MCP-MMLoP generates well-structured embeddings with compact same-class clusters and clear gaps between different discharge categories.

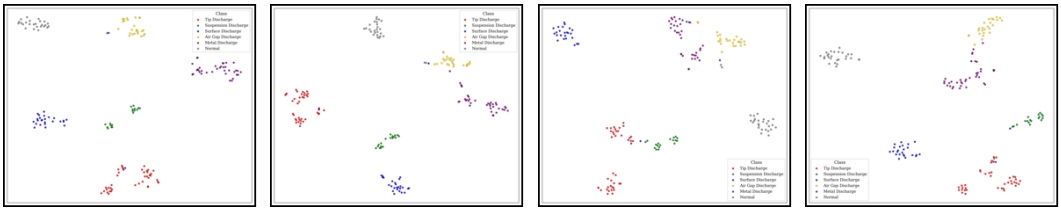
MCP-MMLoP attains the highest silhouette score of 0.6761, an increment of 0.1410 over MMLoP (0.5351). Combined t-SNE visualization and silhouette analysis demonstrate the effectiveness of our multi-center prototype module for PRPD feature learning.



(a) Zero-shot CLIP 0.1352 (b) CoOp 0.3352 (c) CoCoOp 0.3647 (d) MaPLe 0.4689



(e) PromptSRC 0.4829 (f) MMA 0.4893 (g) CoPrompt 0.3993 (h) CLIP-LoRA 0.5343



(i) Block-LoRA 0.5417 (j) Comp-LoRA 0.5327 (k) MMLoP 0.5351 (l) MCP-MMLoP 0.6761

Figure 3: t-SNE visualization of test-set features extracted by different methods.

4.4 Ablation Study

To isolate the contributions of the multi-center prototype module, we evaluate four variants using MMLoP as the baseline: (i) MMLoP without the prototype path and regularization terms; (ii) MMLoP with only $\mathcal{L}_{\text{intra}}$; (iii) MMLoP with only $\mathcal{L}_{\text{inter}}$; and (iv) the full MCP-MMLoP with the prototype path and both regularization terms, as shown in Table 2 .

Table 2: Ablation study of the proposed MCP-MMLoP (%).

Module	Tip	Suspension	Surface	Air Gap	Metal	Normal	OA	macro-F1
Baseline (MMLoP (Ghiasvand et al., 2026))	100.0	100.0	96.2	100.0	92.3	100.0	97.7	98.1
w/ $\mathcal{L}_{\text{intra}}$	100.0	100.0	92.3	96.0	87.2	100.0	95.4	95.5
w/ $\mathcal{L}_{\text{inter}}$	100.0	100.0	92.3	92.0	89.7	100.0	95.4	95.5
MCP-MMLoP (w/ $\mathcal{L}_{\text{intra}}$, w/ $\mathcal{L}_{\text{inter}}$)	**100.0**	**100.0**	**96.2**	**100.0**	**97.4**	**100.0**	**98.9**	**99.0**

Table 2 shows that introducing either regularization term alone reduces the overall accuracy (OA) from 97.7% to 95.4%. This indicates that without the guidance of the complete prototype path, a single regularization constraint can interfere with prompt learning. In contrast, jointly introducing the multi-center prototype path together with both $\mathcal{L}_{\text{intra}}$ and $\mathcal{L}_{\text{inter}}$ losses improves the OA to 98.9%, demonstrating the complementarity between the intra-class attraction and inter-class separation mechanisms.

4.5 Few Shot Result

Figure 4 presents the few-shot experimental results. As the number of training samples per class increases from 4 to 32, MCP-MMLoP shows gradually accelerated performance gains and consistently achieves the best performance at the 32-shot setting. This suggests that as the sample quantity rises, multi-center prototypes can learn more representational information to characterize diverse sub-prototypes, so as to realize accurate recognition for PRPD samples with very high intra-class and inter-class similarity.

