IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 27 May 2023, accepted 10 June 2023, date of publication 14 June 2023, date of current version 22 June 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3286308

RESEARCH ARTICLE

Detection and Classification of Power Quality Disturbances Using Stock Well Transform and Improved Grey Wolf Optimization-Based Kernel Extreme Learning Machine

TATIREDDY RAVI AND K. SATHISH KUMAR, (Member, IEEE)
School of Electrical Engineering, Vellore Institute of Technology, Vellore, Tamil Nadu 632014, India

Corresponding author: K. Sathish Kumar (kansathh21@yahoo.co.in)

This work was supported by the Vellore Institute of Technology, Vellore, India, through the Research Fund Research Grant in Engineering, Management & Science (RGEMS).

ABSTRACT Power quality disturbances (PQDs) in modern electrical power systems, caused by the integration of nonlinear power electronic devices and erratic distributed generation (DG), lead to interruptions and significant energy losses for end users. However, conventional methods for PQDs classification face challenges in dealing with noise interference and feature selection. To address these challenges, this research paper proposes a novel approach that combines the stockwell transform (ST) with an improved grey wolf optimization-based kernel extreme learning machine to enhance classification accuracy. The stockwell transform is utilized to extract meaningful features from the power quality (PQ) signals, which are subsequently input into the kernel extreme learning machine (KELM). Furthermore, the parameters of the KELM model are optimized using the improved grey wolf optimization (IGWO) approach to improve the accuracy of classification. To evaluate the performance of the proposed method, real-time implementation is considered by incorporating PQDs data with signal-to-noise ratios (SNR) of 20 dB, 30 dB, and 40 dB into the original synthetic signals. Multiple noise conditions are simulated to assess the proposed model ability to identify and classify disturbance signals. The results demonstrate a detection accuracy of 99.76% under noiseless conditions, indicating the model's high accuracy. Moreover, the proposed method exhibits robustness to noise, achieving accuracies of 98.86%, 98.32%, and 97.3% at SNR levels of 40 dB, 30 dB, and 20 dB, respectively. In the end, this work has performed a comparative study with other previously published work. Compared with other classification methods, the algorithm proposed in this paper has higher accuracy, and it is an efficient and feasible classification method.

INDEX TERMS Power quality disturbances, stockwell transform, improved grey wolf optimization, kernel extreme learning machine, decision tree, confusion matrix.

ABBREVIATIONS

AI	Artificial intelligence.
DAG	Directed acyclic graph.
DG	Distributed generation.
DRST	Double resolution S-transform.
DRST	Double resolution S-transform.
DT	Decision tree.
EEMD	Ensemble empirical mode decomposition.
EMD	Empirical mode decomposition.
EWT	Empirical wavelet transform.
FFT	Fast fourier transform.
FRFT	Fractional fourier transform.
FST	Fast S-transform.
FT	Fourier transform.
HHT	Hilbert-haung transform.
HHT	Hilbert huang transform.

The associate editor coordinating the review of this manuscript and approving it for publication was Diego Bellan.

HT	Hilbert transform.
IEWT	Improved empirical wavelet transform.
IGWO	Improved grey wolf optimization.
IMFs	Intrinsic mode functions.
KELM	Kernel extreme learning machine.
KF	Kalman filter.
KNN	K-nearest neighbour.
PQ	Power quality.
PQDs	Power quality disturbances.
RER	Renewable energy resources.
RKELM	Reduced kernel- extreme learning machine.
SD	Standard deviation.
ST	S-transform.
STF	Strong tracking filter.
STFT	Short-time fourier transform.
SVM	Support vector machine.
VMD	Variational mode decomposition.
VMD	Variational mode decomposition.
WBELM	Weighted bidirectional-extreme learning machine.
WGO	Wild goat optimization.
WPT	Wavelet packet transform.
WT	Wavelet transform.

I. INTRODUCTION

A. MOTIVATION AND INCITEMENT

Power quality is one of the biggest problems in modern power systems. The rapid growth and development of industries such as computing, telecommunications, and electronics manufacturing have led to increased utilization of power electronic equipment, which serves as a primary source of power quality disturbances (PQDs). Power quality disturbances can also result from the increasing integration of distributed generation (DG), as well as from the widespread use of nonlinear loads, faults in the power system, lightning strikes, switching operations of transformers and capacitors, and so on [1], [2]. The detrimental impacts of poor power quality (PQ) are important and include equipment failures, shortened device lifespans, and higher maintenance costs. Conversely, enhancing power quality can yield several advantages, including increased equipment lifespan, improved energy efficiency, and enhanced safety within the electrical system [3]. Addressing these PQ challenges has become imperative to ensure the reliable and efficient operation of power system while mitigating the detrimental effects on end-users and electrical infrastructure. As a result, researchers have been working to develop novel approaches to PQ monitoring and analysis and to enhance the quality of power in modern electrical networks.

Detecting and classifying PQDs is a key part of making electrical power system more reliable, efficient, and stable while reducing downtime and maintenance costs. Recent research has emphasised the significance of PQ analysis in enhancing power system performance. With advancements in sensor technology and data analytics, real-time monitoring of PQDs has become more feasible. Several methods and technologies, such as deep learning (DL) [4], machine learning (ML) [5], and the internet of things (IoT) [6], have been proposed for power quality research. To detect and classify PQ disturbance, a range of techniques have been developed, including time-domain analysis, frequency-domain analysis, statistical methods, and artificial intelligence (AI) techniques. Time-domain techniques focus on analyzing waveforms in the time domain, while frequency-domain techniques assess the frequency content of waveforms. Statistical methods utilize statistical analysis to classify PQDs, and AI techniques, such as neural networks (NN), support vector machines (SVM), random forest, and fuzzy logic, have gained popularity due to their ability to handle complex data patterns and provide accurate classifications. These diverse techniques and technologies contribute to the advancement of PQ analysis and facilitate the development of robust solutions for PQ disturbance detection and classification, ultimately improving the overall performance of electrical power systems.

B. LITERATURE REVIEW

Typically, detecting and classifying PQ disturbance involves a three-step procedure, as depicted in FIGURE 1. In the first stage, signals are collected under a variety of fault conditions with the required sample frequency in order to classify the associated PQ disturbances. The second stage entails acquiring essential features for PQDs classification either directly or indirectly. Usually, signal processing (SP) techniques are used to derive information from the time and frequency domains. In recent times, researchers have introduced multiple algorithms for the analysis of PQDs, which can be broadly categorized into parametric and non-parametric approaches. The parametric approach involves fitting the signal to a model and deriving parameters such as kalman filter (KF) [7] and the strong tracking filter (STF) [8]. While these methods offer real-time signal parameter tracking with high accuracy, they may struggle with accurately extracting parameters due to the complexity of constructing an unknown signal model. To enhance the adaptability of the KF in analyzing multiple PQ disturbance signals, researchers have suggested considering different models for different types of signals [9]. On the other hand, non-parametric methods obtain signal parameters through integral transformations such as the fast fourier transform (FFT), wavelet transform (WT), and stockwell transform (ST). The FFT technique is the standard for harmonic signals; however, it is ineffective for transient disruptions. The short-time fourier transform (STFT) [10] an enhanced method based on FFT, suffers from the Heisenberg uncertainty principle, which limits simultaneous spectral or temporal resolution. Wavelet transform addresses this issue by assessing both steady-state and transient disruptions, but determining the appropriate wavelet basis and number of decomposition layers poses challenges. Wavelet packet transform (WPT) has

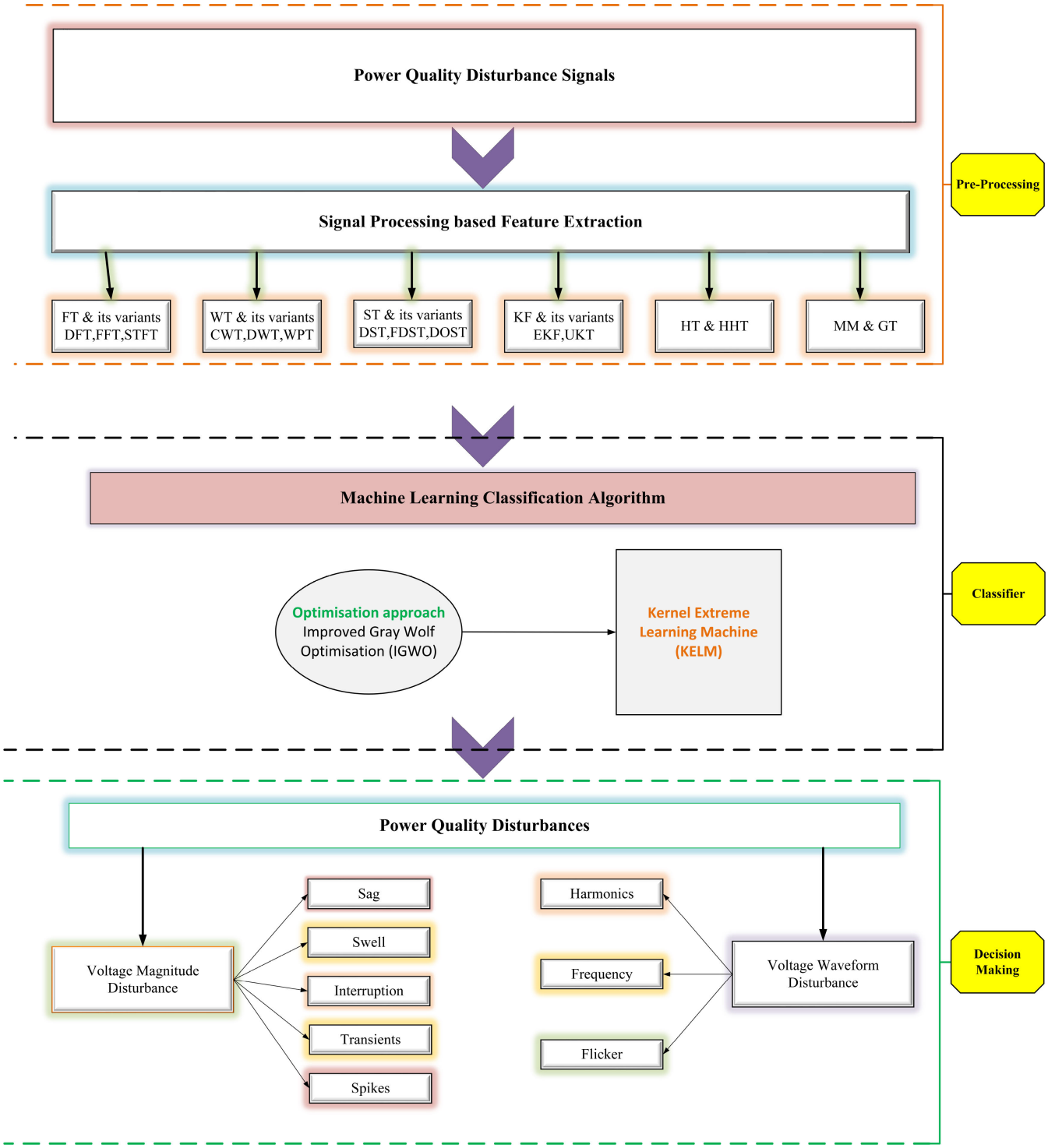

FIGURE 1. Process of PQDs detection and classification.

