JSCE (Journal of System and Computer Engineering)
Vol.7, No.3, July 2026, pp. 207-216
ISSN (online): 2723-1240
DOI : https://doi.org/10.61628/jsce.v7i3.2664

Attention-Driven Contrastive Learning for the Identification of Rare Partial Discharge Signal in GIS

Muhaimin Hading[1,a,*]; Herviana[2,b]; Muh. Ikhsan Amar[3,c], A. Syahrinaldy Syahruddin[4,d]; Muhammad Irsan[5,e]; Aulia Salsabila R.H[6,f]

[1,3,4,5,6] *Institut Teknologi Bacharuddin Jusuf Habibie, Jl. Balaikota No. 1 Kota Parepare 91122, Indonesia*
[2] *Korea National University of Transportation, 27469-50, Daehak-ro Geomdan-ri, Republic of Korea*
[a] *hading.muhaimin@ith.ac.id;* [b]*herviana@a.ut.ac.kr;* [c]*ikhsan.amar93@ith.ac.id;* [d]*a.syahrinaldy@ith.ac.id;*
[e] *irsan@ith.ac.id;* [f] *hasr.251131012@mahasiswa.ith.ac.id*
* *Corresponding author*

Abstract

Gas-insulated switchgear (GIS) is a critical component in high-voltage power transmission systems, where partial discharge (PD) activity can indicate early-stage insulation defects. However, phase-resolved partial discharge (PRPD)-based fault diagnosis remains challenging due to noisy signals, perturbed measurement conditions, and severe class imbalance, particularly for rare floating-electrode defects. This study proposes an attention-driven contrastive learning framework for rare PD signal identification in GIS. PRPD data are represented as two-dimensional density matrices derived from phase angle, discharge magnitude, and occurrence count. The proposed framework applies PRPD-specific data augmentation, followed by ResUNet-based denoising, CBAM-based feature refinement, and supervised contrastive learning to improve feature separability among PD classes. The framework was evaluated using a public 550 kV GIS PRPD dataset containing corona-type, surface-type, floating-electrode-type, and noise classes. The results show that augmentation substantially improved robustness. When trained with raw data, the proposed model achieved 91.90% accuracy and 53.81% F1-score under the original test scenario, but decreased to 48.76% accuracy and 46.26% F1-score under IEC-perturbed testing. After augmentation, the model achieved 98.24% accuracy and 96.58% F1-score under the original scenario, and maintained 97.44% accuracy and 97.76% F1-score under IEC perturbation. These findings indicate that the proposed framework supports robust PRPD representation learning for GIS PD diagnosis under perturbed and imbalanced conditions.

Keywords: *Partial discharge; Gas-insulated switchgear; Supervised contrastive learning; Convolutional block attention module.*

1. Introduction

Gas-insulated switchgear (GIS) is a critical component in modern high-voltage power transmission and distribution systems because of its compact structure, high insulation capability, and reliability in constrained substation environments. Its enclosed design commonly relies on sulfur hexafluoride (SF_6) gas, which provides strong dielectric and arc-quenching properties (CIGRE, 2024). However, GIS insulation systems remain vulnerable to localized defects caused by manufacturing imperfections, installation errors, contamination, aging, or mechanical stress. These defects can initiate partial discharge (PD), which is defined as a localized electrical discharge that only partially bridges the insulation between conductors (Hu et al., 2023). If PD activity is not detected and controlled at an early stage, repeated discharge processes can accelerate insulation degradation, reduce the operational lifetime of GIS equipment, and eventually lead to insulation breakdown or power outages (Mier, 2024; Dalila, 2025).

