Kurdish Studies
May 2024
Volume: 12, No: 4, pp. 1263-1275
ISSN: 2051-4883 (Print) | ISSN 2051-4891 (Online)
www.KurdishStudies.net

DOI: 10.53555/ks.v12i4.3143

Optimizing Power Quality Disturbance Classification With Higher Order Statistics: Social And Economic Impacts

Muhammad Abubakar[1*], Arfan Ali Nagra[2], Junaid Waseem[3], Zain Ali[4], Amber Sultan[5], Hanan Sharif[6], Ali Haider Khan[7]

[1*,2,3,4,6,7]Faculty of Computer Science, Lahore Garrison University, Lahore, 54000, Pakistan
[5]Department of Technology, The University of Lahore, Lahore, 54000, Pakistan

***Corresponding Author:** Muhammad Abubakar
Email: abubakarqazi@lgu.edu.pk

Abstract: In this research, four signal decomposition techniques (Empirical mode decomposition (EMD), Multivariate singular spectral analysis (MSSA), Wavelet Packet Decomposition (WPD), and Discrete wavelet transform (DWT) are studied for the optimal selection of signal processing technique to classify the Power Quality disturbances (PQD). Twelve types of single, multiple and synthetic PQD dataset is simulated from MATLAB R2020b, and real data is acquired from IEEE power quality guideline 1159.3-2019. Statistical parametric analysis for feature decomposition and selection are also explained as well. These statistical parameters are then subdivided into three groups to examine the contribution of each parameter to the selected feature extraction methods. However, MSSA and WPD have the highest accuracy of 99% and 99.9%, with the inclusion of higher-order statistical (HOS) features. Finally, features from each group were fed into a convolutional neural network (CNN) based classifier to classify the power quality disturbances. This study compares the selected techniques for optimal features with and without HOS and highlights the fundamental properties of each method. The proposed method was found to have reliable high classification accuracy under noisy and noiseless conditions.

Keywords: Power quality disturbances, Higher order statistics, Convolutional neural networks.

1. Introduction

Power Quality (PQ) has become a key topic for research due to its significance in the generation and distribution of electric power systems [1], [2]. The PQ abnormalities detection and monitoring capabilities of the system are vital to determine how much electric equipment is affected. On-time action can save the electric system from severe damages [3], [4]. In the modern electric power system, the introduction of the renewable energy source and the increasing adaptation of distributive generation in distribution networks has valued the importance of PQ monitoring and analysis [5], [6]. The state of the fast art equipment takes place the traditional process of visually identifying the abnormalities [7]. The PQ events include sag, swell, harmonics, transients, interruption, flicker, etc. , all of which are present in the nature of PQ events., separated the typical issues in the distributive networks [8]. These abnormalities are presented in the system known as PQ disturbances in the form of single and mostly combinations [9]. PQ analysis generally consists of feature extraction of PQ disturbances, feature selection, and classification.

Essential signal processing techniques that are used for the feature extractions include Discrete Fourier transforms (DFT), Short Time Fourier transforms (STFT), Fast Fourier transforms (FFT), and Fourier transforms (FT) [10], [11], [1- 3]. PQ disturbances are generally non-Gaussian and non-stationary in nature. FT is employed to analyze the stationary signal, and STFT is the extension of DFT but it cannot process the non-stationary signal since it has a fixed window size [10]. In general, FT families are not suitable for handling non-stationary signals. S-transform (ST) is the combination of STFT and wavelet transform (WT), and it is one of the most appropriate tools for non-stationary signal processing. ST is found to be superior because of the excellent ability to localize the signals and modulation of sinusoidal is fixed concerning time. However, the redundant representation of the time-frequency domain and high computational time as compared to WT.

Wavelet transform (WT) is a signal processing technique that provides the analysis of signals with the help of the variable size windows and long windows are used for low frequency bands and short windows are used for high frequency bands. This provides good time-frequency resolution and can be used to identify and classify high and low frequency issues in PQD. However, selecting the appropriate mother wavelet and sampling frequency for WT can be challenging [12]. Wavelet packet decomposition (WPD) is a signal processing technique that is used to decompose a signal into its constituent wavelets at a specific level. It is a variation of wavelet decomposition that allows for more flexibility in analyzing signals by breaking them down into both their frequency and time domains. WPD is often preferred over the discrete wavelet transform (DWT) because it allows for a fixed frequency band. [13].

EMD is also a recently developed signal processing technique and is successfully applied in many engineering applications [14]. EMD is one of the adaptive and suitable decomposition methods for analysing the non-stationary and non-linear PQ

disturbances. It decomposes the time-series signal into another time series component called intrinsic mode functions (IMFs) [15]. Despite the better decomposition results, EMD is sometimes unable to decompose the signal due to the mode mixing leading to the false IMF decomposition. To solve this issue, EEMD is proposed [16]. In the EEMD method, random noise is added to the signal, decomposed using the EMD technique, and the process repeats with different noise series unless it obtains the realistic IMFs. This process may increase the computational burden. It is the severe drawback of EEMD, where data volume is enormous, especially in the case of PQ disturbances [17].
In this study, the effectiveness of four different techniques for the extraction of significant features from PQ events was investigated. The optimal selection of the features from the sub-bands of WPD, DWT, EMD, and MSSA are determined by using six distinct statistical features, including higher-order statistics (HOS) (entropy, energy, standard deviation, skewness, mean, and kurtosis) [18]. The application of a classifier based on a convolutional neural network (CNN) is used for the classification of PQ disturbances. [19]. The statistical features are divided into three groups for detailed analysis of feature extraction techniques, 1), four types of features excluded HOS features. 2), first two, and including HOS features. 3), all six types of components. The performance of these feature extraction techniques has been examined on the IEEE workgroup 1159.2019 and synthetic data generated from MATLAB R2020b.

This study has the following contribution.
1) This research presents the comparative study of feature extraction techniques and explain the pros and cons of each method.
2) A modified architecture of 1D-CNN is employed for the classification. Batch normalization is utilized instead of dropout layer to sort out the overfitting problem.
3) To validate the proposed methods. A comprehensive study on classification accuracy, computational complexity and comparative analysis is presented.
4) Simulated data is generated from the MATLAB and real time data is obtained from the IEEE 1159.2019. The complex or multiple PQ disturbances are also considered in this paper.

2. Feature Extraction

2.1 Discrete Wavelet Transform

PQD are sudden changes in an electrical signal's voltage, current, and magnitude. These disturbances are non- stationary, meaning they are not constant over time [20]. The Discrete Wavelet Transform (DWT) is a signal processing tool that is effective in analyzing and simplifying the complexities of disturbances encountered during computations [21]. It achieves this by decomposing a discrete-time signal x[n] into various levels of wavelet coefficients that carry both time and frequency information. Selecting an appropriate mother wavelet is a crucial consideration in the DWT. The DWT algorithm begins by applying Low Pass (LP) and High Pass (HP) filters to the original level of the discrete time signal, followed by down sampling by a factor of 2. The HP filters extract the high-frequency components of the signal, known as detail coefficients D1, while the LP filters extract the low-frequency components, known as approximation coefficients $A1$. The wavelet transform can be represented in terms of scale $\varphi(t)$and wavelet functions $\omega(t)$ for both the approximation l and detail coefficients h of the signal $x[n]$.

$$\varphi(t) = \sqrt{2}\sum_{n} h(n)\varphi(2t - n) \quad (1)$$

$$\omega(t) = \sqrt{2}\sum_{n} l(n)\omega(2t - n) \quad (2)$$

Secondly, at the next level, the value of A_g is change to $x[n]$ and the value of g_m is raised by one. The process described above will continue until g reached the maximum number of levels that can be selected, as depicted in Figure 1.