been utilized by researchers to improve upon WT by decomposing both high-frequency and low-frequency components. However, the down-sampling technique in WPT fails to accurately express sudden changes in wavelet coefficients across different layers, leading to reduced detection accuracy. Stockwell transform [11], a well-known power quality evaluation algorithm, was created by integrating STFT and WT to overcome some of these limitations. Considering the drawbacks associated with many signal processing-based feature extraction (FE) approaches, the authors of this study propose utilizing the S-transform for PQ disturbance identification. However, the selection of appropriate features

remains a significant challenge, requiring advancements in statistical analysis and ML techniques. Statistical metrics such as energy, entropy, minima, maxima, standard deviation, mean, and root-mean-square (RMS) values are commonly employed to calculate optimal sets of dominant and distinctive feature vectors. After completing the FE phase, the feature selection (FS) procedure comes into play. Feature selection is often seen as a supplementary technique prior to classification tasks, involving the selection of a feature vector consisting of the most relevant features based on their association with the output classes [12]. To address these challenges, researchers focus on the development of efficient technologies for statistical analysis and machine learning to improve PQ disturbance identification accuracy. The feature selection procedure plays a crucial role in selecting the most informative features for classification tasks. By reducing the number of features while preserving their discriminatory power, the FS approach enhances the efficiency of subsequent classification algorithms [13]. Thirdly, an appropriate classifier is used to classify PQ disturbance based on the extracted features. Various classifiers have been proposed in the literature for PQDs classification using features extracted from the S-transform. The choice of kernel, parameter tuning, and extracted features all affect the performance of the classifiers [14]. Some widely used ML techniques for classification include artificial neural networks (ANN) [15], support vector machine (SVM) [16], and decision tree (DT) [13]. However, ANN suffers from slow performance and local minima convergence problems, while SVM requires proper selection of the kernel function and regularization parameter. K-nearest neighbor (KNN) [17] is another method that uses neighboring samples for classification but is susceptible to false positives. Extreme learning machines (ELM) [18], a feed-forward ANN variation with higher generalization capabilities, have recently been used to address classification and regression issues. However, fine-tuning its settings is a difficult process. Metaheuristic optimization methods can be used to optimize parameters. A unique technique combining the S-transform and ELM is suggested for the detection and classification of PQ disturbances [19]. The features extracted from the PQ signals using the S-transform are used to classify seventeen different types of PQ disturbances using ELM. Furthermore, ELM performance is improved by fine-tuning its settings using grey wolf optimization (GWO) [20]. Extensive computer simulations are performed to validate the efficacy of the proposed approach.

Several algorithms for the accurate detection and classification of PQ disturbances have been developed in recent years. Combining the double-resolution S-transform (DRST) and directed acyclic graph support vector machines (DAG-SVMs), Li et al., proposed a novel algorithm. The algorithm high accuracy and efficiency in detecting and classifying PQ disturbances made it appropriate for use in real-time applications [21]. Similarly, Mahela and Shaik developed a method that integrates the S-transform and fuzzy C-means clustering for precise detection and classification of PQDs, even in the presence of noise. Their strategy yielded promising results in addressing power quality disturbances with more accuracy [22]. Singh and Singh proposed a classifier using the fractional fourier transform (FRFT) to classify PQDs precisely. By employing FRFT-based feature extraction, the accuracy of classification was significantly improved in comparison to the S-transform method. The efficacy of the proposed FRFT-based classification method was demonstrated experimentally using actual PQ disturbance data [23]. Chakravorti and Dash used variational mode decomposition (VMD) to detect PQDs and fisher linear discriminant analysis (FDA) to reduce dimensionality. Their VMD-assisted FDA-based feature selection in conjunction with an extreme learning machine with a reduced kernel enabled accurate classification of multiple PQ disturbances [24]. Sahani and Dash presented a real-time method for detecting and classifying PQDs based on the hilbert-huang transform (HHT) and weighted bidirectional extreme learning machine (WBELM). Their method demonstrated its applicability for online power quality monitoring systems by outperforming other classifiers [25]. Combining adaptive filtering and a multiclass support vector machine, Thirumala et al., proposed an automated recognition strategy. Their method demonstrated efficacy, resiliency, and accuracy when managing PQDs individually and in combination [26]. Gao et al., describes a new way to identify PQDs using contributions from adaptive wavelet threshold denoising and deep belief network fusion extreme learning machine (DBN-ELM). The proposed adaptive denoising algorithm effectively reduces noise impact and signal-to-noise ratio (SNR). The new DBN model doesn't require manual feature extraction to get correct disturbance features. But it is not known how well the method works on different data sets and in the real world. It needs a lot of training data and is hard to figure out how to do, which may make it hard to use in real-time systems [37]. Swarnkar et al., introduces an algorithm called multi-variable power quality disturbance identification algorithm (MPQDIA) that combines ST, hilbert transform (HT), and rule based decision tree (RBDT) for identifying and classifying power quality disturbances. The algorithm is evaluated using voltage signals based on IEEE-1159 standards in both noise-free and 20 dB SNR conditions, and its performance is compared to other methods. MATLAB software is used for the study [27]. It is clear from the literature that certain classifiers struggle in the presence of noise, whereas others can successfully detect and categorise PQ disturbances under noiseless conditions using only a small subset of possible PQD combinations. This paper proposes a kernel extreme learning machine (KELM) classifier model for the classification of seventeen PQ disturbances as a solution to this issue.

The objective of this research is to develop an efficient and robust method for the clear and distinct recognition of PQ disturbances, which is essential for designing, planning, and protecting the entire power system. Several SP-based techniques for PQ analysis have been devised, along with a variety of feature extraction methods and

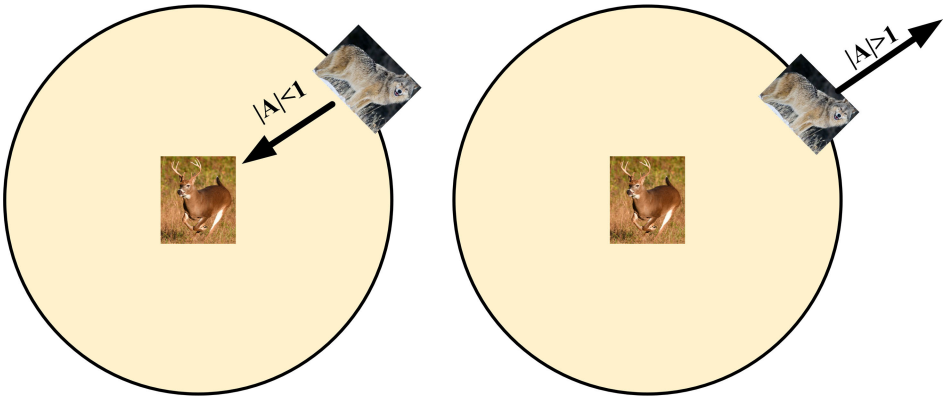

FIGURE 2. R-GWO local search and global search.

classifiers. PQ event classification has utilized machine learning algorithms such as SVM, NN, and DT, as well as kernel-based machine learning algorithms such as KELM, which has gained popularity due to its ability to handle complex and nonlinear data. Additionally, researchers are investigating the use of optimization algorithms, such as improved grey wolf optimization (IGWO), to enhance the efficacy and robustness of PQ recognition and classification. Maintaining the stability and dependability of the power system requires the development of precise and efficient methods for recognizing and classifying PQ disruptions. The application of SP techniques, FE methods, and ML algorithms, such as KELM and optimization algorithms such as IGWO, has the potential to enhance the accuracy and efficacy of PQ recognition and classification.

C. CONTRIBUTION AND ORGANIZATION

The article aims to propose a novel method for identifying and classifying PQDs that emphasizes accuracy. In order to accomplish this objective, this study presents the following key contributions:

1. This paper presents a method for detecting power quality disturbances by employing a feature extractor based on the ST. The S-transform extends both the wavelet and fourier transforms, allowing it to work around issues with fixed-width windows and difficult window function selection.

2. In this study, a total of seventeen different types of PQDs signals are investigated. These signals include both single disturbances (like sag, swell, harmonics, etc.) and multiple disturbances (like sag+harmonics, swell+harmonics, etc.), and they are selected based on the guidelines provided by the IEEE-1159 standard.

3. A 3.2 kHz sampling frequency is employed on ten cycles of distorted waveforms for the feature extraction.

4. The extracted features are applied to a novel improved grey wolf optimization-based kernel extreme learning machine (IGWO+KELM) classifier model for the classification of 17 PQDs.

5. Improved grey wolf optimization is used to improve KELM performance by fine-tuning its parameters. Extensive computer simulations were run to validate the suggested approach efficacy.

The following is how the paper is organized: Section II discusses stockwell transform and its application to PQDs. Section III describes the mathematical model of the IGWO method, which incorporates improvements to the conventional algorithms weight factor and convergence factor. Section IV discusses the origins of KELM as well as its mathematical proofs. Section V discusses the simulation and classification results obtained using the proposed approach. Finally, Section VI presents the conclusions of the study.