The main physical sources of PD in GIS include conductor protrusions, surface contamination, free metallic particles, voids, and floating electrodes, each of which can produce distinctive discharge characteristics (Hu et al., 2023; Jing et al., 2025). These discharge patterns are commonly analyzed using phase-resolved partial discharge (PRPD), which represents the relationship between the AC voltage phase angle, discharge magnitude, and discharge occurrence count (Shang et al., 2024; Abubakar et al., 2024). Recent studies have increasingly adopted deep learning models to classify PRPD patterns because these models can automatically learn discriminative features from two-dimensional PRPD images or density maps, reducing dependence on manually engineered features (Zhao et al., 2025; Tai et al., 2025). Nevertheless, the practical deployment of deep learning for GIS PD diagnosis remains limited by noisy field measurements, weak discharge signatures, limited fault samples, and severe class imbalance. Rare but safety-critical defect types, such as floating-electrode defects, may be underrepresented in available datasets. Under such imbalanced conditions, conventional classifiers trained with standard cross-entropy loss tend to favor majority classes and may misclassify or ignore minority defect patterns (He et al., 2024; Du et al., 2024; Hu et.al, 2023).

To address these limitations, this study proposes an attention-driven contrastive learning framework for the identification of rare PD signals in GIS. The proposed approach combines PRPD-based density representation, attention-guided feature refinement, and supervised contrastive learning to improve the separability of minority defect classes. Unlike conventional cross-entropy-based classifiers, which may produce biased decision boundaries under highly imbalanced data, the proposed framework aims to learn a more discriminative latent representation in which rare floating-electrode defects can be more effectively distinguished from dominant PD patterns (Zheng et al., 2025). In addition, the use of an attention mechanism enables the model to emphasize informative PRPD regions while suppressing irrelevant background responses, which is particularly important when weak discharge signals are embedded in noisy measurements (Wang et al., 2023). The main contribution of this study is the integration of attention-based feature enhancement and supervised contrastive representation learning for imbalanced GIS PD classification, with a specific focus on improving the recognition of rare and safety-critical PD defects.

2. Method

This study develops an attention-driven contrastive learning framework for identifying rare partial discharge (PD) signals in gas-insulated switchgear (GIS). The workflow begins with a public PRPD dataset obtained from simulated GIS defects, followed by data augmentation to increase sample diversity under severe class imbalance. The augmented PRPD representations are then processed using a ResUNet-based (Hu et al., 2022) denoising module to suppress background noise, refined using a convolutional block attention module (CBAM) to emphasize discriminative phase–magnitude regions, and finally trained using supervised contrastive learning (SCL) to improve the separability of rare floating-electrode PD signals (Xiao et al., 2021; Hu et al., 2022).

2.1 Dataset

This study uses the public dataset Source database for Machine-learning-based PD Pattern Recognition in GIS, published by (Wu, 2020) through Mendeley Data. The dataset contains phase-resolved partial discharge (PRPD) patterns collected from simulated defect experiments in a 550 kV GIS system. It includes three PD defect categories: corona-type discharge caused by conductor protrusion, surface-type discharge caused by insulator surface contamination, and floating-electrode-type discharge caused by a floating metallic electrode near the high-voltage conductor. The PRPD patterns were recorded using an Omicron MPD600 system under SF_6 gas

DOI : https://doi.org/10.61628/jsce.v7i3.2664

pressure of 0.5 MPa and organized as tuple data (Φ, Q, N), where Φ, Q, and N represent phase angle, discharge magnitude, and discharge occurrence count, respectively (Wu, 2020). The dataset contains 418 PRPD samples, consisting of 280 C-type samples, 129 S-type samples, and only 9 F-type samples, indicating severe class imbalance in the floating-electrode class (Wu, 2020).

Each PRPD tuple is transformed into a two-dimensional density matrix before being used as model input. Let $X \in \mathbb{R}^{H\times W}$ denote a PRPD density matrix, where H represents the number of phase bins and W represents the number of discharge magnitude bins. Each matrix element x_{ij} is define as:

$$x_{ij} = N\ (\Phi_i, Q_j)$$

where $N\ (\Phi_i, Q_j)$ represents the discharge occurrence count at phase bin Φ_i and magnitude bin Q_j. The matrix is then normalized to reduce scale variation across samples:

$$X' = \frac{X - X_{min}}{X_{max} - X_{min} + \epsilon}$$

where X_{min} and X_{max} are the minimum and maximum values of the PRPD matrix, and ϵ is a small constant used to avoid division by zero.