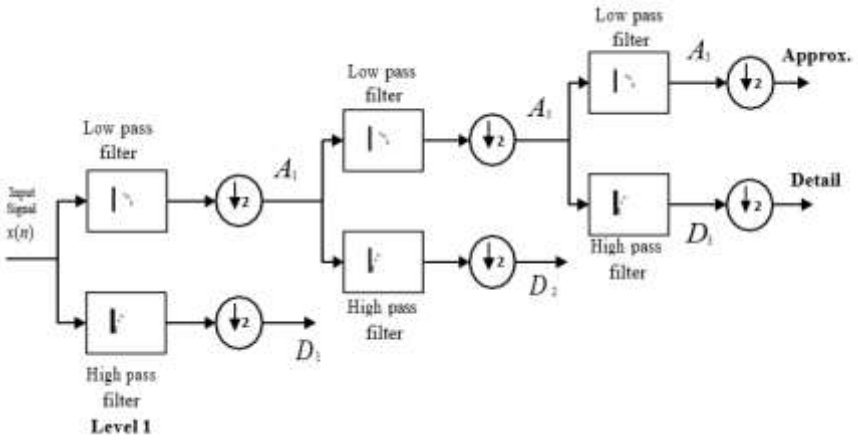


Figure 1: Decomposition of DWT for level two

2.2 Wavelet Packet Decomposition

The Wavelet Packet Decomposition (WPD) is a development of the Discrete Wavelet Transform (DWT). It decomposes the detail coefficients, D_g Moreover, approximation coefficients of the signal in the same way that DWT decomposes. The main difference between the two is that WPD increases the number of wavelet coefficients by a factor of 2^g whereas they are increased by $g + 1$ in DWT. It allows WPD to attain better frequency resolution when decomposing the signal [22]. However, DWT may skip important information in the high frequency components of the signal, as depicted in Figure 2

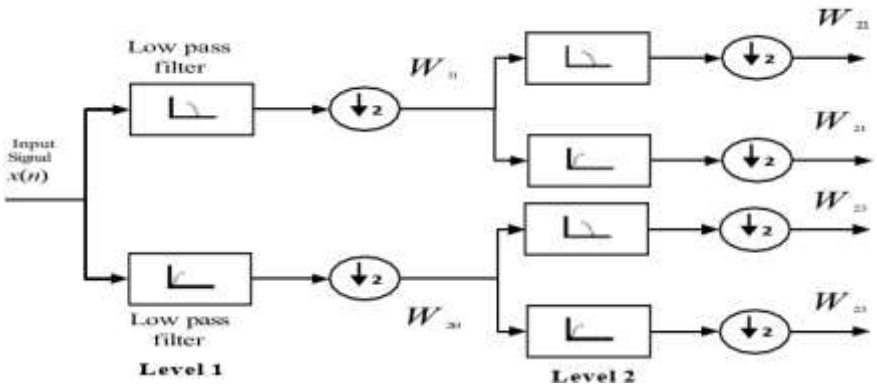


Figure 2: Decomposition of WPD for level two

2.3 Multiclass Singular Spectral Analysis

An alternative approach to decompose non-stationary PQ disturbance signals into features is the multiclass singular spectral analysis (MSSA) method. MSSA is used for decomposing non-stationary signals of PQ disturbances into features. It involves four steps: Embedding, Singular Value Decomposition (SVD), Grouping, and Reconstruction of the signal. The signal is embedded into a higher-dimensional space using a sliding window approach, followed by SVD decomposition to obtain the singular vectors and singular values. The singular vectors are then grouped based on their similarity, and the features are extracted to reconstruct the original signal [23].

Assuming that the PQ disturbance signal is a uniformly sampled one-dimensional signal stored in a vector array, A is defined as $A = [a_1, a_2, \ldots, a_N]^T \in \Re^N$ of length N where N is $b - \frac{1}{2}W \leq i \leq b + \frac{1}{2}W.$ where b is the position and window size W ($1 < W < N$). The trajectory matrix C of the matrix A can be formulated as

$$C = \begin{pmatrix} a_1 & a_2 & \cdots & a_K \\ a_2 & a_3 & \cdots & a_{K+1} \\ \vdots & \vdots & \ddots & \vdots \\ a_L & a_{L+1} & \cdots & a_N \end{pmatrix} \quad (3)$$

$$C = (c_1, c_2, \ldots, c_K) \quad (4)$$

Where each column of matrix C is mapped to a lagged vector K, which is $c_K = [a_k, a_{k+1}, \ldots, a_{k-L+1}]^T \in \Re^L$, where $k \in [1, K]$ and $K = 1 - W + N$. The covariance matrix stores complete information about the input signals.

Trajectory matrix, $S = CC^T$, is used to obtain the lagged covariance matrix, S. Its eigenvalues are calculated and sorted in descending order as, ($\lambda_1 \geq \lambda_2 \geq \cdots \lambda_L \geq 0$), and equivalent eigenvectors are ($U_1, U_2, \ldots, U_L$), and the trajectory matrix after SVD is presented as

$$C = C_1 + C_2 +, \ldots, C_d$$

Where rank of the matrix is $d \leq W$. For simplification, it is considered as $d = W$. As noted, the trajectory matrix, C, is a composite of multiple matrices. Each matrix $X_l | l \in [1, W]$ is referred as an elementary matrix and since it is equivalent to its respective eigenvalue, the elementary matrix defines the decomposition of the signal is defined by

$$X_l = \sqrt{\lambda_l} u_l^b \left(v_l^b\right)^T \quad (5)$$

Where v_l^b is define as

$$v_l^b = \frac{X^T u_l^b}{\sqrt{\lambda_l}} \quad (6)$$

The matrix of empirical orthogonal functions is referred to as the U matrix, and the matrix of principal components is referred to as the V matrix. Both matrices are defined relative to a particular position of b.

$$\boldsymbol{U} = \left(\boldsymbol{u_1^b}, \boldsymbol{u_2^b}, \ldots, \boldsymbol{u_L^b}\right) \in \boldsymbol{\Re^{L\times L}} \quad (7)$$

$$\boldsymbol{V} = \left(\boldsymbol{v_1^b}, \boldsymbol{v_2^b}, \ldots, \boldsymbol{v_L^b}\right) \in \boldsymbol{\Re^{L\times L}} \quad (8)$$

2.4 Empirical Mode Decomposition

EMD is another appropriate method to decompose the non-stationary and non-linear signals. Like other decomposition methods, EMD does not deteriorate the signals into sets of coefficients but decomposes the time signals into another kind of time signal known as intrinsic mode functions (IMFs) [14]. These are the following steps to demonstrate the IMFs of a signal $x(t)$.