II. S-TRANSFORM FOR FEATURE EXTRACTION

In 1996, Stockwell proposed the stockwell transform or S-transform (ST) method, facilitating multi-resolution analysis of time-varying signals. Unlike other techniques, the output of ST is not affected by the noise signal input. The S-transform is a time-frequency analysis approach that extends the wavelet and fourier transforms. It does this by using a Gaussian window that is scalable in both time and frequency. This enables the S-transform to work around the limitations of fixed-width windows, such as the inability to precisely represent both time and frequency information simultaneously. Also, with S-transform, the user doesn't have to choose a window function, which can be challenging and time-consuming. So, the S-transform is a powerful tool for time-frequency analysis that can be used in a variety of applications [28]. Because of this, ST is the best way to get local phase information and a resolution in the time-frequency domain that changes with frequency. Stockwell transform employs a fixed relative bandwidth for filtering signals through multi-resolution analysis. Continuous wavelet transform (CWT) uses a mother wavelet that stays the same, but ST uses a mother wavelet that changes to get information about local phase [29].

A. CONTINUOUS S-TRANSFORM

The CWT for a signal $x(t)$ is defined by

$$\omega(\tau, d) = \int_{-\infty}^{\infty} x(t) w(t - \tau, d) dt \tag{1}$$

where d is scale parameter and τ is wavelet position. The S transform of $x(t)$ is a CWT multiplied by the phase factor.

$$S(\tau, d) = \int_{-\infty}^{\infty} x(t)g(t-\tau, d)e^{-j2\pi f\tau}dt \tag{2}$$

In contrast to the CWT approach, the mother wavelet (window function) in the ST is chosen as a function of the signal frequency content rather than scale d. This is given as

$$g(\tau, d) = \frac{1}{\sigma(f)\cdot\sqrt{2\pi}} e^{\frac{-t^2}{2\sigma^2}} e^{j2\pi ft} \tag{3}$$

where $\sigma(f) = \frac{1}{a+b|f|}$ represents Gaussian window width. From equations (2) and (3) for $a = 0$, the ST can be rewritten as

$$S(\tau, d) = \int_{-\infty}^{\infty} x(t)\frac{b|f|}{\sqrt{2\pi}} e^{-\frac{(\tau-t)^2 f^2 b^2}{2}} e^{-j2\pi ft}dt \tag{4}$$

Mathematically, the S-transform as represented by the fourier transform is

$$S(\tau, d) = \int_{-\infty}^{\infty} X(\alpha + f)e^{\frac{-t^2 f^2}{2}} e^{-j2\pi ft} \tag{5}$$

By using the fast fourier transform (FFT) and the convolution theorem together, you can get the discrete form of the S-transform.

B. DISCRETE S-TRANSFORM

By setting T as the sampling interval, the continuous PQ signal $x(t)$ is discretized as $x(KT)$. Below the equation is a representation of the sampled signals discrete fourier transform (DFT) for $K = 0$ to $N-1$.

$$X\left[\frac{n}{NT}\right] = \sum_{K=0}^{N-1} \frac{1}{N} x(KT)e^{j2\pi\frac{nk}{N}} \tag{6}$$

where $n = 1, 2, \ldots\ldots N-1$. By using DFT and the inverse discrete fourier transform (IDFT), the ST of a discrete-time series $x[n]$ for $\tau = jT$ and $f = \frac{n}{NT}$ can be written as

$$S\left[jT, \frac{n}{NT}\right] = \sum_{K=0}^{N-1} X\left[\frac{m+n}{NT}\right] G(m, n)e^{\frac{j2\pi mk}{N}} \tag{7}$$

where $G(m, n) = e^{\frac{-j\pi^2 m^2}{n^2}}$

S-transform amplitude and phase can be written as in equation (8) and equation (9), respectively.

$$\text{Amplitude } = A(\tau, f) = \left|s\left[jT, \frac{n}{NT}\right]\right| \tag{8}$$

$$\text{Phase} = \phi(\tau, f) = \tan^{-1}\left(\frac{\text{imag}(S[jT, n/NT])}{\text{real}(S[jT, n/NT])}\right) \tag{9}$$

The high-resolution time-frequency information provided by the S-transform makes it the preferred feature extraction technique for power quality detection and classification. This is especially true for non-stationary signals such as power quality disturbances. Compared to fourier or wavelet transforms, the S-transform offers a more detailed analysis of power quality signals, is computationally efficient, and has been shown to be effective in detecting and classifying power quality disturbances.

TABLE 1. Simulation results of test functions.

TF	PSO		GWO		IGWO	
	average	SD	average	SD	average	SD
$f_1(x)$	70.26	18.51	7.36	6.12	2.76	3.68
$f_2(x)$	0.302	0.186	$1.76e^{-9}$	$1.25e^{-9}$	$1.49e^{-13}$	$1.02e^{-13}$

Note: TF-test function; PSO- particle swarm optimizer; GWO- grey wolf optimizer; IGWO- improved grey wolf optimizer.; SD- standard deviation.

III. BACKGROUND THEORY OF IMPROVED GRAY WOLF OPTIMIZATION ALGORITHM

A. STANDARD GRAY WOLF OPTIMIZATION ALGORITHM

Gray wolf optimization (GWO) is a metaheuristic optimization approach that mimics wolves behavioral patterns and hunting mechanisms. Compared to traditional optimization algorithms, GWO has a simple structure, requires fewer parameters, and has excellent robustness, making it an effective global optimization method across various research domains. Gray wolf optimization has been proven superior to traditional optimization algorithms and offers a reliable alternative. A gray wolf pack typically has 5 to 12 members and follows a rigid social order. The GWO algorithm quantitatively models gray wolf seeking and hunting activity. The pack leader, wolf α, with the greatest fitness score, has the highest rank, followed by Wolf β, who has the second-highest fitness score. Wolf δ ranks third, while the remaining wolves, named ω, have the lowest fitness ratings [30].

For a defined d-dimensional issue with a population size of N, each member in the population begins at random as follows:

$$X_i = x_i^1, x_i^2, x_i^3, \ldots, x_i^d \tag{10}$$

where $i = 1, 2, 3, 4, \ldots N$. The team-based hunting behavior of grey wolves is an interesting topic to explore, particularly their three-step hunting mechanism: tracking, driving, and attacking prey. In the tracking stage, the grey wolf collection will encircle and look at the prey, followed by driving and outflanking it until the prey has no escape. Finally, the group will attack the prey. This hunting procedure can be mathematically expressed through the model that follows (11), as described below.

$$\begin{cases} D = C \cdot X_p(k) - X(k) \\ X(k+1) = X_p(k) - A \cdot D \end{cases} \tag{11}$$

In the mathematical model, k denotes the present iteration number, $X_p(k)$ represents the prey position vector, and $X(k)$ represents the gray wolf position vector. The coefficients A and C are defined as follows (Equation 12):

$$\begin{cases} A = (2r_1 - 1)a \\ C = 2r_2 \end{cases} \tag{12}$$

During the prey capture process, the wolves hierarchy plays a crucial role, with higher-ranked wolves having more influence. The lower-ranked ω wolves update their positions

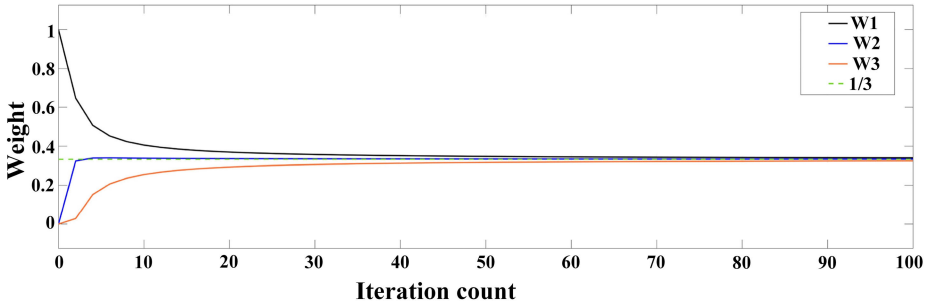

FIGURE 3. Iteration number versus weight curve.

based on the locations of the top three wolves (α, β and δ) using Equation 13 in each iteration of the hunting process.

$$D_g = \left|C.X_g(k) - X(k)\right| v \tag{13}$$

where $g = \alpha, \beta, \delta$ and C is random vector.
In the GWO algorithm, the position of the alpha, beta, and delta wolves (represented as X_α, X_β,and X_δ) are used to update the position of other wolves. The updated position of the current wolf (represented as X) is determined by the approximate distance between the current wolf ω and the alpha, beta, and delta wolves. The estimated distance between the current wolf ω and the α, β, and δ wolves determine the updated location of the current wolf (shown as X). The updated position can be expressed as (Equation 14):

$$X(k+1) = \frac{1}{3}\sum_{g=\alpha}^{\delta}(X_g(k) - A \cdot D_g) \tag{14}$$

In Equation (12), r_1 and r_2 are both random vectors, with random values in [0, 1]. A is determined by the convergence factor a.As shown in FIGURE 2, A governs the wolves search behavior, with values larger than one suggesting global search and values less than one implying local search. The convergence factor a, which is derived as the fraction of wolves doing global and local searches, determines the number of wolves performing global and local searches.

$$a = 2\left(1 - \frac{k}{M}\right) \tag{15}$$

where k represents the current stage of iteration and M is the maximum number of iterations. FIGURE 2 depicts the representative-based grey wolf optimizer (R-GWO) [31], an upgraded version of the GWO that solves its shortcomings and performs both global and local searches.