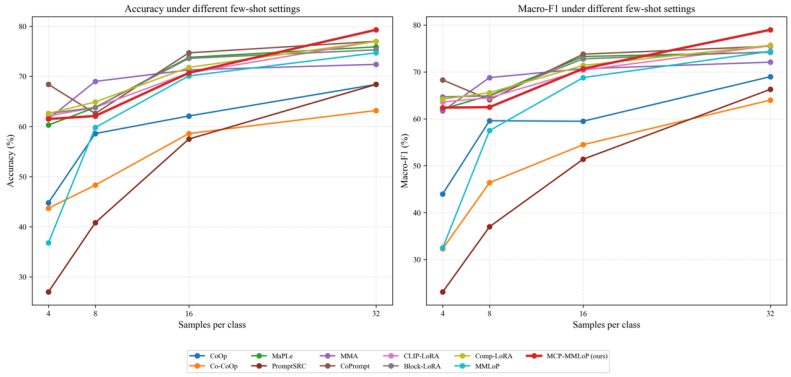
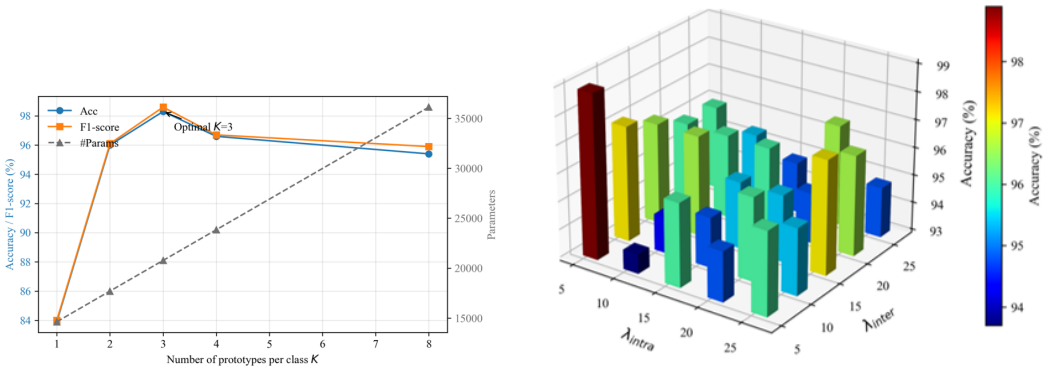


Figure 4: Comparison of OA and macro-F1 under different few-shot settings.

4.6 Hyperparameter Sensitivity

Figures 5 analyze the number of sub-prototypes and prototype regularization weights. Tuning these hyper-parameters changes prototype representation capacity and geometric constraints on the learned feature space. The setting $K = 3$, $\lambda_{\text{intra}} = 5.0$, and $\lambda_{\text{inter}} = 5.0$ achieves the best trade-off between recognition performance and model complexity. Increasing K enables finer modeling for heterogeneous sample distributions yet introduces extra trainable parameters; excessively large regularization coefficients restrict prompt learning and degrade overall accuracy. Table 3 further demonstrates that $\eta = 0.99$ yields the best EMA performance.



(a) (b)

Figure 5: Hyperparameter sensitivity analysis. (a) Sensitivity analysis of the number of sub-prototypes K.; (b) Joint sensitivity analysis of regularization weights λ_{intra} and λ_{inter}. Best: $\lambda_{\text{intra}} = 5.0, \lambda_{\text{inter}} = 5.0, \text{Acc} = 98.9\%$

Table 3: Effect of different EMA momentum coefficients η (%).

η	0.1	0.3	0.5	0.7	0.9	0.95	0.99
OA	96.0	94.8	95.4	96.6	96.6	97.7	98.9
macro-F1	96.5	95.1	95.9	96.8	97.0	98.1	99.0

5 Conclusion

Due to the high cost and difficulty of data acquisition for PRPD partial discharge patterns, the size of our collected dataset is relatively limited. Meanwhile, there are no publicly available benchmark datasets released by existing related works. For this reason, all experiments are conducted on our self-collected dataset rather than public datasets.