1) Begin by setting $h(t) = x(t)$.
2) Use spline interpolation to find the upper and lower envelope by connecting the local maxima and minima. Calculate the mean $m(t)$ of the upper and lower envelope.
3) Subtract the $m(t)$ from $h(t)$ to obtain $h_O(t)$.
4) Check if $h_O(t)$ satisfies the conditions for an IMF. If yes, consider $h_O(t)$ as the first IMF, if not replace $h(t)$ with $h_O(t)$.
5) Repeat the above steps k times until an IMF component is isolated from the data.

Furthermore, Subtract the IMF from the original signal to obtain the remainder. Use the reminder to extract another IMF, until the note becomes a monotonic function. In the end, the original signal is presented as

$$x(t) = \sum_{i=1}^{n} d_i + r_n \quad (9)$$

Where r_n is the remainder of the signal $x(t)$, d_i is the ith IMF extracted from $x(t)$.

2.5 Feature Extraction using Statistical parameters

For accurate classification, the optimal feature extraction is a critical stage [24]. The details of statistical parameters can be seen in the literature [25]. DWT, WPD, and MSSA decomposition techniques represent the PQ disturbances coefficients datasets, while EMD generates IMFs of equal length as the original input signal. The selection of features from statistical parameters significantly reduces the sets of features and characterize the behavior of the PQ disturbances signal. The selected statistical parameters were estimated for the classification of PQ disturbances. Equations for each features are stated in Table 1.

If $X\{x_1, x_2, \ldots, x_N\}$ are the sub-bands (or IMF), where N is the length or number of samples in sub-band or decomposition. i = 1,2,3, … , k , are the number of decompositions at level k. However, the higher-order statistics (HOS) features such as skewness and kurtosis are considered and three manually selected combinations of these features are examined in this study. Each combination of different feature set belongs to the separate experiment. Moreover, the detailed experiments are described in subsection 2.4.

Table 1: Equations of the features for each of the sub bands.

Sr. No	Features Name	Equations
1.	Energy	$E_{ki} = \sum_{j=1}^{N} \left(\lvert X_{ij}\rvert^2\right)$
2.	Entropy	$ET_{ki} = -\sum_{j=1}^{N} {X_{ij}}^2 log\left({X_{ij}}^2\right)$
3.	Standard Deviation	$\sigma_{ki} = \left(\frac{1}{N}\sum_{j=1}^{N}\left(X_{ij} - \mu_i\right)^2\right)^{\frac{1}{2}}$
4.	Mean	$M_{ki} = \frac{1}{N}\sum_{j=1}^{N} X_{ij}$
5.	Kurtosis	$KT_{ki} = \sqrt{\frac{N}{24}}\left(\frac{1}{N}\sum_{j=1}^{N}\left(\frac{X_{ij}-\mu_i}{\sigma_i}\right)^4 - 3\right)$
6.	Skewness	$SN_{ki} = \sqrt{\frac{1}{6N}}\sum_{j=1}^{N}\left(\frac{X_{ij}-\mu_i}{\sigma_i}\right)^3$

3 1-D Convolutional Neural Network (CNN)

A novel approach for classification of PQ disturbances using CNN based soft-max classifier has been presented. In general, CNN is applied for feature extraction and classification, However, it has recently shown better results in many fields like object detection, image classification, face recognition, and vision tasks [26], [27], [28]. CNN provides improved classification accuracy in vision application [29], as compared to the deep neural network (DNN) and commonly used support vector machine (SVM) [27].

3.1 Proposed architecture of 1D-CNN:

The proposed architecture of 1-D CNN can be seen in Figure 3. In generally, three layers are stacked to extract the features which consist of convolution, pooling and batch normalization layers. However, in this paper, the external feature extraction techniques are considered and the convolution layer is not utilized. The fully connected and softmax layers are used for the classification. The pooling layer along with the extracted feature matrix is inserted to reduce the dimensions and prominent the features of power disturbances. The dropouts or batch normalization along with the fully connected layer can be adopted to reduce the overfitting and enhance the ability of generalization. In this research, batch normalization provides better training speed and more efficient to cater the overfitting problem. The PQDs are 1-D signal and completed different form the images. The modified architecture use the1D arrays instead of the 2D matrix.

The architecture of the 1-DCNN model that are utilized in this paper is consists of pooling with batch normalization layer, and fully connected soft-max layers, as shown in Figure 3. In this research, the 1-DCNN model consists of six max- pool layers and three dense or fully connected layers.

1) Pooling Layer

The primary role of the pooling layer is to both map the features and decrease the dimensionality of the data. While average pooling is more prone to being affected by noise, max pooling generally performs better. Both average pooling and max pooling return either the average or maximum value to the activation function.

$$z_n^l = f_n\left(w_n^l\right) \quad (9)$$

where z_n^l is the output and f_n is the non-linear function. The rectified linear units (ReLU) is deployed as activation function in this research. It perform better as compared to other activation function such as tanh. ReLU restricts all output neuron to become active and accept only the positive value and convert the negative into null value. It also help to increase the efficiency and reduce the computational burden. The soft-max layer is expressed as

$$\delta(z)_n = \frac{e^{z_n}}{\Sigma_j^J e^{z_j}}, \; n = 1,\ldots,J \quad (10)$$

Where, z_n represents the input received from a fully connected layer, while J represents the number of classes or the number of soft-max layer units.

