IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 13 August 2025, accepted 29 September 2025, date of publication 9 October 2025, date of current version 28 October 2025.

Digital Object Identifier 10.1109/ACCESS.2025.3619762

RESEARCH ARTICLE

Clarity-Optimized Wavelet With Autoencoder-ReliefF Ranking for Enhanced UHF PD Signal Feature Extraction

DHRUBA KUMAR[1], S. M. KAYSER AZAM[1], (Member, IEEE), WONG JEE KEEN RAYMOND[1], (Senior Member, IEEE), HAZLEE AZIL ILLIAS[1], (Senior Member, IEEE), MOHAMADARIFF OTHMAN[1], (Member, IEEE), TARIK ABDUL LATEF[1], (Senior Member, IEEE), AHMAD ABABNEH[2], A. K. M. ZAKIR HOSSAIN[3], (Senior Member, IEEE), AND MASRULLIZAM MAT IBRAHIM[3]

[1]Department of Electrical Engineering, Faculty of Engineering, Universiti Malaya, Kuala Lumpur 50603, Malaysia
[2]College of Engineering and Technology, American University of the Middle East, Egaila 54200, Kuwait
[3]Centre for Telecommunication Research and Innovation, Fakulti Teknologi dan Kejuruteraan Elektronik dan Komputer, Universiti Teknikal Malaysia Melaka, Durian Tunggal, Melaka 76100, Malaysia

Corresponding authors: S. M. Kayser Azam (kayser@um.edu.my), Mohamadariff Othman (mohamadariff@um.edu.my), and A. K. M. Zakir Hossain (zakir@utem.edu.my)

This work was supported in part by the Universiti Malaya Research Excellence Grant 2024 under Grant UMREG013-2024, in part by the Universiti Malaya (UM) Malaysia through the UM International Collaboration Grant under Grant IMG003-2023, and in part by the Universiti Malaya Centre of Research Grant (UMCoRG) 2025 under Grant CORG005-2025.

ABSTRACT This article investigates advanced signal processing methodologies, with a focus on wavelet-based techniques, for the analysis of time-domain partial discharge (PD) signals captured using ultra-high frequency (UHF) sensors. The raw signals are systematically processed through a sequence of operations including bandpass filtering, wavelet-based denoising, DC offset removal, and pulse extraction. Each processing stage is critically examined in both time and frequency domains to ensure signal integrity and noise suppression. Emphasis is placed on the optimization of wavelet parameters alongside extraction of key temporal and amplitude-based features such as time difference, charge difference, pulse height, rise time, and pulse width. To address the challenge of identifying the most discriminative features, this work integrates advanced feature ranking algorithms, namely, auto-encoder-based ranking and the ReliefF method. Their effectiveness is evaluated through criteria including training convergence, reconstruction error, and relevant quality metrics. The methodological novelty lies in the systematic fusion of optimized wavelet-based signal conditioning with robust feature selection frameworks, enabling a comprehensive assessment of feature importance and clarity. A total of 7 wavelet transformations with various decomposition levels, 17 wavelet families with various sub-categories and 2 non-wavelets are involved in this study. Comparative analysis between wavelet-based and conventional non-wavelet methods demonstrates the superior performance of the former in terms of feature extraction fidelity and signal enhancement. The findings establish that the proposed clarity-importance ranking framework significantly advances the accuracy and efficiency of UHF PD signal processing. This contributes to enhanced interpretability and reliability in PD diagnostics, thereby supporting more effective monitoring and maintenance of high-voltage insulation systems in real-world applications.

INDEX TERMS Autoencoders, denoising, feature extraction, UHF sensor, partial discharge, wavelet transform.

The associate editor coordinating the review of this manuscript and approving it for publication was Pavlos I. Lazaridis.

I. INTRODUCTION

Relevance and motivation for this research are grounded in the limitations of traditional PD signal analysis approaches,

such as pulse sequence analysis (PSA) and time-resolved partial discharge (TRPD). Electrical treeing, a major degradation mechanism in polymeric insulation, was analyzed through PD measurements in epoxy resin needle-plane samples under 0.1 Hz very low frequency excitation in [1], revealing branch-type structures and extended breakdown times compared to industrial frequencies. Nonlinear time series analysis (NLTSA) demonstrated more refined clustering than traditional pulse sequence analysis (PSA). Defect-free insulation in gas-insulated switchgear (GIS) was investigated in [2], where UHF sensors and synchronized high-speed camera imaging were employed to characterize PD signals resulting from free-moving particles under varied DC voltage and particle configurations, with PSA correlating signal amplitudes and repetition rates to particle dynamics. A high-resolution PSA framework integrating dual-sensor discrete-time integration was proposed in [3] for defect classification and severity quantification. A deep learning approach combining PSA with mathematical morphology and Bi-LSTM networks was introduced in [4] for automated defect identification, achieving 98.76% accuracy. In [5], PD characteristics in high-power, high-frequency transformers under modular multilevel converter waveforms revealed a decrease in PD inception voltage and a 50% PD reduction at 10 kHz for five-level configurations. PD behaviour in SF6 and SF6/N2 gas mixtures under negative DC was examined in [6], showing variation in PD inception voltage, amplitude, and frequency depending on gas composition and electric field distortion. PD analysis in power electronic transformers using twisted wire pair models under pulse voltage excitations (50 Hz–20 kHz) in [7] indicated that higher frequencies enhance insulation life due to lower PD amplitudes and increased discharge times. In [8], gel-encapsulated power modules showed distinct PD groups, with humidity-sensitive discharges distinguishable via time-resolved PD measurements. In [9], PD recognition in GIS was enhanced through the fusion of phase- and time-resolved PD patterns using Dempster–Shafer theory and dual neural sub-networks, improving diagnostic credibility under real defect scenarios. These studies primarily utilize PSA and TRPD frameworks.

Recent research [10], [11], [12], [13], [14] demonstrates that wavelet-based approaches offer substantial advantages over PSA and TRPD. In [10], PD signals measured using UHF planar monopole antennas were effectively filtered using stationary discrete wavelet transforms (SDWT), preserving time delays and improving PD localization even under low SNR and high pulse repetition rates. A detection framework in [11] employed wavelet time-reassigned synchro-squeezed transforms with enhanced Long short-term memory-fully connected network (LSTM-FCN) for PD sound recognition, achieving high detection accuracies in converter valve halls with edge-based and GPU-supported implementations. Convolutional autoencoders and continuous wavelet transform (CWT) time-frequency images were used in [12] for automated PD source identification in MV switchgear, demonstrating strong classification performance. In [13], deep scalogram features stacked via stacking deep scalogram features with an enhanced generative adversarial network (SDSF-EGAN) enabled 99.1% accurate PD identification using refined VGG19, ResNet101, InceptionV2, and VGG16 models. A hybrid denoising algorithm, ADTCWT, combining dual-tree complex wavelet transform and adaptive singular value decomposition, was proposed in [14], outperforming existing techniques in noise suppression while preserving signal arrival time and localization.

PD sensing and antenna technologies are further explored in [15], [16], and [17]. In [15], a tapered planar spiral antenna optimized for circular polarization (0.5–3.5 GHz) showed enhanced signal amplitude and minimal polarization-induced amplitude loss. A lotus-shaped UHF antenna with a compact equivalent circuit model was presented in [16] for early weak PD detection in high voltage direct current and high voltage alternating current (HVDC/HVAC) systems, achieving high gain and sensitivity. A comprehensive review in [17] identified design strategies, performance parameters, and future research directions in printable UHF antennas for contactless PD diagnostics.

Despite extensive research, time-domain PD signal analysis remains comparatively underexplored due to its inherent complexity and the need for advanced temporal pattern recognition techniques, which have only recently become viable through modern machine learning tools. Nevertheless, time-domain analysis offers the distinct advantage of capturing transient behaviour and temporal dependencies, often lost in frequency-domain representations.

Feature analysis methods like Neighbourhood Component Analysis (NCA) [18] and ReliefF [19] have been employed across various domains and can be applied to features extracted via TRPD, PSA, or wavelet-based transformations. However, a comprehensive integration of these techniques with wavelet-processed PD signals remains limited.

Therefore, this research is driven by the following objectives:

- The clarity of signals is to be enhanced through the application of wavelet-based denoising techniques, improving overall quality by reducing noise interference.
- Optimal parameters for wavelet-based methods are to be determined to maximize signal processing efficiency, ensuring the best performance of these techniques.
- A comparative analysis between wavelet and non-wavelet methods is to be conducted to validate the superiority of wavelet methods in feature extraction and signal analysis.

The novelty of this research is reflected in:

- Systematic optimization of wavelet families and decomposition levels for denoising and feature extraction in PD signals has not been comprehensively studied before.
- The merging of wavelet-based feature extraction with ranking techniques such as auto-encoder and ReliefF has not been explored for enhancing partial discharge (PD) signal analysis.

- A thorough comparison of wavelet-based approaches with traditional PSA and TRPD methods, assessed via advanced metrics like reconstruction error and feature quality, has not been undertaken yet.
- The integration of time series analysis with wavelet transforms, and machine learning for PD signals remains largely unexamined, offering possibilities for capturing temporal dependencies and transient PD phenomena.

