IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 21 July 2023, accepted 8 August 2023, date of publication 11 August 2023, date of current version 16 August 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3304375

RESEARCH ARTICLE

High-Precision Identification of Power Quality Disturbances Based on Discrete Orthogonal S-Transforms and Compressed Neural Network Methods

MUHAMMAD ABUBAKAR[1], ARFAN ALI NAGRA[1], MUHAMMAD FAHEEM[2], MUHAMMAD MUDASSAR[3], AND MUHAMMAD SOHAIL[4]

[1]Department of Computer Science, Lahore Garrison University, Lahore 54000, Pakistan
[2]School of Technology and Innovations, University of Vaasa, 65200 Vaasa, Finland
[3]Department of Technology, The University of Lahore, Lahore 54590, Pakistan
[4]Department of Computer Software Engineering, MCS, National University of Sciences and Technology, Islamabad 24090, Pakistan

Corresponding author: Muhammad Faheem (muhammad.faheem@uwasa.fi)

The work of Muhammad Faheem was supported in part by the University of Vaasa, and in part by the Academy of Finland.

ABSTRACT Power quality disturbances (PQDs) occur as the use of non-linear load and renewable-based micro-grids increases. This paper presents a new algorithm that consists of the discrete orthogonal S-transform (DOST) in the feature extraction stage, compressive sensing (CS) in the feature reduction stage, and a deep stacking network (DSN) for the automatic classification of single and multiple PQDs. It compresses the extracted feature matrix (orthogonal S-matrix coefficients) to minimize the computational process and provide more diversified features. Firstly, PQDs data is generated from a modified IEEE 13 bus system with wind grid integration, both synthetically and in real time. Moreover, compressive measurements of 24 types of multiple PQDs events and nine types of single PQDs events of synthetic and real data, and 12 type of three-phase single and multiple PQDs from the modified IEEE wind grid integration are fed to a proposed DSN classifier for PQD recognition. The DOST-based CS feature extraction technique achieves good robustness and time-frequency localization while retaining useful information. The DSN classifier method utilizes a Batch-mode gradient as a fine-tune, which has less noise gradient and improved efficiency of PQD classification. A noise level of 20 dB to 50 dB is considered. Other models, such as k-Nearest Neighbor (KNN), Multiclass Support Vector Machine (MSVM), and ensemble learner methods, are also developed to compare the efficiency. The high classification results demonstrate that the DOST-CS feature extraction and the DSN classifier have high precision in identifying multiple power quality events, even in noisy conditions.

INDEX TERMS Multiple power quality disturbances identification, compressed sensing, discrete orthogonal S-transform, deep neural network, wind-grid distribution.

I. INTRODUCTION

As electricity demand increases rapidly over time, the issue of power quality (PQ) has become the primary concern of modern power grids. The use of non-linear loads, power electronics components, renewable energy-based microgrids and solid-state switching devices presents a challenge to poor power quality. Many kinds of research have been conducted on the assessment and monitoring of PQ disturbances [1]. The problem associated with PQ disturbances is swell, interruption, notch, flicker, sag, harmonics, transients, and various forms of multiple PQ disturbances [2]. In general,

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

the process of identification and classification of multiple PQ disturbances requires the extraction, compression, selection and classification of features.

Several signal processing techniques have been presented in recent years for suitable feature extractions, such as S-transform (ST), Fourier rapid transform (FFT), Short-Time Fourier transform (STFT), Fourier transform (FT), and Wavelet transforms (WT) [3], [4], [5]. The STFT is an extended form of discrete Fourier transforming (DFT) but does not apply to non-linear and non-stationary signals [4]. WT has resolved this STFT downside by using variable windows for high and low-frequency components and offers excellent time-frequency resolution [6]. The selection of the appropriate function of the mother wavelet and the sampling frequency is very important for the extraction of the frequency component; otherwise, deceptive information may be obtained [7]. ST is a combination of discrete WT and STFT, making it the most potent tool [8]. The key features of ST are that it retains the phase signal information, combines the resolution based on frequencies, and preserves the time-frequency illustration of time series data [9]. At the same time, the features of ST can be used for efficient detection and classification of PQDs. However, the redundant time-frequency and computational times of ST, in contrast to WT, are major drawbacks of PQD [5], [9]. In [10], authors have introduced a discreet orthogonal S-transform (DOST), which is an orthogonal set of the basis function of S-transform. This helps to minimize the computational burden as it takes M-point time series to M-point frequency components.

Traditionally, the Nyquist theorem is used for signal sampling. The Nyquist theorem stated that, in order to sample the signal without error, the sample rate should be twice the highest frequency component presented in the signal [11]. Compressive sensing (CS) is a recently developed technique. In general, the CS theory defines that the sampling rate is lower than that of the original samples of the signals [12], [13]. Bertocco et al. [14] successfully applied CS to minimize the dimensions of the signal, which decreased the computational load on the classifier and improved the accuracy. CS is one of the most recent prominent research in the signal application theory, such as hyperspectral and medical images. In this study, CS was used to evaluate the data processing time of the PQDs and reduce the sampling rate and storage space [15].

In a real power network, many disturbances can affect the system, including capacitor switching, the use of power electronic circuits, and power failure, which are among the major causes of disturbances. In recent times, various methods based on the neural network classifier have been explored for the identification and classification of PQ disturbances. However, few have investigated the combination of various disturbances [16]. In recent years, most of the research work focuses on a single type and a combination of just two types of disturbances, such as harmonics with flickers, notches, swells and sags [17]. In [18], only flicker, sag, interruption, and swell with harmonics were investigated using the ST and artificial neural network (ANN) classifiers. In [19], six signal types, eight doublet types and four PQ triplet types were classified by means of a single channel Independent Component Assessment (SCICA) as well as neural system-based classification system. In [20], six types of combined PQ disturbances were investigated using the Ensemble Empirical Mode Decomposition (EEMD) based feature extraction method and the ranking wavelet Support Vector Machine (SVM) classifier. Only four types of multiple disturbances: transient, sag-and swell-combined harmonics were classified using Tunable-Q Wavelet Transform (TQWT) and dual multiclass SVM in [21]. In [22], seven types of single PQ events and thirteen types of multiple PQ events were considered. The Kalman Filter with Maximum Likelihood (KFML) was used for feature extraction, and the Deep Belief Network (DBN) was employed for classification.

In recent years, a variety of classifiers have been introduced, such as ANN, the probability neural network (PNN), SVM, the decision tree approach (DTA), the fuzzy expert (FES), the Discrete Stochastic Configuration Network (DSCN), and the Simple Gated Recurrent Network (SGRN) [23], [24], [25], [26], [27], [28], [29], [30], [31], [32]. The ANN is a very important classification method with a simple structure, but it involves a lot of training samples and may trigger long training times [29]. However, PNN has slow processing in case of a combination of PQDs, but is much quicker and more reliable than ANN under normal conditions, and needs more storage space than ANN [33]. SVM is also simple and easy to implement, but more accurate and fast classifiers are required for combined PQDs [34], [35].

Singh and Singh proposed a precise classification method for PQDs by utilizing the fractional Fourier transform (FRFT). Their approach, which employed FRFT-based feature extraction, achieved significantly improved classification accuracy compared to the s-transform method. The effectiveness of their FRFT-based classification method was validated through experimental analysis using actual PQ disturbance data [36]. Chakravorti and Dash adopted the variational mode decomposition (VMD) technique for PQD detection and employed fisher linear discriminant analysis (FDA) to reduce dimensionality. By combining VMD-assisted FDA-based feature selection with an extreme learning machine using a reduced kernel, they achieved accurate classification of various PQ disturbances. Their method was tested and evaluated, demonstrating its effectiveness in handling multiple PQDs [37].

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 approach showcased its applicability in online power quality monitoring systems, outperforming other classifiers. The proposed method proved to be effective, resilient, and accurate in PQD detection and classification [38]. Thirumala et al. proposed an automated recognition strategy for PQDs by combining adaptive filtering and a multiclass support vector machine. Their method demonstrated efficacy, resiliency, and accuracy when managing PQDs individually as well as in combination. The proposed strategy provided an automated approach for PQD recognition [39].

Gao et al. described a novel method for identifying PQDs, utilizing adaptive wavelet threshold denoising and deep belief network fusion extreme learning machine (DBN-ELM). The proposed adaptive denoising algorithm effectively reduced noise impact and improved signal-to-noise ratio (SNR). The DBN model eliminated the need for manual feature extraction, enhancing the accuracy of disturbance feature identification. However, further research is needed to assess the performance of this method on different datasets and in real-world scenarios. It requires a substantial amount of training data and may be challenging to implement in real-time systems [40]. Swarnkar et al. introduced a multi-variable power quality disturbance identification algorithm (MPQDIA) that combined the s-transform (ST), Hilbert transform (HT), and rule-based decision tree (RBDT) for identifying and classifying power quality disturbances. The algorithm was evaluated using voltage signals based on IEEE-1159 standards under both noise-free and 20 dB SNR conditions, and its performance was compared to other methods. MATLAB software was used for the study [41].

It is evident from the literature that the majority of it utilized simulated data and did not address more than twenty PQ disturbance problems. Only a few studies employed three different types of datasets to validate their research. Additionally, most of the studies neglected to discuss the computational complexity of their models. Typically, the Nyquist theorem is employed for signal sampling, resulting in a doubling of the sample rate and an increase in the computational complexity of the model. However, this research developed a model based on CS that produces a smaller sample than the original signal while maintaining good computational speed. Another recent cause of PQ disturbance is the integration of renewable resources. This paper also incorporates wind grid distribution data to thoroughly examine this particular issue.

In this paper, a new approach for the recognition of PQ disturbances based on feature extraction, reduction, and classification is proposed. Three types of datasets are used: synthetic datasets based on mathematical models, real datasets taken from IEEE 1159.3, and simulated datasets from a modified IEEE 13-bus system with wind grid integration. The proposed hybrid technique is based on DOST-CS for the feature extraction and reduction, and the DSN for the classification of nine kinds of single, thirteen kinds of doublet, six kinds of triplet, two kinds of quartet forms and an unbalanced three phase PQ disturbances. The computational complexity of DOST is NlogN, which is lower than that of the traditional ST method, and it requires less memory for the calculation. DOST is the enhanced version of ST, aiming to increase the computational speed and classification accuracy of multiple PQ disturbances.

The key characteristic of CS is that it allows for the reduction of the dimension of the PQD without any loss of valuable information. This paper presents a feature extraction and classifier method that combines DOST, CS and Deep stacking network (DSN) [42]. The sigmoid kernel is multilayer perception kernel that originates from the field of neural networks. Usually, it is most suited for the neural network-based classifier. The sigmoid kernel function used in the SVM model is equal to a two-layer perception neural network [43]. DOST was used to extract the features of PQDs, and the compression of extracted data was performed using CS. This down-sampled data was finally fed to the DSN classifier for the identification of PQDs. In the final stage, the proposed model with highest classification accuracy is compared with K- nearest neighbors (KNN), SVM, and ensemble learners for the selected features. The performance of the proposed model is also tested under noisy conditions and complex PQ events. Moreover, the proposed model is also compared with recently published articles most of which focus on addressing the doublet, triplet and quartet PQ problems.

The key contribution of this paper are as follows.

- The DOST method is introduced, which exhibits lower computational complexity compared to the traditional ST method.
- A total of 33 types of single, doublet, triplet and quartet PQ events are considered for classification.
- Three different datasets are utilized: a synthetic dataset, a real dataset, and a dataset derived from a modified IEEE 13-bus system with wind grid integration simulation.
- Compressive sensing is employed to reduce the dimensionality of the data by a factor of four without compromising the essential information.
- Comparative analyses are conducted with MSVM, KNN and ensemble learners to assess the performance of the proposed model.
- The proposed model is also tested under various noise levels to validate its effectiveness.