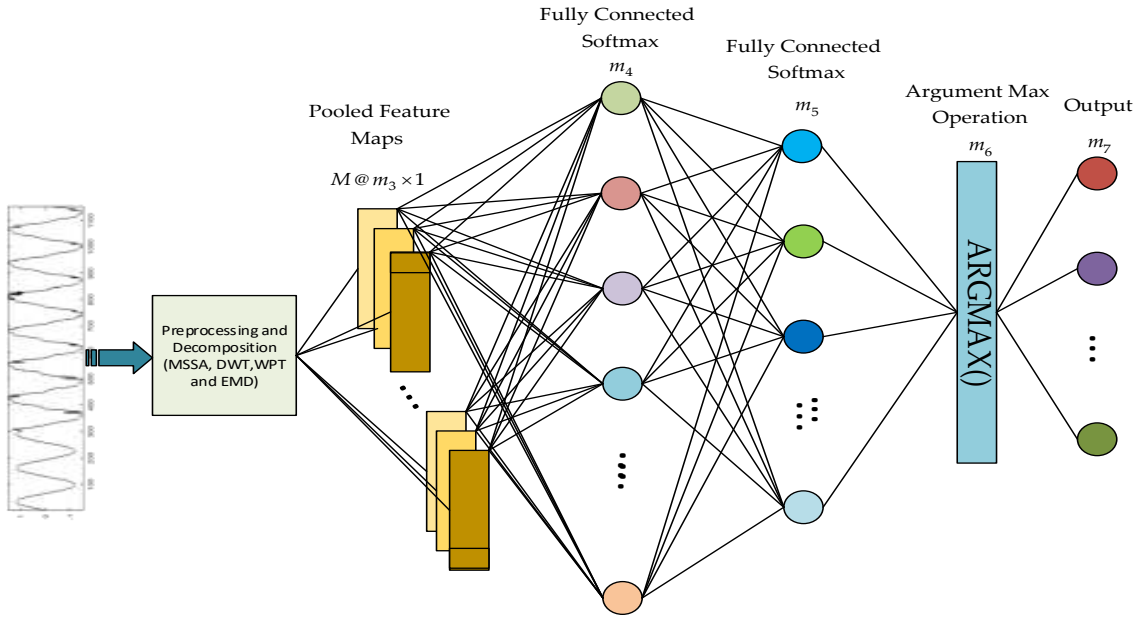


Figure 3: The complete structure of 1-D Convolutional neural network

2) Batch-Normalization

In generally, dropout layer is utilized to overcome the overfitting problem due to complex nature of the data. However, Batch-normalization (BN) is more advance and feasible method to overcome this problem. It use the normalization operator to solve the gradient disappearing problem. The max-pooling layer with BN is utilized to detect errors. The mean square error of the output can be computed as

$$E_q = \sum_{l=1}^{N_c} \left(z_l^I - t_l^q \right)^2 \quad (11)$$

Where "l" represents the number of classes, "q" is the input vector, "t_l^q" is the corresponding target, and "$\left[z_1^I, \ldots, z_q^I\right]$" is the output vector.

3) Dense and Softmax layer:

Fully connected layer or dense layer has the lth fully connected layer and softmax is the activation function and connected with the dense layer. The softmax function predict the value of given class of each category. The category with highest probability will be the output.

This research separates the DWT, WPD, MSSA, and EMD with 1DCNN based algorithms into three parts. 1) Feature extraction, 2) Statistical parameters, and 3) classification. This section evaluated four feature extraction methods (DWT, WPD, EMD, and MSSA) to choose the best feature extraction technique. Four level decomposition is used for all four methods. The statistical parameters are introduced to examine the change in the results of feature extraction methods. Table 2 described the detailed information about the features chosen for different experiments. The optimal features from different experiments are fed to the IDCNN classifier to examine the capability of these feature extraction methods for the classification of PQ disturbances. The detail of the DWT, WPD, MSSA, and EMD with IDCNN based method is shown in Figure 4.

Figure 4: Block diagram of the DWT, WPD, MSSA, and EMD with IDCNN based scheme

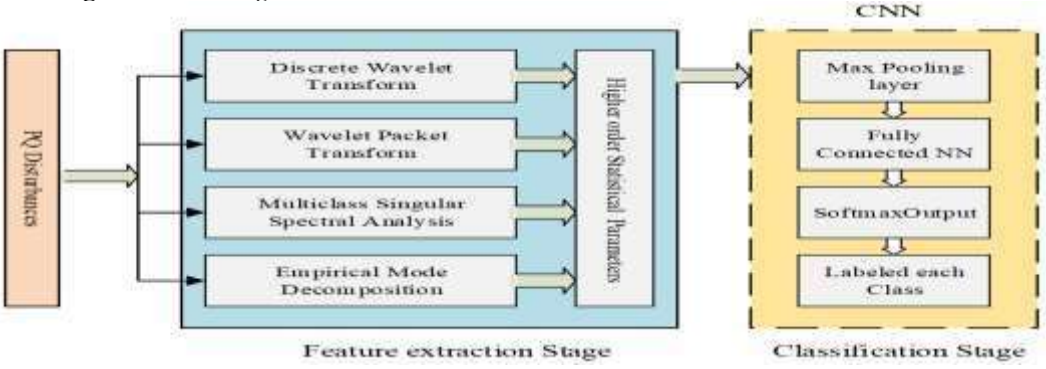


Table 2: Detail of the number of features in each decomposition technique

Decomposition Method	No. of Sub-bands	No. of Features	
		Exp 1-2	Exp 3
DWT	5	20	30
WPD	16	64	96
EMD	5	20	30
MSSA	5	20	30

Table 3: Four level frequency bands of WPD, DWT, and MSSA

4 Experiment

4.1 Generation of Dataset

A dataset of single and multiple PQDs was generated in PSCAD/EMTDC power simulation software for distribution network. In this study, 12 types of single and numerous PQDs were considered. The line faults such as a line-to-line double and single line to ground faults generated the sag, interruption, and swell PQDs. The uncontrolled speed of motor and capacitor switching generated oscillatory, impulsive, and flicker types of PQDs. Other causes of PQDs were nonlinear load, power electronics devices, tripping of the circuit breaker, and overheating of the neutral conductor. The synthetic PQ disturbances were generated from the parametric model and were simulated in MATLAB R2020b. Two thousand and four hundred waveforms were generated for 12 types of PQDs. Two hundred samples, ten cycles, fundamental frequency 50 Hz, the sampling frequency of 10 kHz for each PQ disturbance was considered. Some of the real waveforms are shown in Figure 5. The parametric equation model is shown in Table A1. The noise levels introduced in this experiment are 0dB, 20dB, and 50dB, respectively. The real PQD waveform has been taken from the IEEE power quality workgroup 1159.3-2019. The synthetic PQ disturbance waveforms are shown in Figure 6.

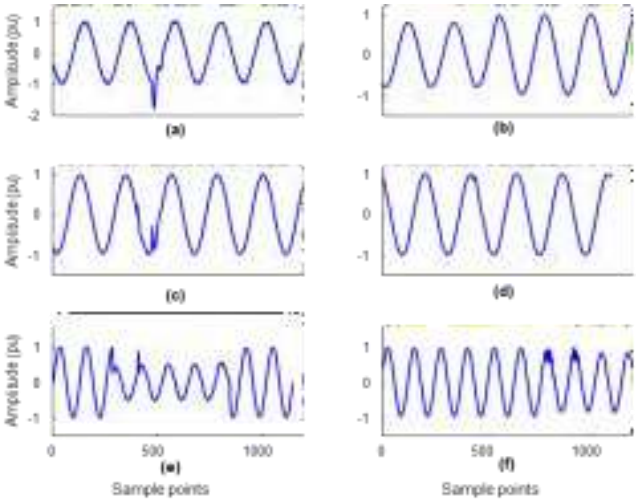


Figure 5: Real time PQ waveforms of different disturbances such as sag, flickers, notch etc.