AI USE STATEMENT

In this work, all research ideas, problem formulation, model design, experimental schemes, dataset construction, result analysis, and core manuscript content are fully conceived, written, and finalized by the authors. The proposed MCP-MMLoP framework is an original contribution of this paper, and all cited references correspond to real and published works.

Large-language model (AI-assisted writing tool) was **only used for language polishing and text refinement.** Given that several word choices and phrasing in the initial draft were sub-optimal, we leveraged AI tools to revise wording, improve readability, smooth sentence transitions, and enhance the overall fluency of the manuscript. No AI-generated ideas, model architectures, experimental designs, interpretations of experimental results, or logical arguments were adopted. All technical contents, technical claims, conclusions, and viewpoints remain entirely the authors' own work. No part of the core technical content was created or rewritten by artificial intelligence.

All experimental results presented in this paper originate from our own implementations and real-world experimental runs; AI was not involved in data generation, simulation, result fabrication, or plotting. All literature cited has been manually verified by the authors.

REPRODUCIBILITY STATEMENT

We attach the corresponding code in this submission to facilitate reproducibility for reviewers. Ready-to-run scripts are provided for all baseline methods used in our experiments. Users can reproduce the main experimental results reported in this paper with one-click execution without tedious manual configuration. All key hyper-parameter settings are documented in the appendix of this manuscript, including model configurations, training parameters, data augmentation strategies, and optimizer hyperparameters. The dataset-splitting protocol is also described in detail, allowing readers to strictly follow our experimental setup for fair comparisons. The codebase is equipped with clear annotations and usage instructions to lower the barrier for reproduction and further algorithm improvement. All experiments are fully reproducible. To promote reproducible research and facilitate follow-up studies on phase-resolved partial-discharge recognition, we will release the full source-code implementation of our proposed **MCP-MMLoP** framework upon paper acceptance.

REFERENCES

Aliyu Abubakar and Christos Zachariades. Phase-resolved partial discharge (prpd) pattern recognition using image processing template matching. *Sensors*, 24(11), 2024. ISSN 1424-8220. doi: 10.3390/s24113565. URL https://www.mdpi.com/1424-8220/24/11/3565.

Kelsey Allen, Evan Shelhamer, Hanul Shin, and Joshua Tenenbaum. Infinite mixture prototypes for few-shot learning. In Kamalika Chaudhuri and Ruslan Salakhutdinov (eds.), *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*, pp. 232–241. PMLR, 09–15 Jun 2019. URL https://proceedings.mlr.press/v97/allen19b.html.

Vigneshwaran Basharan. Recognition of multiple partial discharge patterns by multi-class support vector machine using fractal image processing technique. *IET Science, Measurement & Technology*, 12(8):1031–1038, 2018. doi: 10.1049/iet-smt.2018.5020.

Sana Chouaibi, Mossaad Ben Ayed, Yessine Hadj Kacem, Shaya Abdullah AlShaya, and Fares Tounsi. Fused cnn–lstm for within-cycle partial discharge classification from prpd images and current waveforms. *IEEE Access*, 13:206504–206517, 2025.

Lei Feng, Jiaqi Lv, Bo Han, Miao Xu, Gang Niu, Xin Geng, Bo An, and Masashi Sugiyama. Provably consistent partial-label learning. In H. Larochelle, M. Ranzato, R. Hadsell, M.F. Balcan, and H. Lin (eds.), *Advances in Neural Information Processing Systems*, volume 33, pp. 10948–10960. Curran Associates, Inc., 2020. URL https://proceedings.neurips.cc/paper_files/paper/2020/file/7bd28f15a49d5e5848d6ec70e584e625-Paper.pdf.