B. NONLINEAR CONVERGENCE FACTOR

Based on the above study, it is clear that the iteration of the typical GWO algorithm is controlled by the convergence factor, which decreases linearly. This indicates that the algorithm does global searches at first, then turns to local searches as the search advances. However, the linear reduction in the convergence factor does not properly represent the GWO's

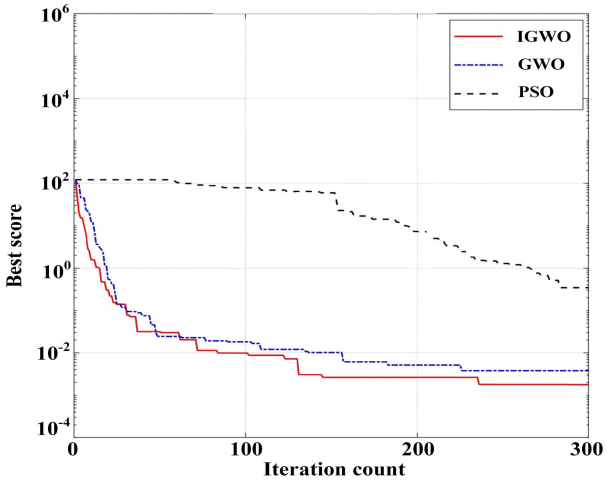

FIGURE 4. Simulation results of test function $f_1(x)$.

exploration kernel generation process. As a result, the method has a low convergence accuracy and is prone to local optimum solutions. This work presents a unique strategy to improving convergence accuracy by altering the convergence factor to meet this issue. The suggested solution increases the algorithm's global and local search performance, eliminating local optimization issues [32]. The suggested convergence factor nonlinear variation mode is stated in the following equation-

$$a = 2\left(-\frac{2k}{M}\right)\left(1 - \frac{k^2}{M^2}\right) \tag{16}$$

According to equation (16), as the number of iterations grows, the attenuation rate of the convergence factor decreases progressively. This, in turn, improves the ability and accuracy of local searches.

C. VARIABLE WEIGHT GRAY WOLF ALGORITHM

Examining equation (14), it is clear that the traditional GWO method counts the optimal, suboptimal, and third-optimal solutions as equally meaningful, which is irrational given that wolves at different levels have differing hunting capacities. To overcome this issue, this study

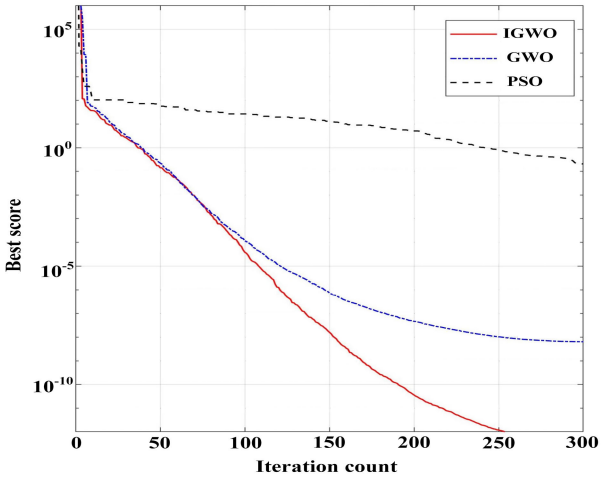

FIGURE 5. Simulation results of test function $f_2(x)$.

improves wolf ω position update mechanism by incorporating weight factors w_1, w_2, and w_3 to reflect the various roles played by wolves of various levels in the process of hunting prey. As a result, equation (14) is rewritten as equation (17):

$$X(k+1) = \sum_{i=1}^{3} w_i X_i, \sum_{i=1}^{3} w_i = 1 \tag{17}$$

According to the weight allocation method, the wolf closest to the prey at the start of the hunt, wolf α, has a weight approximately equal to one, whereas wolves β and δ wolf have weights close to zero. Because all of the wolves surround the prey at the end of the hunt, the position weights of wolves α,β,and δ wolf are deemed identical, as illustrated in Equation (17). As a result, the weight of wolf α decrease gradually with the number of iterations, from 1 to $\frac{1}{3}$, while the weight of wolves β and δ wolf grows gradually from 0 to $\frac{1}{3}$, making them roughly equal. The condition $w_1 \geq w_2 \geq w_3$ governs the change process. The weight is calculated using equation (18):

$$\begin{cases} w_1 = \cos \emptyset \\ w_2 = \frac{1}{2} \sin \emptyset \cdot \cos \gamma \\ w_3 = 1 - w_1 - w_2. \end{cases} \tag{18}$$

Equation (19) defines the angles ϕ and γ in equation (18):

$$\begin{cases} \gamma = \frac{1}{2} \arctan(k) \\ \varnothing = \frac{4}{\pi} \arccos \frac{1}{3} \cdot \gamma \end{cases} \tag{19}$$

FIGURE 3 displays the convergence factor A and changes in positional weights of the three dominant wolves as the total number of iterations increases, indicating conformity with grey wolf social status and hunting strategy. FIGURE 3 shows the relationship between the number of iterations and both the convergence factor A and the position weights of the three main wolves. At the initial stage of hunting the weight of wolf α position is close to 1 ($w_1 = 1$) and the weights of wolf β and δ position are close to 0 (w_2=w_3=0). As the iteration number increases the weight of wolf α decrease gradually from 1 to value 1/3 indicated by green dotted line which is the optimal position to attack pray. While the weights of wolf β and δ increases gradually from 0 to 1/3 with respect to iteration count. At one stage of the hunting, the weight position of leader wolf α and subordinate wolves β and δ are almost equal. This means that all the wolves are gathering around the pray to attack. FIGURE 3 shows very clearly that grey wolves have these characteristics based on social rank and their hunting behavior. This research has so far concentrated on refining the traditional GWO model. To evaluate the modified model's performance, two frequently employed standard test functions ($f_1(x)$ and $f_2(x)$) were chosen to be compared to the traditional GWO strategy and the particle swarm optimization (PSO) method. The functions are represented by equation (20)-

$$\begin{cases} f_1(x) = \sum_{i=1}^{n} \left[x_i^2 - 10\cos(2\pi x_i) + 10\right] \\ f_2(x) = \sum_{i=1}^{n} |x_i| + \prod_{i=1}^{n} |x_i| \end{cases} \tag{20}$$

where $f_1(x)$ involves a combination of cosine and quadratic functions, while $f_2(x)$ is a combination of the absolute sum of x and the sum of the absolute values of x. The two benchmark test functions used in the study, $f_1(x)$ and $f_2(x)$, are commonly used in optimization research and are representative of different types of optimization problems. By testing the IGWO on these two functions and comparing its performance with that of the standard grey wolf algorithm and particle swarm optimization, the authors were able to demonstrate that the IGWO outperformed both GWO and PSO in terms of convergence speed and accuracy. TABLE 1 shows the test function simulation results, while FIGURE 4 and 5 illustrate the optimal solution and iteration number. Based on the results provided in TABLE 1 and FIGURE 4 and 5, the IGWO method suggested in this study has a faster convergence time and higher accuracy than both the standard GWO algorithm and the standard PSO algorithm for the two test functions. This confirms the efficacy and monopoly of the modified GWO method described in this study.

IV. KERNEL EXTREME LEARNING MACHINE (KELM)

After studying extreme learning machine (ELM) [33], it was found that randomly generated input weights and hidden layer bias in the establishment of ELM model may result in reduced stability and generalization ability due to the inability to guarantee learning effectiveness. To overcome this issue, Huang G. B., introduced the kernel function into ELM and proposed kernel ELM [34] to optimize initial weights and bias. To further enhance the model's generalization, Equation (21) is introduced to minimize both the training error and the norm of the output weight.

$$\min : \|H\alpha - T\|^2 \text{ and } \|\alpha\| \tag{21}$$

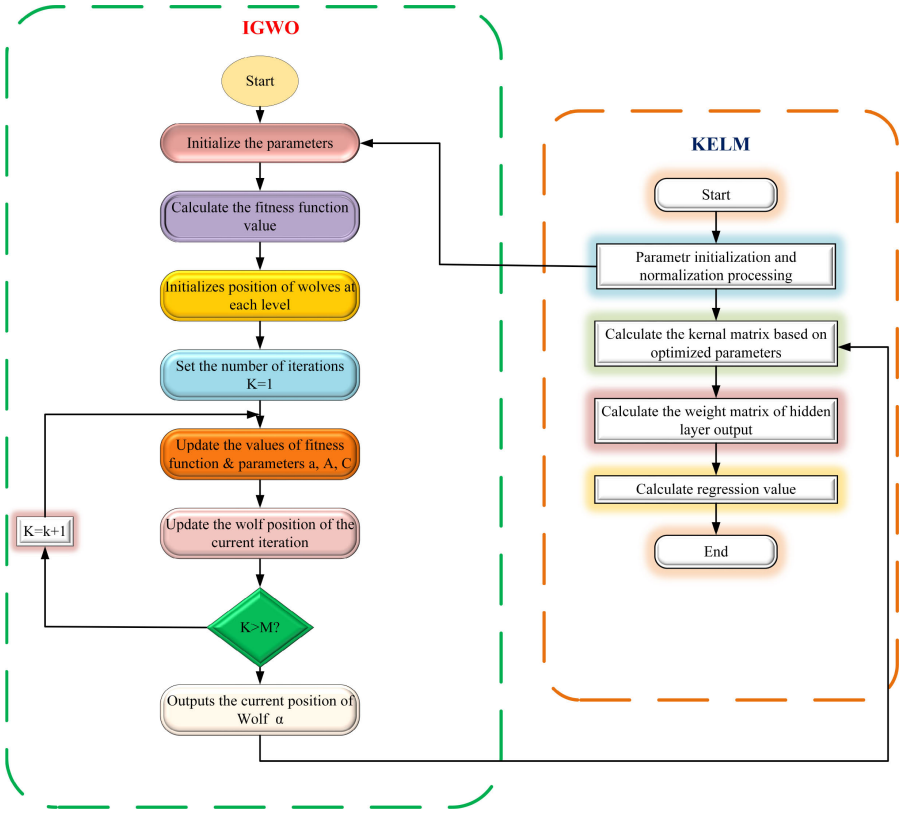

FIGURE 6. Flow chart of proposed IGWO based KELM.