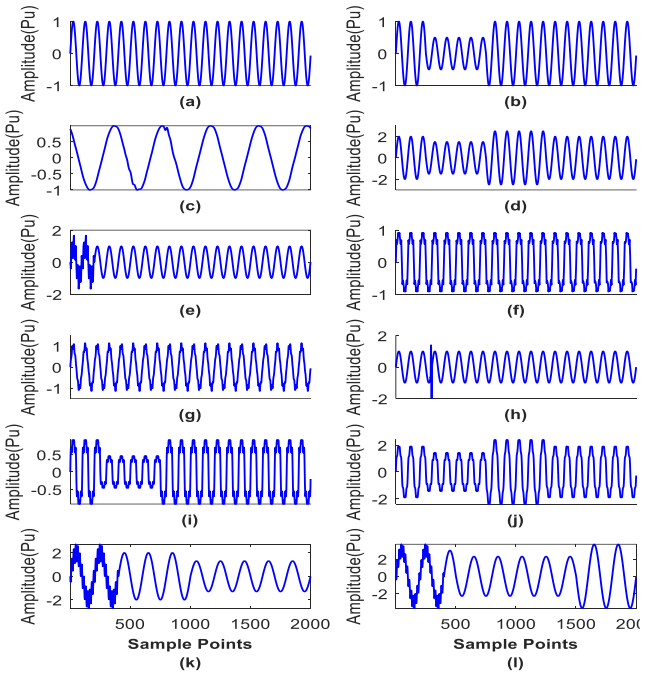


Figure 6: Simulated PQ disturbances waveforms using Matlab such as , (**a**) Normal; (**b**) Sag; (**c**) Notch; (**d**) Sag and swell; (**e**) Oscillatory transient; (**f**) Harmonics; (**g**) Flicker; (**h**) Impulsive transient; (**i**) Sag and harmonics; (**j**) Sag, swell and harmonics; (**k**) Sag and oscillatory transients; (**l**) Sag, swell and oscillatory ransient

4.2 Feature Selection and the Parameters

In this study, four different methods of feature extraction are evaluated and compared. The wavelet technique which is employed is based on the multiresolution analysis. The analysis is done on four number of wavelet decomposition levels. The frequency range for PQ disturbances is considered as 50 Hz. The details of the frequency range for the four-decomposition level and the approximation signal of WPD, MSSA, and DWT are presented in Table 3. Sys4 is used for decomposition and approximation signals, which produced superior results. The other wavelet functions also performed very well. Selecting suitable wavelet functions for a specific application depends upon the trial-and-error approach.

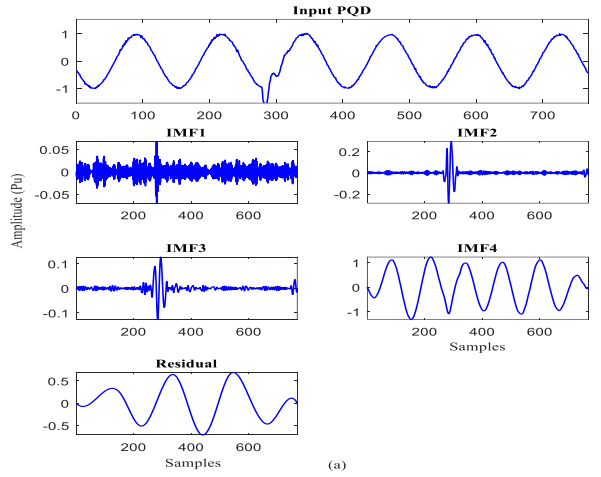
After normalizing the PQ disturbances, the DWT, EMD, MSSA, and WPD are applied to decompose the PQ disturbance signals into the sub-bands. MSSA also decomposed the primary signals into four levels and approximation signals. EMD disintegrated the signals into four IMFs and residual signals. MSSA decomposed the PQDs into four groups, which provided four detailed coefficients ($D_1, D_2, \dots, D_4$)and an approximation coefficient A_4.

However, the WPD decomposition technique utilized four decomposition levels, which generated 16 sub-bands due to $2^g = 2^4$. Table 3 provides information on the frequencies contained in each sub-band of WPD, and their corresponding relationships with DWT sub-bands. The number of features used in each experiment and decomposition method are shown in Table 2. WPD generates eleven additional sub-bands compared to the other decomposition techniques, which results in a difference in the number of features. Figure 7 illustrates the decomposition of PQD signals using DWT, EMD, WPD, and MSSA.

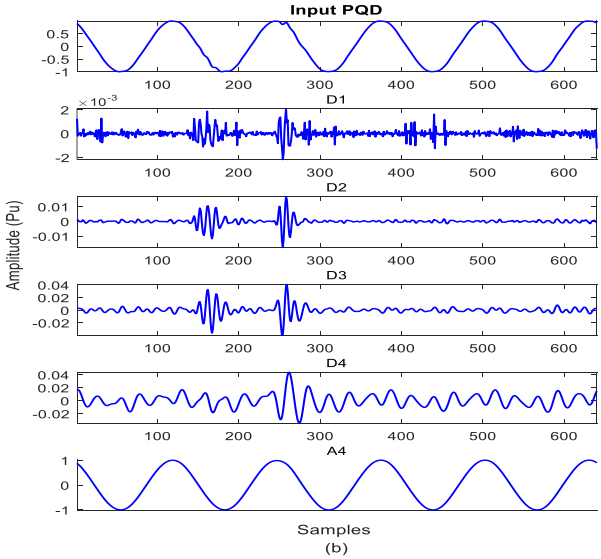
Table 3: Four level frequency bands of WPD, DWT, and MSSA

Sub-band No.	Decomposition Signal	Frequencies (Hz)	DWT/ WPD /MSSA level
1	SB40	0-4.2	APX.4
2	SB 41	4.2-5.9	DTCo.4
3	SB 42	6.0-9.0	DTCo.3
4	SB 43	9.1-13.0	
5	SB 44	13.1-16.0	DTCo .2
6	SB 45	16.1-19	
7	SB 46	19.1-22	
8	SB 47	22.1-25.9	
9	SB 48	26.0-28.9	DTCo .1
10	SB 49	29.0-32.1	
11	SB 4A	32.2-35.1	
12	SB 4B	35.2-38.1	
13	SB 4C	38.2-41.0	
14	SB 4D	41.1-44.0	
15	SB 4E	44.1-47.0	
16	SB 4F	47.1-50.0	

Experiment 1 (Features without HOS), Experiment 2 (a combination of the first two features and HOS), and Experiment 3 (All selected features) are described in section 2.5. These algorithms are tested more than 50 times to confirm the results. In this paper, the data set is separated into training and testing sets for 70-30% respectively.