Marek Florkowski. Classification of partial discharge images using deep convolutional neural networks. *Energies*, 13(20), 2020. ISSN 1996-1073. doi: 10.3390/en13205496. URL https://www.mdpi.com/1996-1073/13/20/5496.

Sajjad Ghiasvand, Haniyeh Ehsani Oskouie, Mahnoosh Alizadeh, and Ramtin Pedarsani. Mmlop: Multi-modal low-rank prompting for efficient vision-language adaptation, 2026. URL https://arxiv.org/abs/2602.21397.

Neil Houlsby, Andrei Giurgiu, Stanislaw Jastrzebski, Bruna Morrone, Quentin de Laroussilhe, Andrea Gesmundo, Mona Attariyan, and Sylvain Gelly. Parameter-efficient transfer learning for nlp. In *Proceedings of the 36th International Conference on Machine Learning*, volume 97, pp. 2790–2799, 2019.

Edward J. Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. Lora: Low-rank adaptation of large language models. In *International Conference on Learning Representations*, 2022.

AIM Jamil, MNKH Rohani, ASA Nasir, AS Rosmi, NDA Ghazata, ASA Rahman, and MHHA Majid. Benchmarking cnn and vision transformer models for small-dataset prpd classification in highvoltage rotating machines. In *2026 13th International Conference on Electrical and Electronics Engineering (ICEEE)*, pp. 517–522. IEEE, 2026.

Bingjian Jia, Qing Sun, Weiwei Guo, Mingzheng Wang, Qian Wang, and Hongfeng Zhao. Gis partial discharge fault diagnosis based on multi-source feature fusion and resnet-mlp. *Energies*, 19(4):1073, 2026a.

Menglin Jia, Luming Tang, Bor-Chun Chen, Claire Cardie, Serge Belongie, Bharath Hariharan, and Ser-Nam Lim. Visual prompt tuning. In Shai Avidan, Gabriel Brostow, Moustapha Cissé, Giovanni Maria Farinella, and Tal Hassner (eds.), *Computer Vision – ECCV 2022*, pp. 709–727, Cham, 2022. Springer Nature Switzerland. ISBN 978-3-031-19827-4.

Shiyuan Jia, Xiaosheng Peng, Yuncong Cao, Ji Wu, Xin Yu, Yinge Li, and Jian Gao. Statistical-prior-guided grad-cam with capsule networks for few-shot partial discharge pattern recognition in pp cables. *IEEE Transactions on Dielectrics and Electrical Insulation*, 2026b.

Muhammad Uzair Khattak, Hanoona Rasheed, Muhammad Maaz, Salman Khan, and Fahad Shahbaz Khan. Maple: Multi-modal prompt learning. In *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 19113–19122, 2023a. doi: 10.1109/CVPR52729.2023.01832.

Muhammad Uzair Khattak, Syed Talal Wasim, Muzammal Naseer, Salman Khan, Ming-Hsuan Yang, and Fahad Shahbaz Khan. Self-regulating prompts: Foundational model adaptation without forgetting. In *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 15144–15154, 2023b. doi: 10.1109/ICCV51070.2023.01394.

Songyuan Li, Xiaopeng Wang, Yue Han, Junji Feng, Zhen Yin, Jiawang Yang, Weipeng Luo, and Jun Xie. Partial discharge data enhancement and pattern recognition method based on a cae-acgan and resnet. *Symmetry*, 17(1):55, 2025. doi: 10.3390/sym17010055.

Ximing Li, Yuanzhi Jiang, Changchun Li, Yiyuan Wang, and Jihong Ouyang. Learning with partial labels from semi-supervised perspective. In *Proceedings of the AAAI conference on artificial intelligence*, volume 37, pp. 8666–8674, 2023.

Tingliang Liu, Jing Yan, Yanxin Wang, Yifan Xu, and Yiming Zhao. Gis partial discharge pattern recognition based on a novel convolutional neural networks and long short-term memory. *Entropy*, 23(6):774, 2021. doi: 10.3390/e23060774.