This equation is converted into an equivalent objective function (22) using the Lagrange multiplier.

$$\begin{cases} \min: \quad L = \frac{1}{2}||\alpha||^2 + \frac{1}{2}C \quad \sum_{i=1}^{N} e_i^2, \\ \text{s.t.} \ : h(x_i)\,\alpha = y_i - e_i, \quad i = 1, 2, \ldots .N \end{cases} \tag{22}$$

The error between the network output and true value is denoted by $e_i = [e_{i1}, e_{i2}, \ldots, e_{im}]$. utilizing KKT theorem, training ELM is equivalent to solving a dual optimization problem (23)

$$\min: \quad L = \frac{1}{2}||\alpha^2|| + \frac{C}{2}\sum_{i=1}^{N} e_i^2 - \sum_{i=1}^{N} \mu_i \left(h(x_i)\,\alpha - y_i + e_i\right) \tag{23}$$

where i is the Lagrangien multiplier of instance i, C is a non-negative constant, and the equation (24) expresses optimality condition.

$$\begin{cases} \frac{\partial L_{ELM}}{\partial \alpha} = 0 \rightarrow \alpha = \sum_{i=1}^{N} \mu_i h(x_i)^T = H^T \mu, \\ i = 1, 2, \ldots, N \\ \frac{\partial L_{ELM}}{\partial e_i} = 0 \rightarrow \mu_i = Ce_i, i = 1, 2, \ldots, N \\ \frac{\partial L_{ELM}}{\partial \mu_i} = 0 \rightarrow h(x_i)\,\alpha - y_i + e_i = 0, \quad i = 1, 2, \ldots, N. \end{cases} \tag{24}$$

where $\mu = [\mu_1, \mu_2, \ldots, \mu_N]^T$

Equation (25) can be formed by substituting first and the second sub-equations of Equation (24) into Equation (22).

$$(\frac{1}{C} + HH^{\intercal})\mu = Y \tag{25}$$

By solving equation (24) and (25) we can get equation (26):

$$\alpha = H^T(\frac{1}{C} + HH^T)^{-1}Y \tag{26}$$

Equation (27) can be produced by submitting equation (26) into the output function of the generalized ELM, $f(x) = h(x)\alpha$:

$$f(x) = \text{sign}\left(h(x)H^T\left(HH^T + \frac{1}{C}\right)^{-1}Y\right) \tag{27}$$

Upon analyzing Equation (27), it becomes apparent that both HH^T and $h(x)H^T$ can be written in the form of inner products. According to Mercer conditional theory, the kernel function can also be expressed in terms of an inner product. As a result, it is feasible to define the kernel matrix for ELM as Equation (28):

$$K_{ELM} = HH^T \rightarrow K_{ELMi,j} = h(x_i) \cdot h(x_j) = K(x_i, x_j) \tag{28}$$

TABLE 2. Types of PQ disturbances by class.

Class	Type of PQDs
C1	Normal sinusoidal
C2	Sag
C3	Swell
C4	Interruption
C5	Harmonics
C6	Flicker
C7	Oscillatory Transient
C8	Notch
C9	Harmonics+Sag
C10	Harmonics+Swell
C11	Harmonics+interruption
C12	Harmonics+Flicker
C13	Harmonics+Oscillatory Transient
C14	Harmonics+Notch
C15	Flicker+Sag
C16	Flicker+Swell
C17	Flicker+interruption

TABLE 3. Confusion matrix for binary classification problem.

		Prediction	
Truth		TP	FN
		FP	TN

This allows us to express the output function of ELM using equation (29):

$$f(x) = \begin{bmatrix} K(x, x_1) \\ \vdots \\ K(x, x_N) \end{bmatrix}^T \cdot (C^{-1} + K_{ELM})^{-1} Y \qquad (29)$$

Selecting the appropriate kernel function and its parameters in kernel extreme learning machine classification is problem-specific and there is no one-size-fits-all solution. Commonly used kernel functions in KELM classification include Gaussian (RBF), polynomial, sigmoid, and Laplacian. Among them, the Gaussian kernel function is used in precent work because of its high flexibility and generalization ability when appropriately tuned with its parameter sigma (σ). Nevertheless, the performance of KELM is influenced by the selection of kernel function and its parameters, and a careful selection process should be based on the characteristics of the data and the specific problem. The flowchart shown in FIGURE 6 represents the KELM optimization based on IGWO algorithm [35].

V. RESULTS ANALYSIS

A. DATASET

The effectiveness of any proposed method must be demonstrated through rigorous testing and evaluation. In this study, authors utilized a vast database of seventeen classes (C1-C17) of distinct and composite PQD signals for classification, as stated in TABLE 2. FIGURE 7 depicts the ST contour of PQDs signals, demonstrating that distortion in the ST contour occurs whenever a disturbance happens. Sampling at a frequency of 3.2 kHz and with a fundamental frequency of 50 Hz, each PQ disturbance signal was of size 10 cycles, resulting in a database of 1700 samples with

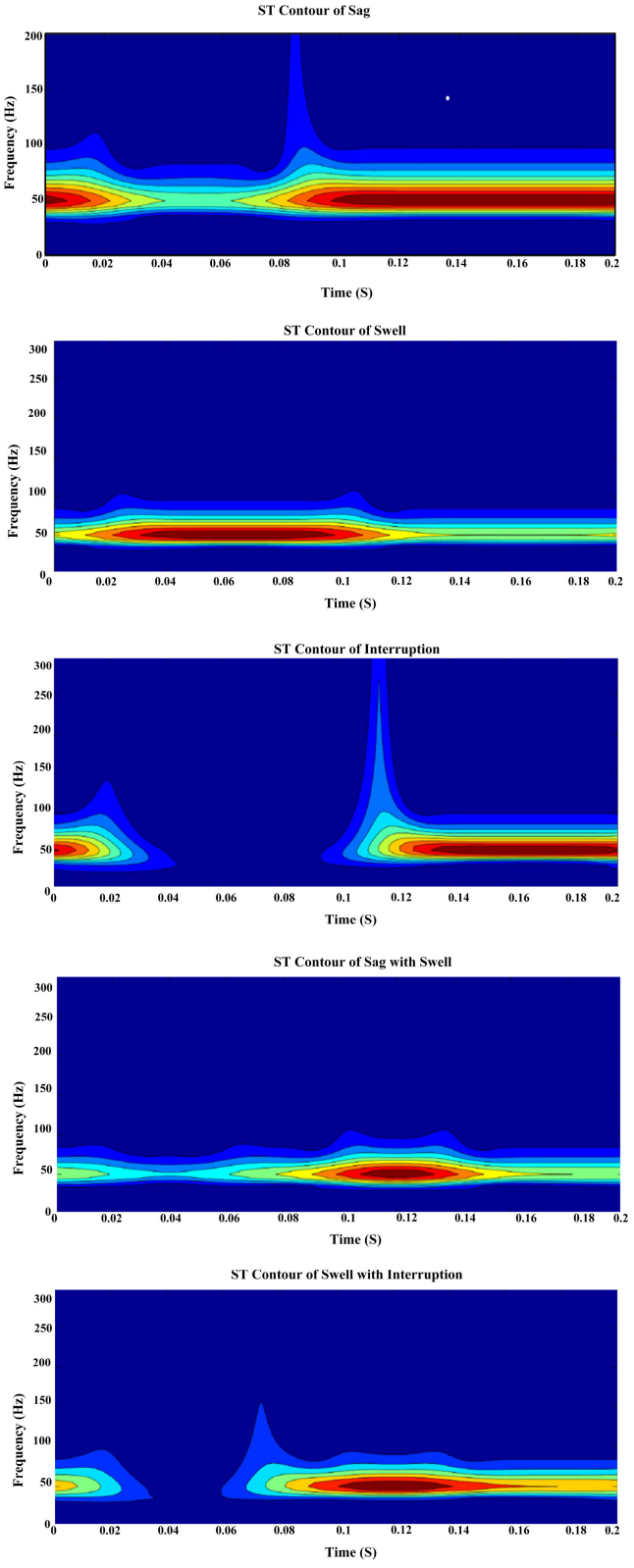

FIGURE 7. ST contour of PQDs.

TABLE 4. Confusion matrix during noiseless condition.

	True PQDs	Predicted PQDs during noise free																
		C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16	C17
Confusion matrix	C1	40																
	C2		40															
	C3			40														
	C4				40													
	C5					40												
	C6						40											
	C7							40										
	C8								40									
	C9									40								
	C10										39		1					
	C11											40						
	C12										1		39					
	C13													40				
	C14														40			
	C15															40		
	C16																40	
	C17																	40
Classification report	Precision	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.98	1.00	0.97	1.00	1.00	1.00	1.00	1.00
	Recall	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.97	1.00	0.98	1.00	1.00	1.00	1.00	1.00
	F-Score	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.98	1.00	0.98	1.00	1.00	1.00	1.00	1.00
	accuracy (%)	100	100	100	100	100	100	100	100	100	98	100	98	100	100	100	100	100

100 samples for each disturbance. To ensure the proposed method's robustness, authors also performed evaluations under noise conditions with noise levels of 40, 30, and 20 dB. The dataset was split into a training set of 60% and a testing set of 40% for each test case. All samples are randomly divided into 1020 samples of training dataset and 680 samples of test dataset.

The data set used for analysis in this study was generated in a simulated environment. Prior to examination, the data undergoes a process known as normalization preprocessing. The goal of normalization is to transform variables of varying magnitudes into a uniform range, usually [0, 1] or [-1, 1]. In the current study, the data samples are mapped to the interval [-1, 1]. The normalization procedure enhances the accuracy and uniformity of analysis, especially for machine learning algorithms that are sensitive to the scale of input characteristics. To ensure unbiased classification results, k-fold cross-validation (CV) was employed in this study. Specifically, a 10-fold CV was utilized to evaluate the performance of the proposed algorithm. However, evaluating a 10-fold CV only once can result in an inaccurate evaluation. To address this issue, the 10-fold CV was executed ten times to generate more reliable and robust results. To find the best classification results, different penalty parameters and kernel parameters were used to choose KELM parameters. The IGWO algorithm suggested in this research is employed to optimize the value of the parameter and the optimal parameter pair to 16 and 0.5, respectively. That is, C is 16 and σ is 0.5.