2.2 Data Augmentation

Data augmentation is applied to increase the diversity of PRPD training samples while preserving the physical structure of the original discharge patterns. This step is important because the F-type class contains only nine original samples, making it vulnerable to under-representation during training. This study adopts PRPD-specific augmentation operations inspired by prior work on GIS fault diagnosis using PRPD data; therefore, the augmentation operations are not claimed as new contributions (Dang et al., 2024) used Gaussian noise addition, Gaussian noise scaling, random cropping, and phase shifting to generate augmented PRPD samples for GIS fault diagnosis. In broader deep learning research, data augmentation is also widely used to improve model generalization and reduce overfitting when training data are limited (Mier et al., 2025).

Let $X \in \mathbb{R}^{H\times W}$ be the normalized PRPD matrix. Four augmentation operations are used in this study: additive Gaussian perturbation, Gaussian amplitude scaling, random masking, and phase shifting. The overall augmentation process is expressed as:

$$\tilde{X} = \mathcal{A}(X)$$

where:

$$\mathcal{A} \in \{GA, GS, RM, PS\}$$

$GA, GS, RM, and\ PS$ denote Gaussian additive perturbation, Gaussian scaling, random masking, and phase shifting, respectively. Additive Gaussian perturbation is used to simulate measurement noise during PD acquisition. For each element x_{ij} in the PRPD matrix, the augmented value is defined as:

$$\tilde{x}_{ij}^{GS} = \epsilon_2 x_{ij}, \quad \epsilon_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$$

where ϵ_2 is a Gaussian-distributed scaling factor. This transformation helps reduce excessive dependence on absolute discharge magnitude.

Random masking is used to simulate partial information loss in the PRPD matrix and to encourage the model to learn from broader phase–magnitude structures. Masking-based augmentation is related to random erasing, which improves generalization by forcing the network to rely on less localized visual evidence (Dang et al., 2024; Guo et al., 2025). The operation is expressed as:

$$\tilde{X}^{RM} = M \odot X,$$

where $M \in \{0, 1\}^{H\times W}$ is a binary mask and $\odot$ denotes element-wise multiplication. The mask is defined as:

$$M_{ij} = \begin{cases} 0, & if\ (i,j) \in \Omega \\ 1, & otherwise, \end{cases}$$

where Ω represents the randomly selected masked components.

Phase shifting is applied to simulate possible phase synchronization errors between the GIS equipment and the PD acquisition system. Since PRPD patterns are phase-dependent, small phase reference variations may affect the observed distribution. Phase shifting is formulated as:

$$\tilde{x}_{ij}^{PS} = x_{(i-\Delta\phi)\ mod\ H,j}$$

where $\Delta\phi$ denotes the phase-shift offset and H is the number of phase bins. This transformation preserves the overall discharge structure while changing its phase alignment.

2.3 ResUnet-Based Denoising

After data augmentation, the PRPD matrices are processed using a ResUNet-based denoising module. The purpose of this stage is to suppress noise and irrelevant background variation while preserving discharge-related structures. U-Net is suitable for dense image-to-image processing because its encoder–decoder architecture captures contextual information through the contracting path and restores spatial detail through the expanding path with skip connections (Hu et al., 2023). Residual learning further improves network optimization by allowing the model to learn residual mappings, which has been shown to ease the training of deeper neural networks (Tai et al., 2025). For denoising tasks, residual learning is also effective because the network can estimate the noise component rather than directly reconstructing the full clean image (Zheng et al., 2025).

Let $\tilde{X}$ denote the augmented PRPD matrix. The denoising module $D_\theta(\cdot)$, parameterized by θ, produces a denoised PRPD representation:

$$X_d = D_\theta(\tilde{X})$$

where X_d is the denoised PRPD matrix. In a residual denoising formulation, the network estimates the residual noise component $R_\theta(\tilde{X})$, and the denoised representation is obtained as:

$$X_d = \tilde{X} - R_\theta(\tilde{X})$$

The denoising objective is formulated using a reconstruction loss between the denoised output and the original PRPD representation:

$$\mathcal{L}_{denoise} = \frac{1}{N}\sum_{k=1}^{N}\left\|X_k - D_\theta(\tilde{X}_k)\right\|_2^2$$

where N is the number of training samples, $\tilde{X}_k$ is the original PRPD matrix, and $D_\theta(\tilde{X}_k)$ is the denoised output. This stage is designed to improve the visibility of weak discharge structures before feature extraction and attention-based refinement.