This paper is structured into the following sections. Section II provides brief information about feature extraction, reduction, and classification techniques, as well as the proposed algorithm. Simulation and experimentation are also discussed in Section II. Section III presents the classification results, discussion, and a comparison with a recently discussed method. Finally, Section IV concludes the paper.

II. MATERIALS AND METHODS

The Methodology for Feature Extraction and Classification of PQDs is represented in this section.

A. STOCKWELL TRANSFORM

ST is an extended form of STTFT and DWT. It can be studied as a phase-corrected DWT or a frequency-dependent

STFTS-transform. It is very useful for non-stationary signals because it provides almost the same pattern information as the original data (disturbance type). Compared to DWT, the lossless transition of the time-domain signal to the time-frequency and back to the time domain is one of the primary advantages of ST. ST works similarly to FT [44]. The S-transform can be described as a phase-rectified DWT, as the phase-corrected factor can be multiplied by a selected mother wavelet. ST also uses a basic visual inspection for classifications and the DWT requires multi-resolution analysis. The DWT $\omega(t)$ of the function $h(t)$ is defined in (1)

$$W(\tau, d) = \int_{-\infty}^{+\infty} h(t)\,\omega(\tau - t, d)\,d\tau \tag{1}$$

where w(t-τ,d) represents the equivalence to the basic mother wavelet and d represents dilation, which is the inverse of frequency f. It provides information about the width of wavelets and limits the resolution. Thus, the ST that is obtained from DWT is defined as the special mother wavelets multiplied by the function h(t) with the phase factor as follows:

$$S(\tau, f) = X(\tau, d)\,e^{i2\pi f\tau} \tag{2}$$

And, by rearranging the equation,

$$S(\tau, f) = \frac{|f|}{\sqrt{2\pi}} \int_{-\infty}^{\infty} h(t)\,e^{\frac{-(t-\tau)^2 f^2}{2}} e^{-i2\pi ft} \tag{3}$$

The discrete S-transform in term of power quality signal is represented as Let h[k]=h(kT), k=0,1,…,N-1. Where N is the number of samplings and T is the time interval of sample points. The discrete-time series h(kT) is formed as

$$H\left[\frac{n}{NT}\right] = \frac{1}{N}\sum_{k=0}^{N-1} h[kT]e^{\frac{-i2\pi nk}{N}} \tag{4}$$

where n=0,1,…,N-1. h[kT] is the discrete-time series of S-transform, and it is obtained by the (3) (letting f→n/NT and $\tau \to$ jT), the discrete S-transform equation becomes as follows [5], [9]:

$$S\left[jT, \frac{n}{NT}\right] = \sum_{m=0}^{N-1} H\left[\frac{m+n}{NT}\right] G(m, n)\,e^{\frac{i2\pi mj}{N}} \tag{5}$$

where, $G(m,n) = e^{\frac{-2\pi^2 m^2 \alpha^2}{n^2}}$, α is used to vary the shape and different resolution of the Gaussian window, and it is equal to "$1/b$" and where the value of "b" is varying from 0.33 to 5. N = Total number of samples $n = 1, 2, \ldots, N-1$; $j = m = 0, 1, 2, \ldots, N-1$. In the case of higher frequency analysis, the value of α would be higher, and a narrow window is selected. Conversely, for lower frequency analysis, a wider window is chosen. Discrete Orthogonal S-transform (DOST) is a less complicated and more efficient illustration of S-transform used in the application of PQ analysis.

B. DISCRETE ORTHOGONAL S-TRANSFORM (DOST) FOR FEATURE EXTRACTION

Discrete orthogonal S-Transform (DOST) was introduced by Stockwell [45]. DOST is used in various applications such as ECG data compression [46], Face recognition [47], image restoration [48], and image denoising [49]. DOST is an efficient illustration of the ST. There are many benefits to the orthogonal representation of the transform as it brings the M-point time series to the M-point frequency illustration. In addition, the computational burden will be minimized as it processes fewer data compared to the S transform. As a result, overall performance and efficiency are improved. DOST is basically an orthogonal representation of the matrix, which means that the inverse of the matrix is equal to the complex conjugate transpose. Each point is linearly independent of any other points. This strategy has been summarized in Figure 1. However, the block diagonal nature of matrix T presents the opportunity to determine the coefficients in a block-wise fashion, thereby enhancing the efficiency of multiplying this sparse matrix.

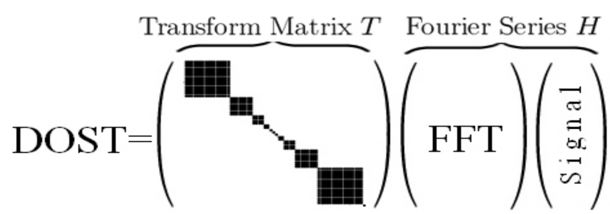


FIGURE 1. Description of DOST [50].

DOST is described as the inner products between time series $h[kT]$ and the basis functions, which are defined as the function of $[kT]$.

- v represents the frequency variable, indicating the center frequency associated with the wavelet transform.
- β represents the dimension of the frequency band, providing information about the width or scale of the wavelet.
- τ refers to the time location, indicating the position or time shift of the wavelet transform.

$$D\{j[pT]\} = D\left(\tau T, \frac{u}{MT}\right) = \sum_{p=0}^{p=M-1} h[pT]\,S_{[u,\alpha,t]}[pT] \tag{6}$$

where t = 0,1,2,…., $\beta-1$. The basis function $D_{[v,\beta,\tau]}$ [kT] is defined as

$$D_{[v,\beta,\tau]}[kT] = \frac{ie^{-i\pi\tau}}{\sqrt{\beta}} \frac{\left\{e^{-i2\pi(\frac{k}{N}-\frac{\tau}{\alpha})(u-\frac{\beta}{2}-\frac{1}{2})} - e^{-i2\pi(\frac{k}{N}-\frac{\tau}{\beta})(u+\frac{\beta}{2}-\frac{1}{2})}\right\}}{2sin\left[\pi(\frac{k}{N}-\frac{\tau}{\beta})\right]} \tag{7}$$

At this instant, the (7) has not determined the sampling of time-frequency space. To examine the orthogonality of time-space below said rules must be applied

- The samples of Fourier frequency should use for once the chosen parameters of v and β.
- The implicit feature of the S transform is the phase correction that distinguishes the wavelet transforms.

If there are more samples in time (large τ), then frequency resolution would be broader (large β) for each voice.

- Each voice equals to one or more local time samples (τ) and up to frequency resolution β.

S transforms and wavelets transform hold the progressive resolution property. In the case of octave sampling, the increase in voice doubles the bandwidth of the voice and the introduction of a new variable p which corresponds the octave number $p = 0, 1, 2, .. \log_2(N) - 1$, by rearranging all these parameters now the equation is:

$$\begin{aligned} &\text{For } k > 1, \text{ we have :} \\ &p = 2, \ldots, log_2(N) - 1 \\ &v = 2^{(p-1)} + 2^{(p-2)} \\ &\beta = 2^{(p-1)} \\ &\tau = 0, 1, \ldots, 2^{(p-1)} - 1 \end{aligned} \tag{8}$$

For the case p=0, then v=0, $\beta = 1$, and $\tau = 0$. Also, when p=1, then v=1, $\beta = 1$, and $\tau = 0$. The octave sampling of a time series for the DOST basis function is presented by substituting (8) into (7).

$$D_{[p,\tau]}[pT] = \frac{ie^{-i\pi\tau}}{\sqrt{2}^{(p-1)/2}} \times \frac{\left\{e^{-i2\pi(\frac{k}{N}-\tau/2^{(p-1)})(2^{p-1}-1/2)} - e^{-i2\pi(\frac{k}{N}-\frac{\tau}{2^{p-1}})(2^{p}-1/2)}\right\}}{2sin\left[\pi(\frac{k}{N}-\frac{\tau}{2^{p-1}})\right]} \tag{9}$$

The reason for selecting this method is that the computational complexity of DOST is NlogN, which is lower than that of traditional ST. Additionally, it requires less memory for the calculations.

Both ST and DOST provide time-frequency information for time series data. However, a significant difference between DOST and ST is the non-redundancy of DOST, which represents a one-dimensional image of the time-frequency space. In DOST, both the input and output are "N" point data, and the sampling rate decreases at lower frequencies and vice versa [10], [48], [51].

C. COMPRESSIVE SENSING FOR FEATURE REDUCTION

CS is a modern paradigm for signal processing and information theory. It measures and reconstructs a sparse signal from the small set of non-adaptive and linear measurements. Let $z \in \mathbb{R}^{M\times 1}$ be a signal [52]. Given an orthonormal basis matrix $B \in \mathbb{R}^{M\times M}$, whose columns are the basis elements $\{b_i\}_{i=1}^{M}$, where z is represented in term of basis as

$$z = \sum\nolimits_{i=1}^{M} a_i b_i \tag{10}$$

More compactly, $Z = B\alpha$, where α is a $M \times 1$ column sparse vector with only K nonzero elements, that is $\|\alpha_0\| = K$, this measurement process is described as

$$x = \emptyset B\alpha = A\alpha \tag{11}$$

where $x \in \mathbb{R}^{N\times 1}$ is a column vector of the compressive measurement. $A = \phi B$ is the sensing matrix. A represents the dimension reduction and is $N \times M$ where N is much lesser than M, and the sensing matrix A should satisfy the restricted isometry property of order K. There are many algorithms used for sparse representation, such as Basis Pursuits, Matching Pursuits and Orthogonal Matching Pursuits [52], [53].

In this paper, CS is described as the reduction of dimensions and the selection of functions. By applying CS to the classification of PQD events, it is possible to improve efficiency and reduce the computational burden. Features of the time-frequency analysis (DOST) also alleviate the computational load as discussed earlier. Many techniques are already used to minimize the dimension of data like kernel PCA, kernel LDA [54], [55], locality preserving projection, generalized discriminant analysis [56], [57], linear discriminant analysis (LDA) [57], independent component analysis (ICA), PCA and CS [58]. The key characteristic of CS is that it allows for the reduction of the dimensionality of the PQD signal without any loss of valuable information. Figure 2 describes the feature and dimension reduction of PQD signal using CS. Consider $x \in \mathbb{R}^{M\times 1}$ is the extracted feature from DOST, where $A \in \mathbb{R}^{N\times M}$, M is lesser than N and $x \in \mathbb{R}^{M\times 1}$.

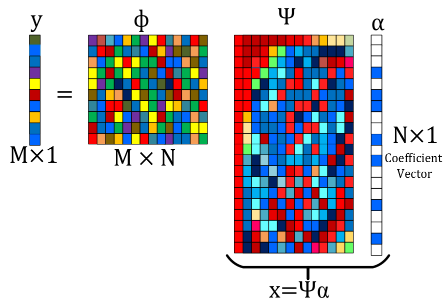


FIGURE 2. Dimension reduction using compressive sensing.

D. SELECTED MACHINE LEARNING CLASSIFIERS

Various machine learning classifiers are available in the literature. In this paradigm, the following classifiers are selected: K-Nearest neighbours (KNN), MSVM, DSN, Subspace KNN and Adaboost. These methods are briefly explained in the following section.

1) K-NEAREST NEIGHBOURS (KNN)

KNN is a non-parametric and popular technique among the machine learning methods. It has various advantages for example, KNN is relatively simple as it only requires the tuning of parameters k. It does not rely heavily on prior information about the data distribution, and it can handle large training input data effectively. In recent studies, researchers have focused on finding the optimal value of k, exploring different distance measurement techniques, and reducing the computational burden in KNN. The location of data points can be determined using distance metrics such as Minkowski, Euclidean, and others.

2) MULTICLASS SUPPORT VECTOR MACHINE (MSVM)

The SVM is supervised machine learning method for the classification. SVM determines the data points of one class from the data points of the different class. For three or more classes the common methods used are one versus all (OVA), one versus one (OVO), all at once (AAO) and error correcting output code (ECOC). ECOC is simple, powerful and effective method as compared to other methods including OVA and OVO SVM [59]. The choice of kernel function can affect the classification accuracy.