Novelty of the current method is shown in Figure 1. The proposed signal processing and ranking methods are elaborated in Section II, with comparative evaluations detailed in Section III. Concluding remarks are presented in Section IV.

II. DESCRIPTION OF SIGNAL PROCESSING AND RANKING FRAMEWORK

Based on experiments, different types of signal condition techniques are theorized here. The experiment involves three PD defects: a void, a surface, and a free wire [16] as shown in Fig. 2. Tests are conducted under both HVAC and HVDC. The setup for measuring radiated PD signals is shown in Fig. 2. High voltage is applied to a defect through points 1 and 2. For HVDC tests, a 25 nF capacitor smooths the rectified DC output, ensuring a steady voltage. This capacitor only stabilizes the HVDC supply and does not influence the electromagnetic emission from the PD defects.

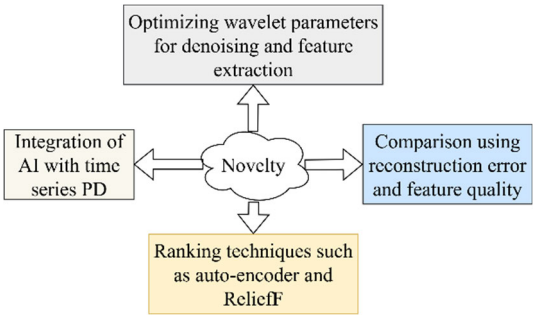

FIGURE 1. Novelty of the proposed method.

A. OUTLINE OF THE PROPOSED WORK

In Figure 3 (a), a signal processing pipeline is depicted, designed for optimizing wavelet parameters and feature extraction. The process is initiated with raw data, which is subjected to band-pass filtering. Subsequently, wavelet parameters for filtering are optimized, and wavelet filtering along with other pre-processing steps are applied. Basic algorithms, specifically TRPD and PSA, are then performed. Further optimization of wavelet parameters is conducted to extract features that surpass those obtained from TRPD and PSA. The outcomes of TRPD, PSA, and wavelet transforms are visualized for analytical purposes. Performance ranking of these methods is executed using an auto-encoder, and an alternate verification of this ranking is carried out through the application of ReliefF. This structured sequence ensures that signal features are effectively extracted and evaluated, with each step represented in a flowchart using interconnected rectangular blocks.

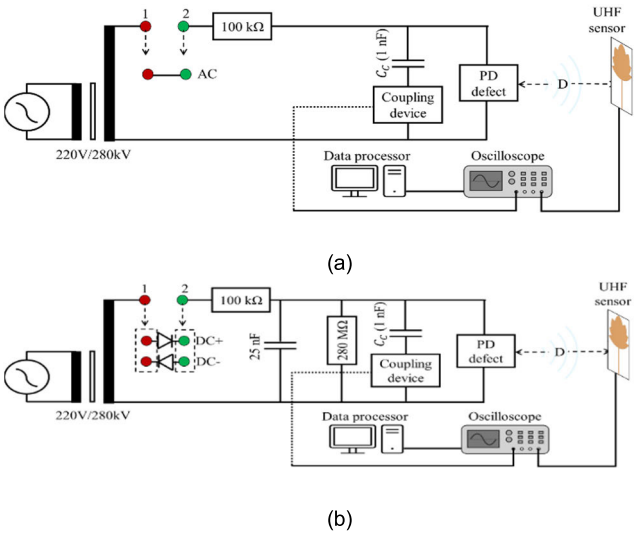

FIGURE 2. Experimental setup for UHF PD signal sensing; (a) HVAC condition (b) HVDC+ and HVDC- conditions.

B. PROCESSING THE RAW SIGNAL

In Figure 3 (b), a comprehensive flowchart for adaptive wavelet denoising and signal processing is illustrated, with an emphasis on noise removal and signal enhancement. The process is commenced by the initialization of constants, where the lower cutoff is set at 300 MHz, and the higher cutoff at is set at 3 GHz. Bandpass filter coefficients are calculated using a 6th-order Butterworth filter, which is then applied to the data. Adaptive wavelet denoising is performed on the filtered signal, employing the Optimum Wavelet and Level Method with the Universal Threshold Rule set to "Soft." Following this, the DC offset is removed from the signal, and the noise level is determined as the 50th percentile of the absolute value. An adaptive threshold is established at three times the noise level, and a pulse mask is created for absolute values exceeding this threshold. The pulse mask is then extended through dilation by structuring elements to enhance its coverage. Subsequently, the mask is applied to the filtered signal by multiplying it with the pulse signal, effectively isolating the desired signal components. The maximum absolute value of the resulting signal is calculated, and normalization is performed to finalize the process. This detailed workflow is presented in a flowchart, where each step is encapsulated within rectangular blocks connected by arrows, ensuring a clear and sequential representation of the signal processing methodology. The integration of specific parameters and advanced techniques underscores the precision required for effective noise reduction and signal enhancement, making the process robust and adaptable to varying signal conditions.

C. THEORETICAL JUSTIFICATION OF OPTIMIZING WAVELET PARAMETERS FOR TWO DIFFERENT PURPOSE

In the context of UHF partial discharge signal processing, wavelet transforms are employed for two distinct purposes:

denoising and feature extraction. For the purpose of denoising, the optimization of wavelet parameters is conducted based on energy considerations. This approach is justified by the fact that noise typically possesses significantly less energy compared to the partial discharge signals. By focusing on energy, the wavelet parameters can be tuned to effectively separate the low-energy noise from the high-energy PD components, thus enhancing the signal quality. Conversely, when the objective is feature-extraction, the optimization of wavelet parameters is performed with a focus on variance. This is because variance is a critical factor in extracting meaningful features from the signals.

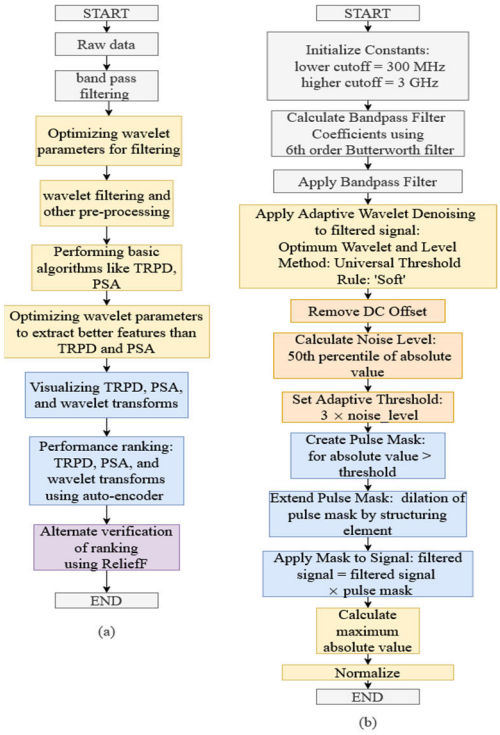

FIGURE 3. Flowcharts (a) Overall Process, (b) Filtering and Pre-processing.

By maximizing variance, the wavelet transform can capture the most significant variations in the signal, which are essential for distinguishing between different types of partial discharges or identifying specific patterns. Thus, the dual optimization strategy ensures that the wavelet transforms are tailored to the specific requirements of each task, leveraging energy for denoising and variance for feature extraction. This methodological distinction underscores the importance of parameter optimization in achieving optimal performance for both signal denoising and feature extraction in UHF partial discharge analysis.

The parameters which are to be optimized are inscribed in Table 1 which outlines the different transformation types along with the range of wavelet families and other settings to be optimized. These settings, which include families such as Daubechies, Symlets, Coiflets, and various others like Biorthogonal, Reverse Biorthogonal, Discrete Meyer, etc., are specified with decomposition levels ranging from 1 to 15 for discrete wavelets and parameters $[\beta\gamma]$, where $\beta \in [2, 4]$ and $\gamma \in [7, 25]$ for continuous wavelets. This comprehensive list of parameters is intended to be optimized prior to the results presented in Table 3, providing a foundation for selecting the most effective configurations for the wavelets.

The non-wavelet and wavelet features are described below:

- Temporal and Amplitude-Based Metrics: These include time difference, charge difference, pulse height, rise time, and pulse width. These metrics, derived from time-domain analysis using methods like PSA and TRPD, quantify the timing and magnitude of PD events, essential for distinguishing defect types (e.g., void, surface, free wire).
- Wavelet-Based Metrics: Wavelet transforms (e.g., DWT, WPT, SCALOGRAM) extract multi-resolution features.
- Energy-Based Metrics: Energy distribution across decomposition levels and frequency sub-bands (approximation and detail coefficients), reflecting the signal's spectral energy concentration.
- Coefficient-Based Metrics: Wavelet coefficients versus sample index, capturing variations in signal structure across different groups.
- Statistical Metrics: Obtained through the optimization process, which focuses on variance for feature extraction, suggesting metrics like mean, variance, or entropy of wavelet coefficients to characterize PD signal patterns.

These features are mentioned in Table 2. These metrics are processed through the signal conditioning pipeline, involving bandpass filtering, wavelet denoising, DC offset removal, pulse extraction, and normalization, to enhance signal clarity and feature quality for subsequent ranking.

