STU
SLOVAK UNIVERSITY OF TECHNOLOGY IN BRATISLAVA

Journal of Electrical Engineering, Vol. 76, No. 6, 2025, pp. 565-573
https://reference-global.com/journal/JEE

Partial discharge defect classification in cast-resin transformers using machine learning-based algorithms

Gyeong-Yeol Lee[1], Gyung-Suk Kil[2*], Sung-Wook Kim[3*]

Cast-resin transformers have been widely installed in a power system due to excellent arc-extinguishing characteristics, easy installation, and low expense of maintenance. From the perspective of condition monitoring, identification and classification techniques of internal defects based on partial discharge (PD) measurements are getting more important. This paper studies PD defect classification using two kinds of machine learning (ML) algorithms, random forest (RF) and artificial neural network (ANN) models. Four typical PD defect models were designed: metal protrusion, a particle on insulator, delamination, and a void. PD single pulses and phase-resolved partial discharge (PRPD) patterns at each partial discharge inception voltage (PDIV) were measured by the printed circuit board (PCB) based Rogowski-type PD sensor. Various kinds of PD features were extracted from each PD single pulse and PRPD pattern. From the experimental results, the two different ML algorithms, used in this paper, could classify the PD defects with over 90%, and the PD classification rate using RF model was slightly higher than that of the ANN model.

Keywords: cast-resin transformers, PD defect classification, machine learning algorithm, Rogowski-type PD sensor

1 Introduction

Cast-resin transformers have been widely adapted in power apparatus due to their numerous advantages, such as cost-effective installation, low maintenance requirements, and self-extinguishing properties by the flame-retardant epoxy insulation [1-2]. One of the challenges faced by the cast-resin transformers is the occurrence of electrical transient partial discharge (PD) pulses generated from localized defects inside or outside of insulation layers. Since the PDs generated from the insulation defects can progressively degrade the electrical and mechanical properties of insulation systems, leading to insulation breakdown and electrical failures, eventually, it is important to detect and analyze the PD signals at the early stage of the PD mechanism [3].

From the condition monitoring and asset management perspective, PD detection techniques have been effectively adopted for several decades. Each PD signal has its own characteristics depending on the type of PD defect model. Among various PD measurement techniques, PD single pulse analysis and phase-resolved partial discharge (PRPD) pattern analysis methods have been widely applied to diagnose the insulation defects [4-5]. PD single pulse analysis figures out distinguishable PD features including time-based features, such as rising and falling times and pulse widths, frequency-based features, such as fast Fourier transform (FFT) spectra, and physical shape-based features, such as kurtosis and skewness, extracted from the PD pulses with a pulse duration time in a range of less than few microseconds. This analysis method is widely used due to its ease of application and quick and intuitive interpretation without expert judgments. The PRPD pattern analysis interprets the magnitudes and the number of PD pulses in terms of phase angles based on the accumulated PD pulses over one sinusoidal cycle of applied high voltage (HV) signals. Various PD studies have confirmed that different defect types produce different PRPD patterns with respect to the amount of charge, number of charges, and phase distribution [6-7].

Since the introduction of computerized PD detection methods for medium voltage equipment in the 1990s, there has been a lot of research into PD defect classification [8]. Nowadays, in the perspective of PD defect classification, machine learning (ML) algorithms, such as random forest (RF) and artificial neural network (ANN) models, have been receiving a lot of attention. RF algorithm is an ensemble classification model, which relies on multiple decision trees. Each decision tree independently selects the optimal classification result. The RF model is highly effective in processing high-dimensional data and offers several benefits, such as eliminating the need for feature selection, resistance to

[1] Korea Hydro & Nuclear Power Co., LTD (KHNP), Gyeongju 38120, Republic of Korea
[2*] Department of Electrical and Electronics Engineering, National Korea Maritime and Ocean University, Busing 49112, Republic of Korea
[3*] Department of Electrical and Electronics Engineering, Silla University, Busan 46958, Republic of Korea
gyeongyeollee1@gmail.com, kilgs@kmou.ac.kr, number1@silla.ac.kr

https://doi.org/10.2478/jee-2025-0059, Print (till 2015) ISSN 1335-3632, On-line ISSN 1339-309X

overfitting, robustness against noise, and strong noise tolerance [9]. In contrast, the ANN model can increase the classification accuracy by adjusting the input and output weights, along with bias terms, by training on various types of data to establish the mapping between features and categories [10].

In this paper, from the perspective of condition monitoring of the cast-resin transformers, two different kinds of ML algorithms, ANN and RF models, were prepared and compared the accuracy of PD defect classification. Figure 1 summarizes the overall process of PD defect classification using two ML algorithms used in this paper. Four different types of PD defect models were fabricated to simulate typical insulation defects in the cast-resin transformers, and the PD single pulses and PRPD patterns were measured by the printed circuit board (PCB) based Rogowski-type PD sensor at each partial discharge inception voltage (PDIV). Various statistical PD features were systematically extracted from each PD single pulse and PRPD pattern and were applied to both ML algorithms for comparison of the accuracy of PD defect classification. It was confirmed that both algorithms achieved classification accuracies over 90%, while the accuracy using RF model was slightly higher than that of ANN model.

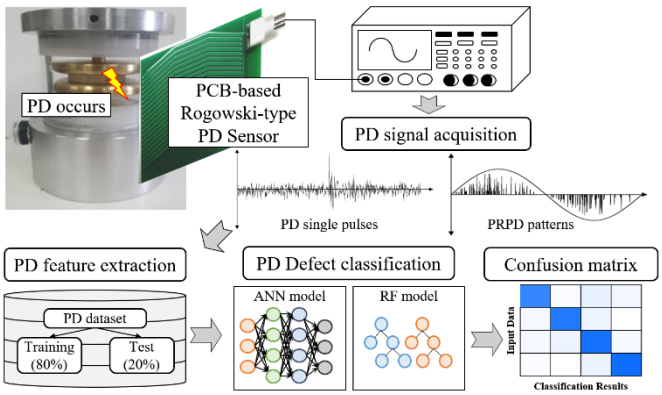


Fig. 1. Overall process of PD defect classification using ML algorithms for the cast-resin transformer

2 PD Classification algorithms

2.1 Random forest (RF) algorithm

RF algorithm is a representative ensemble classifier model which aggregates the predictions of multiple decision trees to achieve improved performance compared to a single classifier. Each decision tree is constructed from bootstrap-resampled subsets of the original PD datasets, while only a randomly selected subset of PD features is considered at each split through the feature subspace method. This two-randomization step ensures that the decision trees do not become overly similar to each other, thereby reducing the correlations among the trees and enhancing the diversity of the ensemble. As a result, the RF model improves the generalization by effectively lowering the variance while maintaining a relatively low bias. Figure 2 shows the conceptual structure of the RF algorithm, where each tree makes its own decisions independently, and the ensemble aggregates them through majority voting in classification or averaging in regression [11-13].

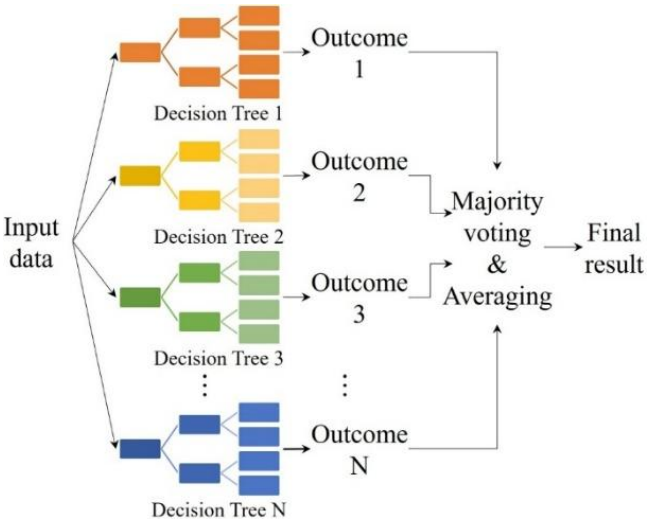


Fig. 2. Conceptual structure of the RF algorithm [14]

The final prediction $\hat{y}$ is determined by the majority voting for classification tasks using Eqn. (1) [15]:

$$\hat{y} = mode\{h_i(x)\}, \quad i = 1, 2, \cdots, n \tag{1}$$

where $h_i(x)$ denotes the prediction of each decision tree trained on a bootstrap sample D_i [16]. This sampling method ensures diversity among the trees, which enhances the overall robustness of the RF model. During the construction of each tree, node splitting is determined by a measure of impurity, such as Gini Index G or entropy H. They can be calculated using Eqns. (2) and (3), respectively [17, 18]:

$$G = 1 - \sum_{k=1}^{K} p_k^2 \tag{2}$$

$$H = -\sum_{k=1}^{K} p_k \log_2 p_k \tag{3}$$

where p_k is the proportion of samples in class k at the node, and K is the total number of classes. The classification performance of the RF model is significantly influenced by several hyperparameters, including the number of estimators, the minimum number of samples required to split per a node, and the maximum depth of each tree [19].

2.2 Artificial neural network (ANN) algorithm

ANN models learn from data, adaptively optimize features, and make accurate predictions. This algorithm is composed of a multi-layer feed forward and backward based on the ANN model. Each layer's weights and thresholds are continuously controlled. Figue 3 shows the conceptual structure of ANN model in this paper [20].

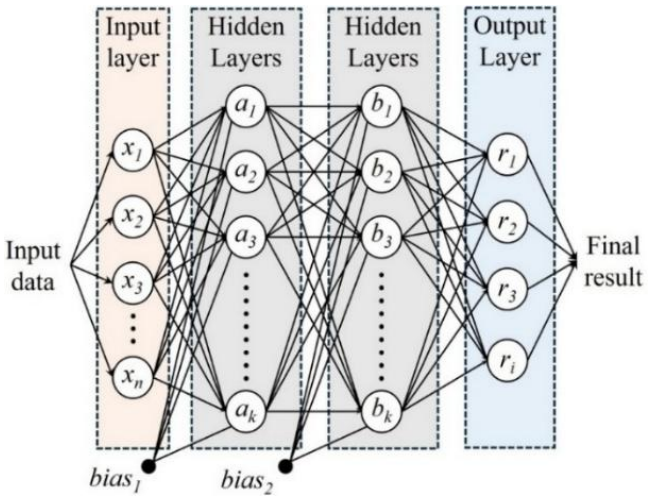


Fig. 3. Conceptual structure of the ANN algorithm

The input layer consists of several neurons that correspond to the PD dataset, labelled as x_1, x_2, ⋯, x_n in Fig. 4. Each neuron in the input layer receives PD features and passes them directly to the hidden layers.

$$a_0 = x = [x_1, x_2, \cdots, x_n] \tag{4}$$

The hidden layers extract high-level features from the received input dataset through the non-linear transformations. Each neuron in the hidden layers, labelled as a_1, a_2, ⋯, a_n in the 1st hidden layer b_1, b_2, ⋯, b_n in 2nd hidden layer, computes the weighted sum Z of the input dataset following Eqn. (5) and the computed output from each hidden layer is expressed as following Eqn. (6). The neurons apply the rectified linear unit (ReLU) which is a common non-linearity activation function in the neural network defined as $f(x)=\max(0, x)$. In general, when the input is positive, the output is equivalent to the input; otherwise, the output is zero as, shown in Eqn. (7) [21].

$$Z_{a_k}, Z_{b_k} = W(a_k, b_k) + k \tag{5}$$

$$f\left(Z_{a_k}, Z_{b_k}\right) = e^{(Z_{a_k}, Z_{b_k})} \tag{6}$$

$$f(x) = \begin{cases} x & (x > 0) \\ \alpha x & (x \le 0, \alpha \ll 1) \end{cases} \tag{7}$$

where W and K indicate the weight and the bias of each neuron, respectively, and Z presents the weight matrix.

The output layer conducts the final prediction. The neurons in the output layer which are labelled as r_1, r_2, ⋯, r_n represent the recognitions of each PD defect type. To normalize the outputs into probabilities, which makes them suitable for multi-layers classification tasks, SoftMax activation is applied as described in Eqn. (8) [22].

$$r_j = \frac{f(Z_{a_k}, Z_{b_k})}{\sum_{j=1}^{k} f(Z_{a_k}, Z_{b_k})} \tag{8}$$

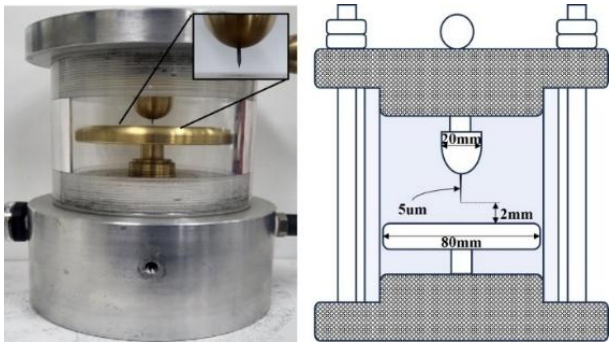
The numerator, $f(Z_{a_k}, Z_{b_k})$, represents the relative importance of the weighted sum, influenced by the positive outputs from the hidden layer after applying the ReLU activation function.

Denominator $\sum_{j=1}^{k} f(Z_{a_k}, Z_{b_k})$ normalizes these weighted scores across all PD defect types, ensuring they are represented as probabilities.

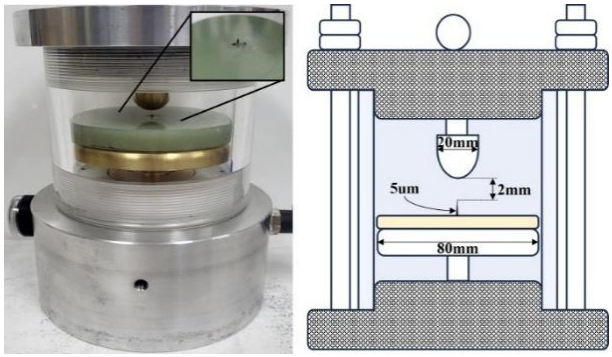
3 Experimental setup and method

3.1 PD defect models

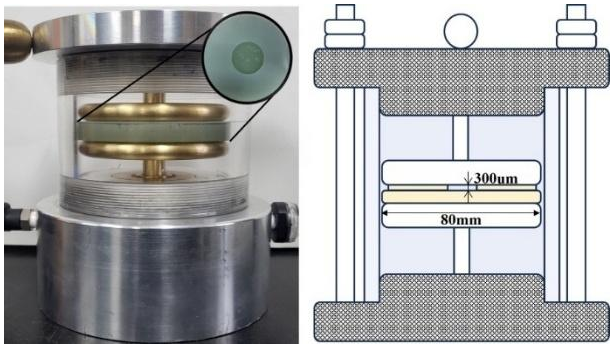
To simulate the typical defects of the cast-resin transformers, four different types of PD defect models were prepared. Figure 4 shows the photographs and their cross-sectional diagrams of PD defect models [23-25]. The defects models are a metal protrusion defect model and a particle on insulator defect model for external discharge models and a delamination defect model and a void defect model for internal discharge models. The metal protrusion and particle on insulator defect models were made of a sharp metal needle with a curvature radius of 5 μm on a HV side hemispherical electrode and an epoxy insulator with a thickness of 10 mm on a grounded plane electrode, respectively. The delamination defect model was fabricated by a stack of thin epoxy specimen with a thickness of 300 μm between HV side plane electrode and epoxy plate with a diameter of 80 mm. The void defect model was made of a void specimen with a diameter of 4 mm fabricated by injecting dry air during the epoxy curing process. All edges of the electrodes and lower and higher covers in the defect models were rounded to minimize the electric field concentration.