3) DEEP STACKING NEURAL NETWORK

The main idea behind the design of DSN is presented in [42] and [60], which is associated with the concept of stacking. Firstly, a simple module of function and classifier is composed. Then they are loaded or stacked on top of each other to settle the complicated classifier and functions. The high level of stacking architecture comes from the unions of low-level classifier output, and the raw input features (W_{rand}), where W_{rand} represents the raw input feature. Originally, DSN was introduced in [42] and was described as a deep convex network. In the primary module of DSN, the hidden layers or units are sigmoidal nonlinear, while the output layer or units are linear. The inclusion of linear units in the output layer ensures higher efficiency, polarizability, and optimization of the network weights.

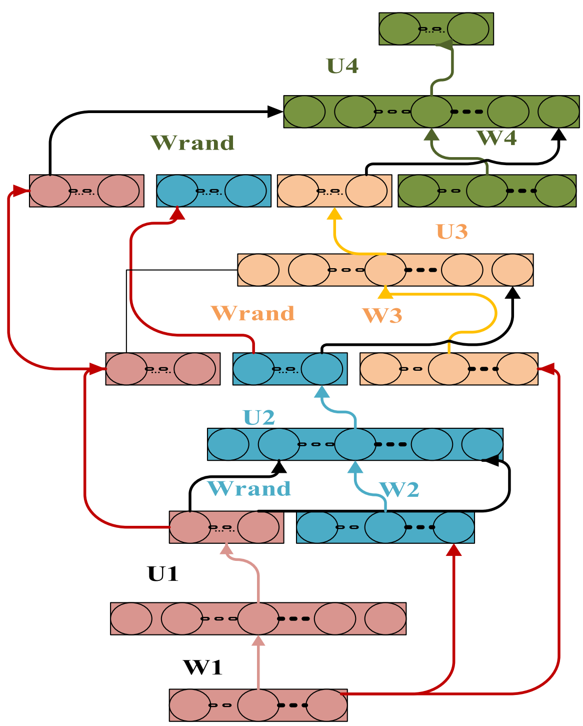


FIGURE 3. Deep stacking network structure with multi-module.

The architecture of DSN is presented in Figure 3, which includes the multi-layered modules. Figure 4 describes the structure of a single module, and this functional neural network module consists of two training sets of weighs and one hidden layer. More specifically, the hierarchy of the lowest module consists of the linear input layer and the hidden output layer with linear and nonlinear sets respectively. The fine-tune algorithm of DSN weight is carried out in each module, rather than propagating it across all layers globally. Once overall training samples in activity matrix H are known in hidden layer after that the upper weight matrix, U can be efficiently learned. Let the training vector denoted by $X = [x_1, \ldots\ldots, x_i, \ldots\ldots, x_N]$ in which each vector x_i is denoted by $x_i = \left[x_1, \ldots, x_{ji}, \ldots, x_{Di}\right]$, D is the size of the input vector, the total number of training samples are denoted by N, L represents the hidden units, and C belongs to the dimension of the output vector.

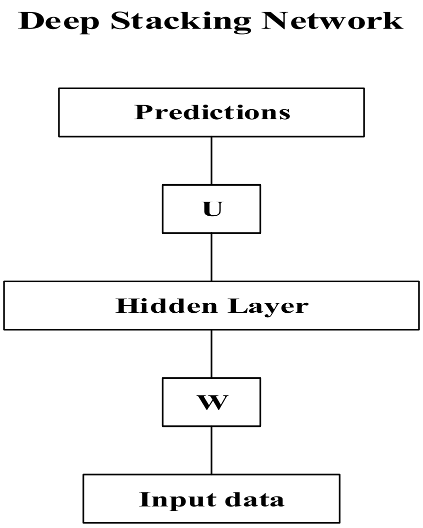


FIGURE 4. The single module of deep stacking network.

In Figure 3, only four modules are discussed, and different colors separate each module. As an example, if DSN is used to recognize the image, then the number of pixels is linked to linear input units (or extracted features) in the model, and the obtained feature is based on intensity value and linked to the respective pixels. In the case of speech recognition, extracted features of speech waveform are the input units. The classification represents the output units in the linear output layer. Same as, if DSN is used for PQDs, then the set of input units is the extracted features from PQDs waveform and then compressed by the factor of 0.25 using CS and converted to the 2D-image.

The lowest input module of DSN has hidden layer which consists of a set of nonlinear units that is denoted by W. The output units are represented by U. The value of weight matrix is comprised from 0 to 1 or weight matrix of PQDs extracted features separately. The output is the linear layer that describes the details of particular classification. In the case of PQDs, classification of output weight matrix will be like sag, swell, harmonics, interruption, and transients etc.

Batch mode gradient computation results in less noise in the gradient and trains more efficiently compared to the conventional backpropagation method [61]. It is because of the exploitation of explicit constraint $U = F(W)$.

4) ENSEMBLE LEARNERS (EL)

In this study, Subspace and Adaboost learners are applied for the classification. EL combined multiple learning algorithms to improve the precision of the classification and become an efficient tool. The model's performance is determined by the majority vote.

E. CLASSIFIERS PERFORMANCE

The K-cross validation method is used to check the performance of aforementioned classifiers. In this method, the data is divided into K equal parts and each fold is used for the test and training purpose. The classification accuracy is calculated from the average accuracy of each folder.

$$Accuracy(\%) = \frac{TP + TN}{TP + TN + FP + FN} \tag{12}$$

The expression criterion in (12) provides information about the accuracy of the selected classifiers. Here, TP represents true positives, TN represents true negatives, FP represents false positives, and FN represents false negatives.

F. PROPOSED ALGORITHM FOR CLASSIFICATION OF PQDS

The flow diagram and procedure of the proposed algorithm for classification of PQ events are shown in Figure 5. The following steps need to be investigated for PQ disturbances.

Aim: The classification of multiple PQDs using DOST, CS, and different machine learning methods (DSN, KNN, MSVM and EL).

Input: The PQD signal is captured using synthetic and standard IEEE 1159.2 and 1159.3 real data [62] and wind grid distribution data. The PQD signal is denoted as $x \in \mathbb{R}^{(1\times n)}$, where n is the length of the signal. The corresponding plots can be found in subsubsection II-G. 2.

Step 1: The input signal is decomposed using multiresolution analysis that is based on DOST and can be calculated from (9). This process yields discrete orthogonal S-matrix coefficients.

Step 2: The resulting plots, shown in Section G.4, are obtained using the DOST-matrix. These plots include time-frequency contour, amplitude-time curve, and the amplitude-frequency curve, which provide precise information on time-frequency localization.

Step 3: The decomposed signals are applied to CS sampling technique to reduce the dimension of the signal without losing the core features. The variable M represents number of samples. Mathematically, the compressive measurements are described as y = Ax, where x is the input signal and A is the sensing matrix and details regarding this process can be found in subsubsection II-G. 5.

Step 4: Statistical Features such as Mean, Skewness, Standard deviation, Energy, Harmonics distortion, Kurtosis, and Frequency estimation, $F_1, F_{2,...,} F_7$ are extracted from this information obtained from DOST-CS. These features serve as inputs to the proposed classifier DSN.

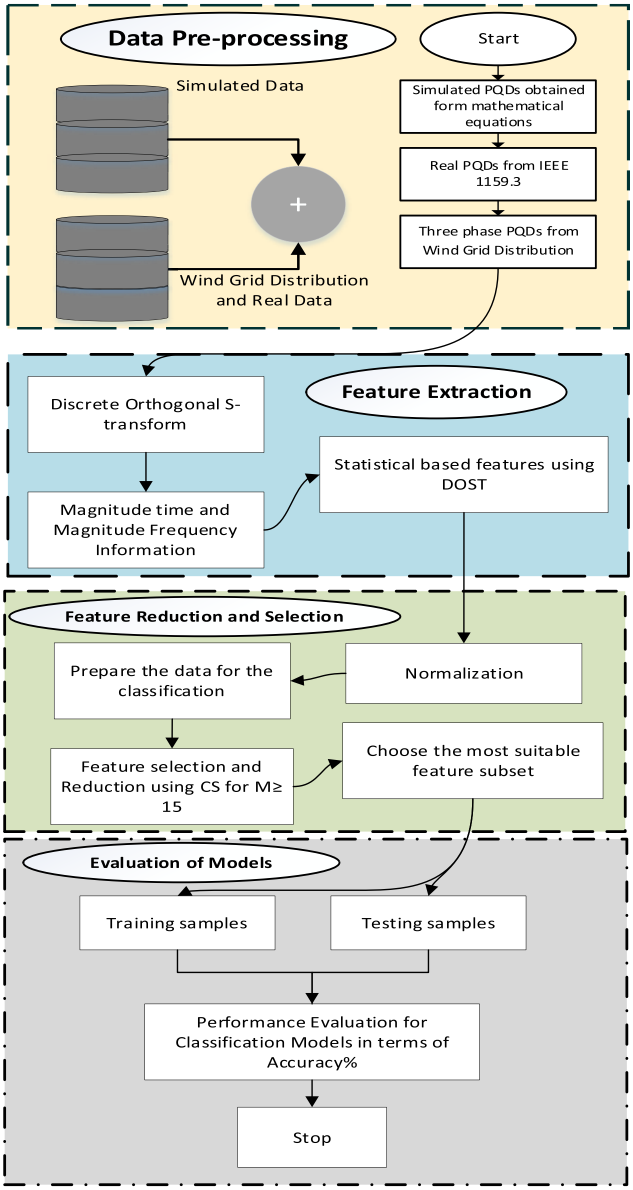


FIGURE 5. Proposed PQDs feature extraction and classification (a) flowchart (b) procedure.

Step 5: A newly developed fine-tune, batch mode gradient descent method is utilized to assist the learning of DSN. In DSN, the goal is to minimize the average square error and optimize the learning of the "W" weights. The batch mode gradient descent computes gradients with reduced noise, resulting in more efficient training compared to the conventional backpropagation method.

Step 6: The optimal feature obtained from the previous step is used to assess and compare the performance of various machine learning methods. At last, these results are compiled, and classification is done by using the proposed DOST, CS, and different machine learning methods (DSN, KNN, MSVM and EL).

Output: labeled each class as $C_1, C_{2,...,} C_n$.

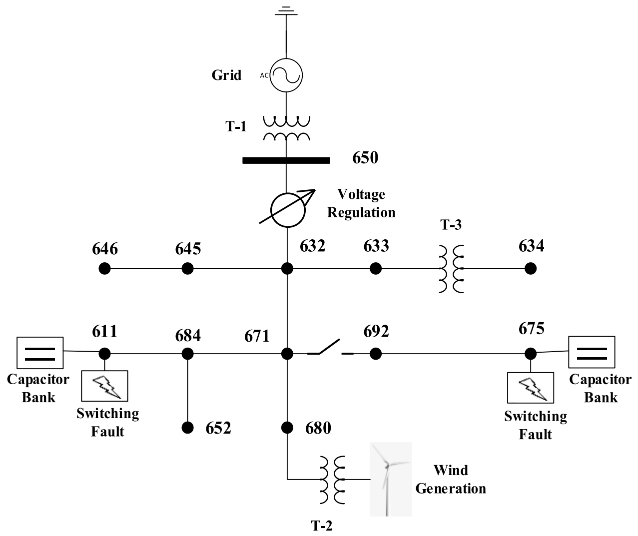


FIGURE 6. Modified IEEE 13 bus system with wind system connected at node.