(a)



(b)

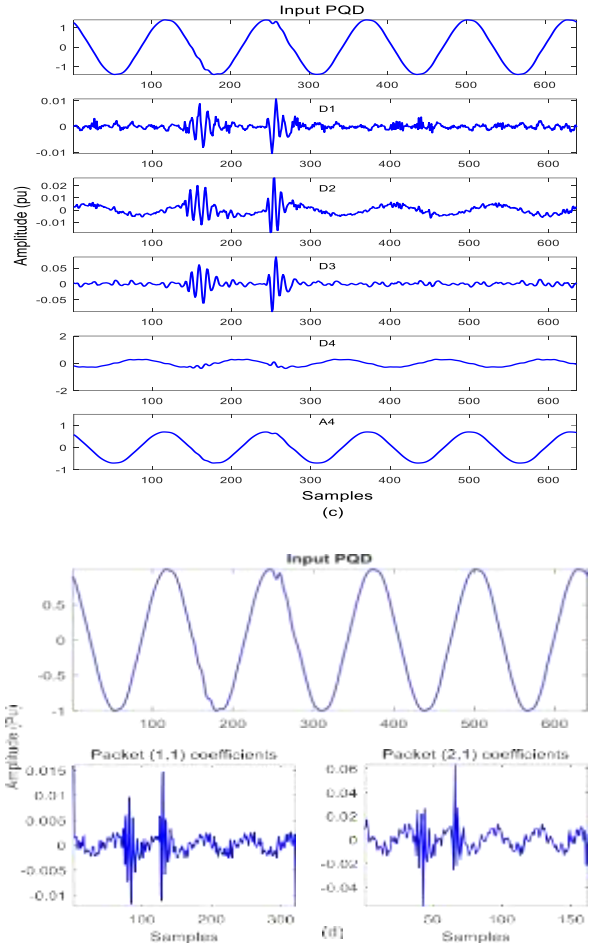


Figure 7: PQ disturbances decomposition (a) EMD four level decomposition (IMF1-IMF4) (b) DWT four level decomposition (D1-D4) (c) MSSA four level decomposition (D1-D4) (d) WPD decomposition coefficient (1,1) and (2,1)

5. Results and Discussion

In this study, synthetic datasets are produced using MATLAB R2020b, and real datasets are taken from the IEEE 1159.2019 task force. In this section, the four feature extraction techniques are separately evaluated, and compared to achieve the best optimal classification accuracy. The classification accuracy can be evaluated from the following equation [30].

$$1\text{DCNN\_accuracy} = \frac{correctely\ classified}{total\ number\ of\ classes} \times 100 \qquad (12)$$

Tables 4-6 show the overall classification accuracy using the comparative decomposition techniques for 12 types of multiple PQ disturbances. The outcomes of all feature combinations (experiment 1-3) for synthetic and real data are presented. The combination of statistical features in experiment 2 outperformed the experiment 1 and 3. However, experiment 1 produced the worst classification accuracy among them. On the other side, EMD has the worst classification accuracy. In contrast, WPD has improved classification accuracy among MSSA, DWT, but MSSA has less computational complexity than other decomposition techniques.

Table 3 provides a detailed explanation of the frequency ranges of decomposition signals of WPD, DWT and MSSA. It can be seen that the most frequently occurring sub-bands are those of MSSA and DWT, specifically in the detail DC1 frequency range of 25 to 28.1 Hz. However, the WPD method produces more features in the frequency range of 25-50 Hz, providing a deeper understanding of the differences between the classes. WPD produces a more finely detailed decomposition of the highest frequencies using lower scale levels than DWT and MSSA.

WPD generates a detailed range of features for the same decomposition levels. In contrast, other decomposition methods require a more significant number of decomposition levels to achieve a similar number of distinctive features. The IMFs generated from PQ disturbance signals do not keep the maximum frequencies, and most of the information is lost during the decomposition. On the other side, the detailed DC1 of MSSA and DWT hold most of the lost information in the EMD.

The classification accuracy may be affected due to the lost information. The use of the Ensemble EMD (EEMD) method solved this problem. EEMD produced the same number of IMFs for the same signal length. However, IMF2-IMF4 holds the lower frequency the same as A4 and D2-D4 in MSSA and DWT.

Effective feature extraction of PQDs in MSSA is influenced by two key parameters. The first parameter is "α," which

determines the number of eigenvalues to be extracted during the decomposition process. The second parameter, known as the eigenvalue group (EVG), determines how the extracted components are grouped for optimal feature extraction. When all eigenvalues are included in the EVG, it may result in the removal of all features in the PQD signal. On the other hand, excluding small eigenvalues from the EVG results in more effective feature extraction, as small eigenvalues tend to contain noise. After multiple simulations, it was found that the optimal value of "α" for PQD decomposition using MSSA is 3.

Another essential comparison in this study is related to three sets of features: Experiment 1, Experiment 2, and Experiment 3 are presented in Tables 4-6. In experiment 1, the classification rate was lowest, and 1-4 features were used. Including higher-order statistical (HOS) features 5 and 6 (skewness and kurtosis) in experiments 2 and 3 may be explained with superior accuracy. Generally, PQ disturbances are non-Gaussian, non-stationary, and nonlinear. However, First and second-order statistics have a significant impact on PQ disturbances signal processing, and somehow very restrictive in analysing the non-stationary, and non-linearity. The superior resolution and dynamic ability to drive the feature of the signal using different kinds of WTs and MSSA decompositions are appropriate for analysing PQ disturbance. The powerful wavelets and MSSA may fail when extracting non-linear behavior within the signal. The inclusion of HOS improved the feature extraction capabilities of decomposition methods and can be seen in Tables 2.4-2.5. HOS cures the abnormalities such as nonlinearity, nonstationary within the signal, which makes HOS and time/frequency methods superior in the analysis of PQ disturbances.

Figure 8 shows how DWT, EMD, MSSA, and WPD perform with different noise levels and real data. The overall classification accuracy decreases with the increase of noise level, which can be observed for all decomposition techniques and experiments. Compared with the DWT and EMD decomposition technique, the decrease in classification accuracy with the increase in noise level is lesser for MSSA and WPD using CNN based classifiers for all experiments. The classification accuracies of WPD and MSSA are very close, but WPD is leading among them. CNN based classifier using WPD has better classification accuracies than DWT, EMD, and MSSA, even at the high noise level of 20 dB. This method is more robust to the high level of noise.

Table 4: Classification accuracy (%) for experiment 1

			DWT		**EMD**		**MSSA**		**WPD**		
Power Quality Disturbances	**Class Labelled**	**Testing /Training sets**	dB								**Real data**
			0	**50**	**0**	**50**	**0**	**50**	**0**	**50**	
Normal	C1	200	96	96	95	94	97	97	98	98	97.7
Sag	C2	200	95	94	94	93	97	96	97	96	95.6
Notch	C3	200	95	94	94	94	96	96	96	95	94.8
Flickers	C4	200	94	93	94	93	95	94	97	96	96
Impulsive Transients	C5	200	96	96	92	91	96	95	96	96	95.5
Oscillatory Transients	C6	200	95	94	94	93	95	95	97	95	94.8
Harmonics	C7	200	95	95	94	94	97	96	97	96	96
Sag with Swell	C8	200	96	95	93	92	96	94	97	96	95.6
Sag with Harmonics	C9	200	97	96	95	94	97	96	98	98	97.9
Harmonics with Sag and Swell	C10	200	98	96	96	95	98	98	98	97	97
Sag with Oscillatory Transients	C11	200	97	97	95	93	98	97	98	97	97
Oscillatory Transients with Swell, and Sag	C12	200	96	96	95	95	97	97	97	96	96
Classification Accuracy (%)			95.8	95.2	94.3	93.4	96.6	95.9	97.2	96.3	96.2

Table 5: Classification accuracy (%) of PQ disturbances for experiment 2

Class Labelled	**Training/Testing sets**	**DWT**		**EMD**		**MSSA**		**WPD**		**Real data**
		dB								
		0	**50**	**0**	**50**	**0**	**50**	**0**	**50**	
C1	200	99	99	98	97	100	99	100	100	100
C2	200	98	96	96	95	100	99	100	99	99
C3	200	97	96	95	94	99	98	100	99	99
C4	200	98	98	96	95	98	97	100	99	98.8
C5	200	98	97	94	93	99	99	100	99	99
C6	200	97	97	95	95	99	98	99	99	98.7
C7	200	98	96	96	96	99	99	99.5	99.1	99
C8	200	98	97	95	94	99	98	100	99.5	99
C9	200	99	98	97	96	98	98	99.7	98.2	97.9
C10	200	98	98	97	97	99	98	99.6	99.2	99
C11	200	99	98	97	96	99	99	99.8	99.3	98.9
C12	200	98	98	96	96	99	98	99.7	99.3	99
Classification Accuracy (%)		98.1	97.3	96	95.3	99	98.3	99.8	99.1	98.9

Table 6: Classification accuracy (%) of PQ disturbances for experiment 3

	Training/	DWT		EMD		MSSA		WPD		
Class Labelled	Testing sets	dB								Real data
		0	50	0	50	0	50	0	50	
C1	200	99	99	97	97	100	99	100	100	99
C2	200	98	97	95	94	100	99	100	99	99
C3	200	96	95	93	93	98	98	99	99	98.8
C4	200	97	97	95	95	98	97	99	98	97.7
C5	200	98	96	92	92	98	97	99	98.2	97.9
C6	200	97	97	93	93	98	98	98	99	98.5
C7	200	98	96	95	95	99	98	99	98.5	98
C8	200	98	97	95	94	99	98	99	99	98.3
C9	200	98	97	96	96	97	97	98.8	98	98.5
C10	200	97	97	97	97	99	98	99	98.3	98
C11	200	98	98	97	96	98	98	99	98.3	98
C12	200	98	97	96	96	99	98	99	98.7	98.2
Classification Accuracy (%)		97.6	96.9	95.1	94.8	98.6	97.9	99	98.7	98.3

Figure 8: Classification performance of all decomposition technique for different noise levels

5.1 Computation Complexity

Figure 9 shows the computational complexity of all decomposition methods for each experiment. The experiment environment comprises MATLAB (R2020b), Intel core i7-6700 CPU @ 3.40 GHz, and 8 GB of RAM.

The increase in the number of features can lead to a rise in computational complexity, and the impact of higher order statistics (HOS) is apparent in experiment 2. For all decomposition techniques, the computational complexity is the lowest in comparison to other experiments. However, WPD has a greater level of complexity due to its intricate frequency sub- bands analysis. In contrast, MSSA has lower complexity due to the optimal value of "α" and the requirement of fewer features than WPD. Considering these factors, it can be concluded that MSSA is the most appropriate decomposition technique if computational speed is a crucial factor.

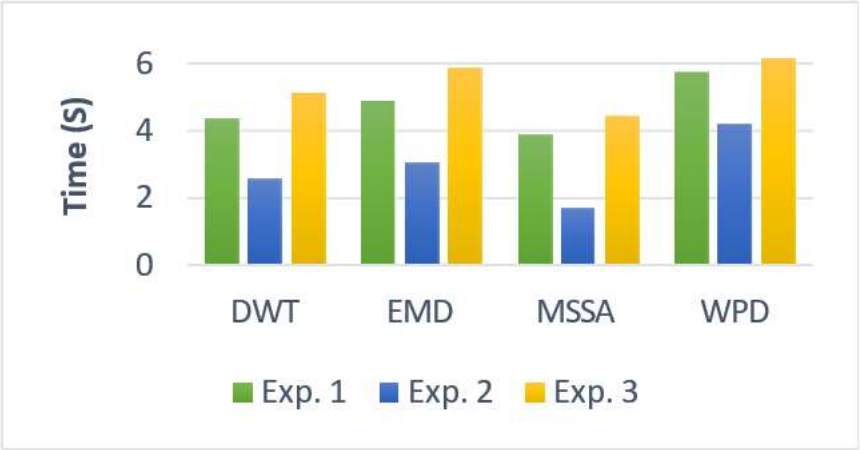


Figure 9: Comparison of computational time(s) between decomposition techniques using CNN based classifiers for noiseless data

5.2 Performance Comparison

Finally, a comparison is made between this research and some recently published articles. These published articles adopted multiple decomposition techniques, as shown in Table 7. In [31], six decomposition techniques were presented for the classification of PQ disturbances, and 99.43% accuracy was achieved. In [32], two decomposition techniques multi-scale morphological gradient filter (MSMGF) and short-time modified Hilbert transform (STMHT) were compared to achieve better classification results. In [7], ordinary ST (OST) and Modified ST (MST) were compared for the classification of nineteen types of PQ disturbances, and the classification accuracy is 97.3%, which is less than the present study. In [33], authors compared second-generation WT (SGWT), maximum overlapping DWT (MODWT), and ST to determine the optimal method for classification, and it was found that ST achieved the highest classification rate of 97.39%. In [34], six decomposition methods were adopted with extreme learning machine (ELM) and Fraction Fourier Transform (FRFT) provided the best classification accuracy of 98.27%. This comparative study shows that the DWT, WPD, MSSA and EMD with IDCNN based study attained higher classification results.

Table 7: Performance Comparison with published articles

Ref.	Classifier	Data Type	Decomposition methods	No. of PQDs	Accuracy (%)
[31]	DT-RF	Simulated	WT, HT, STFT, ST, FT, TTT	16	99.43
[32]	DTB	Real and Simulated	MSMGF, STMHT	10	99.70
[7]	SVM	Simulated	MST, OST	19	97.30
[33]	RF	Real and Simulated	MODWT, SGWT, ST	10	97.39
[34]	ELM	Simulated	FRFT, ELM	10	98.27
-	HOS, CNN	Real and Simulated	DWT, EMD, WPD, MSSA	12	99.90

6. Conclusions

MATLAB and IEEE task force 1159.3-2019 has generated two types of datasets, i.e., synthetic, and real. Four feature extraction techniques have been compared in this study, and optimal features are used to classify the PQ disturbances. Twelve types of multiple PQ disturbances were considered with different noise levels. Six types of statistical parameters were manually rearranged to form three types of groups. Three groups with varying sets of the feature were investigated, and the impact of higher-order statistics (HOS) was analysed. The results showed that the inclusion of HOS significantly affects the classification rate. WPD has higher frequency resolution features than the DWT and MSSA features. The inclusion of HOS also improved the feature extraction capabilities of DWT, EMD, and MSSA. This comparative study is one of the fewer, which presented a detailed analysis of the decomposition methods combined with HOS for the classification of PQ disturbances. However, MSSA has been found to be less computationally intensive than other decomposition methods. Higher classification rates revealed that these methods are sensitive to noise. The DWT, WPD, MSSA, and EMD with IDCNN based model have significant potential to accomplish good classification accuracy. The increased classification accuracy serves as proof that this study has the potential for application in various areas, including but not limited to image analysis, facial recognition, fault detection, classification, recognition, ECG signal detection, and classification.