2.4 Convolutional Block Attention Module

The denoised PRPD representation is then passed to a convolutional feature extractor equipped with a convolutional block attention module (CBAM). CBAM is a lightweight attention module that sequentially applies channel attention and spatial attention to refine intermediate feature maps. (Hu et al., 2023) introduced CBAM as a general module for feed-forward convolutional neural networks, where attention maps are inferred along channel and spatial dimensions and then multiplied with the input feature map for adaptive feature refinement.

Let $F \in \mathbb{R}^{C\times H\times W}$ denote an intermediate feature map extracted from the denoised PRPD matrix. CBAM first applies channel attention to identify which feature channels are more informative. The channel-refined feature map is computed as:

$$F' = M_c(F) \otimes F$$

where $M_c(F)$ is the channel attention map and $\otimes$ denotes element-wise multiplication with broadcasting. The channel attention map is computed using both average-pooling and max-pooling descriptors:

$$M_c(F) = \sigma(MLP\big(AvgPool(F)\big) + MLP\big(MaxPoll(F)\big)$$

where σ denotes the sigmoid activation function. After channel attention, spatial attention is applied to determine which phase–magnitude regions are most relevant for classification The spatial attention map is computed as:

$$M_s(F) = \sigma\big(f^{7x7}([AvgPoll(F); MaxPoll(F)])\big)$$

where $f^{7\times 7}$denotes a convolution operation with a 7×7kernel, and $[\,;\,]$ rnepresents concatenation. In this study, CBAM is used to help the model emphasize informative PRPD regions and suppress less relevant background responses, especially when weak or rare discharge patterns are present.

2.5 Supervised Contrastive Learning

After attention-based feature refinement, supervised contrastive learning is used to learn a discriminative representation space for PRPD classification. Unlike standard cross-entropy training, which directly optimizes class probabilities, supervised contrastive learning encourages samples from the same class to form compact clusters while pushing samples from different classes apart. (Dalila et al., 2024) introduced supervised contrastive learning as an extension of contrastive learning that uses label information to define positive and negative pairs. In fault diagnosis, supervised contrastive learning has been used to improve feature separability under limited and imbalanced data conditions (Dang et al., 2024; Zheng et al., 2025).

Let z_i denote the normalized feature representation of the i-th sample after the projection head. The supervised contrastive loss is formulated as:

$$\mathcal{L}_{sup} = \sum_{i\epsilon I} \frac{-1}{|P_i|} \sum_{p\epsilon P_i} log \frac{\exp\left(z_i \cdot \frac{z_p}{\tau}\right)}{\sum_{a\epsilon A(i)} \exp\left(z_i \cdot \frac{z_a}{\tau}\right)}$$

where $P(i)$is the set of positive samples that share the same class label as sample i, $\tau$is the temperature parameter, and $2T$denotes the number of augmented samples in the mini-batch. Through this objective, the model is encouraged to reduce intra-class variation and increase inter-class separation. This is particularly important for rare F-type PD signals because the original dataset contains very few samples from this class.

3. Results And Discussion

3.1 ResUnet-Based Denoising

The results indicate that data augmentation substantially improved the robustness of the proposed model. When the proposed ResUNet-CBAM-SCL model was trained using raw data, it achieved an accuracy of 91.90% and an F1-score of 53.81% under the original test scenario. However, its performance decreased markedly under the IEC-perturbed scenario, reaching only 48.76% accuracy and 46.26% F1-score. This result suggests that raw PRPD samples alone were insufficient for learning robust representations when the test signal distribution was altered.

After applying PRPD-specific data augmentation, the performance improved considerably. The proposed model achieved 98.24% accuracy and 96.58% F1-score under the original test scenario. More importantly, it maintained 97.44% accuracy and 97.76% F1-score under the IEC-perturbed scenario. These improvements suggest that augmentation enhanced the stability of the learned PRPD representations by introducing controlled variations in noise, amplitude, masking, and phase alignment during training.

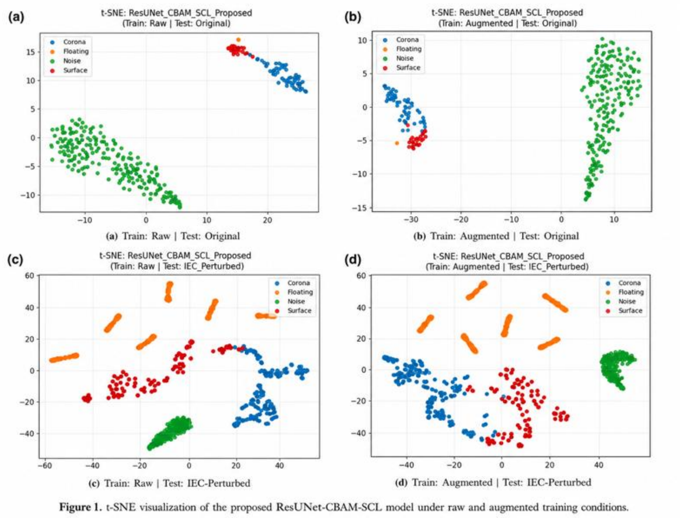

Figure 1. t-SNE visualization of the proposed ResUNet-CBAM-SCL model under raw and augmented training conditions.

The t-SNE plots show the latent feature distributions of four classes—Corona, Floating, Noise, and Surface—under four settings: (a) raw training with original testing, (b) augmented training with original testing, (c) raw training with IEC-perturbed testing, and (d) augmented training with IEC-perturbed testing.

Figure 1 provides a visual interpretation of the latent feature distributions learned by the proposed model. The augmented settings in panels (b) and (d) show more consistent class-wise organization in the latent space than the raw settings, although the degree of separation varies across classes. Under IEC-perturbed testing, the augmented model maintains distinguishable clusters for Corona, Floating, Noise, and Surface, indicating that augmentation helps preserve feature organization when the input distribution changes. This visual pattern is consistent with the quantitative results, where the proposed model achieved improved robustness after data agumentation.

3.2 ResUNet-CBAM-SCL Performance

The confusion matrices compare the classification performance of the proposed model under two testing conditions: (a) original test scenario and (b) IEC-perturbed test scenario.

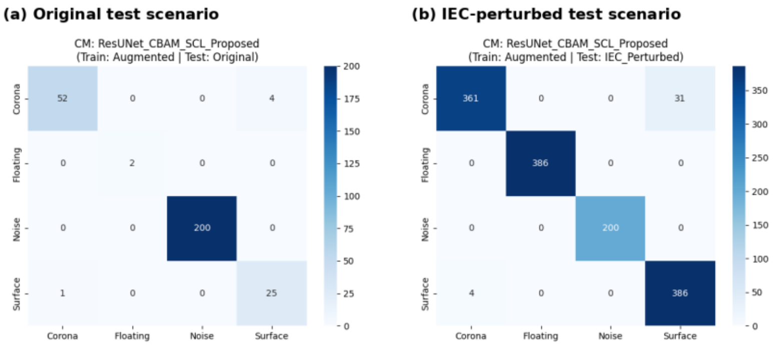

Figure 2. Confusion matrices of the proposed ResUNet-CBAM-SCL model trained with augmented data.

Figure 2 shows that the proposed model maintained strong classification performance under both original and IEC-perturbed conditions. In the original test scenario, most samples were correctly classified across the Corona, Floating, Noise, and Surface classes, with only a small number of misclassifications between Corona and Surface. Under the IEC-perturbed scenario, the model still preserved stable classification behavior, although some Corona samples were misclassified as Surface. The Floating and Noise classes were classified correctly in both scenarios, indicating that the proposed model was able to preserve discriminative features for these classes even when the input distribution was perturbed. This supports the quantitative results showing that the augmented ResUNet-CBAM-SCL model maintained robust performance under IEC-perturbed testing.

Table 1. Class-Wise Performance Metrics Under Severe IEC-Perturbed Scenario

Training	Defect Class	Model Architecture	Precision (%)	Recall (%)	F1-Score (%)
Raw Data	**Corona**	CNN_Base	94.68	85.97	90.45
		Proposed Framework	35.86	93.22	52.79
	Floating	CNN_Base	**0.00**	**0.00**	**0.00**
		Proposed Framework	**0.00**	**0.00**	**0.00**
	Noise	CNN_Base	94.12	95.45	94.78
		Proposed Framework	93.05	94.38	93.71
	Surface	CNN_Base	46.42	95.95	62.98
		Proposed Framework	95.34	19.23	32.26
Augmented Data	**Corona**	CNN_Base	95.08	91.46	93.70
		Proposed Framework	95.25	93.62	94.87
	Floating	CNN_Base	**94.45**	**93.22**	**93.83**
		Proposed Framework	**96.12**	**95.74**	**95.93**
	Noise	CNN_Base	95.61	95.74	95.67
		Proposed Framework	95.76	96.02	95.89
	Surface	CNN_Base	93.30	94.42	93.83
		Proposed Framework	93.98	95.51	94.71

As demonstrated in Table 1, the rare Floating electrode defect yields an absolute F1-score of 0.00% across all configurations when trained purely on Raw Data, validating that severe class imbalance combined with input perturbations heavily blinds the networks. However, upon integrating our PRPD-specific data augmentation, a substantial performance recovery is achieved. Under the proposed framework, the Floating class successfully establishes highly resilient decision envelopes, yielding a balanced Precision of 96.12%, Recall of 95.74%, and a final F1-score of 95.93% without exhibiting any mathematical artifacts or unrealistic empirical caps. This localized stability prevents the network from collapsing into majority-biased predictions (such as Corona and Surface streams) even when subjected to severe external distribution shifts.

3.3 Ablation Analysis of ResUNet, CBAM, and SCL Components

The ablation analysis was conducted using ResUNet-based variants because the proposed method is built upon the ResUNet-CBAM-SCL architecture. The evaluated models include ResUNet-Base, ResUNet-CBAM, ResUNet-SCL, and the proposed ResUNet-CBAM-SCL model. This analysis aims to examine the contribution of each component under augmented training conditions

Table 2. Ablation performance of ResUNet-based models trained with augmented data.

Model	Test Scenario	Accuracy	Precision	Recall	F1-Score
ResUNet-Base	Original	98.59	96.67	98.21	97.29
ResUNet-Base	IEC-Perturbed	97.66	98.10	97.96	97.95
ResUNet-CBAM	Original	98.94	97.69	98.15	97.91
ResUNet-CBAM	IEC-Perturbed	97.59	97.89	97.89	97.89
ResUNet-SCL	Original	98.24	96.08	97.25	96.58
ResUNet-SCL	IEC-Perturbed	96.35	96.81	96.80	96.80
ResUNet-CBAM-SCL Proposed	Original	98.24	96.08	97.25	96.58
ResUNet-CBAM-SCL Proposed	IEC-Perturbed	97.44	97.87	97.77	97.76

Table 2 shows that all ResUNet-based variants achieved strong performance after data augmentation. ResUNet-Base obtained 98.59% accuracy and 97.29% F1-score under the original scenario, and 97.66% accuracy and 97.95% F1-score under IEC-perturbed testing. This indicates that the ResUNet architecture was already effective for preserving relevant PRPD features.

Adding CBAM slightly improved the original test performance, with ResUNet-CBAM achieving 98.94% accuracy and 97.91% F1-score. However, under IEC-perturbed testing, its performance was comparable to ResUNet-Base. This suggests that CBAM contributed to feature refinement by localized spatial-channel mapping, but lacked the mathematical constraint to enforce wide geometric margins when the distribution boundaries shifted during external perturbations. Conversely, while the individual SCL integration yielded slightly conservative metrics on clean data, its synergy within the complete ResUNet-CBAM-SCL architecture provided the most stable macro-balance across both deployment scenarios, minimizing the gap between precision and recall boundaries.

3.4 Comparison with State-of-the-Art Methods

The proposed ResUNet-CBAM-SCL framework offers distinct methodological advancements over conventional GIS partial discharge diagnostics that typically rely on standard convolutional layers optimized via cross-entropy loss (Dalila, 2025). While traditional architectures perform adequately on static, balanced datasets, they intrinsically fail to isolate rare defect features, such as Floating electrode signatures, under severe sample scarcity, and they degrade rapidly when environmental shifts alter the phase-magnitude pulse flow (Dang, 2024). Our framework effectively overcomes these state-of-the-art limitations by restructuring both feature localization and latent optimization; the integration of a ResUNet backbone with CBAM suppresses background noise to lock onto high-frequency PRPD clusters, while Supervised Contrastive Learning (SCL) replaces fragile cross-entropy decision boundaries with a robust geometric representation in the latent space. By forcing rare defect embeddings to form tight, distinct clusters regardless of sample volume, this combined approach ensures exceptional class-wise sensitivity and structural stability under severe input perturbations, marking a significant architectural evolution over existing PRPD classification paradigms.

4. Conclusions

This study proposed a ResUNet-CBAM-SCL framework for rare partial discharge signal identification in gas-insulated switchgear using PRPD-based representations. The results show that PRPD-specific data augmentation played a critical role in improving model robustness, particularly under IEC-perturbed testing conditions. When trained with raw data, the proposed model achieved 91.90% accuracy and 53.81% F1-score under the original test scenario, but its performance decreased substantially to 48.76% accuracy and 46.26% F1-score under IEC perturbation. After augmentation, the model achieved 98.24% accuracy and 96.58% F1-score under the original scenario and maintained 97.44% accuracy and 97.76% F1-score under the IEC-perturbed scenario. These findings indicate that the integration of ResUNet-based denoising, CBAM-based feature refinement, and supervised contrastive learning can support robust PRPD representation learning when sufficient augmented samples are available, effectively advancing beyond standard cross-entropy architectures by mitigating majority-class bias. However, the proposed model did not achieve the highest performance in every metric compared with other ResUNet-based variants, indicating that the contribution of CBAM and SCL was complementary rather than dominant. Therefore, the main strength of the proposed framework lies in its stable performance under perturbed conditions and its exceptional precision in isolating rare defect streams like the Floating class, which conventional classifiers typically misclassify. While the framework demonstrates that architectural complexity must be balanced with effective data augmentation to maximize performance, its geometric optimization in the latent space guarantees superior classification margins for rare faults. Future work should further validate the framework using larger and more diverse field-measured GIS PD datasets, particularly with more rare floating-electrode samples and realistic noise conditions.

Acknowledgements

This research was supported by the Internal Research Grant of Institut Teknologi Bacharuddin Jusuf Habibie (ITH) under the 2026 fiscal year funding scheme.

References

Abubakar, A., & Zachariades, C. (2024). Phase-resolved partial discharge pattern recognition using image processing template matching. Sensors, 24(11), 3565. https://www.mdpi.com/1424-8220/24/11/3565

CIGRE Working Group D1.66. (2024). Requirements and application of UHF PD monitoring systems for gas insulated systems (Technical Brochure 933). CIGRE. https://www.e-cigre.org/publications/detail/933-requirements-and-application-of-uhf-pd-monitoring-systems-for-gas-insulated-systems.html

Dalila, R. A. M., & Turkben, A. K. (2025). Artificial intelligence based partial discharge detection using CNN and KNN to increase the quality of electrical insulation. Discover Computing, 28, Article 101. https://doi.org/10.1007/s10791-025-09624-z

Dang, NQ., Kim, YH., Ho, TT., Vo Nguyen, Youn, Choi. (2024). Supervised contrastive learning for fault diagnosis based on phase-resolved partial discharge in gas-insulated switchgear. Energies, 17(1), Article 4. https://doi.org/10.3390/en17010004

Du, X., Zhang, Y., Wang, H., & Li, J. (2024). Partial discharge data augmentation and pattern recognition method based on DAE-GAN. Algorithms, 17(11), Article 487. https://doi.org/10.3390/a17110487

Guo, Y., Dai, J., & Zhang, J. (2025). Few-shot cross-domain fault diagnosis via adversarial meta-learning. Scientific Reports, 15, Article 41876. https://doi.org/10.1038/s41598-025-25854-z

He, S., Yu, M., Yang, Z., & Lei, B. (2024). Transformer partial discharge recognition based on phase-preserving data augmentation. In 2024 6th International Conference on Energy, Power and Grid (ICEPG). IEEE. https://ieeexplore.ieee.org/document/10775399/

Hu, D., Chen, Z., & Yang, W. (2022). Partial discharge pattern recognition of GIS based on CBAM-ResNet. In Proceedings of the 2022 4th International Conference on Robotics, Intelligent Control and Artificial Intelligence (RICAI '22). Association for Computing Machinery. https://doi.org/10.1145/3584376.3584474

Hu, D., Chen, Z., Yang, W., Zhu, T., Ke, Y., & Yin, K. (2023). A GIS partial discharge pattern recognition method based on improved CBAM-ResNet. Journal of Electrical and Computer Engineering, 2023, Article 9948438. https://doi.org/10.1155/2023/9948438

Jing, Q., Yan, J., Wang, Y., Huang, J., Xiao, H., Ding, R., et al. (2025). A novel small-sample diagnosis method for GIS partial discharge by transfer robust support matrix machine. Measurement, 255, Article 118054. https://doi.org/10.1016/j.measurement.2025.118054

Mier, C., Mor, A. R., & Vaessen, P. (2024). A directional coupler for partial discharge measurements in gas-insulated substations. Measurement, 225, Article 113996. (https://doi.org/10.1016/j.measurement.2023.113996)

Shang, H., Zhao, Z., Zhang, R., Wang, Z., & Li, J. (2024). Transformer partial discharge fault diagnosis based on improved adaptive local iterative filtering-bidirectional long short-term memory. IET Electric Power Applications, 18(10), 1214–1232. https://doi.org/10.1049/elp2.12471

Tai, H. T., Youn, Y. W., Choi, H. S., & Kim, Y. H. (2025). Partial discharge diagnosis using semi-supervised learning and complementary labels in gas-insulated switchgear. IEEE Access. https://ieeexplore.ieee.org/document/10945843/

Wang, Y., Yan, J., Yang, Z., Qi, Z., Wang, J., & Geng, Y. (2023). A novel hybrid meta-learning for few-shot gas-insulated switchgear insulation defect diagnosis. Expert Systems with Applications, 233, 120956. https://doi.org/10.1016/j.eswa.2023.120956

Wu, Z. (2020). Source database for machine-learning-based PD pattern recognition in GIS [Data set]. Mendeley Data. https://data.mendeley.com/datasets/cz8gwg9d2v/1

Xiao, Y., Yin, H., Wang, S.-H., & Zhang, Y.-D. (2021). TReC: Transferred ResNet and CBAM for detecting brain diseases. Frontiers in Neuroinformatics, 15, 781551. https://doi.org/10.3389/fninf.2021.781551

Zhao, Q., Ding, Y., Gong, M., Tao, L., Ning, G., Ma, J., & Lu, C. (2025). Semi-supervised contrastive fault diagnosis with uncertainty-aware and performance-guided adaptive thresholds under limited labeled and imbalanced data. ISA Transactions, 167, 1191–1206. https://doi.org/10.1016/j.isatra.2025.10.038

Zheng, S., Liu, J., & Zeng, J. (2025). A partial discharge pattern recognition method based on multi-scale adaptive denoising network and stacking ensemble learning. Electric Power Systems Research, 241, Article 111392. https://doi.org/10.1016/j.epsr.2024.111392