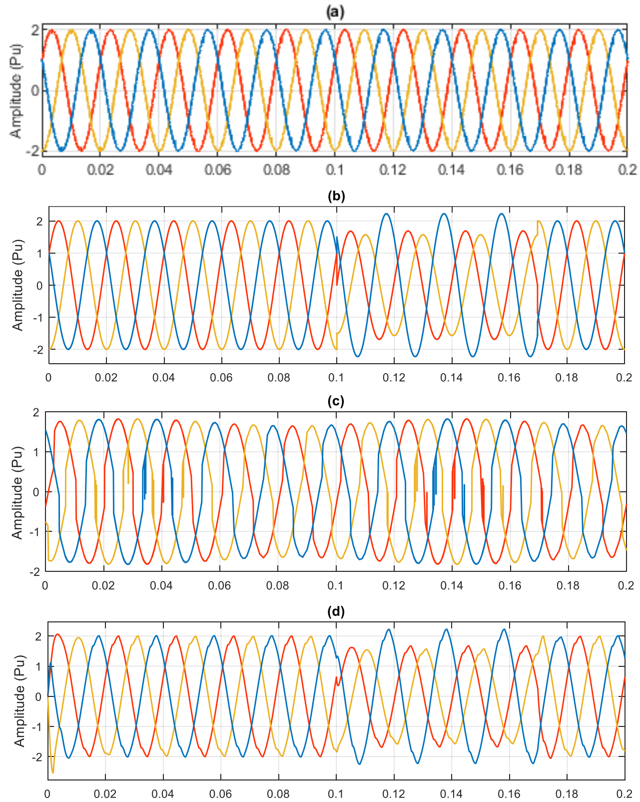


FIGURE 7. Three-phase PQDs generation from wind grid integration model (a) voltage notch (b) voltage sag and swell (c) voltage flicker (d) harmonics with sag and swell.

G. SIMULATION AND EXPERIMENT

1) GENERATION OF PQ DISTURBANCES

The PQ Generation Using Modified IEEE 13 Bus System is represented in Figure 6. 12 types of three phase single and multiple PQ disturbances were generated with the modified IEEE wind-grid model and simulated in MATLAB/Simulink as shown in Figure 6. The detail of this model can be seen in the literature [63]. Some of the three phase PQ disturbances were presented in Figure 7.

Three-phase voltage sag and swell are generated by multistage faults, line to line faults, and line to line faults between two phases at generation end. The three-phase voltage notch, oscillatory transients, harmonics, flickers, and impulsive transients are generated by the three-phase nonlinear load, three-phase capacitor bank, arc furnace, and lightning at distributive end respectively. Multiple PQ disturbances are generated from their combinations.

2) GENERATION OF SYNTHETIC AND REAL PQ DISTURBANCES

Based on mathematical models of PQ disturbances as stated in Annex. A [64], [65], 33 types of single and combined PQ disturbances were simulated using MATLAB and PSCAD/EMTDC software in a power network. The simulations were conducted under both noiseless and noisy conditions [62]. The different types of events and faults were applied to obtain the PQ disturbances. The line-to-line faults, double line to ground faults, three phase faults and line to ground faults are the key root of generating the sag, swell and Interruption. The voltage flickers, harmonics, and notch were generated by using the nonlinear converters based on power electronics, and variations of the speed of the motor. Impulsive transients and oscillatory transients were generated by applying the large capacitor banks switching. They were used to maintain the power factor of the system.

Multiple disturbances were generated by combining the aforementioned faults, converters based on power electronics, and capacitor banks. Additionally, factors such as overheating of the neutral conductor, circuit breaker trips, and power electronics-based equipment can also impact the normal voltage waveforms. The real PQ disturbances used in the study were sourced from the IEEE taskforce based on 1159.3. The signal to noise ratio (SNR) of the PQ events was observed to be approximately 45 dB. The amplitude of PQ signals was converted into per units (PU) to ease the experiment. Some of the Real PQ disturbances are presented in Figure 8. The detail of PQ disturbance can be studied in the literature [66], [67].

3) GENERATION OF DATASETS

In this study, the first dataset was based on the modified IEEE 13 bus system with wind-grid integration. 5590 samples waveforms were considered for training and testing the proposed classifier. The fundamental and sampling frequency was 50 Hz and 10 kHz respectively. For the second type of dataset, a total of more than 9900 samples were included for classification. Specifically, there were 300 samples available for each PQ event. The training and testing samples of data

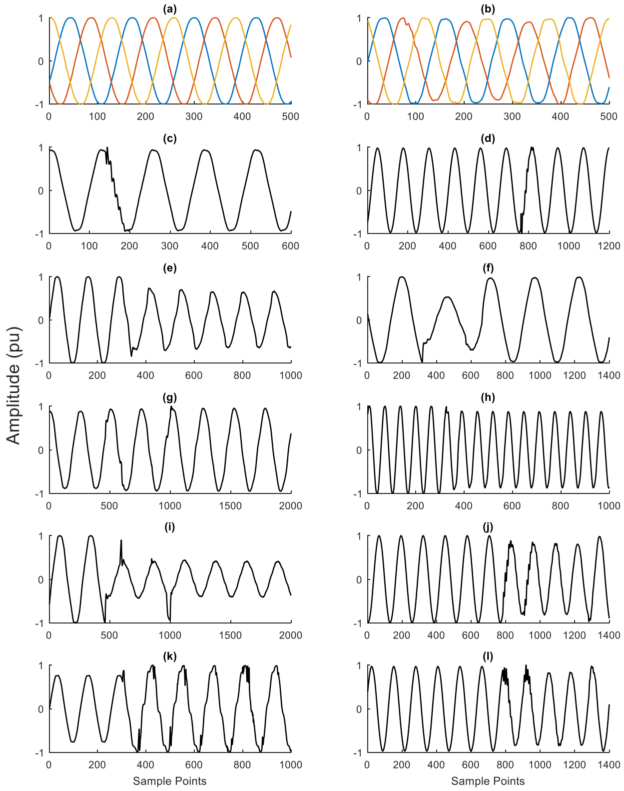


FIGURE 8. Real PQ disturbances waveforms [29]. (a) normal waveform (3-Phase);; (b) Unbalanced waveform (3- Phase); (c) Harmonics with oscillation transient; (d) Impulsive transient with harmonics; (e) Flicker with sag; (f) Sag with impulsive transient; (g) Notch with harmonics; (h) Sag with notch; (i) Sag and harmonics; (j) Harmonics with oscillation transient with sag; (k) Swell with impulsive transient with flickers with harmonics; (l) Harmonics with sag with oscillatory transient with notch.

sets were divided into 70% and 30% respectively. Nine types of single disturbances, fifteen types of the doublet disturbances, six types of triplet disturbances, two types of the quartet disturbance, and one unbalanced three phase PQ disturbances were considered. The variation in the fundamental frequency ranged from 49 Hz to 51 Hz. Some synthetic waveforms from this dataset are shown in Figure 9, with a sampling frequency of 10 kHz and ten cycles. The fundamental frequency was 50 Hz. PQ disturbances were classified into two groups based on single and combined PQDs. The Gaussian noise level was considered 0dB, 20 dB and 50dB for each signal.

4) IMPLEMENTATION OF DOST-CS FEATURE EXTRACTION AND REDUCTION

Figure 10 demonstrates the normalized and centered representations of PQD signals using FFT and DOST. The time-frequency representation of DOST for PQ disturbances reveals visually distinctive features that are encircled in the discussed figure. These visually distinctive features make it easier for the proposed machine learning methods to solve recognition problems based on these visual characteristics.

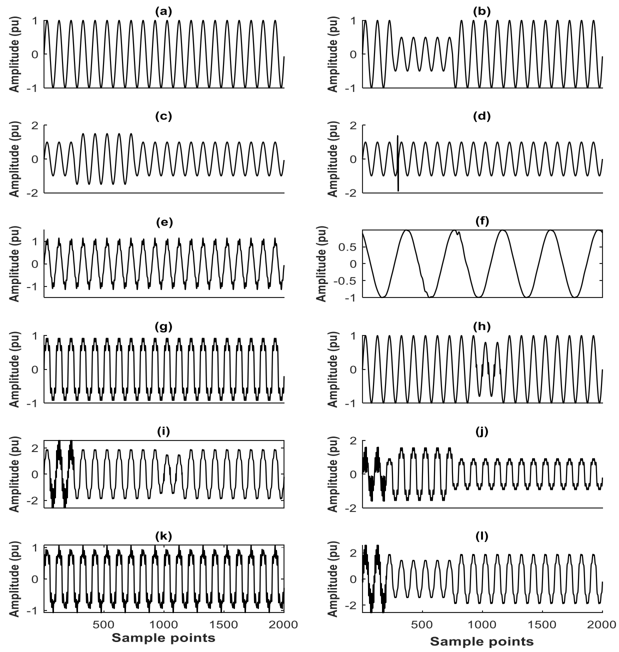


FIGURE 9. Simulated PQDs. (a) Normal; (b) Sag; (c) Swell; (d) Impulsive transient; (e) Voltage flickers; (f) Voltage notch; (g) Harmonics; (h) Sag with notch; (i) Notch with oscillatory transient with sag with harmonic; (j) Harmonics with oscillatory transient with swell; (k) Harmonics and flickers; (l) Harmonics with sag with oscillatory transient with notch.

5) FEATURE REDUCTION USING COMPRESSIVE SENSING

Figure 11 illustrates the process of CS (Compressive Sensing) sampling in the context of PQD (Power Quality Disturbance) classification. The input PQD signal undergoes pre-processing, followed by a feature extraction technique. The selection of an appropriate set of features is crucial for accurate PQD classification. The feature selected from the proposed method (DOST-CS) is fed to the respective machine learning classifiers. The primary aim of applying time-frequency analysis (DOST) and compressive technique is to simplify input data and reduce the vector dimension.

The extraction of features is important for the accuracy of the classification of PQ indices. PQDs are produced using MATLAB R2020a software in accordance with IEEE standards P1159.2 and 1159.3, real data is taken from the IEEE workgroup [56] and thirty-three types of steady-state and non-stationary PQ disturbances are considered in this analysis.

Using the proposed technique (DOST-CS), many features of the PQDs can be extracted, such as phase angle, start time, amplitude variation, frequency and many more. Relevant features have been selected to reduce the computational burden.

These features are drawn from the (DOST-CS) output. They are listed as follows:

$$F1 = \max(Y) + \min(Y) - \max(Z) - min(Z).$$

In the context of the S-matrix, Y and Z represent disturbances and the absence of disturbances, respectively, in the amplitude versus time information.

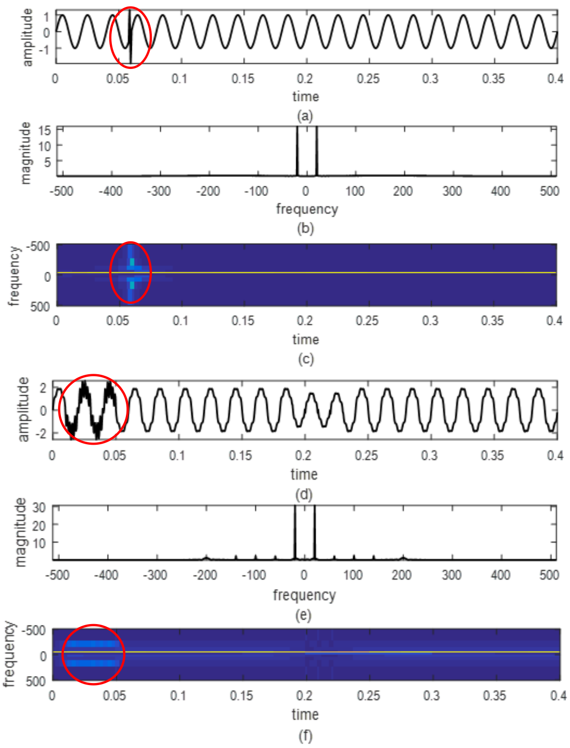


FIGURE 10. **(a) Transient PQD signal (b) normalized and centered FFT of Transient PQD signal (c) DOST coefficients of Transient PQD signal (d) Harmonics with sag with oscillatory transient with notch PQD signal (e) normalized and entered FFT of Sag PQD signal (f) DOST coefficients of Sag PQD signal.**

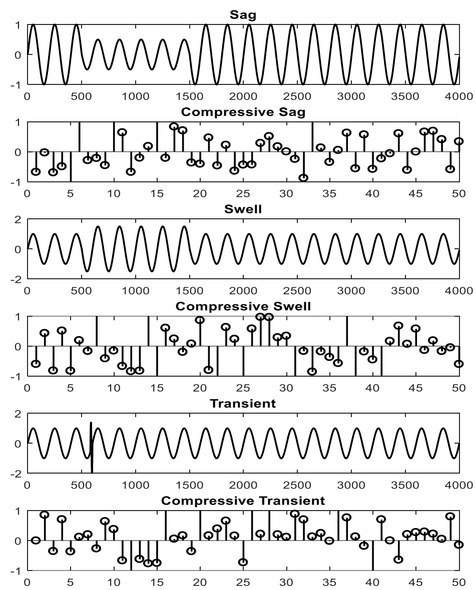


FIGURE 11. **Compressive sensing sampling and compression.**

F2 = skewness of a phase curve derived from DOST, and it is determined by the relationship $s = \frac{Y(m-\mu)^3}{\sigma^3}$, where m is the amplitude versus time signal data, μ is mean of m and Y is the expected value of the quantity.