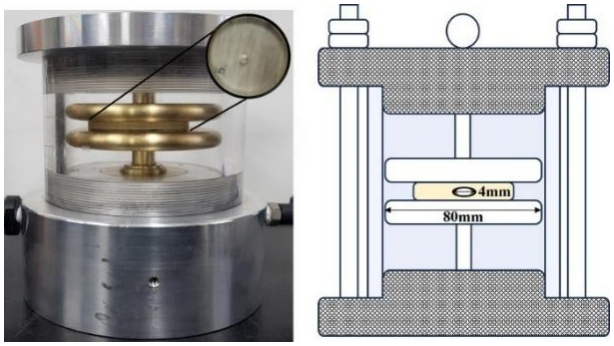
(a) Metal protrusion



(b) Particle on insulator



(c) Delamination

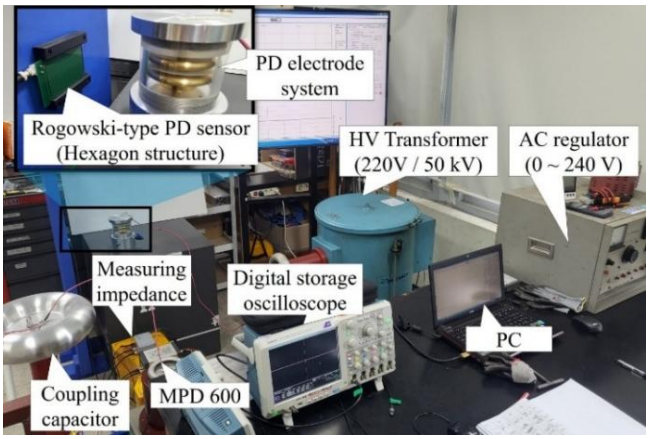


(d) Void

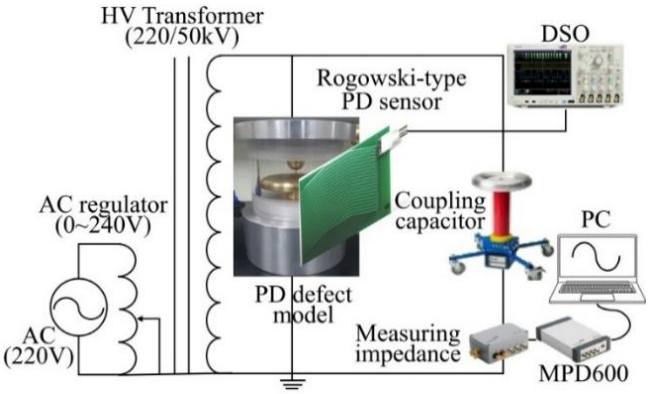
Fig. 4. PD electrode systems

3.2 Experimental setup and measuring equipment

Figure 5 presents the photograph and configuration of the experimental setup used in this study. An oil-immersed HV transformer with maximum outputs of 50 kV at 5 kVA (Model 770-5, Hipotronics) was used to apply high voltage and the voltage was controlled by the AC regulator. The Rogowski-type PD sensor was positioned 25 mm away from the PD defect models to detect the PD single pulses. The single pulses were recorded using a digital storage oscilloscope (MSO 5204B, Tektronix) with a maximum sampling rate of 10 GS/s. A high-precision measurement device (MPD 600, Omicron Corp.) with a 1 nF coupling capacitor was used to compare PD signals as the conventional detection method. Once PDs occur within the defects, a transient current flows, generating a magnetic field. The Rogowski-type PD sensor was designed to measure the magnetic field formed by PD current flowing within the defects according to Faraday's law. Based on previous experimental results [26], the winding structure was designed in a hexagonal configuration, which had the highest output sensitivity among various structures.



(a) Photograph



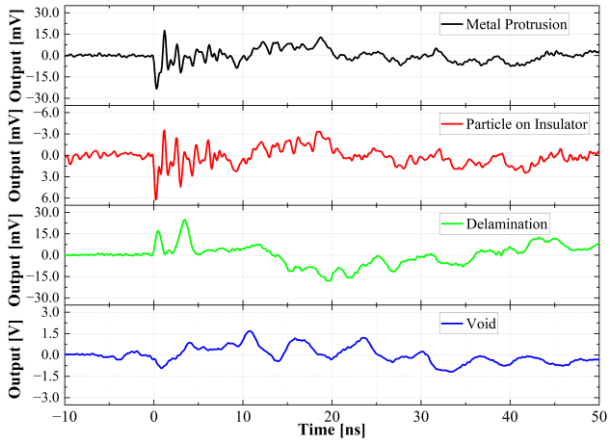
(b) Configuration

Fig. 5. Experimental setup

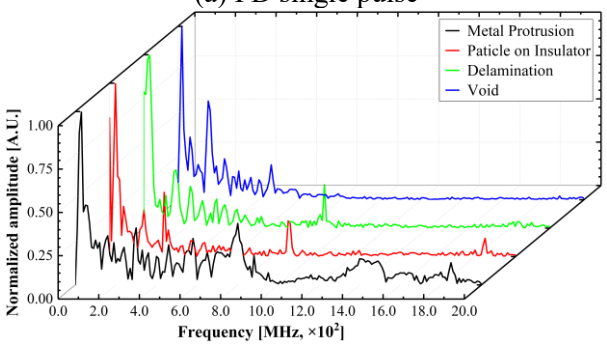
4 Results and analysis

4.1 PD signal acquisition

PD single pulses and PRPD patterns at each PDIV were measured by the Rogowski-type PD sensor in accordance with four types of PD defect models. Figure 6 shows the representative PD single pulses and their corresponding FFT spectra, respectively. With the relatively fast PD development at external discharges, the metal protrusion and particle on insulator defect models exhibited shorter rise times than delamination and void defect models, while the delamination defect model showed the longest rising time. The frequency bands of the four PD defect models were commonly distributed within the frequency range of 20 to 35 MHz.



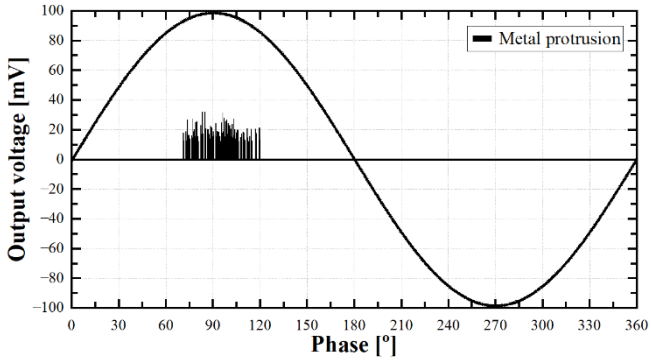
(a) PD single pulse



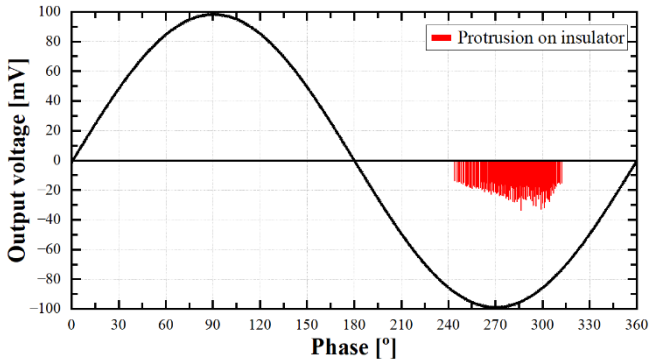
(b) FFT spectra

Fig. 6. PD single pulse and their FFT spectra

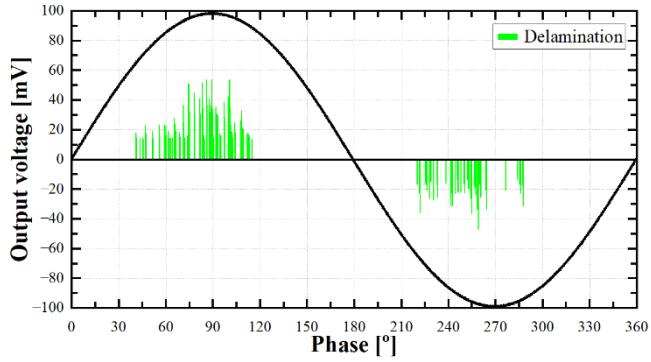
Figure 7 shows the PRPD patterns measured at each PDIV in accordance with four types of PD defect model, respectively. The PD pulses of the metal protrusion defect model were distributed from 71° to 120° in the positive polarity half-cycle, while those of the protrusion on insulator were distributed from 244° to 312° in the negative polarity half-cycle. They commonly exhibited polarity-dependent characteristics. On the other hand, for the internal discharge models, the PD pulse of the delamination and void defect models were distributed in both positive and negative polarity half-cycles. The pulses were measured from 41° to 115° in the positive polarity half-cycle and from 220° to 288° in the negative polarity half-cycle for the delamination defect model, while those were measured from 6° to 152° in the positive polarity half-cycle and from 184° to 329° in the negative polarity half-cycle for the void defect model. Among four defect models, the output voltages measured from the void defect model were about 10 to 20 times larger than other defect models.



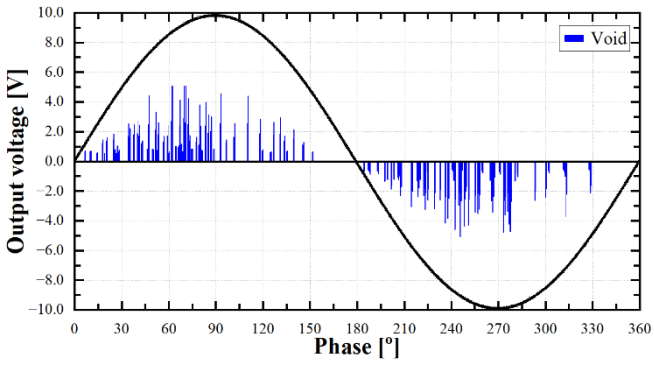
(a) Metal protrusion



(b) Particle on insulator



(c) Delamination



(d) Void

Fig. 7. PRPD patterns

4.2 PD feature extraction

Distinguishable PD feature extraction process was conducted under controlled consistent conditions to ensure that the derived features accurately represent the intrinsic characteristics of each defect model. The extracted statistical PD features were as follows: rising and falling times are the time intervals between 10 to 90% and 90 to 10% of the peak magnitude and pulse widths are the time intervals between 50% of the peak magnitude extracted from the PD single pulses as time-based features. Peak frequencies are the frequency bands extracted from the FFT spectra as frequency-based features. Pulse distribution ratios are PD pulse count ratios in every 30-degree interval extracted from the PRPD patterns as distribution-based features, and kurtosis and skewness of both positive and negative polarity half-cycles from the PRPD patterns as physical-based features [27, 28]. A total of 20 distinguishable PD features were systematically extracted from 500 samples of PD single pulses and PRPD patterns in accordance with PD defect models at each PDIV.

Based on the systematically pre-processed and extracted features, PD dataset provides a reliable foundation for evaluating and comparing ML models in PD defect classification. Tables 1 and 3 summarize the average values of each PD feature for the four PD defect models. PD dataset was randomly divided into training (80%) and test (20%) sets to ensure reliable model evaluation [29].

Table 1. Average values of time and frequency-based feature extracted from PD single pulses and FFT spectra

PD features	Metal protrusion	Particle on insulator	Delami-nation	Void
Rising time (ns)	0.73	0.80	13.03	8.70
Falling time (ns)	1.66	1.24	1.83	1.61
Pulse width (ns)	1.21	1.02	7.46	5.15
Peak frequency (MHz)	34.36	30.03	24.64	29.98

Table 2. Average values of distribution and physical-based features extracted from PRPD patterns

PD features	Metal protrusion	Particle on insulator	Delami-nation	Void
1°~30° (%)	0	–	0	9.3
31°~60° (%)	0		11.8	12.9
61°~90° (%)	41.2		28.6	16.9
91°~120° (%)	58.8		23.5	5.3
121°~150° (%)	0		0	6.2
151°~180° (%)	0		0	0.9
Kurtosis	1.86		5.53	5.91
Skewness	1.79		2.45	2.41
181°~210° (%)	–	0	0	9.8
211°~240° (%)		0	12.6	11.6
241°~270° (%)		32.82	18.5	14.2
271°~300° (%)		52.67	5.0	8.4
301°~330° (%)		14.5	0	4.4
331°~360° (%)		0	0	0
Kurtosis		–0.64	9.37	4.10
Skewness		–1.01	–3.13	–2.16

4.3 PD defect classification

PD defect classification using the RF algorithm was evaluated through a systematic exploration of various combinations to optimize the hyperparameters. The hyperparameters were set within the ranges of 2-10 for the minimum sample split, 100-500 for the number of estimators, and none and 10-50 for the maximum depth. The robust performance of the RF model is attributed to bootstrap sampling and random feature selection, which can enhance diversity among decision trees and mitigate overfitting. This increases the reliability and efficiency of the RF algorithm, especially for the classification tasks of PD defect types on high-dimensional.

Table 3 shows the accuracies of PD defect classification in accordance with combinations of key hyper-parameters. The RF model in this paper achieved the PD defect classification accuracy of 93%. It was confirmed that the number of minimum sample split had a significant impact on the PD defect classification accuracy higher than the number of estimators and the maximum depth. These results were attributed to the relatively simple structure of the PD feature dataset, which allowed us to achieve the sufficient classification accuracy even at low number of estimators and shallow maximum depths. Figure 8 shows the confusion matrix of PD defects using the RF model with the optimized key hyper-parameters.

Table 3. Average accuracies of PD defect classification using RF algorithm

Hyperparameters			Accuracy (%)
Minimum sample split	Number of estimators	Maximum depth	
20	100	20	93.0
2, 4, 6	100, 200	None, 20, 50	92.8
2, 4, 6, 8, 10	100, 200, 300, 400, 500	None, 10, 20, 30, 40, 50	92.5
8	200	50	92.0
2	100	10	91.8

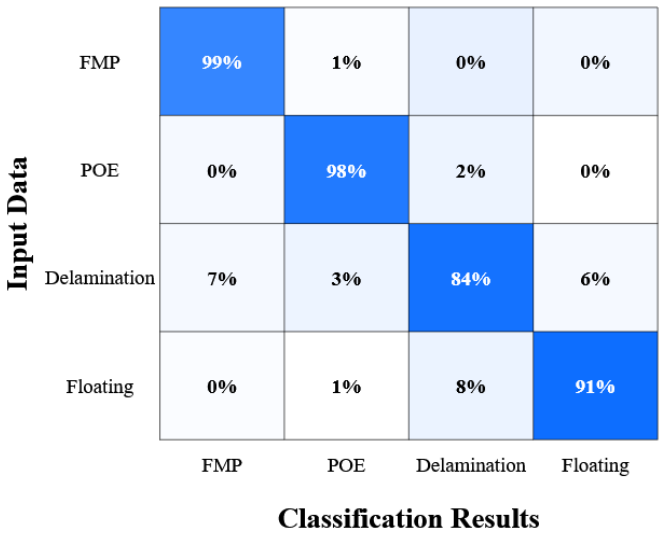


Fig. 8. Confusion matrix using RF algorithm

The ANN model was applied to the PD feature dataset to classify the types of PD defects effectively. The exploring combinations of the number of hidden layers and nodes per each layer was included in the optimization process. In this study, the ANN model with 12 input nodes, hidden layers ranging from 1 to 3 with nodes per layer varying between 1 to 256, and 4 output layers were applied. The extracted PD feature dataset were trained over 25 iterations on the input layers. The output layers were derived and compared to the target value of 1 during supervised learning processes, using backpropagation to optimize the weights and biases iteratively. To optimize the ANN architecture, the number of hidden layers and nodes per layer was systematically adjusted while using a fixed learning rate of 0.001. This optimization enabled the ANN to achieve high classification reliability, demonstrating its effectiveness as a practical tool for PD defect classification. The architecture of the ANN significantly influenced classification accuracy, with deeper networks and higher nodes yielding better performance.

A configuration of three hidden layers with 64 nodes per layer achieved the highest accuracy of 92.5%, as shown in Tab. 4. Figure 9 presents the confusion matrix of PD defect classification using the ANN model with the optimized key hyperparameters.

Table 4. Average accuracies of PD defect classification using ANN model

Hyperparameters			Accuracy (%)
Epoch	Number of hidden layers	Number of nodes per layer	
20	3	64	92.5
20, 40, 60, 80, 100	1, 2, 3	32, 64, 128	91.8
		256, 512	91.2
20, 40	1, 2	32, 128	89.8
60, 80, 100	1, 2, 3	16, 32, 64	87.8

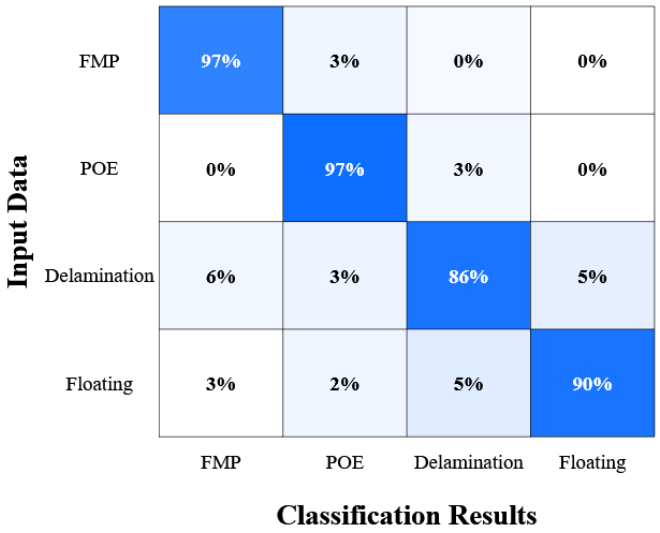


Fig. 9. Confusion matrix using ANN algorithm

5 Discussion and conclusions

This paper presents PD defect classification based on two ML algorithms for the identification and classification of PD defect types in the cast-resin transformers. PD single pulses and PRPD patterns from four defect types at each PDIV were measured using the PCB-based Rogowski-type PD sensor. A total of 20 kinds of distinguishable PD features, including time, frequency, distribution, and physical-based features, were systematically extracted. The experimental results showed that both algorithms achieved high recognition accuracies exceeding 90%. In terms of performances, the RF algorithm achieved the highest accuracy of 93% due to its faster training and greater robustness, while the ANN model achieved 92.5% accuracy due to its deeper and wider architecture, but it required longer training time and greater computational resources. The trade-off of classification accuracy and computational requirements highlights the distinct advantages of each algorithm. The RF model has advantages when rapid training, robustness against noise, and consistently high accuracy are required. In contrast, the ANN model has comparable accuracy while offering better adaptability to complex and nonlinear PD features. However, since the PD single pulses and PRPD patterns in this paper were obtained in the controlled laboratory environment and consistent conditions, the external noise signals need to be considered to enhance the classification accuracy of insulation defects of the cast-resin transformers.

In addition, future studies will investigate advanced machine learning and deep learning approaches such as long short-term memory (LSTM) networks and convolutional neural networks (CNNs) to enhance the extraction of temporal and spatial characteristics of PD signals. In addition, lightweight models will be examined to enable real-time monitoring in field environments where computational resources are constrained. Further efforts will also be directed toward incorporating additional PD defect types to improve the robustness and generalizability of the recognition framework.

References

[1] G.-M. Ma, H.-Y. Zhou, C. Shi, Y.-B. Li, Q. Zhang, C.-R. Li, "Distributed Partial Discharge Detection in a Power Transformer Based on Phase-Shifted FBG," *IEEE Sensors Journal*, vol. 18, no. 7, pp. 2788-2795, 2018.

[2] M. Sacha, T. Stefan, F. Kurt, "Detection and location of partial discharges in power transformers using acoustic and electromagnetic signals," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 15, no. 6, pp. 1576-1583, 2008.

[3] M. Mondal, G.B. Kumbhar, "A technique based on an Archimedean copula for the localization of partial discharge in a transformer winding," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 23, no. 5, pp. 2908-2916, 2016.

[4] L. Shengtao, L. Jianying, "Condition monitoring and diagnosis of power equipment review and prospective," *High voltage*, vol. 2, no. 2, pp. 82-91, 2017.

[5] R. Bartnikas, "Partial discharges. Their mechanism, detection and measurement," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol 9, no. 5, pp. 763-808, 2002.

[6] F.H. Kreuger, E. Gulski, A. Krivda, "Classification of partial discharges," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 28, no. 6, pp. 917–931, 1993.

[7] W.J.K. Raymond, H.A. Illias, A.H.A. Bakar, H. Mokhlis, "Partial discharge classifications: Review of recent progress," *Measurement*, vol. 68, pp. 164–181, 2015.

[8] E. Gulski, "Computer-aided Measurement of Partial Discharges in HV Equipment," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 28, no. 6, pp. 763-808, 2002.

[9] R. Yao, J. Li, M. Hui, L. Bai and Q. Wu, "Feature Selection Based on Random Forest for Partial Discharges Characteristic Set," *IEEE Access*, vol. 8, pp. 159151-159161, 2020.

[10] S. Lu, H. Chai, A. Sahoo and B. T. Phung, "Condition Monitoring Based on Partial Discharge Diagnostics Using Machine Learning Methods: A Comprehensive State-of-the-Art Review," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 27, no. 6, pp. 1861-1888, 2020.

[11] L. Breiman, "Random Forests," *Machine Learning*, vol. 45, pp. 5–32, 2001.

[12] R. Hussein, K.B. Shaban, A.H. El-Hag "Robust Feature Extraction and Classification of Acoustic Partial Discharge Signals Corrupted with Noise," *IEEE Transactions on Instrumentation and Measurement*, vol. 66, no. 3, pp. 405–413, 2017.

[13] B. Dai, C. Gu, E. Zhao, X. Qin, "Statistical model optimized random forest regression model for concrete dam deformation monitoring," *Structural Control and Health Monitoring*, vol. 25, no. e2170, pp. 1–15, 2018.

[14] Z. Sun, G. Wang, P. Li, H. Wang, M. Zhang, X. Liang, "An improved random forest based on the classification accuracy and correlation measurement of decision trees," *Expert Systems with Applications*, vol. 237, no. B, pp. 121549, 2024.

[15] B. Predić, G. Dimić, D. Rančić, P. Štrbac, N. Maček, P. Spalević, "Improving final grade prediction accuracy in blended learning environment using voting ensembles," *Computer Applications in Engineering Education*, vol. 26, no. 6, pp. 2294–2306, 2018.

[16] Z. Khan, N. Gul, N. Faiz, A. Gul, W. Adler, B. Lausen, "Optimal Trees Selection for Classification via Out-of-Bag Assessment and Sub-Bagging," *IEEE Access*, vol. 9, pp. 28591-28607, 2021.

[17] Z. He, J. Wang, M. Jiang, L. Hu, Q. Zou, "Random subsequence forests," *Information Sciences: an International Journal*, vol. 667, no. C, pp. 120478, 2024.
[18] Z. Jiang, Y. Zhou, Y. Zhang, G. Dong, Y. Chen, Q. Zhang, L. Zou, Y. Cao, "Classification of Depression Using Machine Learning Methods Based on Eye Movement Variance Entropy," *IEEE Access*, vol. 12, pp. 146107-146120, 2024.
[19] K. Olcay, S.G. Tunca, M.A. Özgür, "Forecasting and Performance Analysis of Energy Production in Solar Power Plants Using Long Short-Term Memory (LSTM) and Random Forest Models," *IEEE Access*, vol. 12, pp. 103299-103312, 2024.
[20] R. Sahoo, S. Karmakar, "Investigation of electrical tree growth characteristics and partial discharge pattern analysis using deep neural network," *Electric Power Systems Research*, vol. 220, pp. 109287, 2023.
[21] H. Kazuyuki, S. Daisuke, S. Hayaru, "Analysis of function of rectified linear unit used in deep learning," *2015 International Joint Conference on Neural Networks (IJCNN),* 12-17 Jul. 2015, Killarney, Ireland.
[22] M. Kuhn, K. Johnson, "Measuring Performance in Classification Models" in *Applied Predictive Modeling*, NY. Springer, 2023.
[23] S. Sharifinia, M. Allahbakhshi, T. Ghanbari, A. Akbari, and H. R. Mirzaei, "A New Application of Rogowski Coil Sensor for Partial Discharge Localization in Power Transformers," *IEEE Sensors Journal*, vol. 21, no. 9, pp. 10743-10751, 2021.
[24] M. H. Samimi, A. Mahari, M. A. Farahnakian, H. Mohseni, "The Rogowski Coil Principles and Applications: A Review," *IEEE Sensors Journal*, vol. 15, no. 2, pp. 651-658, 2015.
[25] M. Erdogan, M.K. Eker, "A Comparative Analysis of Partial Discharge in 13 Combined Insulation Structures of 11 Materials Used in Cast-Resin Dry-Type Transformers." *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 29, no. 6, pp. 2330-2339, 2022.
[26] G.-Y. Lee, S.-W. Kim, G.-S. Kil, "Design and Fabrication of Rogowski-type Partial Discharge Sensor for Insulation Diagnosis of Cast-Resin Transformers." *Journal of the Korean Institute of Electrical and Electronic Material Engineers*, vol. 35, no. 6, pp. 594-602, 2022.
[27] G.-Y. Lee, G.-S. Kil, "Insulation Defect Diagnosis Using a Random Forest Algorithm with Optimized Feature Selection in a Gas-Insulated Line Breaker," *Electronics*, vol. 14, no. 10, pp. 1940, 2025.
[28] S.-W. Kim, "Identification of Partial Discharge Defects based on Back-Propagation Algorithm in Eco-friendly Insulation Gas," *Journal of Information and Communication Convergence Engineering*, vol. 21, no. 3, pp. 233-238, 2023.
[29] S. Seitz, T. Götz, C. Lindenberg, R. Tetzlaff, S. Schlegel, "Towards Generalizable Classification of Partial Discharges in Gas-Insulated HVDC Systems Using Neural Networks: Protrusions and Particle," *IEEE Transactions on Power Delivery,* vol. 39, no. 3, pp. 1491-1499, 2024.

Gyeong-Yeol Lee received the B.Eng., M.Eng., and Ph.D. degrees in electrical engineering and high voltage engineering from National Korea Maritime and Ocean University, Korea in 2015, 2020, and 2025, respectively. He has worked in Korea Hydro & Nuclear Power Co., Ltd. (KHNP) since 2016. His research interests are partial discharge diagnosis and analysis of gas-insulated switchgears and high-voltage transformers and precision high voltage and large current measurements.

Gyung-Suk Kil received the B.Eng., M.Eng., and Ph.D. degrees from Inha University, Korea in 1984, 1987, and 1996, respectively. Since then, he has been a professor at the Division of Electrical and Electronics Engineering, National Korea Maritime and Ocean University, Korea (https://sites.google.com/view/hvlab/home). He worked as a visiting professor at Cardiff University in 2003. His research interests are in high voltage engineering, condition monitoring and diagnosis of power facilities, and lightning protection.

Sung-Wook Kim received the B.Eng., M.Eng., and Ph.D. degrees in electrical engineering and high voltage engineering from National Korea Maritime and Ocean University, Korea in 2008, 2010 and 2017, respectively. He worked for R&D center of Hyosung Corporation and was one of the members in charge of DGA project for 9 years. He is currently a professor at Silla University, Korea. His research interests are partial discharge and DGA measurements of high-voltage transformers and gas-insulated switchgears.

Received 25 September 2025