D. DESCRIPTION OF THE AUTOENCODER RANKING ALGORITHMS

The methodology depicted in Figure 4 (a) illustrates a comprehensive autoencoder-based approach for feature selection and importance quantification. The process is initiated through the establishment of result arrays for storing compression ratios, training times, latent dimensions, feature weights, loss histories, and maximum weights. Current features are extracted and down-sampled, followed by normalization and conversion to GPU arrays for computational efficiency. The autoencoder architecture is determined dynamically, where the hidden layer size is expressed as a function of the input feature dimensionality. The latent dimensions and compression ratios are subsequently stored for analysis. Data partitioning is performed using an 80/20 split for training and validation sets. The autoencoder architecture follows a symmetric encoder-decoder structure (1):

$$\text{Input} \rightarrow \text{FCh} \rightarrow \text{ReLU} \rightarrow \text{FCnfeatures} \rightarrow \text{Output} \quad (1)$$

Where FC is fully connected layer, ReLU is an activation layer, n_{features} denotes the total number of input features.

TABLE 1. Different transformation types, wavelet families and other settings.

Sl. No.	Transformation Type	Range of Wavelet Families and other Settings
1	Visual representation of the Continuous Wavelet Transform (SCALOGRAM); Curvelet Transform (CURVELET); Ridgelet Transform (RIDGELET); Dual-Tree Complex Wavelet Transform (DTCWT); Adaptive Wavelet Transform (AWT); Discrete Wavelet Transform (DWT); Wavelet Packet Transform (WPT).	**For discrete wavelets:** Daubechies family (db1-db45); Symlets (sym2-sym45); Coiflet wavelet (coif1-coif5); biorthogonal (bior1.1-bior6.8); Reverse Biorthogonal Wavelet Pairs (rbio1.1-rbio6.8); Discrete Meyer (dmey); Gaussian Wavelets (gaus1-gaus8); Mexican Hat (mexh); Morlet (morl); Complex Gaussian (cgau1-cgau8); Shannon Wavelets (shan); Frequency B-Spline (fbsp); Complex Morlet (cmor); Fejer-Korovkin (fk4-fk22). Decomposition level: 1-15. **For continuous/hybrid wavelets:** Morse, Bump, and Amor for AWT, and SCALOGRAM. Wavelet parameters [β γ]: β ∈ [2,4], γ ∈ [7,25].

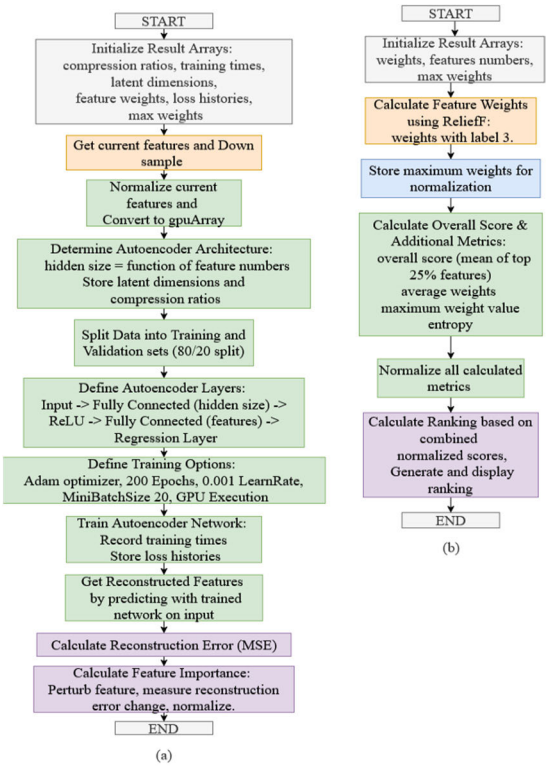

FIGURE 4. Flowchart of Ranking Algorithms (a) Auto-encoder, (b) Relieff.

Training is conducted using Adam optimization with specified hyperparameters: 200 epochs, learning rate of 0.001, and mini-batch size of 20, utilizing GPU execution for enhanced computational performance. Upon completion of training, reconstructed features are generated through forward propagation of the input data. The reconstruction error is quantified using Mean Squared Error (MSE). Feature importance is subsequently calculated through perturbation analysis. This methodology enables normalized quantitative assessment of each feature's contribution to the overall reconstruction quality. The autoencoder-based ranking methodology is implemented through a multi-metric evaluation framework that quantifies feature extraction quality across multiple dimensions. Four primary metrics are computed and normalized for comparative analysis: reconstruction quality (MSE), compression efficiency (ratio between input feature dimensionality and latent space dimensions), distribution similarity (Kullback-Leibler divergence), and training speed. The final ranking is determined through a weighted combination of these normalized metrics, mathematically expressed as (2):

$$\begin{aligned}\text{Overall Score} = {} & 0.4\times\text{Reconstruction} + 0.2\times\text{Compression} \\ & + 0.3\times\text{Distribution} + 0.1\times\text{TrainingSpeed}\end{aligned} \tag{2}$$

Methods are ranked in descending order based on their calculated overall scores, with higher scores indicating superior feature extraction performance.

Figure 4 (b) presents sophisticated multi-criteria feature selection framework that integrates multiple evaluation metrics for comprehensive feature ranking. The process commences with the initialization of result arrays containing weights, feature numbers, and maximum weight values. The feature weight calculation employs ReliefF algorithm, specifically utilizing weights associated with label 3, which can be mathematically expressed as (3):

$$W_{\text{ReliefF}} = \sum_{i=1}^{n} 1/\text{m}\left[\text{diff}\left(A, R_i, H_i\right) - \text{diff}\left(A, R_i, M_i\right)\right] \tag{3}$$

Where, W_{ReliefF} represents the ReliefF weight for attribute A, R_i denotes randomly selected instances, H_i and M_i represent nearest hit and miss instances respectively. Maximum weights are stored for normalization purposes, followed by the computation of an overall score incorporating additional metrics. The scoring mechanism evaluates the mean of the top 25% features and calculates average weights, maximum weight values, and entropy measures. The entropy calculation is performed as (4):

$$H\left(X\right) = -\sum_{i=1}^{n} p\left(x_i\right)\log_2 p\left(x_i\right) \tag{4}$$

Where, $H\left(X\right)$ represents the entropy of feature X and $p\left(x_i\right)$ denotes the probability of occurrence for value x_i.

All computed metrics undergo normalization to ensure comparable scales across different measurement units. The final ranking is determined through a combined normalized scoring approach (5):

$$S_{\text{combined}} = a{\cdot}S_{\text{ReliefF}} + b\cdot S_{\text{entropy}} + c{\cdot}S_{\text{weight}} \tag{5}$$

Where, S_{combined} represents the final combined score obtained from summation of reliefF score (S_{ReliefF}), entropy score $\left(S_{\text{entropy}}\right)$, weight score $\left(S_{\text{weight}}\right)$. a, b, c are weighting

coefficients for different metric contributions. The methodology concludes with the generation of a comprehensive feature ranking based on the integrated multi-metric evaluation, facilitating informed feature selection decisions.

E. EXPLANATION OF PERTURBATION-BASED FEATURE WEIGHTING

Section II-D describes the autoencoder-based ranking methodology, where perturbation analysis is used to calculate feature importance for PD signal features. The following step-by-step explanation is derived directly for the current research:

1) Feature Extraction and Preprocessing: Current features (e.g., wavelet coefficients, temporal metrics like pulse height, rise time) are extracted from the processed PD signal. These features are down-sampled and normalized to ensure consistent scales, then converted to GPU arrays for computational efficiency.
2) Autoencoder Architecture Definition: The autoencoder is configured with a symmetric encoder-decoder structure, as specified in (1).
3) Training the Autoencoder: The dataset is split into 80% training and 20% validation sets. The autoencoder is trained using the Adam optimizer with 200 epochs, a learning rate of 0.001, and a mini-batch size of 20, leveraging GPU execution. Training minimizes the Mean Squared Error (MSE) between the input and reconstructed output, ensuring the autoencoder learns to reconstruct PD signal features accurately.
4) Perturbation Analysis for Feature Importance: After training, perturbation analysis is conducted to assess feature importance:
5) Baseline Reconstruction: The original input features are passed through the trained autoencoder to generate a reconstructed output, and the baseline MSE is calculated to measure reconstruction quality without perturbations.
6) Feature Perturbation: Individual features are perturbed (e.g., by adding small noise or scaling), while keeping other features unchanged. The perturbed input is passed through the autoencoder to produce a new reconstructed output.
7) Impact Evaluation: The MSE for the perturbed input is computed. Features whose perturbation causes a significant increase in MSE are deemed more important, as they are critical to accurate signal reconstruction.
8) Weight Assignment: Feature importance is calculated through this perturbation analysis, with weights stored in result arrays. The increase in MSE due to perturbation reflects each feature's contribution to reconstruction quality.
9) Multi-Metric Ranking: Feature importance weights from perturbation analysis are combined with other metrics—reconstruction quality (MSE), compression efficiency (ratio of input to latent dimensions), distribution similarity (Kullback-Leibler divergence), and training speed—using a weighted formula (2). Methods are ranked based on these normalized scores, with higher scores indicating superior feature extraction performance.

The perturbation analysis identifies features critical for reconstructing PD signals. This process ensures that features contributing most to signal reconstruction are assigned higher weights, improving the selection of discriminative features for PD diagnostics.

III. COMPARATIVE EVALUATION OF METHODS AND RANKING

The wavelet and non-wavelet methods are evaluated and ranked in this section.

A. OPTIMIZING THE WAVELET FOR DENOISING

Optimizing the Wavelet for Denoising The DWT should be optimized first for signal conditioning. The optimization for wavelet denoising is illustrated in Figure 5, where the relationship between combination number and energy is depicted in Figure 5 (a), and the variation across wavelet families and decomposition levels is shown in Figure 5 (b). The energy levels are analyzed to identify the optimal combination for denoising, with a marked peak indicating the most effective parameter set. The heatmap in Figure 5 (b) highlights the performance (in terms of objective function value from energy) of different wavelet families across various decomposition levels, aiding in the selection of the best denoising configuration. The optimum point (data1) is circled red.

B. STEP-BY-STEP RESULTS OF SIGNAL PROCESSING

Figure 6 is presented to illustrate the sequential processing of a raw signal, with each subplot depicting a distinct stage of signal conditioning applied to partial discharge (PD) data. The raw signal is initially shown in the time domain Figure 6 (a) and frequency domain Figure 6 (b), where amplitude variations and a broad frequency spectrum are observed, respectively. Bandpass filtering is then applied, with the time domain representation Figure 6 (c) revealing a refined signal focused within the designated frequency range of 300 MHz to 3 GHz, and the corresponding frequency domain Figure 6 (d) displaying a narrowed spectral response. Subsequently, wavelet denoising is implemented, resulting in a cleaner time domain signal Figure 6 (e) with reduced noise, while the frequency domain Figure 6 (f) exhibits a more defined magnitude distribution. Following this, DC offset removal is performed, yielding a time domain signal Figure 6 (g) centered around zero amplitude, with the frequency domain Figure 6 (h) showing minimal low-frequency components. Pulse extraction is next applied, where the time domain Figure 6 (i) highlights isolated significant pulses, and the frequency domain Figure 6 (j) reflects a concentrated magnitude at relevant frequencies. Finally, amplitude normalization is conducted, producing a time domain signal Figure 6 (k) with scaled amplitudes for consistency

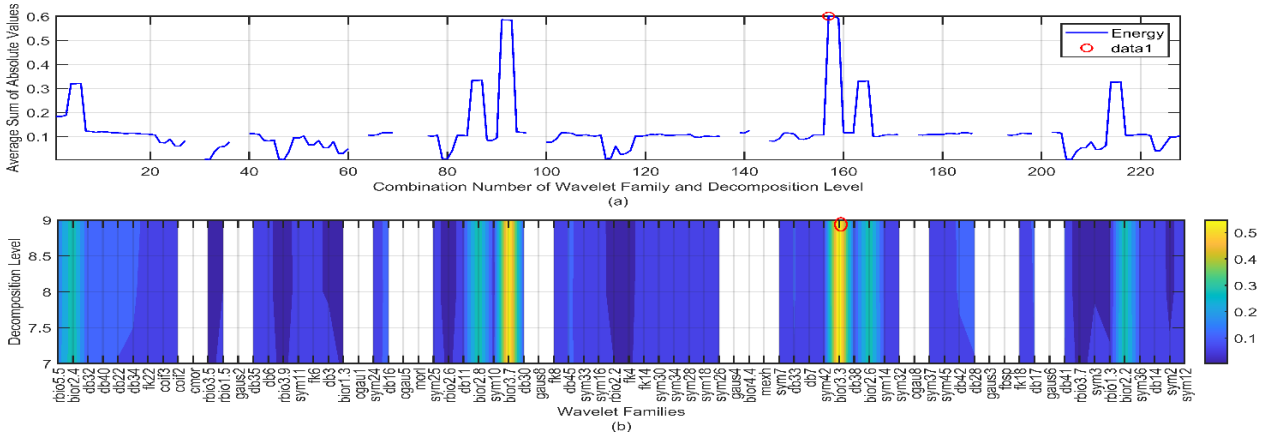

FIGURE 5. Optimization For Wavelet Denoising (a) Combination number vs energy, (b) Wavelet families vs decomposition level.

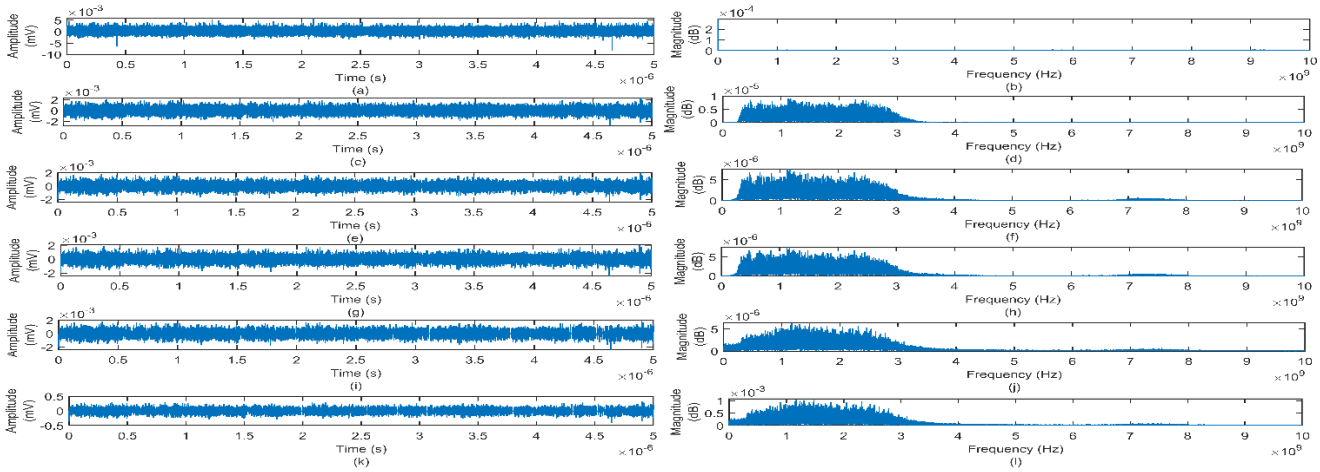

FIGURE 6. Processing the signal (a) Raw signal (T), (b) Raw signal (F), (c) After bandpass filtering (T), (d) After Bandpass filtering (F), (e) After wavelet denoising (T), (f) After wavelet denoising (F), (g) After DC offset removal (T), (h) After DC offset removal (F), (i) After pulse extraction (T), (j) After pulse extraction (F), (k) After normalization (T), (l) After normalization (F).

across cases, and a frequency domain Figure 6 (l) maintaining the spectral characteristics with adjusted magnitude levels. Bandpass filtering is executed using a sixth-order Butterworth filter targeting the specified frequency range. Wavelet denoising employs an optimum bior3.3 wavelet at level 9 for noise reduction. DC offset is removed by subtracting the mean, while pulse extraction involves an adaptive threshold set at three times the noise level, followed by morphological dilation to preserve pulse shapes. Normalization is achieved by scaling the signal amplitudes relative to the maximum absolute value, ensuring relative amplitude preservation across cases. These steps collectively enhance signal quality for subsequent transient PD or phase-resolved PD analysis.

C. OPTIMIZATION RESULTS FROM THE WAVELETS FEATURE EXTRACTION

The optimization for wavelet feature extraction is presented in Figure 7, where Figure 7 (a) displays the relationship between combination number and key parameters such as total variance, mean energy, and mean entropy. The optimal combination (data1) is highlighted by a peak in total variance, indicating the most effective parameter set for feature extraction. Figure 7 (b) illustrates the performance across various wavelet families and decomposition levels through a heatmap, aiding in the identification of the best configuration for extracting meaningful features.

The results of the optimization technique are summarized in Table 3, where various transformation types are evaluated with their respective optimum parameters. This table provides a clear reference for selecting the most effective configuration for signal processing tasks based on the optimized settings.

D. FEATURES OBTAINED FROM NON-WAVELETS (PSA, TRPD) AND WAVELETS (DWT, RIDGELET, SCALOGRAM, ETC.)

The conventional methods are evaluated first. The time series partial discharge signal is processed to generate the PSA depicted in Figure 8, where the relationship between time

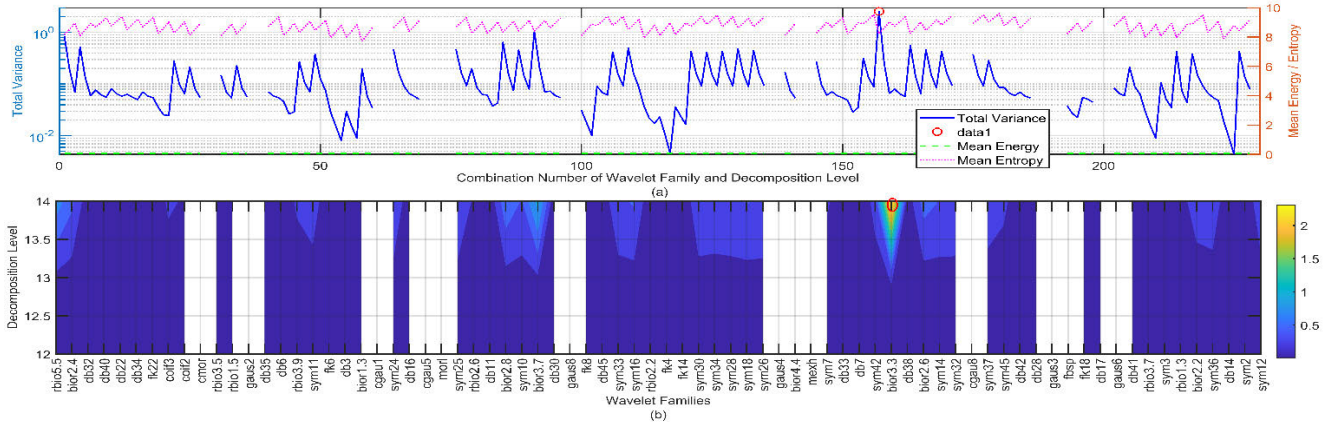

FIGURE 7. Optimization for wavelet feature extraction (a) Combination Number vs parameters, (b) Wavelet families vs decomposition level.