F3 = standard deviation total magnitude to time spectrum obtained from DOST matrix.

$$SD = \left(\frac{1}{N}\sum_j (S(j) - \frac{1}{N}\sum_j S(j))^2\right)^{1/2} \tag{13}$$

F4=Energy in the DOST-CS output is $= |S(n, j)|^2$

F5=total harmonics distortion $= \frac{\sqrt{\sum_{n=2}^{N}[V_n]^2}}{|V_1|}$, where N belongs to the number of points in DOST and Vn is the *nth* harmonics component of the DOST.

F6=Kurtosis in the DOST-CS output is $k = \frac{E(x-\mu)^4}{\sigma^4}$

F7=Estimated frequency obtained from the output of DOST-CS matrix at maximum amplitude.

The feature vector can be written as $F_s = \{F_1, F_2, F_3, F_4, F_5, F_6, F_7\}$. However, 7 features were extracted for each PQDs and 231 features were extracted for all selected PQDs input. One of the main concerns of this research was to lessen the computation burden. The classification rate in Figure 12 shows the precise relationship between the number of measurements (M) and the classification rate. Here M is the length feature vector. Since the nature of PQDs is complex, it is crucial to select the value of M within a specific margin in order to successfully reconstruct the signal. The simulation was conducted with various numbers of compressive measurements, including 5, 10, 15, 20, 30, 40, 50, and 60. Accurate classification results were achieved when the number of measurements, M, was equal to or greater than 15. As M increases, the accuracy of PQ event classification also improves. Additionally, the data dimension was reduced by a factor of four, indicating a compression ratio of 1:4.

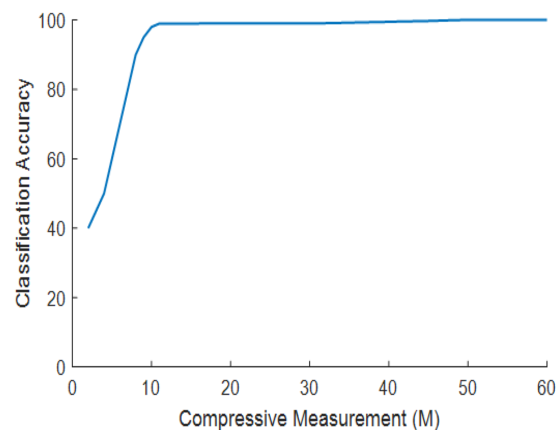


FIGURE 12. **Relationship between compressive sensing (M) and classification accuracy.**

III. RESULTS AND DISCUSSIONS

A. CLASSIFICATION PERFORMANCE

After the feature reduction stage, the data, which consists of seven features and has been reduced by a factor of four, is used as an input for various classifiers. The reduced data comprises of 33 types of PQ events. The features are utilized to train different machine learning models, including KNN,

TABLE 1. Selected classifiers and their parameters.

Classifiers	Parameters	Details
DSN	Number of hidden neurons for DCNN ranged between 1 to 50.	Deep stacking neural network was trained using the hidden layer with neuron number ranging from 3 to 50, the weight matrix value was 0 to 1, biases, sigmoid kernel, and output was the linear layer.
KNN	The value of KNN is 1,3,5,7 and 9	The KNN was trained for different nearest neighbors.
Subspace KNN	The learning numbers were 1,3,5, …, 39	Subspace KNN learner was trained with 1 to 39 nearest neighbors and subspace dimension was 1.
MSVM	Four kernels were used such as linear, polynomial, Gaussian and with automatic scale of 3.5	Four kernels (linear, polynomial, Gaussian and sigmoid) were trained for one vs one method.
Adaboost	The learning numbers were 1,3,5, …, 39	Adaboost learner was trained for the 40 numbers with maximum number of splits was 2 and the learning rate was 0.05.

MSVM, DSN, and ensemble learners. Each model has its own set of parameters for training and testing the data. Table 1 provides the specific parameters for each classification model. The simulation is performed over 60 iterations, and the performance of each model is evaluated in terms of classification accuracy.

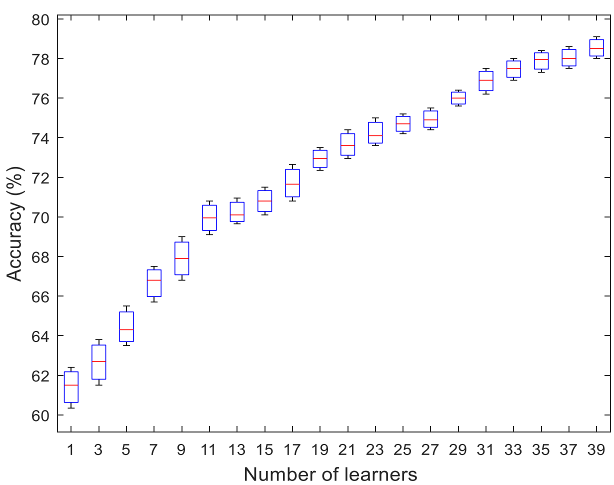


FIGURE 13. The performance of Adaboost classifier for the number of learners.

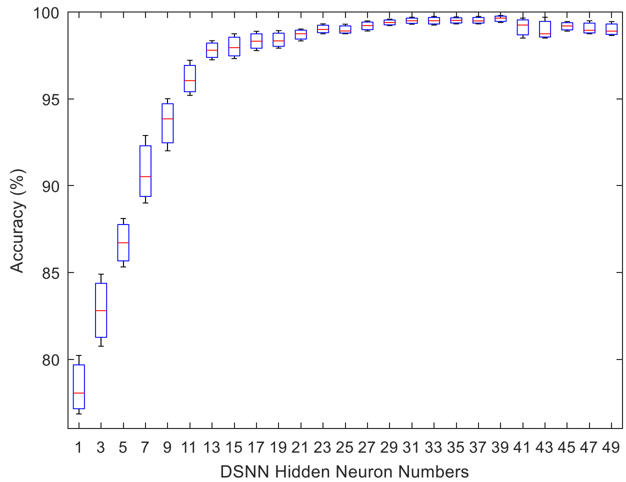


FIGURE 14. The performance of DSN classifier for the hidden neurons.

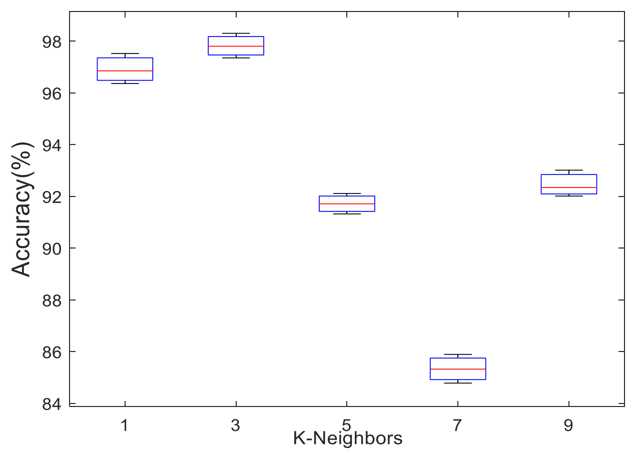


FIGURE 15. The performance of KNN classifier for the selected K neighbors.

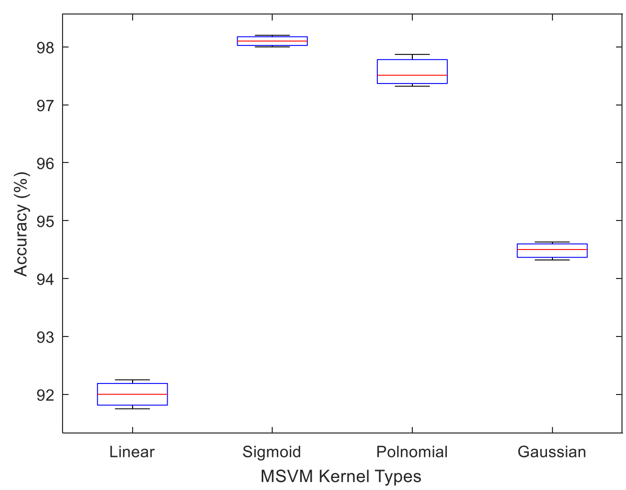


FIGURE 16. The performance of MSVM classifier for different types of kernels.

In Figures (13-17), The performance of selected classifiers are discussed. Out of four primary kernels (linear, radial basis function, polynomial, and sigmoid), the sigmoid kernel was selected as it is most advanced and compatible with

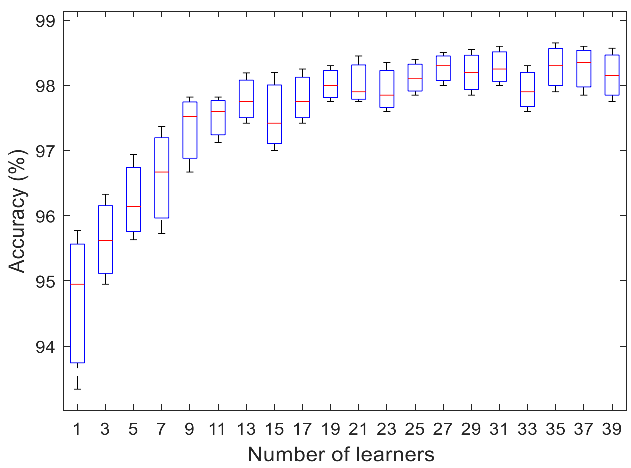


FIGURE 17. The performance of subspace KNN classifier for the number of learners.

TABLE 2. The classification accuracy of the classifiers with optimal parameters.

Classifiers	Parameters	Accuracy (%)
KNN	K=3	98.56
DSN	Hidden neurons are 50, kernel= Sigmoid	99.60
MSVM	Kernel= sigmoid	98.26
Adaboost	Leaners number= 39	79.30
Subspace KNN	Leaners number= 37	98.25

neural network classifier. The range of hidden neurons ranged between 3 to 50 and best classification accuracy was obtained at the hidden neuron number 43. It was observed that increasing the number of hidden neurons resulted in higher accuracy; however, it also led to an increase in computational burden. For MSVM, polynomial kernel function had the best accuracy among the selected four kernel functions. Among Adaboost and subspace KNN, the subspace KNN method achieved the best accuracy. The KNN classifier exhibited the highest accuracy when the value of K was set to 3. Furthermore, it is noteworthy that increasing the value of K leads to a decrease in the classifier's performance. Table 2 displays the classification accuracies for all the classification methods. It is evident that the DSN model achieves the highest classification accuracy among all the methods.

Under a noiseless environment, the proposed algorithm was implemented. Different sets of trained samples, consisting of 50, 100, and 200 samples, were utilized for each type of PQ disturbance. The automatic classification results for the different training sample sizes are presented in Table 3. The table clearly demonstrates that as the number of training samples increased, the classification accuracy of PQ disturbances also improved.