Xiaosheng Peng, Fan Yang, Ganjun Wang, Yijiang Wu, Lee Li, Zhaohui Li, Ashfaque Ahmed Bhatti, Chengke Zhou, Donald M. Hepburn, Alistair J. Reid, Martin D. Judd, and Wah Hoon Siew. A convolutional neural network-based deep learning methodology for recognition of partial discharge patterns from high-voltage cables. *IEEE Transactions on Power Delivery*, 34(4):1460–1469, 2019. doi: 10.1109/TPWRD.2019.2906086.

Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, Gretchen Krueger, and Ilya Sutskever. Learning transferable visual models from natural language supervision. In Marina Meila and Tong Zhang (eds.), *Proceedings of the 38th International Conference on Machine*

Learning, volume 139 of *Proceedings of Machine Learning Research*, pp. 8748–8763. PMLR, 18–24 Jul 2021. URL https://proceedings.mlr.press/v139/radford21a.html.

Shuvendu Roy and Ali Etemad. Consistency-guided prompt learning for vision-language models. In *International Conference on Learning Representations*, 2024.

N.C. Sahoo, M.M.A. Salama, and R. Bartnikas. Trends in partial discharge pattern classification: a survey. *IEEE Transactions on Dielectrics and Electrical Insulation*, 12(2):248–264, 2005. doi: 10.1109/TDEI.2005.1430395.

Tohid Shahsavarian, Yue Pan, Zhousheng Zhang, Cheng Pan, Hadi Naderiallaf, Jim Guo, Chuanyang Li, and Yang Cao. A review of knowledge-based defect identification via prpd patterns in high voltage apparatus. *IEEE Access*, 9:77705–77728, 2021. ISSN 2169-3536. doi: 10.1109/ACCESS.2021.3082858.

Greg C Stone. Partial discharge diagnostics and electrical equipment insulation condition assessment. *IEEE Transactions on Dielectrics and Electrical Insulation*, 12(5):891–904, 2005.

Yuanyuan Sun, Shuo Ma, Shengya Sun, Ping Liu, Lina Zhang, Jun Ouyang, and Xianfeng Ni. Partial discharge pattern recognition of transformers based on mobilenets convolutional neural network. *Applied Sciences*, 11(15):6984, 2021. doi: 10.3390/app11156984.

Yanxin Wang, Jing Yan, Zhou Yang, Zhenkang Qi, Jianhua Wang, and Yingsan Geng. Novel metric-based meta-learning model for few-shot diagnosis of partial discharge in a gas-insulated switchgear. *ISA Transactions*, 134:268–277, 2023. doi: 10.1016/j.isatra.2022.08.009.

Yanze Wu, Jing Yan, Zhuofan Xu, Guoqing Sui, Meirong Qi, Yingsan Geng, and Jianhua Wang. Subdomain adaptation capsule network for partial discharge diagnosis in gas-insulated switchgear. *Entropy*, 25(5):809, 2023.

Jing Yan, Yanxin Wang, Yuxin Lin, Yingsan Geng, Jianhua Wang, Hanyan Xiao, and Ran Ding. A novel domain adaptive network guided by expert experience and active learning for field gas insulated switchgear insulation defect diagnosis. *Measurement Science and Technology*, 36(11):115101, 2025.

Lingxiao Yang, Ru-Yuan Zhang, Yanchen Wang, and Xiaohua Xie. Mma: Multi-modal adapter for vision-language models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 23826–23837, 2024.

Maxime Zanella and Ismail Ben Ayed. Low-rank few-shot adaptation of vision-language models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, pp. 1593–1603, 2024.

Kaiyang Zhou, Jingkang Yang, Chen Change Loy, and Ziwei Liu. Conditional prompt learning for vision-language models. In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 16795–16804, 2022a. doi: 10.1109/CVPR52688.2022.01631.

Kaiyang Zhou, Jingkang Yang, Chen Change Loy, and Ziwei Liu. Learning to prompt for vision-language models. *International Journal of Computer Vision*, 130(9):2337–2348, 2022b. ISSN 1573-1405. doi: 10.1007/s11263-022-01653-1. URL https://doi.org/10.1007/s11263-022-01653-1.

A Supplementary Experimental Details

A.1 Dataset Statistics and Representative Patterns

Table 4 reports the class-wise split used in all experiments. Stratified splitting is adopted to maintain the original class distribution across training, validation and test partitions and mitigate the influence of class imbalance. Figure 6 provides representative PRPD patterns from the six discharge classes..

Table 4: PRPD dataset statistics.

Class	Typical physical cause	Training	Validation	Test
Tip Discharge	Local field concentration at metal protrusion	307	38	39
Metal Discharge	PD induced by free moving metal particles	307	38	39
Surface Discharge	Creeping discharge along oil-pressboard interface	196	24	26
Suspension Discharge	Charge accumulation on floating metal parts	134	16	18
Air Gap Discharge	Gas-void breakdown inside solid insulation	200	25	25
Normal	Background noise without insulation defect	209	26	27
Total		1353	167	174

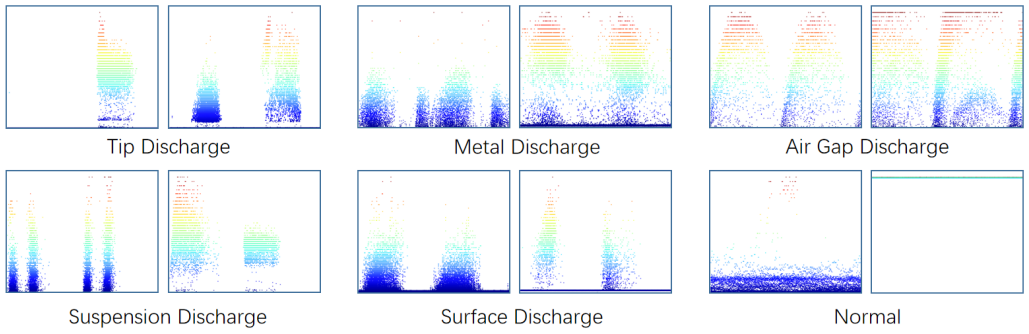


Figure 6: Examples of PRPD patterns for six discharge classes, with two samples shown per class. Notice that patterns within the same discharge category exhibit high morphological variability, alongside the fact that features from distinct discharge types overlap with one another, leading to recognition ambiguity. For instance, patterns from Metal Discharge and Air Gap Discharge exhibit visually similar distributions, posing challenges for discriminative learning.

A.2 Implementation Details and Hyperparameters

For reproducibility, the key implementation hyperparameters are summarized in Table 5. We adopt the parameter-efficient adaptation paradigm: the full CLIP vision-language backbone (including both image and text encoders) is frozen throughout training to preserve the general vision-text alignment knowledge pre-trained on large-scale external datasets. Only lightweight low-rank prompt factors and multi-center prototype parameters are updated via back-propagation. This strategy limits its learnable parameters to a small subset of the whole model, alleviating the accuracy-computation trade-off while exploiting the powerful pre-trained representations of CLIP.

We employ CLIP ViT-B/16 with frozen image and text encoders. Input images are resized to 224×224 and normalized following CLIP's statistical parameters. Considering the physical semantics of PRPD patterns, horizontal and vertical flipping are disabled. We adopt mild brightness jitter, Gaussian noise ($\sigma \leq 0.01$), and small-angle rotation ($\leq 5^\circ$) for data augmentation. Following the base configuration of MMLoP, we set $T = V = 4$ prompt tokens injected into the first 9 Transformer layers, and use rank-1 ($r = 1$) factorization with a shared up-projection matrix. The hyperparameters are configured as $\lambda_1 = 25$, $\lambda_2 = 10$, $K = 3$, $\tau = 0.07$, $m = 0.2$, and $\eta = 0.99$. AdamW optimizer is adopted, where the learning rates are set to 1×10^{-3} for prompt parameters and 1×10^{-4} for prototype parameters. We use cosine annealing scheduler with weight decay of 5×10^{-4}. The batch size is 32, and the maximum training epoch is 100 with early stopping (patience

Table 5: Summary of hyperparameters.

Module	Hyperparameter	Value
Low-rank prompting	Rank r / prompt length $T = V$ / number of layers L	1 / 4 / 9
Multi-center prototypes	Number of sub-prototypes K / temperature τ / margin m / EMA momentum η	3 / 0.07 / 0.2 / 0.99
Loss weights	λ_1 / λ_2 / λ_{intra} / λ_{inter} / β	25 / 10 / 5.0 / 5.0 / 0.5
Training	Learning rate (prompt / prototype) / batch size / epochs	1e−3 / 1e−4 / 32 / 100

= 10). All experiments are implemented in PyTorch with mixed precision training on an NVIDIA RTX 3090 GPU (24 GB memory).

B Training Algorithm

The training procedure of MCP-MMLoP is summarized in Algorithm 1.

Algorithm 1 MCP-MMLoP Training and Inference

Require: Training set $\mathcal{D}$, frozen CLIP encoders f, g, rank r, prompt depth $L = 9$, number of sub-prototypes K, loss weights $\lambda_{\text{SCL}}, \lambda_{\text{intra}}, \lambda_{\text{inter}}$, fusion coefficient β.

Ensure: Low-rank factors $\{\mathbf{U}^{(l)}, \mathbf{V}_v^{(l)}, \mathbf{V}_t^{(l)}\}$ and multi-center prototypes $\{\mathbf{P}_c\}$.

Initialize $\mathbf{U}^{(l)}, \mathbf{V}_v^{(l)}, \mathbf{V}_t^{(l)} \sim \mathcal{N}(0, 0.05)$ for $l = 1, \ldots, L$.
Initialize $\{\mathbf{p}_{c,k}\}$ by applying K-means to the training features of each class.
for each training iteration **do**
  Sample a mini-batch $\{(\mathbf{x}_n, y_n)\}$.
  Construct low-rank prompts using Eq. (1) and obtain $\tilde{\mathbf{f}}_p$ and $\tilde{\mathbf{g}}_{p,k}$.
  Apply UDC correction using Eq. (3) to obtain $\hat{\mathbf{g}}_k$.
  Compute $s_c = \text{sim}(\tilde{\mathbf{f}}_p, \hat{\mathbf{g}}_c)$.
  Compute $g_c = \max_k \text{sim}(\tilde{\mathbf{f}}_p, \mathbf{p}_{c,k})$ and $\ell_c = s_c + \beta g_c$.
  Compute $\mathcal{L}_{\text{CE}}$ and $\mathcal{L}_{\text{SCL}}$.
  Compute entropy and sample weights using Eqs. (8)–(9).
  Compute $\mathcal{L}_{\text{intra}}$ and $\mathcal{L}_{\text{inter}}$ using Eqs. (11)–(12).
  Update the low-rank factors by minimizing $\mathcal{L} = \mathcal{L}_{\text{CE}} + \lambda_{\text{SCL}}\mathcal{L}_{\text{SCL}} + \lambda_{\text{intra}}\mathcal{L}_{\text{intra}} + \lambda_{\text{inter}}\mathcal{L}_{\text{inter}}$.
  Update prototypes using EMA according to Eq. (10).
end for
Inference: $\hat{y} = \arg\max_c (s_c + \beta g_c)$.

C Additional Analysis

C.1 Training Convergence

Figure 7 shows stable convergence: validation accuracy rises rapidly in the early epochs and stabilizes later, while the validation loss decreases and converges with only minor fluctuations. This observation indicates that the joint optimization of low-rank prompt factors and multi-center prototypes works well under our entropy-weighted EMA update rule. No severe oscillation or divergence is observed throughout the training process. After rapid improvement in the initial training phase, both metrics gradually settle, which demonstrates that the proposed dual-module framework can reach a well-behaved optimum without additional optimization tricks.

C.2 Additional Visualization and Error Analysis

Figure 8 shows that the remaining errors are concentrated among these visually similar classes. Such misclassifications correspond to the intrinsically ambiguous samples of the PRPD task, where fault patterns exhibit high visual resemblance across different discharge categories. Figure 3 compares

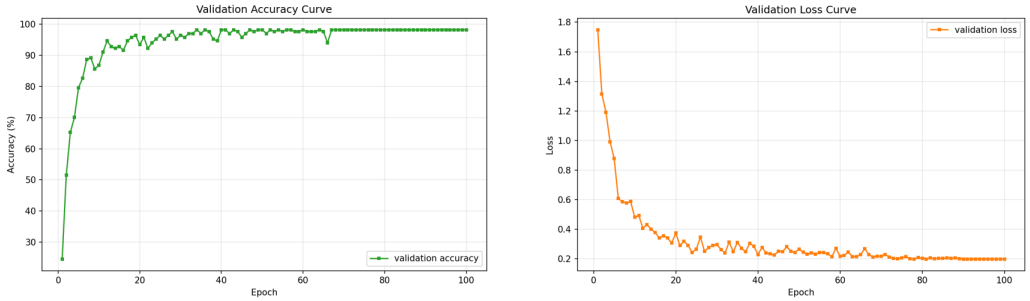


Figure 7: Validation accuracy and loss curves over training epochs.

the feature distributions of 12 methods via t-SNE projection. MCP-MMLoP yields clearer class separation in the latent feature space, particularly for the morphologically similar Metal and Surface Discharge classes, which are challenging to distinguish for competing approaches.

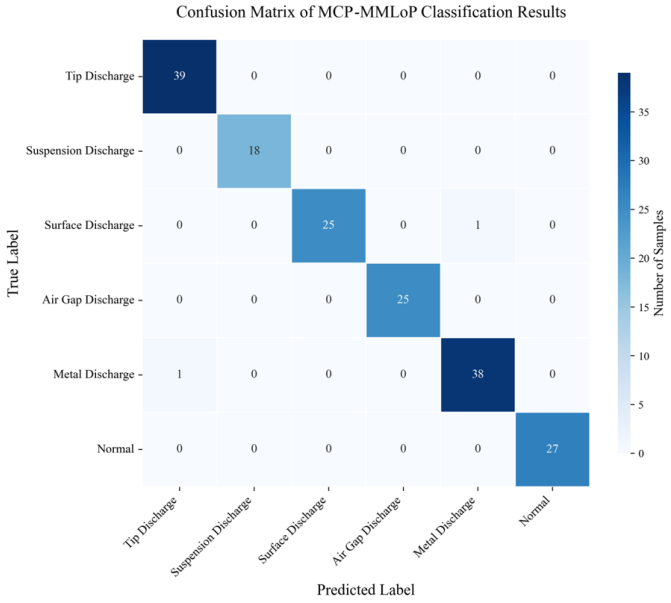


Figure 8: Confusion matrix of classification results obtained by MCP-MMLoP on the test set.

C.3 Parameter Analysis

As shown in Figure 1(c), MCP-MMLoP achieves a good balance between computational efficiency and recognition accuracy, and obtains the best overall accuracy with a small number of parameters. Compared with prompt-learning baselines such as MaPLe, CoPrompt and MMA, our method achieves superior performance with fewer parameters. Relative to MMLoP, the multi-center prototype module introduces around 9.2K extra parameters and adds only one matrix multiplication during inference while delivering the best performance. The results demonstrate that the prototype module improves recognition accuracy and model interpretability with negligible extra parameter and computational overhead.