References

1. R. Moreno, N. Visairo, C. Núñez, and E. Rodríguez, "A novel algorithm for voltage transient detection and isolation for power quality monitoring," *Electr. Power Syst. Res.*, vol. 114, pp. 110–117, 2014.
2. S. Kasa, P. Ramanathan, S. Ramasamy, and D. P. Kothari, "Effective grid interfaced renewable sources with power quality improvement using dynamic active power filter," *Int. J. Electr. Power Energy Syst.*, vol. 82, pp. 150–160, 2016.
3. B. Milešević, I. Uglešić, and B. Filipović-Grčić, "Power quality analysis in electric traction system with three-phase induction motors," *Electr. Power Syst. Res.*, vol. 138, pp. 172–179, 2016.
4. A. F. Zobaa and S. H. E. A. Aleem, "A new approach for harmonic distortion minimization in power systems supplying

nonlinear loads," *IEEE Trans. Ind. Inform.*, vol. 10, no. 2, pp. 1401–1412, 2014.
5. M. H. Bollen and I. Y. Gu, *Signal processing of power quality disturbances*, vol. 30. John Wiley & Sons, 2006.
6. C. Muscas, "Power quality monitoring in modern electric distribution systems," *IEEE Instrum. Meas. Mag.*, vol. 13, no. 5, pp. 19–27, 2010.
7. H. Shamachurn, "Assessing the performance of a modified S-transform with probabilistic neural network, support vector machine and nearest neighbour classifiers for single and multiple power quality disturbances identification," *Neural Comput. Appl.*, vol. 31, no. 4, pp. 1041–1060, 2019.
8. P. Thakur and A. K. Singh, "Unbalance voltage sag fault-type characterization algorithm for recorded waveform,"
9. *IEEE Trans. Power Deliv.*, vol. 28, no. 2, pp. 1007–1014, 2013.
10. W. Zhao, F. Rusu, K. Wu, and P. Nugent, "Automatic identification and classification of Palomar Transient Factory astrophysical objects in GLADE," *Int. J. Comput. Sci. Eng.*, vol. 16, no. 4, pp. 337–349, 2018.
11. G. T. Heydt, P. S. Fjeld, C. C. Liu, D. Pierce, L. Tu, and G. Hensley, "Applications of the windowed FFT to electric power quality assessment," *IEEE Trans. Power Deliv.*, vol. 14, no. 4, pp. 1411–1416, 1999.
12. D. Griffin and J. Lim, "Signal estimation from modified short-time Fourier transform," *IEEE Trans. Acoust. Speech Signal Process.*, vol. 32, no. 2, pp. 236–243, 1984.
13. 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, 2008.
14. J. Wang and C. Wang, "A classification method of power quality disturbance based on wavelet packet decomposition," in *2004 IEEE Region 10 Conference TENCON 2004.*, 2004, vol. 100, pp. 244–247.
15. N. E. Huang *et al.*, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proc. R. Soc. Lond. Ser. Math. Phys. Eng. Sci.*, vol. 454, no. 1971, pp. 903–995, 1998.
16. Z. Lu, J. S. Smith, Q. H. Wu, and J. Fitch, "Empirical mode decomposition for power quality monitoring," in *2005 IEEE/PES Transmission & Distribution Conference & Exposition: Asia and Pacific*, 2005, pp. 1–5.
17. Z. Wu and N. E. Huang, "Ensemble empirical mode decomposition: a noise-assisted data analysis method," *Adv. Adapt. Data Anal.*, vol. 1, no. 01, pp. 1–41, 2009.
18. Y. Zhang and Z. Liu, "Application of EEMD in power quality disturbance detection," *Electr. Power Autom. Equip.*, vol. 31, no. 12, pp. 86–91, 2011.
19. J. M. Mendel, "Tutorial on higher-order statistics (spectra) in signal processing and system theory: Theoretical results and some applications," *Proc. IEEE*, vol. 79, no. 3, pp. 278–305, 1991.
20. 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, 2019.
21. 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, 2018.
22. C. E. Heil and D. F. Walnut, "Continuous and discrete wavelet transforms," *SIAM Rev.*, vol. 31, no. 4, pp. 628–666, 1989.
23. M. Unser and A. Aldroubi, "A review of wavelets in biomedical applications," *Proc. IEEE*, vol. 84, no. 4, pp. 626–638, 1996.
24. R. Vautard, P. Yiou, and M. Ghil, "Singular-spectrum analysis: A toolkit for short, noisy chaotic signals," *Phys. Nonlinear Phenom.*, vol. 58, no. 1–4, pp. 95–126, 1992.
25. H. Erişti, Ö. Yıldırım, B. Erişti, and Y. Demir, "Optimal feature selection for classification of the power quality events using wavelet transform and least squares support vector machines," *Int. J. Electr. Power Energy Syst.*, vol. 49, pp. 95–103, 2013.
26. C.-Y. Lee and Y.-X. Shen, "Optimal feature selection for power-quality disturbances classification," *IEEE Trans. Power Deliv.*, vol. 26, no. 4, pp. 2342–2351, 2011.
27. P. Sermanet and Y. LeCun, "Traffic sign recognition with multi-scale convolutional networks," in *the 2011 International Joint Conference on Neural Networks*, 2011, pp. 2809–2813.
28. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks,"
29. *Adv. Neural Inf. Process. Syst.*, vol. 25, 2012.
30. R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2014, pp. 580–587.
31. I. Sutskever and G. E. Hinton, "Deep, narrow sigmoid belief networks are universal approximators," *Neural Comput.*, vol. 20, no. 11, pp. 2629–2636, 2008.
32. R. Ahila, V. Sadasivam, and K. Manimala, "Particle swarm optimization-based feature selection and parameter optimization for power system disturbances classification," *Appl. Artif. Intell.*, vol. 26, no. 9, pp. 832–861, 2012.
33. S. Jamali, A. R. Farsa, and N. Ghaffarzadeh, "Identification of optimal features for fast and accurate classification of power quality disturbances," *Measurement*, vol. 116, pp. 565–574, 2018.
34. T. Chakravorti, R. K. Patnaik, and P. K. Dash, "Detection and classification of islanding and power quality disturbances in microgrid using hybrid signal processing and data mining techniques," *IET Signal Process.*, vol. 12, no. 1, pp. 82–94, 2018.
35. S. Upadhyaya, S. Mohanty, and C. N. Bhende, "Hybrid methods for fast detection and characterization of power quality disturbances," *J. Control Autom. Electr. Syst.*, vol. 26, no. 5, pp. 556–566, 2015.