In general, the signals in real electric power systems are contaminated with noise, specifically Additive White Gaussian Noise (AWGN), which is added to PQ events. The proposed algorithm was tested under different signal-to-noise ratios (SNRs) ranging from 20 to 50 dB. In this paper, instead of using raw data, optimally chosen features were utilized. These selected features, including skewness, standard deviation, energy, total harmonic distortion, kurtosis, and estimated frequency, were used to approximate the performance of the chosen deep learning methods.

The DSN model demonstrates the highest overall classification accuracy of 99.82% for single PQ events and 99.60% for multiple PQ events. These models were evaluated using synthetic, real, and simulated datasets. Table 4 presents the classification accuracy rates for single PQ events, indicating 100% accuracy for noiseless data, 99.84% for 20 dB noise, 99.81% for 30 dB noise, 99.80% for 50 dB noise, 99.60% for real data, and 99.88% for simulated data. In the case of multiple PQ events, Table 5 showcases the classification rates of 99.83%, 99.66%, 99.62%, 99.61%, 99.33%, and 99.43% for noiseless, 20 dB, 30 dB, 50 dB, real data, and simulated data, respectively.

In Annex. B, the high classification rate for combined events and noisy conditions indicates that the proposed algorithm was robust to noise and capable of accurately classifying complex PQ events. In general, the amplitude of sag typically falls around 0.95 PU, so the classifier needs to effectively operate within the range of 0.94 to 0.96 PU for accurate classification. However, the classifier may face challenges in distinguishing events occurring near this boundary. Similar challenges may arise in the case of swells. When harmonics, impulsive transients, and oscillatory transients are combined with sags or swells, the system may encounter difficulties due to the high-frequency and low-magnitude nature of these transients.

Based on the results, it is evident that the proposed classifier performs accurately for single PQ events. However, in the case of complex PQ events, misclassification frequently occurred. For instance, sag with harmonics was correctly classified, but sag with swell and harmonics were often confused with each other. This is attributed to the complexity arising from the combination of multiple disturbances, particularly the occurrence of sag with swell for a short period, which the model may erroneously identify as harmonics.

In this study, the selected features were extracted from DOST-CS. Compressive sensing was employed to compress the dimension of the feature vector, and the selected features were used as input to the best-selected DSN classifier. Figure 18 depicts the relationship between the number of training samples and the classification error. It suggests that as the number of training samples increases, the classification error has the potential to decrease. This implies that having a larger training dataset can contribute to improved classification accuracy.

TABLE 3. Impact of the number of training samples on classification accuracy under noiseless condition.

Class Labeled	Training Samples			Class Labeled	Training Samples			Class Labeled	Training Samples			Class Labeled	Training Samples		
	30	**50**	**100**		**30**	**50**	**100**		**30**	**50**	**100**		**30**	**50**	**100**
C0	100	100	100	C8	96	99	100	C16	94	98	99	C24	96	99	100
C1	98	99	100	C9	97	98	99	C17	95	99	100	C25	97	98	99
C2	99	99	100	C10	98	99	100	C18	96	98	99	C26	96	97	99
C3	98	98	100	C11	96	97	99	C19	98	99	100	C27	97	98	99
C4	97	99	100	C12	98	98	99	C20	98	99	100	C28	98	99	99
C5	99	100	100	C13	97	98	99	C21	96	98	99	C29	97	98	99
C6	98	99	100	C14	98	99	100	C22	98	99	99	C30	98	99	100
C7	97	99	100	C15	95	97	99	C23	97	99	100	Accuracy%	97.2	98.5	99.8

TABLE 4. Classification accuracy of single pq disturbances for simulated and real data under different noise level.

PQ Disturbances	Class labeled	Synthetic Data				Real Data	Simulated Data	Overall (%)
		0 dB	20 dB	30 dB	50 dB			
Normal	C0	100	100	100	100	100	100	100
Sag	C1	100	99.96	99.92	99.92	99.87	100	99.95
Swell	C2	100	99.98	99.95	99.95	99.85	-	99.95
Interruption	C3	100	99.90	99.91	99.85	99.67	-	99.87
Impulsive transient	C4	100	99.87	99.83	99.83	99.28	99.81	99.77
Oscillatory transient	C5	100	99.80	99.75	99.75	99.37	99.7	99.73
Flicker	C6	100	100	100	100	99.83	100	99.97
Notch	C7	100	99.28	99.32	99.32	99.15	99.51	99.43
Harmonic	C8	100	99.76	99.63	99.61	99.47	99.68	99.69
Accuracy (%)		100	99.84	99.81	99.80	99.60	99.88	99.82

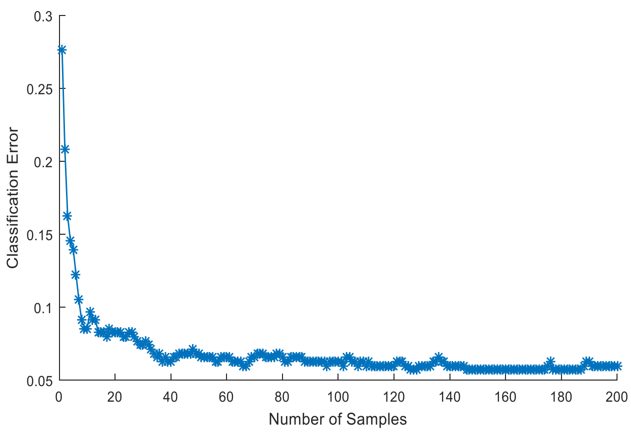


FIGURE 18. Classification error with training samples.

B. COMPARISON OF COMPUTATIONAL TIME

Extensive PQ event data requires a very high computational cost. The hardware utilized in this study is Intel core i7-6700 CPU @ 3.40 GHz, and 16 GB of RAM. The proposed technique is productive to reduce the features that represent the essential properties of the signal without dropping off its distinctive properties. By employing the optimal method of DOST for feature extraction, utilizing CS for feature dimension reduction, and employing the DSN classifier for PQ event classification, the proposed technique successfully reduces computational time, memory space, and pre-processing requirements while improving classification accuracy. Table 6 shows the estimated time assessment between the DSN, KNN, subspace KNN, MSVM, and Adaboost classifiers The proposed method is more optimized as compared to the other selected methods. The DSN model is found to be one of the fastest classification methods for the classification of PQ disturbances.

C. PERFORMANCE COMPARISON WITH PUBLISHED ARTICLES

Finally, to evaluate the effectiveness and practicability of the proposed method, classification results were compared with other newly introduced tested techniques. The evaluation consists of various criteria, such as, the total number of single and multiple PQ disturbances, types of datasets, and accuracy. Table 7 illustrates the comparative results of the proposed and other selected latest classifiers. Optimal multi-resolution fast S-transform (OMFST) and cart algorithm (CA) based algorithm were presented in [68]. The algorithm aimed to classify five types of simulated single PQ disturbances

TABLE 5. Classification accuracy results of multiple pq disturbances for simulated and real data under different noise level.

PQ Disturbances	Class labeled	Synthetic Data				Real Data	Simulated Data	Overall
		0 dB	20 dB	30 dB	50 dB			
Sag with harmonics	C9	100	100	99.97	100	99.79	99.65	99.90
Sag with Swell with harmonics	C10	99.93	99.91	99.87	99.84	99.71	99.5	99.79
Notch with harmonics	C11	99.77	99.67	99.43	99.33	99.27	-	99.50
Flicker with harmonics	C12	99.93	99.81	99.83	99.81	99.86	-	99.85
Oscillatory transient with harmonics	C13	99.75	99.52	99.47	99.49	99.65	-	99.58
Impulsive transient with harmonics	C14	99.25	99.15	99.35	99.05	98.95	-	99.15
Sag with flicker	C15	100	99.93	99.96	99.89	99.57	-	99.87
Swell with flicker	C16	100	99.94	99.95	99.94	99.64	-	99.89
Interruption with flicker	C17	99.91	99.83	99.83	99.77	98.97	-	99.66
Swell with notch	C18	100	99.88	99.87	99.91	99.07	-	99.75
Sag with notch	C19	99.53	99.25	99.37	99.42	99.41	-	99.40
Sag with impulsive transient	C20	99.71	99.56	99.35	99.33	99.05	-	99.32
Swell with impulsive transient	C21	99.74	99.70	99.63	99.76	98.92	-	99.55
Sag with oscillatory transient	C22	99.85	99.36	99.21	99.23	99.1	99.25	99.33
Sag and Swell with oscillatory transient	C23	99.92	99.23	99.12	99.17	98.93	99.3	99.28
Swell with notch with harmonics	C24	99.94	99.84	99.84	99.75	99.01	-	99.68
Swell with oscillatory transient with harmonics	C25	99.95	99.89	99.87	99.77	99.63	-	99.82
Swell with impulsive transient with harmonics	C26	99.97	99.79	99.68	99.71	99.1	-	99.65
Sag with notch with harmonics	C27	99.57	99.47	99.42	99.42	99.60	-	99.50
Sag with impulsive transient with harmonics	C28	99.69	99.36	99.30	99.28	99.25	-	99.38
Oscillatory transient with impulse transient with harmonics	C29	99.90	99.85	99.85	99.87	99.40	-	99.77
Harmonics with sag with oscillatory transient with notch	C30	99.83	99.74	99.79	99.69	99.15	-	99.64
Harmonics with Flickers with impulsive transient with swell	C31	99.65	99.05	99.05	99.15	99.00	-	99.18
Three phase unbalance voltage	C32	100	100	100	100	100	-	100
Accuracy (%)		99.83	99.66	99.63	99.61	99.33	99.4	99.60

TABLE 6. Computational time (s) comparison of the proposed method with the SVM based method.

Classifiers	Computational Time (sec)
DSN	0.1976
SVM	2.685
KNN	1.065
Subspace KNN	0.734
Adaboost	4.971

and seven types of simulated double PQ disturbances. The achieved classification rate for these simulated PQ disturbances was reported as 99.59%. However, it is worth noting that this paper did not evaluate real data and the number of evaluated PQ disturbances was relatively limited. In [69], the automatic classification of PQ disturbances using the DWT-PNN (Discrete Wavelet Transform-Probabilistic Neural Network) classifier was discussed. The study focused on ten types of single PQ disturbances and six types of multiple PQ disturbances under noiseless conditions. The reported classification accuracy of the DWT-PNN classifier in the noiseless condition was 99.875%. However, it was noted that in the presence of noise, the accuracy of the DWT-PNN classifier decreased to 98.6%. This lower accuracy in noisy conditions suggests a need for improvement in the classifier's performance to enhance its robustness and accuracy when dealing with noisy PQ signals. In order to improve the classification accuracy, a study [70] explored the use of Sparse Signal Decomposition (SSD) along with a Rule-Based (RB) classifier. The approach was tested on a dataset consisting of eleven types of single PQ disturbances and twenty-one types of multiple PQ disturbances. The reported classification rates for single PQ disturbances and multiple PQ disturbances were 99.87% and 96%, respectively.

In [71], the classification of PQ disturbances was carried out using Multidimensional Analysis (MA) and a Neuro-Tree (NT) based classifier. However, it is important to note that the classification specifically focused on six types of single PQ disturbances, ten types of double PQ disturbances, and four types of triplet PQ disturbances. No consideration was given

TABLE 7. Proposed method comparison with recently published methods.