B. PERFORMANCE ANALYSIS METRICS

The results of classification models are discrete, so we need a way to compare discrete classes. Classification metrics measure how well a model works and tell you how good or bad the classification is, but each of them does this in a different way. So, to evaluate classification models various performance metrics are discussed here. The confusion matrix links truth labels to what the model says will happen. In confusion matrix, each row is made up of predicted class instances, while each column is made up of real class instances. The confusion matrix is not a performance statistic but a base for others. The confusion matrix for binary classification problem is shown in TABLE 3.

The confusion matrix provides an evaluation of the classifier, with each cell representing a different factor. The True Positive (TP), True Negative (TN), False Positive (FP),

TABLE 5. Confusion matrix during 20 dB noise condition.

	True PQDs	Predicted PQDs during 20 dB noise																
		C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16	C17
Confusion matrix	C1	40																
	C2		39		1													
	C3			40														
	C4		1		39													
	C5					32				1			6		1			
	C6						40											
	C7							40										
	C8								40									
	C9					1				36		3						
	C10										40							
	C11					1				1		37				1		
	C12					1							39					
	C13													40				
	C14														40			
	C15											1				40		
	C16																40	
	C17																	40
Classification report	Precision	1.00	0.97	1.00	0.97	0.82	1.00	1.00	1.00	0.91	1.00	0.93	0.98	1.00	1.00	1.00	1.00	1.00
	Recall	1.00	0.97	1.00	0.97	0.92	1.00	1.00	1.00	0.95	1.00	0.92	0.87	1.00	1.00	1.00	1.00	1.00
	F-Score	1.00	0.97	1.00	0.97	0.87	1.00	1.00	1.00	0.93	1.00	0.92	0.92	1.00	1.00	1.00	1.00	1.00
	accuracy (%)	100	97.5	100	97.5	81.0	100	100	100	92.9	98	93.0	92.2	100	100	100	100	100

and False Negative (FN) values are included in the confusion matrix. The accuracy of the classifier can be calculated using equation (30).

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{30}$$

1) PRECISION

Precision is the ratio of true positives to predicted total positives.

$$Precision = \frac{TP}{TP + FP} \tag{31}$$

2) RECALL

Recall, also called Sensitivity or true positive rate, is a measure of how well a classifier finds positive examples. It is measured by the number of true positive predictions (instances that were correctly labelled as positive) out of all the real positive instances in the dataset. In other words, sensitivity is a measure of how many positive cases the model correctly identified as positive. Sensitivity can be shown mathematically as:

$$Recall = \frac{TP}{TP + FN} \tag{32}$$

3) F1-SCORE

The F1-score metric considers both precision and recall. The F1 score is the harmonic mean of equation (31) and (32). The formula is

$$F_1 - score = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{33}$$

C. PERFORMANCE DURING NOISELESS CONDITIONS

For this study, a dataset was created for training and testing a classifier. The dataset consisted of multiple classes of signals, and a random selection of 60% of samples from each class was used for training the classifier. The remaining 40% of samples from each class were used for testing the classifier, resulting in a total of 40 samples per class used for testing purposes. The performance of the classifier was evaluated using a confusion matrix, which provides a detailed breakdown of the accuracy of each class in the form of a 17×17 matrix. TABLE 4 is a representation of the confusion matrix that occurs when there is no noise; based on this, we can see that one of the harmonics with swell signals is wrongly classified as a harmonic with flicker signal. One of the harmonics with flicker signal has been incorrectly identified as a harmonic with swell signal. The classifier made accurate determinations about the classification of each

TABLE 6. Confusion matrix during 30 dB noise condition.

	True PQDs	Predicted PQDs during 30 dB noise																
		C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16	C17
Confusion matrix	C1	40																
	C2		39		1													
	C3			40														
	C4		1		39													
	C5					39										1		
	C6						39			1								
	C7							40										
	C8								40									
	C9									38		1			1			
	C10										40							
	C11					1						38	1					
	C12					1				1			37	1				
	C13													40				
	C14														39			1
	C15					1										39		
	C16																40	
	C17																	40
Classification report	Precision	1.00	0.97	1.00	0.97	0.98	1.00	1.00	1.00	0.95	1.00	0.95	0.93	1.00	0.97	0.98	1.00	1.00
	Recall	1.00	0.97	1.00	0.97	0.93	1.00	1.00	1.00	0.95	1.00	0.98	0.98	1.00	0.98	0.97	1.00	1.00
	F-Score	1.00	0.97	1.00	0.97	0.95	0.99	1.00	1.00	0.95	1.00	0.97	0.96	1.00	0.97	0.98	1.00	1.00
	accuracy (%)	100	97.5	100	97.5	95.0	99.0	100	100	95.0	100	96.4	95.9	100	97.0	98.0	100	100

and every other signal. The confusion matrix presented in TABLE 4 includes the accuracy of all classes, labelled as C1, C2, C3, and so on, up to C17. The results show that the proposed approach achieved an overall accuracy of 99.76%. These findings suggest that the classifier performed very well and can accurately classify signals from each of the classes in the dataset.

D. PERFORMANCE DURING 20 DB NOISE CONDITIONS

In real power networks, the signals extracted often include accompanying noise signals, posing a challenge to extract desired information accurately. To address this, the proposed approach has been tested in a noise environment with varying levels of noise, as measured by signal-to-noise ratio (SNR) levels of 40 dB, 30 dB, and 20 dB. To evaluate the proposed method's performance, a pre-trained network using the dataset mentioned in the previous section was used to test the classifier on three individual datasets, each with a different SNR value. TABLE 5 shows the confusion matrix that results when the noise level is 20 dB. It was seen that one sag signal was incorrectly classified as an interruption. Likewise, 4 harmonics with sag signals have been misidentified as 1 harmonic and 3 harmonics with interruption. 1 interruption was misclassified as sag. In the same way, 3 harmonics with interruption were wrongly classified as 1 harmonic, 1 sag with harmonics, and 1 flicker with sag. 8-harmonic signals are misclassified as 6 harmonics with flicker, 1 flicker with sag, and 1 harmonic with sag signals. One signal with a flicker was incorrectly classified as a harmonic signal. The suggested classifier was able to correctly classify all the other signals. The results of detection accuracy for each class are presented in TABLE 5, with an overall detection accuracy of 97.3% achieved with a 20 dB signal. These findings demonstrate the proposed approach's ability to effectively detect PQDs even in noise environments, highlighting its potential for real-world applications.

E. PERFORMANCE DURING 30 DB NOISE CONDITION

The efficacy of the proposed method, which combines the ST feature extraction method with the IGWO+KELM classifier, was evaluated through experiments using PQDs datasets with 30 dB of noise. The confusion matrix during the 30 dB noise condition presented in TABLE 6 shows that 1 sag was misclassified as interruption, 1 flicker with sag was misclassified as harmonic, 1 interruption was misclassified as sag, 1 harmonic was misclassified as flicker with sag, 1 flicker was misclassified as harmonic with sag, 1 harmonic with a notch was misclassified as flicker with interruption,

TABLE 7. Confusion matrix during 40 dB noise condition.

	True PQDs	Predicted PQDs during 40 dB noise																
		C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16	C17
Confusion matrix	C1	40																
	C2		40															
	C3			40														
	C4				39								1					
	C5				1	37							1			1		
	C6						40											
	C7							40										
	C8								40									
	C9									38		2						
	C10										40							
	C11									1		39						
	C12					1							39					
	C13													40				
	C14														40			
	C15															40		
	C16																40	
	C17																	40
Classification report	Precision	1.00	1.00	1.00	0.97	0.93	1.00	1.00	1.00	0.95	1.00	0.98	0.97	1.00	1.00	1.00	1.00	1.00
	Recall	1.00	1.00	1.00	0.97	0.97	1.00	1.00	1.00	0.97	1.00	0.96	0.95	1.00	1.00	1.00	1.00	1.00
	F-Score	1.00	1.00	1.00	0.97	0.95	1.00	1.00	1.00	0.96	1.00	0.97	0.96	1.00	1.00	1.00	1.00	1.00
	accuracy (%)	100	100	100	97.5	93.5	100	100	95.2	100	98	97.5	97.0	100	100	100	100	100

2 harmonics with sag were misclassified as 1 harmonic with notch and 1 harmonic with interruption, and 3 harmonics with flicker were misclassified as 1 harmonic, 1 harmonic with sag and 1 harmonic with oscillatory transient. The results shown in Table 6 demonstrate that the overall detection accuracy with the 30 dB signal was 98.32%. These findings demonstrate the approach's robustness and high accuracy. The ST feature extraction and IGWO+KELM classifier effectively identified and classified PQDs, even under challenging conditions. The approach remained highly accurate despite significant noise, indicating its robustness. These results suggest that the approach could enhance the reliability and efficiency of power systems.

F. PERFORMANCE DURING 40 DB NOISE CONDITION

The proposed approach for PQDs detection and classification was evaluated by conducting tests on PQ disturbance data with a 40 dB noise addition. TABLE 7 is a representation of the confusion matrix that occurs when there is a noise level of 40 dB; based on this, we can see that one of the interruption signals is wrongly classified as a harmonic with a flicker signal. One harmonic with interruption signals has been incorrectly identified as a harmonic with sag signal. There is one instance of the harmonic with flicker signal being incorrectly classified as the harmonic signal. 2 harmonics with sag signals were incorrectly classified as 2 harmonics with interruption signals. 3 harmonic signals were misclassified as 1 interruption, 1 harmonic with flicker, and 1 flicker with sag. Results of the classification report are presented in Table 7, indicating the approach's robustness, reliability, and effectiveness in classifying PQDs under various conditions. The overall detection accuracy obtained with a 40 dB signal was found to be 98.86%. These findings demonstrate the potential of the proposed approach to enhance the safety, comfort, and efficiency of power systems. As a result, the approach is likely to be of significant interest to stakeholders in the field of power quality and contribute to advancing research and development in this area.

G. COMPARISON WITH OTHER METHODS

In this section, a comparative analysis of the proposed method is presented against other contemporary works in the field of power quality disturbance detection and classification. The evaluation is based on several criteria, including the type of PQDs studied, the total number of signals tested, sampling rates, and accuracy. TABLE 8 summarizes the results of this comparative study, which shows that the proposed

TABLE 8. Comparing effectiveness of the suggested strategy in view of other recently published articles.

Reference	Methodology	Number of classes	Number of features	Sampling frequency	Accuracy in % during noise-free	Accuracy in % during noise		
						40 dB	30 dB	20 dB
[21]	DRST+DAG-SVM	9	9	5 kHz	99.38	-	-	97.77
[22]	ST+FCM	10	14	3.2 kHz	99.2	-	-	98.5
[23]	FRFT+DT	15	9	6.4 kHz	99.93	99.57	-	94.27
[24]	VMD+RKELM	15	36	16 kHz	98.82	-	-	-
[25]	HHT+WBELM	16	4	3.2 kHz	99	95.6	-	91.5
[26]	EWT+SVM	16	6	6.4 kHz	95.56	-	-	-
[37]	ST+DT	9	5	3.2 kHz	99.78	-	-	-
[27]	DBN+ELM	21	12	3.2 kHz	-	98.7	98.2	95.8
[38]	KI-EMD+SOS-ELM	12	4	3.2 kHz	99.6	-	-	99.52
[28]	ST+HT+RBDT	21	4	3.2 kHz	99.52	-	-	99.02
[39]	ST+WGOELM	9	12	3.2 kHz	99.68	-	-	-
Proposed	**ST+IGWO+KELM**	**17**	**9**	**3.2 kHz**	**99.76**	**98.86**	**98.32**	**97.3**

Note: DRST: double resolution S-transform; DAG: directed acyclic graph; FRFT: fractional fourier transform; VMD: variational mode decomposition; RKELM: reduced Kernel- extreme learning machine; WGO: wild goat optimization; HHT: hilbert-haung transform; WBELM: weighted bidirectional based extreme learning machine; KI-EMD: kriging interpolation based empirical mode decomposition; SOS-ELM: symbiotic organism search based ELM; DBN+ELM: deep belief network fusion extreme learning machine; HT-hilbert transform; RDBT-rule based decision tree.

approach outperforms other methods in terms of accuracy and number of PQ disturbance classes studied. The proposed approach employs ST for feature extraction and IGWO based KELM approach for classification. The comparison of the proposed approach with several recent studies indicates that it achieves better accuracy and considers a larger number of PQ disturbance classes. The proposed approach achieved a classification accuracy of 99.76% for noiseless signals and 98.86%, 98.32%, and 97.3% for signals with SNR of 40 dB, 30 dB, and 20 dB, respectively. The proposed approach for PQDs detection and classification outperformed other approaches in terms of accuracy. Previous studies, such as [39] and [21], generated fewer classes of PQDs and had lower detection accuracy compared to the proposed approach. Li et al., proposed a DAG-SVM classifier to classify PQDs with an accuracy of 99.38%. Samanta et al., proposed a PQ disturbance classifier based on ST and GWOELM with 99.68% classification accuracy [21]. Another study in [22] used ST and FCM to classify ten PQ disturbance classes and achieved a lower accuracy value of 99.2% than the proposed method in both noiseless and noise conditions. To further evaluate the effectiveness of the proposed approach, it was compared with other approaches that tested more classes, such as FRFT + DT [23], VMD + RKELM [24], HHT + WBELM [25], and EWT + SVM [26]. In 2017, Singh and Singh proposed a classifier based on FRFT and DT with an overall accuracy of 99.93% during noise-free conditions. However, during noise, 40 dB and 20 dB, the accuracy is 99.57% and 94.27%, respectively, which is less compared to the proposed approach. In 2018, Chakravorti and Dash proposed a VMD and RKELM-based classifier for 15 PQ disturbance classes with 98.82% classification accuracy. Sahani and Dash proposed a classifier based on HHT and WBELM with an accuracy of 99%. In 2019, Thirumala et al., proposed a PQ disturbance classifier based on EWT and SVM with 95.56% classification accuracy. To classify PQDs, Gao et al. suggested a classifier based on DBN and ELM, which achieve 98.7%, 98.2%, and 95.6% accuracy at SNRs of 40, 30, and 20 dB [37]. The ST and HT based RBDT classifier presented by Swarnkar et al., has a 99.52% accuracy rate [27]. Despite testing more PQ disturbance classes and different sampling frequencies in noise environments, the proposed approach outperformed all other approaches. Overall, these results demonstrate that the proposed approach is highly effective and superior to other existing approaches for PQDs detection and classification.

H. LIMITATIONS OF THE ALGORITHM

In this subsection, the limitations of the proposed method are examined. The analysis of TABLE 9 reveals that certain

TABLE 9. Overall performance metrices during noise free, 40 dB, 30 dB, and 20 dB noise.

S.No	Type of PQDs	Classification accuracy(%)			
		Noiseless condition	40 dB	30 dB	20 dB
1	Normal sinusoidal (C1)	100	100	100	100
2	Sag (C2)	100	100	97.5	97.5
3	Swell (C3)	100	100	100	100
4	Interruption (C4)	100	97.5	97.5	97.5
5	Harmonics (C5)	100	93.5	**95.0**	**81.0**
6	Flicker (C6)	100	100	99.0	100
7	Oscillatory Transient (C7)	100	100	100	100
8	Notch (C8)	100	100	100	100
9	Harmonics+Sag (C9)	100	95.2	**95.0**	92.9
10	Harmonics+Swell (C10)	**98.0**	100	100	100
11	Harmonics+interruption (C11)	100	97.5	96.4	93.0
12	Harmonics+Flicker (C12)	**98.0**	97.0	95.9	92.2
13	Harmonics+ Oscillatory Transient (C13)	100	100	100	100
14	Harmonics+Notch (C14)	100	100	97.0	100
15	Flicker+Sag (C15)	100	100	98.0	100
16	Flicker+Swell (C16)	100	100	100	100
17	Flicker+interruption (C17)	100	100	100	100
18	**overall accuracy (%)**	**99.76**	**98.86**	**98.32**	**97.3**

signals cannot be identified accurately. In this regard, in the case of class 5 with a signal-to-noise ratio of 20 dB, the detection accuracy is the lowest at 81.0%, and 19.0% of the data are wrongly classified as class 9, class 12, and class 15. This misclassification happens due to the similarity of features between class 5 and the remaining classes caused by the 20 dB SNR. Similarly, under a 40 dB SNR, class 5 (harmonics) achieves the lowest detection accuracy of 93.5%. Under a 30 dB SNR, the detection accuracy of class 9 is 95.0%, but 5.0% of the data is incorrectly classified as class 11 and class 14. In noiseless conditions, the model obtains a detection accuracy of 98.0% for classes 10 and 12, which is the lowest value in this scenario. Additionally, the proposed method can only simultaneously identify two distinct types of disturbances. This limitation highlights the need to consider additional classes that incorporate multiple types of PQ disturbances in order to improve the classification accuracy. However, the proposed method requires the tuning of several parameters, which may be time-consuming and computationally expensive. The proposed method uses the S-transform for feature extraction and the IGWO+KELM method for classification. Future research will therefore concentrate on addressing these limitations and identifying methods to improve the classification of PQ disturbances, particularly by incorporating classes containing more than two PQDs.

VI. CONCLUSION

In this article, a new hybrid method based on the stockwell transform and an improved grey wolf optimization-based kernel extreme learning machine (ST+IGWO+KELM) is presented for detecting and classifying power quality disturbances. The stockwell transform can be employed for the extraction of useful features from the disturbance signal. The S-transform is a time-frequency analysis approach that resolves the limitations of fixed-width windows and requires

no window function selection. It is more computationally effective than the wavelet transform. The KELM classifier was employed to classify the PQ disturbances. To improve the accuracy of classification, the parameters of the KELM classifier were tuned using IGWO, which is a good meta-heuristic optimization method. The main findings of the study are: (i) A detection accuracy of 99.76% under noiseless conditions; (ii) Robustness to noise with an accuracy of 98.86%, 98.32%, and 97.3% for signals with SNR of 40 dB, 30 dB, and 20 dB, respectively; and (iii) The performance of the proposed approach is compared to other contemporary approaches such as FRFT + DT, VMD + RKELM, HHT + WBELM, EWT + SVM, DBN+ELM, and ST+HT+RBDT. The comparison results reveal that the proposed method has great ability for the detection and classification of PQDs, even in noisy conditions. However, the proposed method requires the tuning of several parameters, which may be time-consuming and computationally expensive. The suggested method could be useful for real-time applications, and future research could include testing it with real-time PQD data, analyzing PQ disturbances in the context of renewable energy resources, and looking into how deep learning methods can be used in more complex situations.

AUTHOR CONTRIBUTIONS

Conceptualization, writing-original draft preparation, and editing: Tatireddy Ravi; conceptualization, review, editing, and supervision: K. Sathish Kumar.

ACKNOWLEDGMENT

The authors would like to thank the management of the Vellore Institute of Technology (VIT), Vellore, Tamil Nadu, India. Also, this work is carried out in the Digital Simulation Laboratory (DSL), School of Electrical Engineering (SELECT), VIT-Vellore.

REFERENCES

[1] P. Rodríguez-Pajarón, A. Hernández, and J. V. Milanović, "Probabilistic assessment of the impact of electric vehicles and nonlinear loads on power quality in residential networks," *Int. J. Electr. Power Energy Syst.*, vol. 129, Jul. 2021, Art. no. 106807.

[2] T. Ravi and K. S. Kumar, "Analysis, monitoring, and mitigation of power quality disturbances in a distributed generation system," *Frontiers Energy Res.*, vol. 10, pp. 1–31, Nov. 2022.

[3] T. Jin, Y. Chen, J. Guo, M. Wang, and M. A. Mohamed, "An effective compensation control strategy for power quality enhancement of unified power quality conditioner," *Energy Rep.*, vol. 6, pp. 2167–2179, Nov. 2020.

[4] J. Wang, D. Zhang, and Y. Zhou, "Ensemble deep learning for automated classification of power quality disturbances signals," *Electr. Power Syst. Res.*, vol. 213, Dec. 2022, Art. no. 108695.

[5] F. Long, F. Liu, X. Peng, Z. Y. Yu, X. H. Huan, H. X. Xu, and J. L. Li, "Power quality disturbance identification and optimization based on machine learning," *Distrib. Gener. Alternative Energy J.*, vol. 37, pp. 159–174, Oct. 2021.

[6] T. Karthick, S. C. Raja, J. J. D. Nesamalar, and K. Chandrasekaran, "Design of IoT based smart compact energy meter for monitoring and controlling the usage of energy and power quality issues with demand side management for a commercial building," *Sustain. Energy, Grids Netw.*, vol. 26, Jun. 2021, Art. no. 100454.

[7] S. Swain and B. Subudhi, "Grid synchronization of a PV system with power quality disturbances using unscented Kalman filtering," *IEEE Trans. Sustain. Energy*, vol. 10, no. 3, pp. 1240–1247, Jul. 2019.

[8] S. He, K. Li, and M. Zhang, "A new transient power quality disturbances detection using strong trace filter," *IEEE Trans. Instrum. Meas.*, vol. 63, no. 12, pp. 2863–2871, Dec. 2014.

[9] H. K. Sahoo and P. K. Dash, "Robust estimation of power quality disturbances using unscented H∞ filter," *Int. J. Electr. Power Energy Syst.*, vol. 73, pp. 438–447, Dec. 2015.

[10] J.-H. Yoo, S.-K. Shin, J.-Y. Park, and S.-H. Cho, "Advanced railway power quality detecting algorithm using a combined TEO and STFT method," *J. Electr. Eng. Technol.*, vol. 10, no. 6, pp. 2442–2447, Nov. 2015.

[11] W. Zhao, L. Shang, and J. Sun, "Power quality disturbance classification based on time-frequency domain multi-feature and decision tree," *Protection Control Mod. Power Syst.*, vol. 4, no. 1, pp.1–6, Dec. 2019.

[12] M. V. Reddy and R. Sodhi, "A modified S-transform and random forests-based power quality assessment framework," *IEEE Trans. Instrum. Meas.*, vol. 67, no. 1, pp. 78–89, Jan. 2018.

[13] M. Biswal and P. K. Dash, "Measurement and classification of simultaneous power signal patterns with an S-transform variant and fuzzy decision tree," *IEEE Trans. Ind. Informat.*, vol. 9, no. 4, pp. 1819–1827, Nov. 2013.

[14] O. P. Mahela, A. G. Shaik, and N. Gupta, "A critical review of detection and classification of power quality events," *Renew. Sustain. Energy Rev.*, vol. 41, pp. 495–505, Jan. 2015.

[15] I. S. Samanta, P. K. Rout, S. Mishra, K. Swain, and M. Cherukuri, "Fast TT transform and optimized probabilistic neural network-based power quality event detection and classification," *Electr. Eng.*, vol. 104, no. 4, pp. 2757–2774, Aug. 2022.

[16] W.-M. Lin, C.-H. Wu, C.-H. Lin, and F.-S. Cheng, "Detection and classification of multiple power-quality disturbances with wavelet multiclass SVM," *IEEE Trans. Power Del.*, vol. 23, no. 4, pp. 2575–2582, Oct. 2008.

[17] S. Zhang, X. Li, M. Zong, X. Zhu, and R. Wang, "Efficient kNN classification with different numbers of nearest neighbors," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 5, pp. 1774–1785, May 2018.

[18] G. Feng, G.-B. Huang, Q. Lin, and R. Gay, "Error minimized extreme learning machine with growth of hidden nodes and incremental learning," *IEEE Trans. Neural Netw.*, vol. 20, no. 8, pp. 1352–1357, Aug. 2009.

[19] P. Mohapatra, S. Chakravarty, and P. K. Dash, "An improved cuckoo search based extreme learning machine for medical data classification," *Swarm Evol. Comput.*, vol. 24, pp. 25–49, Oct. 2015.

[20] F. Mohanty, S. Rup, B. Dash, B. Majhi, and M. N. S. Swamy, "A computer-aided diagnosis system using Tchebichef features and improved grey wolf optimized extreme learning machine," *Appl. Intell.*, vol. 49, no. 3, pp. 983–1001, Mar. 2019.

[21] J. Li, Z. Teng, Q. Tang, and J. Song, "Detection and classification of power quality disturbances using double resolution S-transform and DAG-SVMs," *IEEE Trans. Instrum. Meas.*, vol. 65, no. 10, pp. 2302–2312, Oct. 2016.

[22] O. P. Mahela and A. G. Shaik, "Recognition of power quality disturbances using S-transform based ruled decision tree and fuzzy C-means clustering classifiers," *Appl. Soft Comput.*, vol. 59, pp. 243–257, Oct. 2017.

[23] U. Singh and S. N. Singh, "Application of fractional Fourier transform for classification of power quality disturbances," *IET Sci., Meas. Technol.*, vol. 11, no. 1, pp. 67–76, Jan. 2017.

[24] T. Chakravorti and P. K. Dash, "Multiclass power quality events classification using variational mode decomposition with fast reduced kernel extreme learning machine-based feature selection," *IET Sci., Meas. Technol.*, vol. 12, no. 1, pp. 106–117, Jan. 2018.

[25] M. Sahani and P. K. Dash, "Automatic power quality events recognition based on Hilbert Huang transform and weighted bidirectional extreme learning machine," *IEEE Trans. Ind. Informat.*, vol. 14, no. 9, pp. 3849–3858, Sep. 2018.

[26] K. Thirumala, S. Pal, T. Jain, and A. C. Umarikar, "A classification method for multiple power quality disturbances using EWT based adaptive filtering and multiclass SVM," *Neurocomputing*, vol. 334, pp. 265–274, Mar. 2019.

[27] N. K. Swarnkar, O. P. Mahela, and M. Lalwani, "Multivariable signal processing algorithm for identification of power quality disturbances," *Electr. Power Syst. Res.*, vol. 221, Aug. 2023, Art. no. 109480.

[28] R. G. Stockwell, L. Mansinha, and R. P. Lowe, "Localization of the complex spectrum: The S transform," *IEEE Trans. Signal Process.*, vol. 44, no. 4, pp. 998–1001, Apr. 1996.

[29] S. Ventosa, C. Simon, M. Schimmel, J. J. Danobeitia, and A. Manuel, "The *S*-transform from a wavelet point of view," *IEEE Trans. Signal Process.*, vol. 56, no. 7, pp. 2771–2780, Jul. 2008.
[30] D. Ocran, S. S. Ikiensikimama, and E. Broni-Bediako, "A compositional function hybridization of PSO and GWO for solving well placement optimisation problem," *Petroleum Res.*, vol. 7, no. 3, pp. 401–408, Sep. 2022.
[31] M. Banaie-Dezfouli, M. H. Nadimi-Shahraki, and Z. Beheshti, "R-GWO: Representative-based grey wolf optimizer for solving engineering problems," *Appl. Soft Comput.*, vol. 106, Jul. 2021, Art. no. 107328.
[32] Q. Xie, Z. Guo, D. Liu, Z. Chen, Z. Shen, and X. Wang, "Optimization of heliostat field distribution based on improved gray wolf optimization algorithm," *Renew. Energy*, vol. 176, pp. 447–458, Oct. 2021.
[33] S. Ding, H. Zhao, Y. Zhang, X. Xu, and R. Nie, "Extreme learning machine: Algorithm, theory and applications," *Artif. Intell. Rev.*, vol. 44, no. 1, pp. 103–115, Jun. 2015.
[34] G. Huang, H. Zhou, X. Ding, and R. Zhang, "Extreme learning machine for regression and multiclass classification," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 42, no. 2, pp. 513–529, Apr. 2012.
[35] L. Pan, Y. Xiong, Z. Zhu, and L. Wang, "Research on variable pitch control strategy of direct-driven offshore wind turbine using KELM wind speed soft sensor," *Renew. Energy*, vol. 184, pp. 1002–1017, Jan. 2022.
[36] N. M. Khoa and L. Van Dai, "Detection and classification of power quality disturbances in power system using modified-combination between the Stockwell transform and decision tree methods," *Energies*, vol. 13, no. 14, p. 3623, Jul. 2020.
[37] Y. Gao, Y. Li, Y. Zhu, C. Wu, and D. Gu, "Power quality disturbance classification under noisy conditions using adaptive wavelet threshold and DBN-ELM hybrid model," *Electr. Power Syst. Res.*, vol. 204, Mar. 2022, Art. no. 107682.
[38] I. S. Samanta, P. K. Rout, K. Swain, M. Cherukuri, and S. Mishra, "Power quality events recognition using enhanced empirical mode decomposition and optimized extreme learning machine," *Comput. Electr. Eng.*, vol. 100, May 2022, Art. no. 107926.
[39] I. S. Samanta, P. K. Rout, and S. Mishra, "Power quality events recognition using S-transform and wild goat optimization-based extreme learning machine," *Arabian J. Sci. Eng.*, vol. 45, no. 3, pp. 1855–1870, Mar. 2020.

TATIREDDY RAVI received the B.Tech. degree in electrical and electronics engineering and the M.Tech. degree in power electronics and drives from the Sri Venkatesa Perumal College of Engineering and Technology, JNTU Anantapur, Andhra Pradesh, India, in 2012 and 2017, respectively. He is currently a Research Scholar with the School of Electrical Engineering (SELECT), Vellore Institute of Technology (VIT), Vellore. His research interests include power systems, power electronics, machine learning, deep learning, and power quality.

K. SATHISH KUMAR (Member, IEEE) received the Ph.D. degree from VIT University, Vellore, India. He has 14 years of total teaching experience and in ten years of his research experience, he has worked in various research teams to develop new applications of evolutionary computing algorithms for solving various power system problems like unit commitment, economic dispatch, emission reduction, smart grids, power system reconfiguration, and restoration. He is currently with the School of Electrical Engineering as an Associate Professor. He has published 51 research papers in different journals and conferences of international repute and authored a book on power quality. His current research interests include studying interconnection problems in linkages of HVDC and HVAC (765 kV) transmission lines with existing 230 kV high voltage lines, nano additives for high voltage XLPE cables, and the optimization of smart grids. He is a member of IEEE-PES and SSI. He is a reviewer for various SCI journals, such as *International Journal of Electrical Power & Energy Systems*, *ISA Transactions*, and AAL.