TABLE 2. Different types of features.

Sl. No.	Feature Type	Short Description
1	Temporal and Amplitude-Based Metrics	time difference, charge difference, pulse height, rise time, and pulse width
2	Wavelet-Based Metrics	Wavelet transforms (e.g., DWT, WPT, SCALOGRAM, etc.)
3	Energy-Based Metrics	Energy distribution across decomposition levels and frequency sub-bands
4	Coefficient-Based Metrics	Wavelet coefficients versus sample index, capturing variations
5	Statistical Metrics	mean, variance, or entropy of wavelet coefficients

TABLE 3. Results of optimization techniques.

Sl. No.	Transformation Type	Optimum Parameters
1	SCALOGRAM	Morse, Wavelet Parameters: [2 20]
2	CURVELET	Bior3.3, level 5
3	RIDGELET	Bior3.3, level 14
4	DTCWT	Bior3.3, level 13
5	AWT	Morse, Wavelet Parameters: [3 8]
6	DWT	Bior3.3, level 14
7	WPT	Bior3.3, level 7

difference and charge difference is illustrated for various groups. The data points are plotted on a logarithmic scale, with distinct colors representing different groups (e.g., Group 1-10: Surface DC+, Group 11-20: Surface DC-, Group 21-30: Void DC+, Group 31-40: Void DC-, Group 41-50: Free wire AC), highlighting the distribution and clustering patterns observed across the dataset.

The TRPD results are presented in Figure 9, where the relationships between pulse height, time, rise time, and pulse width (50%) are analyzed across different groups. The scatter plots and distribution histograms reveal distinct patterns, with pulse height versus time Figure 9 (a) and rise time Figure 9 (c) showing varied clustering, while the pulse height distribution Figure 9 (b) and pulse width (50%) versus pulse height Figure 9 (d) illustrate the frequency and spread of discharge

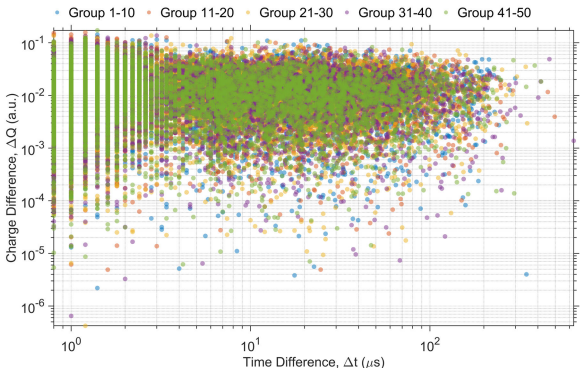

FIGURE 8. Time difference vs charge difference for PSA of different groups.

characteristics. These visualizations, color-coded by groups (e.g., Group 1-10: Surface DC+, Group 11-20: Surface DC-, Group 21-30: Void DC+, Group 31-40: Void DC-, Group 41-50: Free wire DC+), provide insights into the temporal and amplitude variations of the discharge signals.

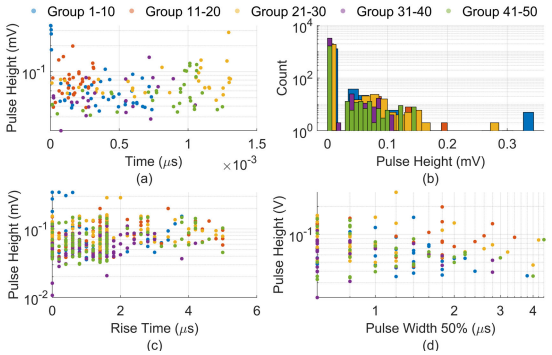

FIGURE 9. TRPD (a) Pulse height vs time, (b) Pulse height distribution, (c) Rise Time vs pulse height, (d) Pulse width (50%) vs Pulse Height.

The features of the signal are extracted using wavelet transforms, as shown in Figure 10, where energy versus sample in Figure 10 (a) and wavelet coefficients versus sample in Figure 10 (b) are plotted for different groups (e.g., Group 1-3: Free wire AC with DTCWT, Group 4-6:

Free wire DC+ with AWT, Group 7-9: Free wire DC- with CURVELET). The energy distribution and coefficient patterns are visualized on a logarithmic scale, revealing distinct variations across the sample indices. These representations provide valuable insights into the temporal and frequency characteristics of the signal.

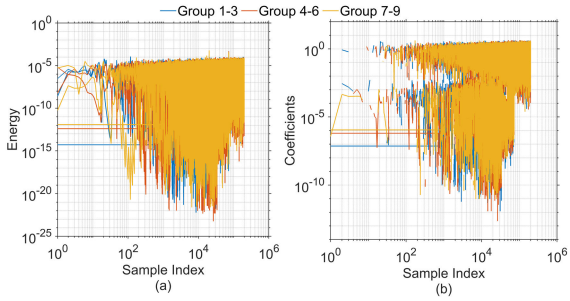

FIGURE 10. Wavelet features (a) Energy vs sample index, (b) Wavelet coefficients vs sample index.

E. AUTOENCODER RESULTS ANALYSIS

The autoencoder-based evaluation results demonstrate distinct performance patterns across nine feature extraction methods, with clear differentiation between wavelet and non-wavelet approaches. The training convergence patterns, as illustrated in Figure 11 (a), reveal that all methods achieve stable convergence within 200 epochs, with PSA and TRPD (non-wavelet methods) exhibiting slightly more volatile training trajectories compared to the wavelet-based approaches. The discrete wavelet methods display more consistent convergence behavior, while the continuous wavelet methods show intermediate stability.

Feature importance analysis, presented in Figure 11 (b), reveals significant disparities in discriminative power across methods. The comprehensive quality metrics evaluation in Figure 11 (c) provides normalized scores across four critical dimensions: reconstruction fidelity, compression efficiency, distribution preservation, and training speed. RIDGELET emerges as the superior method, achieving the highest combined score through excellent reconstruction quality and optimal compression ratios. CURVELET and SCALOGRAM follow in the ranking, demonstrating strong performance in reconstruction and distribution metrics. The continuous wavelet methods, despite their extensive feature spaces, show competitive performance in reconstruction quality but suffer from computational inefficiency penalties. The reconstruction error versus feature dimensionality analysis in Figure 11 (d) reveals an inverse relationship between feature count and reconstruction error, with methods possessing larger feature sets generally achieving lower reconstruction errors.

F. RELIEFF RESULTS ANALYSIS

The ReliefF-based evaluation presents a comprehensive assessment of feature discriminative power across the nine methods, revealing distinct performance characteristics between continuous and discrete approaches. The feature weight distribution analysis in Figure 12 (a) demonstrates that continuous wavelet methods (SCALOGRAM and WPT) exhibit highly irregular and sparse weight patterns across their extensive feature spaces, with most features showing minimal discriminative value.

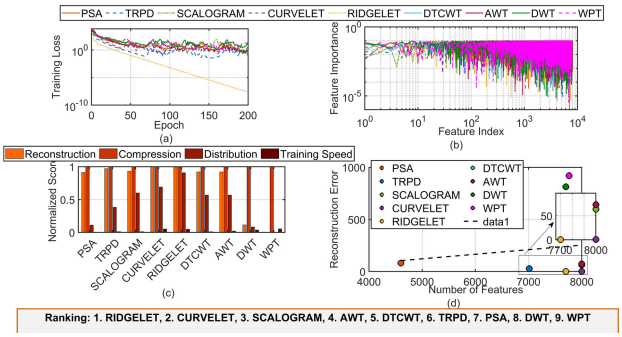

FIGURE 11. Auto-encoder based ranking (a) Training convergence, (b) Feature Importance Across Index, (c) Feature quality metrics by method, (d) Reconstruction error vs feature dimensionality.

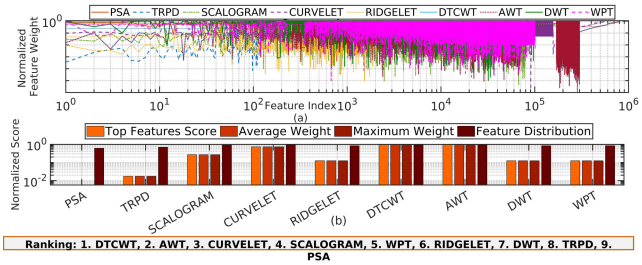

FIGURE 12. ReliefF based ranking of Wavelet and Non-wavelet methods (a) Feature weights across methods, (b) Feature quality metrics by method.

The discrete wavelet methods display more concentrated and consistent weight distributions, indicating better feature quality despite smaller dimensionalities. The multi-metric evaluation in Figure 12 (b) incorporates four normalized criteria: top features score, average weight, maximum weight, and feature distribution entropy. DTCWT achieves the highest ranking through superior performance across all evaluation metrics, particularly excelling in top features score and maximum weight values. AWT follows closely, demonstrating strong discriminative capabilities and balanced feature distributions.

G. JUSTIFICATION OF DIFFERENT RANKING AND COMPARATIVE ANALYSIS: AUTOENCODER VS RELIEFF

While the theoretical view favored autoencoders, there are critical practical limitation: autoencoders suffer from the curse of dimensionality and become computationally intractable with very large feature sets due to their neural network architecture requiring matrix operations that scale poorly with feature count. Also, the down sampling has

reduced the performance of autoencoder. ReliefF, being a statistical algorithm that evaluates features through nearest neighbor comparisons, can handle much higher dimensional datasets more efficiently. The autoencoder's memory requirements and training complexity grow exponentially with feature dimensionality, often making it impractical for datasets with tens of thousands of features, whereas ReliefF maintains reasonable computational complexity even with very high-dimensional data.

H. COMPUTATIONAL LOAD AND RESOURCE USAGES

In this section the computational load and resource usage of autoencoder and ReliefF are compared. Autoencoders, as neural network-based architectures, involve matrix operations that scale poorly with feature dimensionality, leading to high computational complexity. The current research notes that their memory requirements and training complexity grow exponentially with feature count, making them less practical for datasets with tens of thousands of features. Key factors contributing to this computational load include training time and resource usage.

The autoencoder training process, as described in Section II-D, involves 200 epochs with a learning rate of 0.001 and a mini-batch size of 20, executed on a GPU. The training time is influenced by the input feature dimensionality, the symmetric encoder-decoder architecture (1), and the need for forward propagation and perturbation analysis to compute feature importance. Figure 11(a) shows that autoencoders achieve stable convergence within 200 epochs, but the training trajectory is more volatile for non-wavelet methods (PSA, TRPD) compared to wavelet-based methods. For high-dimensional datasets (e.g., continuous wavelet methods like SCALOGRAM and WPT with extensive feature spaces), the computational burden increases significantly due to the large number of parameters and matrix computations. While exact training times are not important, the current research implies that autoencoders require substantial time for large feature sets, potentially taking hours or more on standard GPU hardware for datasets with thousands of features.

Autoencoders demand significant memory due to the need to store weights for fully connected layers (FC), activations (ReLU), and latent representations. The architecture's hidden layer size is a function of input feature dimensionality, which amplifies memory usage as feature count grows. For instance, with tens of thousands of features (common in continuous wavelet transforms), the memory requirements for storing and manipulating matrices during training can exceed several gigabytes, potentially overwhelming standard GPU memory (8–16 GB). Current research also mentions down-sampling to mitigate this, but this reduces performance, suggesting a trade-off between computational feasibility and accuracy. The use of GPU arrays for computational efficiency indicates reliance on high-performance hardware, which may not be universally accessible.

Next, the computational load of ReliefF is considered. ReliefF, a statistical feature selection algorithm, evaluates features through nearest-neighbor comparisons, offering better scalability for high-dimensional data. The current research states that ReliefF maintains reasonable computational complexity even with very high-dimensional datasets, making it more efficient than autoencoders for large feature sets.

ReliefF's computational process, outlined in Section II-D and Figure 4(b), involves calculating feature weights based on nearest hit and miss instances (3), followed by multi-metric evaluations (4) and (5). Unlike autoencoders, ReliefF does not require iterative training epochs or backpropagation, significantly reducing computation time. The algorithm's complexity is primarily driven by the number of instances and features, but it scales linearly or near-linearly with feature count, as it evaluates features independently through statistical comparisons. For a dataset with tens of thousands of features, ReliefF can typically complete its computations in minutes to tens of minutes on standard CPU hardware, compared to hours for autoencoders on GPUs. The emphasis on ReliefF's efficiency suggests that its runtime is orders of magnitude lower than that of autoencoders for high-dimensional PD signal datasets.

ReliefF has modest memory requirements, as it primarily stores feature weights, instance distances, and metric scores (e.g., entropy, top feature scores). The memory footprint scales linearly with the number of features and instances, requiring significantly less memory than autoencoders. For example, a dataset with 10,000 features and 1,000 instances might require only a few hundred megabytes of memory for ReliefF, compared to gigabytes for autoencoders. ReliefF's computations can be efficiently performed on standard CPUs, eliminating the need for specialized GPU hardware, which enhances its practical applicability in resource-constrained environments. Quantitative Comparison can be made in this regard:

Training time of autoencoder scales approximately as $O\left(N_f^2 \times N_i \times E\right)$, where N_f is the number of features, N_i is the number of instances, and E is the number of epochs (200 in this case). For a dataset with 10,000 features and 1,000 instances, training on a GPU takes several hours (2)–10 hours) depending on hardware and optimization. Computation time of ReliefF scales approximately as $O\left(N_f \times N_i \times k\right)$, where k is the number of nearest neighbors. For the same dataset, ReliefF completes in 5–30 minutes on a standard CPU, reflecting its linear scaling and lack of iterative training.

Memory usage of autoencoder scales with $O\left(N_f^2\right)$ due to matrix operations and weight storage. A dataset with 10,000 features requires 4–16 GB of GPU memory, potentially

exceeding available resources for larger feature sets. Memory usage of ReliefF scales with $O\left(N_f \times N_i\right)$, is requiring 100–500 MB for a similar dataset, easily manageable on standard CPUs.

The discussion in Section III-G underscores that autoencoders become computationally intractable for very large feature sets due to exponential growth in memory and training time, whereas ReliefF's statistical approach allows it to handle high-dimensional data efficiently. For PD signal analysis, where continuous wavelet transforms (e.g., SCALOGRAM, WPT) generate extensive feature spaces, ReliefF's ability to process these datasets quickly and with lower resource demands makes it preferable for practical applications. The autoencoder's higher computational load is justified only when its reconstruction-based feature ranking (emphasizing reconstruction fidelity and compression efficiency) provides unique insights not captured by ReliefF's statistical metrics.

I. AUTOENCODER VS RELIEFF

The research highlights that autoencoders and ReliefF offer cross-verification of feature rankings. By applying both methods, the current research ensure that the selected features are not only significant in terms of signal reconstruction (autoencoder) but also effective for distinguishing PD defect types (ReliefF). This dual approach mitigates the risk of overfitting to a single method's biases—autoencoders might prioritize features irrelevant to classification, while ReliefF might overlook features critical for signal integrity. Autoencoders provide a deep learning perspective, emphasizing features that preserve the overall signal structure, while ReliefF offers a statistical perspective, focusing on features that maximize class separability. This combination ensures that selected features are both representative of the signal's intrinsic properties and effective for diagnostic classification, critical for PD applications in high-voltage systems.

Based on Figure 12, a complementary use of autoencoder and ReliefF is shown in Figure 13. A simplified block diagram for the Autoencoder vs. ReliefF decision process is presented, starting with an input block for PD signal features. A decision diamond, assessing feature dimensionality, directs high-dimensional data to the ReliefF process and lower-dimensional data to the Autoencoder process. Each process block outlines key steps: ReliefF calculates weights, entropy, and a combined score (3)–(5), while the Autoencoder involves training, perturbation analysis, and metric scoring (1)–(2). Both paths converge to an output block, delivering the feature ranking, with DTCWT highlighted as the top-ranked method.

J. COMPARISON WITH OTHER METHODS

The current method is compared with other methods in Table 4. Various methods, including PSA, Deep learning PSA, basic wavelets, Scalogram, DTCWT, and the current method, are listed and evaluated. The normalized ranking index is obtained after optimization and evaluating the ranking algorithm (RefiefF). Remarks are provided for each method, indicating that preliminary techniques for PD are utilized, deep learning performance is limited by initial PSA features, and more improvements are needed for Basic Wavelets. Additionally, it is noted that better wavelet options and parameter optimization are required for Scalogram and DTCWT, while the current method is described as having optimized wavelet parameters and ranking highly with significantly improved DTCWT.

The algorithm could use DFT and DCT for comparison, but their utility is limited compared to wavelet transforms for PD signal analysis due to the transient, non-stationary nature of the signals. They could serve as supplementary tools for frequency-domain analysis or feature extraction but are unlikely to replace wavelets given in the findings (e.g., DTCWT's ranking index of 63.97, Table 4).

TABLE 4. Comparison summary.

Sl. No.	Method	Ranking Index (Out of 100)	Remarks
1.	PSA [1], TRPD [7]	21.7, 27.05 respectively	Preliminary/naïve techniques for PD
2.	Deep learning PSA [4]	Same as PSA (21.17)	Deep learning can't perform better if the initial PSA features are not good.
3.	Basic Wavelets [10]	49.88 (unoptimized)	More improvements needed
4.	Scalogram [12], [13]	58.81 (unoptimized)	Better wavelet is available, and the parameters are not optimized
5.	DTCWT [14]	61.17 (unoptimized)	The parameters could be optimized
6.	This method	DTCWT (rank 1) has improved to 63.97.	Wavelet parameters are optimized and ranked. Much improved DTCWT is obtained.

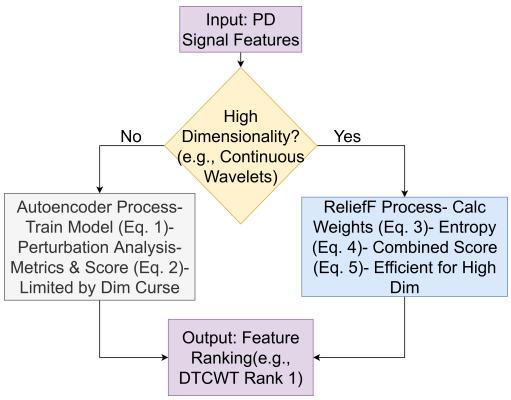

FIGURE 13. Complementary approach of autoencoder and ReliefF.

IV. CONCLUSION

This research effectively addresses key challenges in partial discharge (PD) signal analysis by establishing a systematic framework for signal conditioning and optimized wavelet-based feature extraction. In this study, a total of 7 wavelet transformations, each with distinct decomposition levels, along with 17 wavelet families, categorized into various sub-groups, and 2 non-wavelet methods,

have been incorporated. The absence of structured methodologies for selecting suitable wavelet parameters, critical for enhancing denoising and feature quality, was resolved through rigorous optimization procedures. The integration of wavelet-extracted features with advanced ranking algorithms, specifically autoencoder and ReliefF, presents a novel contribution to PD diagnostics, offering refined feature importance assessment not previously explored in this domain. The ReliefF ranks as the algorithm DTCWT, AWT, CURVELET, SCALOGRAM, WPT, RIDGELET, DWT, TRPD and PSA with ranking indices of 63.97, 60.1, 59.95, 59.22, 57.79, 52.7, 50.11, 27.05 and 21.7 respectively. A major computational barrier arising from high-dimensional feature spaces was also overcome, enabling tractable processing and model training without compromising diagnostic accuracy. The methodologies developed herein demonstrate significant improvements in signal clarity and feature relevance, directly contributing to more accurate and efficient PD diagnostics in practical applications such as high-voltage substations, switchgear, and transformers. The work lays a solid basis for implementing the intelligent condition monitoring system in future high-voltage industries. The final outcome and future scopes are highlighted below:

- Established a systematic framework for PD signal conditioning and optimized wavelet-based feature extraction, resolving the lack of structured methodologies for wavelet parameter selection.
- Integrated wavelet-extracted features with autoencoder and ReliefF for novel, refined feature importance assessment in PD diagnostics, overcoming high-dimensional feature space computational barriers.
- Achieved significant improvements in signal clarity and feature relevance, leading to more accurate and efficient PD diagnostics in high-voltage applications.
- Lays a solid basis for implementing intelligent condition monitoring systems in future high-voltage industries.
- Poisson-Gaussian noise are naturally present in the PD signals. Such noise is plausible due to sensor limitations and background electromagnetic interference in UHF PD measurements. The algorithm's wavelet-based denoising could partially address Poisson-Gaussian noise, though specific preprocessing (e.g., variance-stabilizing transforms) would enhance performance. Special type of noise (e.g., Speckle noise) can be considered as a future work.

REFERENCES

[1] P. Donoso, R. Schurch, J. Ardila, and L. Orellana, "Analysis of partial discharges in electrical tree growth under very low frequency (VLF) excitation through pulse sequence and nonlinear time series analysis," *IEEE Access*, vol. 8, pp. 163673–163684, 2020, doi: 10.1109/ACCESS.2020.3022292.

[2] P. Wenger, M. Beltle, S. Tenbohlen, U. Riechert, and G. Behrmann, "Combined characterization of free-moving particles in HVDC-GIS using UHF PD, high-speed imaging, and pulse-sequence analysis," *IEEE Trans. Power Del.*, vol. 34, no. 4, pp. 1540–1548, Aug. 2019, doi: 10.1109/TPWRD.2019.2909830.

[3] K. Zhang, A. Reid, D. Clark, M. Michelarakis, and A. M. Haddad, "Diagnostics of partial discharge measurements utilizing multi-sensor temporal pulse sequence analysis," *IEEE Access*, vol. 12, pp. 88992–89001, 2024, doi: 10.1109/ACCESS.2024.3419082.

[4] R. Das, A. K. Das, B. Chakraborty, S. Chatterjee, S. Dalai, B. Chatterjee, and K. Bhattacharyya, "Mathematical morphology-aided bidirectional long short-term memory network-based partial discharge pulse sequence classification," *IEEE Trans. Plasma Sci.*, vol. 51, no. 12, pp. 3454–3461, Dec. 2023, doi: 10.1109/TPS.2023.3330891.

[5] R. Shao, J. Jiang, W. Chen, and Z. Shen, "Partial discharge characteristics of polyimide under the high-frequency multilevel stepped waveform," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 32, no. 2, pp. 1113–1122, Apr. 2025, doi: 10.1109/TDEI.2024.3469157.

[6] X. Han, X. Zhang, H. Wang, Y. Sun, Y. Zhou, and J. Li, "Partial discharge characteristics of SF_6 and SF_6/N_2 in nonuniform electric field under negative DC voltage," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 30, no. 1, pp. 131–138, Feb. 2023, doi: 10.1109/TDEI.2022.3208512.

[7] J. Jiang, B. Zhang, Z. Li, P. Ranjan, J. Chen, and C. Zhang, "Partial discharge features for power electronic transformers under high-frequency pulse voltage," *IEEE Trans. Plasma Sci.*, vol. 49, no. 2, pp. 845–853, Feb. 2021, doi: 10.1109/TPS.2021.3053960.

[8] M. Sherriff, A. Griffo, C. Jia, C. Ng, Z.-Q. Zhu, and D. A. Stone, "Partial discharge in silicone gel on power module substrates in high-humidity conditions," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 32, no. 1, pp. 484–493, Feb. 2025, doi: 10.1109/TDEI.2024.3404367.

[9] L. Li, J. Tang, and Y. Liu, "Partial discharge recognition in gas insulated switchgear based on multi-information fusion," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 22, no. 2, pp. 1080–1087, Apr. 2015, doi: 10.1109/TDEI.2015.7076809.

[10] G. V. R. Xavier, R. de A. Coelho, H. S. Silva, A. J. R. Serres, E. G. da Costa, and A. S. R. Oliveira, "Partial discharge location through application of stationary discrete wavelet transform on UHF signals," *IEEE Sensors J.*, vol. 21, no. 21, pp. 24644–24652, Nov. 2021, doi: 10.1109/JSEN.2021.3116491.

[11] Y. Rui, L. Yunpeng, G. Jianghai, W. Yimiao, L. Yuling, and Z. Lin, "Detection of partial discharge acoustic signals in converter valve Halls using multivariate time–frequency slices," *IEEE Sensors J.*, vol. 25, no. 12, pp. 21836–21846, Jun. 2025, doi: 10.1109/JSEN.2025.3560991.

[12] S. Barrios, D. Buldain, M. P. Comech, and I. Gilbert, "Partial discharge identification in MV switchgear using scalogram representations and convolutional AutoEncoder," *IEEE Trans. Power Del.*, vol. 36, no. 6, pp. 3448–3455, Dec. 2021, doi: 10.1109/TPWRD.2020.3042934.

[13] B. Vigneshwaran, R. Manimala, and M. M. Iqbal, "Stacking deep scalogram features with the enhanced GAN for defect recognition in power transformers with highly imbalanced partial discharge signals," *IEEE Trans. Dielectr. Electr. Insul.*, p. 1, 2025, doi: 10.1109/TDEI.2025.3566933.

[14] M. Ghorat, G. B. Gharehpetian, H. Latifi, and M. A. Hejazi, "A new partial discharge signal denoising algorithm based on adaptive dual-tree complex wavelet transform," *IEEE Trans. Instrum. Meas.*, vol. 67, no. 10, pp. 2262–2272, Oct. 2018, doi: 10.1109/TIM.2018.2816438.

[15] Y. R. Yadam, S. Ramanujam, and K. Arunachalam, "Study of polarization sensitivity of UHF sensor for partial discharge detection in gas insulated switchgear," *IEEE Sensors J.*, vol. 23, no. 2, pp. 1214–1223, Jan. 2023, doi: 10.1109/JSEN.2022.3224475.

[16] S. M. K. Azam, M. Othman, H. A. Illias, T. A. Latef, D. Fahmi, W. J. K. Raymond, W. N. L. Wan Mahadi, A. K. M. Z. Hossain, M. Z. A. A. Aziz, and A. Ababneh, "Unknown PD distinction in HVAC/HVDC by antenna-sensor with pulse sequence analysis," *Alexandria Eng. J.*, vol. 91, pp. 457–471, Mar. 2024, doi: 10.1016/j.aej.2024.02.013.

[17] S. M. K. Azam, M. Othman, H. A. Illias, T. A. Latef, M. T. Islam, and M. F. Ain, "Ultra-high frequency printable antennas for partial discharge diagnostics in high voltage equipment," *Alexandria Eng. J.*, vol. 64, pp. 709–729, Feb. 2023, doi: 10.1016/j.aej.2022.11.026.

[18] D. Kumar, W. J. K. Raymond, H. A. Illias, H. B. Mokhlis, M. B. Othman, and A. S. B. M. Khairuddin, "Twin network approach for multi-fault classification in power transformers using enriched features and zero synthetic data," *IEEE Trans. Dielectr. Electr. Insul.*, p. 1, 2025, doi: 10.1109/TDEI.2025.3563316.

[19] Z. Fang, D. Hu, R. Zheng, T. Jiang, F. Gao, and J. Cao, "Multiple artifact detection based on adaptive scalp region selection and classifier fusion," *IEEE Sensors J.*, vol. 24, no. 6, pp. 8438–8449, Mar. 2024, doi: 10.1109/JSEN.2024.3358911.

DHRUBA KUMAR received the B.Tech. degree in electrical engineering from the WBUT, India, in 2012, the M.E. degree in electrical and electronics engineering from the Birla Institute of Technology Mesra, Ranchi, India, in 2015, and the Ph.D. degree in electrical engineering from the National Institute of Technology, Durgapur, India.

He was with the Department of Electrical Engineering, National Institute of Technology Patna, Patna, India, as a Faculty Member. He is currently with the Universiti Malaya as a Postdoctoral Research fellow. He has published many journals and transactions related to power systems. His research interests include smart grids, high voltage, application of artificial intelligence, and machine learning in power systems. He is an active Reviewer for many journals and transactions.

S. M. KAYSER AZAM (Member, IEEE) was born in Pabna, Bangladesh. He received the B.Sc. degree in electrical and electronic engineering from Khulna University of Engineering and Technology, Bangladesh, in 2013, the M.Sc. (Research) degree in electronics engineering from the International Islamic University Malaysia, Kuala Lumpur, Malaysia, in 2019, and the Doctor of Philosophy degree in the Department of Electrical Engineering from the Universiti Malaya, Kuala Lumpur, Malaysia, in 2024.

He was the Principal Investigator and a Co-Researcher on various internally and internationally funded research projects. He had experience of working as a Power Engineer with Huawei Technologies Ltd. He has authored and co-authored numerous articles in top-tier journals and reputable conference proceedings, with several receiving Best Paper Awards. He is currently an Academic Staff with the Department of Electrical Engineering, Faculty of Engineering, Universiti Malaya, Malaysia. His research interests include applied electromagnetics, antenna sensors, microwave components, chipless RFID, partial discharge diagnosis, and the IoT sensors. He is an active Reviewer of IEEE TRANSACTIONS ON INSTRUMENTATION AND MEASUREMENT, IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS II: EXPRESS BRIEFS, IEEE ACCESS, and Measurement (Elsevier).

WONG JEE KEEN RAYMOND (Senior Member, IEEE) received the Bachelor of Electrical and Electronics Engineering (Hons.) and the Master of Electrical Engineering degrees from the Universiti Tenaga Nasional, in 2010 and 2013, respectively and the Ph.D. degree from the Universiti Malaya, in 2016.

He was a Product Engineer with Motorola Solutions, from 2010 to 2011, a Senior Lecturer, from 2016 to 2018, and a Program Leader with Tunku Abdul Rahman University College, from 2019 to 2021. Since 2021, he has been a Senior Lecturer with Universiti Malaya. He is currently a Chartered Engineer (C.Eng. U.K.) and a Professional Engineer (P.Eng. BEM), Member of Institution of Engineering and Technology (IET), Member of Institution of Engineers Malaysia (IEM), and a Professional Technologist of Malaysia Board of Technologist (MBOT). His research interests are insulation condition monitoring, partial discharge phenomenon, machine learning, and deep learning for fault localization and classification. He was the Honorary Secretary for IEEE Dielectrics and Electrical Insulation Society (DEIS) Malaysia since 2020.

HAZLEE AZIL ILLIAS (Senior Member, IEEE) received the bachelor's degree in electrical engineering from Universiti Malaya, Kuala Lumpur, Malaysia, in 2006, and the Ph.D. degree in electrical engineering from the University of Southampton, Southampton, U.K., in 2011.

Since 2011, he has been an Academic Staff with the Department of Electrical Engineering, Universiti Malaya, and the Head of UM High Voltage Research Group (UMHVRG). He is currently registered as a Chartered Engineer (C.Eng. U.K.) and a Professional Engineer (P.Eng. Mal). His research interests include modeling and measurement of partial discharge in solid dielectric insulation, lightning overvoltage, transmission line modeling, optimization methods, and artificial intelligence. He is a Fellow of the Institution of Engineering and Technology (IET) and a Corporate Member of the Institution of Engineers Malaysia.

MOHAMADARIFF OTHMAN (Member, IEEE) received the B.Eng. degree (Hons.) in electronic engineering from Multimedia University, Malaysia, in 2006, and the M.Sc. degree in RF and microwave and the Ph.D. degree in antenna and propagation from Universiti Sains Malaysia (USM), Malaysia, in 2008 and 2015, respectively.

He has been with the Department of Electrical Engineering, Universiti Malaya, Malaysia, as a Senior Lecturer, since 2016. Prior to that, he was a Lecturer with UCSI University, Malaysia for almost 1.5 years. His research interests include 5G antennas, dielectric characterization, dielectric resonator antenna design, and antenna design optimization. He is a member of the Board of Engineers Malaysia.

TARIK ABDUL LATEF (Senior Member, IEEE) received the B.Sc. degree in electrical and electronic engineering from the University of Oita, Japan, in 1997, the M.Sc. degree in electrical engineering from the University of Leeds, in 2005, and the Ph.D. degree from The University of Sheffield, U.K., in 2011.

He is currently a Senior Lecturer with the Department of Electrical Engineering, Faculty of Engineering, University of Malaya, Malaysia. His research interests include circularly polarized antennas, conformal antennas, frequency independent antennas, array design, dielectric resonator antennas, superconducting metamaterial antennas, and RCS reduction.

AHMAD ABABNEH was born in Irbid, Jordan. He received the B.S. degree in electrical engineering from Mutah University, Jordan, in 2003, the M.S. degree in electrical engineering from Arizona State University, in 2007, and the Ph.D. degree in electrical engineering from Kansas State University, in 2010.

He was a Professor with the Electrical Engineering Department, Jordan University of Science and Technology. He is currently an Academic Staff with the College of Engineering and Technology, American University of the Middle East, Kuwait. His research interests include Bayesian filtering, distributed detection and estimation, sensor networks, and 5G communication systems.

A. K. M. ZAKIR HOSSAIN (Senior Member, IEEE) received the B.Sc. degree from Rajshahi University of Engineering and Technology (RUET), Bangladesh, in 2007, the master's degree in electronics/telecommunications from the University of Gavle, Gavle, Sweden, in 2013, and the Ph.D. degree in engineering from International Islamic University Malaysia (IIUM), Kuala Lumpur, Malaysia, in 2017.

He is currently a Senior Lecturer with the Fakulti Teknologi dan Kejuruteraan Elektronik dan Komputer, Universiti Teknikal Malaysia Melaka (UTeM), Malacca, Malaysia. He was a Postdoctoral Research Fellow with IIUM, from 2018 to 2019. He has authored/co-authored more than 70 peer-reviewed journal articles and conference papers. His current research interests include microwave passive devices, chipless RFID technology, microwave passive sensors, MIMO antennas, and 5G mobile communication. He is a Certified Engineering Technologist with the Board of Engineers Malaysia (BEM), a member of the International Association of Engineers (IAENG) and the IEEE Society, a Full Member of the Institution of Engineering and Technology (IET), and a Graduate Technologist with Malaysian Board of Technologists (MBOT). He serves as a reviewer for several reputed journals, including those published by IEEE, Elsevier, Springer, Hindawi, and MDPI.

MASRULLIZAM MAT IBRAHIM was born in Perak, Malaysia, in 1978, He received the B.Eng. degree (Hons.) and M.Eng. degree in electrical and electronics engineering from the Universiti Teknologi Malaysia, in 2001 and 2006, respectively, and the Ph.D. degree in electronic engineering from the University of Strathclyde, U.K., in 2014.

He was a Research and Development Engineer with a Telecommunications Services Company, from 2002 to 2003, before joining Universiti Teknikal Malaysia Melaka (UTeM), in 2003. Since 2020, he has been the Dean of the Faculty of Electronics and Computer Technology and Engineering. With over 20 years of academic experience, he has taught more than 20 subjects and supervised more than 100 undergraduate and 25 postgraduate students. His leadership roles include the Head of Department, the Deputy Dean (Academic), and the Dean. His research interests include signal processing, artificial intelligence, computer vision, embedded systems, and the IoT, with over 70 high-impact journal papers, more than 30 conference papers, and more than RM ten million in research grants as PI and co-researcher. As a certified HRDCorp Trainer, he provides training and consultancy in computer vision, the IoT, and academic knowledge transfer. His industry-linked achievements include IoT Labs, competency certification programs, nationwide CUEPACS certification, and the Academy in Industry program, bridging academia and industry.