Ref.	Feature Extraction and Classification Algorithms	Data Type	No. of single PQDs	With Accuracy(%)	No. of multiple PQDs	With Accuracy(%)
[68]	OMFST + CA	Synthetic	5	99.6	7	95.33
[69]	DWT + ABC + PNN	Synthetic and Real	10	99.8	6	99.8
[70]	SSD + RB	Synthetic	11	99.87	21	96
[71]	MA + NT	Synthetic and Real	6	99.86	14	96.96
[72]	ACMP + SVM	Synthetic and Real	12	99.12	11	99.56
[39]	EWT + SVM	Synthetic and Real	9	96.44	12	94.43
[73]	GWO + ELM	Synthetic	9	100	8	99.41
[74]	ST + IGWO + KELM	Synthetic	8	100	9	99.76
[41]	ST + HT + RBDT	Synthetic	10	99.82	11	99.24
[40]	DBN + ELM	Synthetic and Real	7	99.28	14	99.4
-	DOST – CS – DSN	Synthetic and Real	9	100	24	99.81

to quartet PQ disturbances. The reported results indicated that the classification accuracy for single PQ disturbances and some of the double PQ disturbances achieved a perfect accuracy rate of 100%. However, for triplet PQ disturbances and other double PQ disturbances, the classification rate dropped to 95.3%. In [72], an approach combining Adaptive Chirp Mode Pursuit (ACMP), Grasshopper Optimizer (GO), and an SVM-based classifier was presented. The study utilized twelve types of single PQ events and eleven types of multiple PQ events for classification. However, the classification accuracy achieved by this approach was lower than that of the proposed model.

Adaptive chirp mode pursuit (ACMP) and grasshopper optimizer (GO) and SVM based classifier were presented [72] and twelve type of single and 11 type of multiple PQ events were used to classify. But the classification accuracy is less than the proposed model.

In [39], Empirical Wavelet Transform-based Adaptive Filtering (EWT) and SVM-based model were utilized to classify nine types of single PQ events and twelve types of doublet PQ events. The classification accuracy achieved by this model was reported as 96.44% for single PQ events and 94.43% for doublet PQ events. It is important to note that no triplet or quartet PQ events were included in the classification. In [73], Grey Wolf Optimizer (GWO) and Extreme Learning Machine (ELM) were employed for the classification of nine types of single PQ events and eight types of doublet PQ events. The classification accuracy obtained by this model was 100% for single PQ events and 99.41% for doublet PQ events. However, the proposed model has achieved superior classification results for both single and multiple PQ disturbances.

In [74], the authors proposed the ST with improved Grey Wolf Optimizer (IGWO) and Kernel Extreme Learning Machine (KELM) for classification. This algorithm achieved a good classification accuracy of 100% for single PQ events and 99.76% for multiple PQ event problems. However, the authors did not discuss the computational complexity, nor did they test it with real PQ events. In [41], the ST with Hilbert transform (HT) for feature extraction and a rule-based decision tree (RBDT) for classification was proposed. The algorithm attained better classification accuracy of 99.80% for single PQ events and 99.24% for multiple PQ events. However, the authors did not discuss the computational complexity of their model. In [40], wavelet threshold denoising (WTD) and Deep Belief Network fusion with Extreme Learning Machine (DBN-ELM) were proposed to classify PQ events. This model worked very well for the proposed datasets and achieved a high classification accuracy of 99.24% for single PQ events and 99.4% for multiple PQ disturbance classes. However, the authors discussed the problem of a smaller number of PQ events.

Moreover, it is clear from the comparative findings that the presented method attains higher classification results for triplet and quartet PQ disturbances, even under high-intensive noisy conditions. Only a limited number of studies discussed the computational complexity of their models, and neither of them considered thirty-three or more PQ events for the classification. It becomes evident that the proposed model outperforms previous studies in terms of classification accuracy.

For future work, real data of PQDs can be used to test the efficiency of DOST-CS-DSN and this work can be utilized to test the fault detection of transmission lines and transformers.

IV. CONCLUSION

In this research work, a novel algorithm for the detection and classification of PQDs is presented, based on DOST-CS and DSN. However, this model is compared with different machine learning methods, such as KNN, subspace KNN, MSVM, and Adaboost. Synthetic and real datasets comprising thirty-three types of single and multiple PQ disturbances, with varying noise levels, as well as a simulated dataset of 12 types of PQ disturbances using a modified IEEE 13 system wind-grid model, were evaluated. In the case of feature extraction, DOST is measured to be more accurate as compared to S-transform and, it has better time-frequency localization at low frequency and high frequencies respectively. It takes N points and provides N points on the time-frequency spectrum which helps to reduce the computational burden. Seven features were selected for correct identification of multiple PQ disturbance, and CS-based method also helped to reduce the computational burden. The dimension of the data is reduced by the four time as the compression ratio is 1:4. At last, the proposed classifier was examined with synthetic, real, simulated, and different noise level datasets. The results showed that the proposed method achieved the highest classification accuracy even in a noisy environment. The results are further compared with previously published articles that have undertaken classification with a larger set of PQ events. The high classification performance confirms that the proposed method can be applied to a digital signal processing-based PQ monitoring system for future research. Furthermore, these proposed methods can also be utilized in digital image processing, hyperspectral images and voice recognition, and classification.

ACKNOWLEDGMENT

The authors are highly grateful to their affiliated universities and institutes for providing research facilities.

REFERENCES

[1] H. Liu, F. Hussain, Y. Shen, R. Morales-Menendez, M. Abubakar, S. J. Yawar, and H. J. Arain, "Signal processing and deep learning techniques for power quality events monitoring and classification," *Electr. Power Compon. Syst.*, vol. 47, nos. 14–15, pp. 1332–1348, Sep. 2019.

[2] H. Liu, F. Hussain, S. Yue, O. Yildirim, and S. J. Yawar, "Classification of multiple power quality events via compressed deep learning," *Int. Trans. Electr. Energy Syst.*, vol. 29, no. 6, Jun. 2019, Art. no. e12010.

[3] C.-C. Liao, H.-T. Yang, and H.-H. Chang, "Denoising techniques with a spatial noise-suppression method for wavelet-based power quality monitoring," *IEEE Trans. Instrum. Meas.*, vol. 60, no. 6, pp. 1986–1996, Jun. 2011.

[4] F. Jurado and J. R. Saenz, "Comparison between discrete STFT and wavelets for the analysis of power quality events," *Electr. Power Syst. Res.*, vol. 62, no. 3, pp. 183–190, Jul. 2002.

[5] L. Mansinha, R. Stockwell, R. Lowe, M. Eramian, and R. Schincariol, "Local S-spectrum analysis of 1-D and 2-D data," *Phys. Earth Planet. Interiors*, vol. 103, nos. 3–4, pp. 329–336, 1997.

[6] S. Santoso, E. J. Powers, W. M. Grady, and P. Hofmann, "Power quality assessment via wavelet transform analysis," *IEEE Trans. Power Del.*, vol. 11, no. 2, pp. 924–930, Apr. 1996.

[7] M. Uyar, S. Yildirim, and M. T. Gencoglu, "An effective wavelet-based feature extraction method for classification of power quality disturbance signals," *Electr. Power Syst. Res.*, vol. 78, no. 10, pp. 1747–1755, Oct. 2008.

[8] P. K. Dash, B. K. Panigrahi, and G. Panda, "Power quality analysis using s-transform," *IEEE Trans. Power Del.*, vol. 18, no. 2, pp. 406–411, Apr. 2003.

[9] 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.

[10] M. J. B. Reddy, R. K. Raghupathy, K. P. Venkatesh, and D. K. Mohanta, "Power quality analysis using discrete orthogonal S-transform (DOST)," *Digit. Signal Process.*, vol. 23, no. 2, pp. 616–626, Mar. 2013.

[11] A. J. Jerri, "The Shannon sampling theorem—Its various extensions and applications: A tutorial review," *Proc. IEEE*, vol. 65, no. 11, pp. 1565–1596, Nov. 1977.

[12] M. Mishali and Y. C. Eldar, "Sub-nyquist sampling," *IEEE Signal Process. Mag.*, vol. 28, no. 6, pp. 98–124, Nov. 2011.

[13] J. A. Urigiien, Y. C. Eldar, P. Dragotti, and Z. Ben-Haim, "Sampling at the rate of innovation: Theory and applications," in *Compressed Sensing, Theory and Applications*, vol. 148. Israel: The Weizmann Institute of Science, 2012.

[14] M. Bertocco, G. Frigo, C. Narduzzi, and F. Tramarin, "Resolution enhancement by compressive sensing in power quality and phasor measurement," *IEEE Trans. Instrum. Meas.*, vol. 63, no. 10, pp. 2358–2367, Oct. 2014.

[15] M. A. Davenport, P. T. Boufounos, M. B. Wakin, and R. G. Baraniuk, "Signal processing with compressive measurements," *IEEE J. Sel. Topics Signal Process.*, vol. 4, no. 2, pp. 445–460, Apr. 2010.

[16] F. Hussain, H. Liu, and Y. Shen, "Notice of removal: A novel method for analysis of power quality transient disturbances," in *Proc. Int. Comput., Signals Syst. Conf. (ICOMSSC)*, Sep. 2018, pp. 563–566.

[17] H. Liu, F. Hussain, Y. Shen, S. Arif, A. Nazir, and M. Abubakar, "Complex power quality disturbances classification via curvelet transform and deep learning," *Electr. Power Syst. Res.*, vol. 163, pp. 1–9, Oct. 2018.

[18] R. Kumar, B. Singh, D. T. Shahani, A. Chandra, and K. Al-Haddad, "Recognition of power-quality disturbances using S-transform-based ANN classifier and rule-based decision tree," *IEEE Trans. Ind. Appl.*, vol. 51, no. 2, pp. 1249–1258, Mar. 2015.

[19] D. D. Ferreira, J. M. D. Seixas, and A. S. Cerqueira, "A method based on independent component analysis for single and multiple power quality disturbance classification," *Electr. Power Syst. Res.*, vol. 119, pp. 425–431, Feb. 2015.

[20] Z. Liu, Y. Cui, and W. Li, "A classification method for complex power quality disturbances using EEMD and rank wavelet SVM," *IEEE Trans. Smart Grid*, vol. 6, no. 4, pp. 1678–1685, Jul. 2015.

[21] K. Thirumala, M. S. Prasad, T. Jain, and A. C. Umarikar, "Tunable-Q wavelet transform and dual multiclass SVM for online automatic detection of power quality disturbances," *IEEE Trans. Smart Grid*, vol. 9, no. 4, pp. 3018–3028, Jul. 2018.

[22] Y. Xi, Z. Chen, X. Tang, Z. Li, and X. Zeng, "Type identification and time location of multiple power quality disturbances based on KF-ML-aided DBN," *IET Gener., Transmiss. Distrib.*, vol. 16, no. 8, pp. 1552–1566, Apr. 2022.

[23] S. Mishra, C. N. Bhende, and B. K. Panigrahi, "Detection and classification of power quality disturbances using S-transform and probabilistic neural network," *IEEE Trans. Power Del.*, vol. 23, no. 1, pp. 280–287, Jan. 2008.

[24] Z. Wang, Y. Liu, and P. J. Griffin, "A combined ANN and expert system tool for transformer fault diagnosis," in *Proc. IEEE Power Eng. Soc. Winter Meeting*, Jan. 1999, pp. 1261–1269.

[25] P. G. V. Axelberg, I. Y.-H. Gu, and M. H. J. Bollen, "Support vector machine for classification of voltage disturbances," *IEEE Trans. Power Del.*, vol. 22, no. 3, pp. 1297–1303, Jul. 2007.

[26] S. R. Samantaray, "Decision tree-initialised fuzzy rule-based approach for power quality events classification," *IET Gener. Transmiss. Distrib.*, vol. 4, no. 4, pp. 538–551, 2010.

[27] P. Dash, "Power quality monitoring using an integrated Fourier linear combiner and fuzzy expert system," *Int. J. Electr. Power Energy Syst.*, vol. 21, no. 7, pp. 497–506, Oct. 1999.

[28] H. Liu, F. Hussain, and Y. Shen, "Power quality disturbances classification using compressive sensing and maximum likelihood," *IETE Tech. Rev.*, vol. 35, no. 4, pp. 359–368, Jul. 2018.

[29] A. K. Ghosh and D. L. Lubkeman, "The classification of power system disturbance waveforms using a neural network approach," *IEEE Trans. Power Del.*, vol. 10, no. 1, pp. 109–115, Jan. 1995.

[30] K. Cai, B. Alalibo, W. Cao, Z. Liu, Z. Wang, and G. Li, "Hybrid approach for detecting and classifying power quality disturbances based on the variational mode decomposition and deep stochastic configuration network," *Energies*, vol. 11, no. 11, p. 3040, Nov. 2018.

[31] F. Ucar, O. Alcin, B. Dandil, and F. Ata, "Power quality event detection using a fast extreme learning machine," *Energies*, vol. 11, no. 1, p. 145, Jan. 2018.

[32] X. Zu and K. Wei, "A simple gated recurrent network for detection of power quality disturbances," *IET Gener., Transmiss. Distrib.*, vol. 15, no. 4, pp. 751–761, Feb. 2021.

[33] H. Wang, P. Wang, and T. Liu, "Power quality disturbance classification using the S-transform and probabilistic neural network," *Energies*, vol. 10, no. 1, p. 107, Jan. 2017.

[34] F. G. Montoya, A. García-Cruz, M. G. Montoya, and F. Manzano-Agugliaro, "Power quality techniques research worldwide: A review," *Renew. Sustain. Energy Rev.*, vol. 54, pp. 846–856, Feb. 2016.

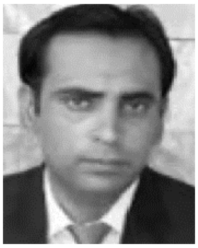
[35] D. De Yong, S. Bhowmik, and F. Magnago, "An effective power quality classifier using wavelet transform and support vector machines," *Expert Syst. Appl.*, vol. 42, nos. 15–16, pp. 6075–6081, Sep. 2015.

[36] 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.

[37] 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.
[38] 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.
[39] 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.
[40] 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.
[41] 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.
[42] L. Deng and D. Yu, "Deep convex net: A scalable architecture for speech pattern classification," in *Proc. Interspeech*, Aug. 2011, pp. 1–4.
[43] H.-T. Lin and C.-J. Lin, "A study on sigmoid kernels for SVM and the training of non-PSD kernels by SMO-type methods," *Neural Comput.*, vol. 3, nos. 1–32, p. 16, Mar. 2003.
[44] T. K. Abdel-Galil, M. Kamel, A. M. Youssef, E. F. El-Saadany, and M. M. A. Salama, "Power quality disturbance classification using the inductive inference approach," *IEEE Trans. Power Del.*, vol. 19, no. 4, pp. 1812–1818, Oct. 2004.
[45] R. G. Stockwell, "A basis for efficient representation of the S-transform," *Digit. Signal Process.*, vol. 17, no. 1, pp. 371–393, Jan. 2007.
[46] C. K. Jha and M. H. Kolekar, "Electrocardiogram data compression using DCT based discrete orthogonal stockwell transform," *Biomed. Signal Process. Control*, vol. 46, pp. 174–181, Sep. 2018.
[47] B. Nithya, Y. B. Sankari, K. Manikantan, and S. Ramachandran, "Discrete orthonormal Stockwell transform based feature extraction for pose invariant face recognition," *Proc. Comput. Sci.*, vol. 45, pp. 290–299, Jan. 2015.
[48] Y. Wang and J. Orchard, "The discrete orthonormal Stockwell transform for image restoration," in *Proc. 16th IEEE Int. Conf. Image Process. (ICIP)*, Nov. 2009, pp. 2761–2764.
[49] R. Upadhyay, P. K. Padhy, and P. K. Kankar, "EEG artifact removal and noise suppression by discrete orthonormal S-transform denoising," *Comput. Electr. Eng.*, vol. 53, pp. 125–142, Jul. 2016.
[50] Y. Wang, "Efficient Stockwell transform with applications to image processing," UWSpace, Canada, Tech. Rep. rpn.18373, 2011.
[51] M. J. Bharata Reddy, B. K. Chandra, and D. K. Mohanta, "A DOST based approach for the condition monitoring of 11 kV distribution line insulators," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 18, no. 2, pp. 588–595, Apr. 2011.
[52] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.
[53] E. J. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006.
[54] G.-C. Feng, P. C. Yuen, and D.-Q. Dai, "Human face recognition using PCA on wavelet subband," *J. Electron. Imag.*, vol. 9, no. 2, pp. 226–233, 2000.
[55] G. Daqi and L. Jie, "Kernel Fisher discriminants and kernel nearest neighbor classifiers: A comparative study for large-scale learning problems," in *Proc. IEEE Int. Joint Conf. Neural Netw.*, Jul. 2006, pp. 1333–1338.
[56] H. Wang, Z. Hu, and Y. Zhao, "An efficient algorithm for generalized discriminant analysis using incomplete Cholesky decomposition," *Pattern Recognit. Lett.*, vol. 28, no. 2, pp. 254–259, Jan. 2007.
[57] J. Yang, A. F. Frangi, J.-Y. Yang, D. Zhang, and Z. Jin, "KPCA plus LDA: A complete kernel Fisher discriminant framework for feature extraction and recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 2, pp. 230–244, Feb. 2005.
[58] L. Yan, J.-S. Pan, and X. Zhu, "Feature extraction based on nearest feature line and compressive sensing," in *Proc. 9th Int. Conf. Intell. Inf. Hiding Multimedia Signal Process.*, Oct. 2013, pp. 354–357.
[59] J. Weston and C. Watkins, *Multi-Class Support Vector Machines*. Princeton, NJ, USA: Citeseer, 1998.
[60] L. Deng, X. He, and J. Gao, "Deep stacking networks for information retrieval," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, May 2013, pp. 3153–3157.
[61] N. Benvenuto and F. Piazza, "On the complex backpropagation algorithm," *IEEE Trans. Signal Process.*, vol. 40, no. 4, pp. 967–969, Apr. 1992.
[62] *IEEE 1159.3 PQDIF Task Force*. Accessed: Jun. 24, 2023. [Online]. Available: https://grouper.ieee.org/groups/1159/3/
[63] Y. Shen, M. Abubakar, H. Liu, and F. Hussain, "Power quality disturbance monitoring and classification based on improved PCA and convolution neural network for wind-grid distribution systems," *Energies*, vol. 12, no. 7, p. 1280, Apr. 2019.
[64] S. Zhang, P. Li, L. Zhang, H. Li, W. Jiang, and Y. Hu, "Modified S transform and ELM algorithms and their applications in power quality analysis," *Neurocomputing*, vol. 185, pp. 231–241, Apr. 2016.
[65] S. Khokhar, A. A. Mohd Zin, A. S. Mokhtar, and N. Zareen, "Automatic pattern recognition of single and multiple power quality disturbances," *Austral. J. Electr. Electron. Eng.*, vol. 13, no. 1, pp. 43–53, Jan. 2016.
[66] M. H. Bollen and I. Y. Gu, *Signal Processing of Power Quality Disturbances*. Hoboken, NJ, USA: Wiley, 2006.
[67] A. Kusko, *Power Quality in Electrical Systems*. New York, NY, USA: McGraw-Hill, 2007.
[68] N. Huang, H. Peng, G. Cai, and J. Chen, "Power quality disturbances feature selection and recognition using optimal multi-resolution fast S-transform and CART algorithm," *Energies*, vol. 9, no. 11, p. 927, Nov. 2016.
[69] S. Khokhar, A. A. M. Zin, A. P. Memon, and A. S. Mokhtar, "A new optimal feature selection algorithm for classification of power quality disturbances using discrete wavelet transform and probabilistic neural network," *Measurement*, vol. 95, pp. 246–259, Jan. 2017.
[70] M. S. Manikandan, S. R. Samantaray, and I. Kamwa, "Detection and classification of power quality disturbances using sparse signal decomposition on hybrid dictionaries," *IEEE Trans. Instrum. Meas.*, vol. 64, no. 1, pp. 27–38, Jan. 2015.
[71] E. G. Ribeiro, T. M. Mendes, G. L. Dias, E. R. S. Faria, F. M. Viana, B. H. G. Barbosa, and D. D. Ferreira, "Real-time system for automatic detection and classification of single and multiple power quality disturbances," *Measurement*, vol. 128, pp. 276–283, Nov. 2018.
[72] S. Z. T. Motlagh and A. A. Foroud, "Power quality disturbances recognition using adaptive chirp mode pursuit and grasshopper optimized support vector machines," *Measurement*, vol. 168, Jan. 2021, Art. no. 108461.
[73] U. Subudhi and S. Dash, "Detection and classification of power quality disturbances using GWO ELM," *J. Ind. Inf. Integr.*, vol. 22, Jun. 2021, Art. no. 100204.
[74] T. Ravi and K. S. Kumar, "Detection and classification of power quality disturbances using stock well transform and improved grey wolf optimization-based kernel extreme learning machine," *IEEE Access*, vol. 11, pp. 61710–61727, 2023.

MUHAMMAD ABUBAKAR received the B.S. degree in electrical engineering from The University of Faisalabad, Faisalabad, in 2008, the M.S. degree in electronics engineering from International Islamic University, Islamabad, and the Ph.D. degree in electrical and information engineering from Jiangsu University, Zhenjiang, China.

He is currently an Assistant Professor with the Department of Computer Science, Lahore Garrison University, Lahore. He has more than ten years of experience in academics and research. He has published many research articles in renowned peer-reviewed journals and reviewer in multiple journals. His research interests include power quality, artificial intelligence, deep learning, and signal processing.



ARFAN ALI NAGRA received the Ph.D. degree from the Department of Computer Science, Jiangsu University, China. He is currently an Assistant Professor with Lahore Garrison University, Lahore. His principal research interests include intelligent computing and intelligent information processing, neural networks, particle swarm optimization, and bioinformatics.

MUHAMMAD FAHEEM received the B.Sc. degree in computer engineering from the University College of Engineering and Technology, Bahauddin Zakariya University Multan (BZU), Pakistan, in 2010, the M.S. degree in computer science from Universiti Teknologi Malaysia (UTM), Johor Bahru, Malaysia, in 2012, and the Ph.D. degree in computer science from the Faculty of Engineering, UTM, in 2021. He was a Lecturer with the COMSATS Institute of Information and Technology, Pakistan, from 2012 to 2014. He was an Assistant Professor with the Department of Computer Engineering, Abdullah Gul University (AGU), Kayseri, Turkey, from 2014 to 2022. He is currently a Researcher with the School of Technology and Innovations, Computer Science, University of Vaasa, Vaasa, Finland. He has numerous publications in journals and international conferences. His research interests include the areas of cybersecurity, industry 4.0, smart cities, smart grid, and underwater sensor networks.

MUHAMMAD MUDASSAR received the B.S. degree in electrical engineering from The University of Faisalabad, Faisalabad, Pakistan, in 2013, and the M.S. degree in energy engineering from the University of Engineering and Technology, Lahore, Punjab, Pakistan, in 2016. He is currently pursuing the Ph.D. degree in energy engineering with the University of the Punjab, Lahore.

He has served multiple institutes as a Lecturer and has more than seven years of experience in research and academics. He is a reviewer of multiple journals and has published more than 15 research articles in well-reputed local and international journals.

MUHAMMAD SOHAIL received the bachelor's and master's degrees in telecommunications and computer networks engineering from the University of Engineering and Technology, Peshawar, Pakistan, in 2010 and 2016, respectively, and the Ph.D. degree in computer application technologies from the School of Computer Science and Telecommunication Engineering, Jiangsu University, China. From 2011 to 2013, he was with Oman Telecom Company Ltd., as a Network Operation Engineer. He is currently with the Department of Computer Software Engineering, MCS, National University of Sciences and Technology, Islamabad, Pakistan. He is also associated with the School of Computer Science, Jiangsu University. His research interests include VANETs, information security, artificial intelligence, and software-defined networks. He served as a TCP Member for IEEE-ITIA, Jiangsu, China, in November 2020. He is serving as a Reviewer for *IET Information Security*, *Neural Computing and Applications*, *EURASIP Journal on Wireless Communications and Networking*, *Peer J Computer Science* journal, IEEE ACCESS, *Sustainable Cities and Society* and IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS.