Journal of Data Acquisition and Processing

ISSN: 1004-9037
https://sjcjycl.cn/
DOI: 10.5281/zenodo.7778197

COMPOSITE FEATURES AND ENSEMBLE LEARNING FOR THE DETECTION AND CLASSIFICATION OF PQDS

C. Krishna Reddy[1*], Dr. B. Mangu[2]
[1]Research Scholar, Dept., of Electrical Engineering, Osmania University, Hyderabad, India.
[2]Professor, Dept., of Electrical Engineering, Osmania University, Hyderabad, India.
Email: ckr.net@gmail.com

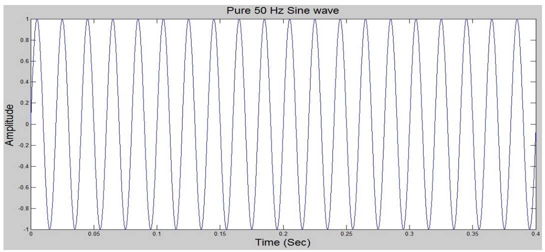
Abstract: Recently, the advances in the technology causes the power systems switch to smart grids, ever-increase of the distributed power generation, and expansion of microgrid causes power quality problems to the consumers. Power Quality Disturbances causes serious damages to the electrical appliances and hence their prior detection classification gained significant interest. Towards such prospect, this paper introduced a new method for PQDs detection and classification based on composite features and ensemble learning. At features extraction, each PQD is signified through two different set of features extracted from S-transform and Statistical methods. S-transform reveals time-frequency characteristics and statistical features reveals the feature independency between inter PQDs. Further at classification, two machine learning algorithms namely Support Vector Machine (SVM) and Kernalized Extreme Learning Machine (KELM) are employed. KELM explores the perfect discrimination between PQDs and is formulated as a combination of polynomial and Radial Basis Function Kernels. Extensive simulations of synthetic PQDs shows the effectiveness of proposed method, especially at the mixed PQDs. The proposed method gained only noticeable improvement at single PQDs classification, but it gained significant improvement at mixed PQDs classifications.

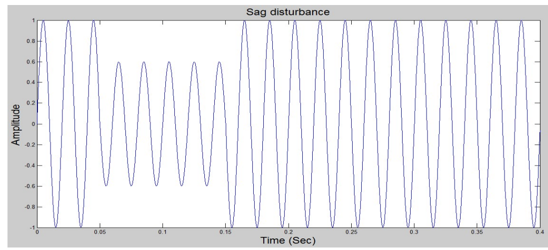
Keywords: power Quality Disturbances, S-transform, Statistical features, Mixed PQDs, KELM, SVM, and Accuracy.

I. Introduction

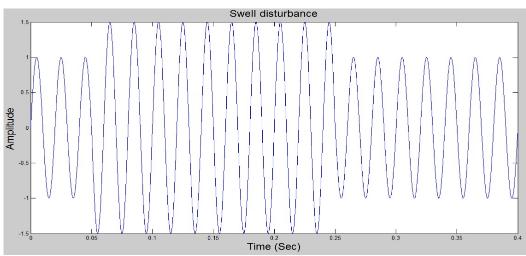
From the past few years, a huge rise in the power consumption is observed due to the population growth and the integration of new equipment with the Power Grid system. Further, to reach the increased demands of consumers, several types of power generation and storage strategies are developed [1]. Moreover, the rapid growth in the solid state switching devices and power electronic equipment in both public sectors and industrial sectors also demands for a continuous supply of electric power. Such kind of integrated devices for distributed power generation, particularly with renewable energy sources (Eolic and Photovoltaic) and the micro grid's consolidation [2] leads to a change in the operational management of power system [3]. Hence, the electric system needs to become smarter and needs to perform the operations in a decentralized fashion [4].

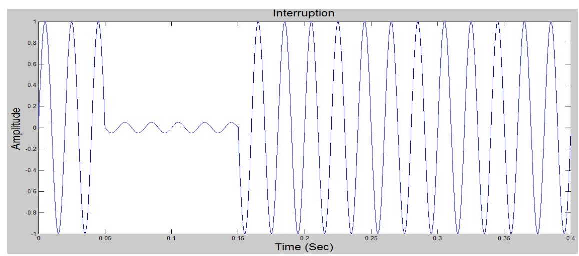
Along with provision of continuous power supply, there must a power quality monitoring unit which assesses the quality of power supplied to home sectors and industrial sectors. Since there exists sensitive electric and electronic devices, the qualitative power needs to be supplied otherwise they will get impacted much seriously. Any deviation in the Original characteristics of power is regarded as Power Quality Disturbance (PQD) [5]. Some examples of such deviations are Interruption, Elevation, Increased outages and under voltages etc. One of the main reasons behind the occurrence of PQDs is load changes according to the consumer's facilities, irrespective of commercial, industrial or residential [6]. Further, the integration of different sources is the major sources of PQDs. PQDs shows huge impact on the experience of consumers, as they are non-linear in nature and cause huge damage to voltage sensitive loads like efficient energetic lighting and computers. In the case of industrial consumers, the PQDs can cause a production line to get stopped [7], [8].



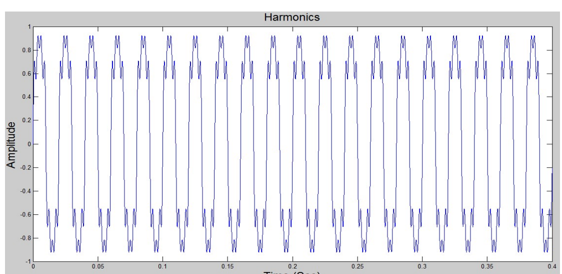


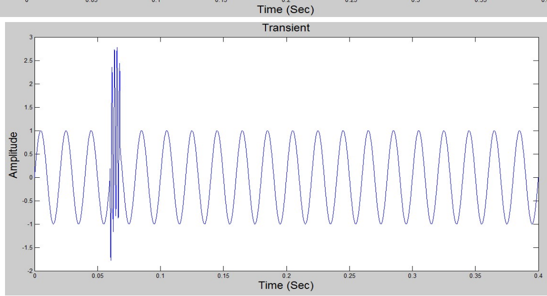
(a) (b)



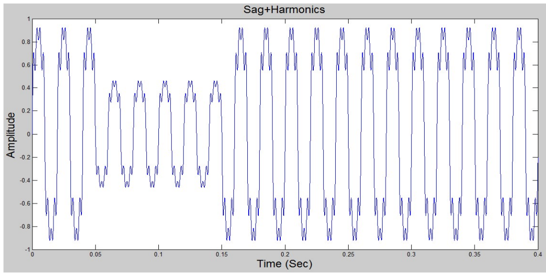


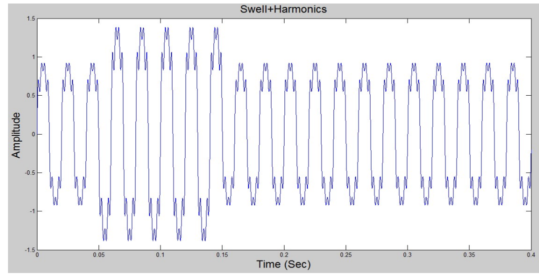
(c) (d)



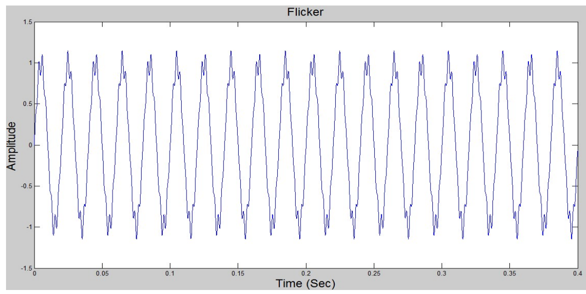


(e) (f)





(g) (h)



(i)

Figure.1 PQDs, (a) Normal, (b) Sag, (c) Swell, (d) Interruption, (e) Harmonics, (f) Oscillatory Transient, (g) Sag with harmonics, (h) Swell with harmonics and (i) Flickr

In general, the power quality is affected due to the occurrence of Voltage Sags/Swells, Harmonics, interruptions, flickers etc. Along with these issues, the PQDs can also signify through the deviations of an electrical signal in its time, amplitude and frequency characteristics. Figure.1 shows some examples of PQD signals. Since the PQDs are totally uncertain in nature, a continuous power quality monitoring unit is required which can identify the deviations instantaneously and identifies the PQDs. Such kind of quick detection is a challenging task for humans and hence, an automatic PQDs detection and classification is the main motivation of this work. In the past, various approaches have been introduced for automatic detection and classification of PQDs through several methods [9-11] including signal processing, machine learning and statistical analysis methods etc. However, most of the earlier methods focused on single PQDs and very less concentrate on the Compound PQDs which are the combinations of multiple and single PQDs. For example, Sag is single PQD and Sag with harmonics is a Compound PQD in which the signal is deviated in multiple characteristics. Moreover, the compound PQDs are non-stationary in nature and the stationary signal processing methods like Fourier transform can't explore the real PQD nature.

To sort out these problems, this paper proposes a new Method for the Detection and Classification of both single and mixed PQDs. The complete methodology involves two feature extraction methods and two machine earning algorithms. S-Transform (ST) and Statistical methods are employed to extract features from PQDs and SVM and KELM are employed for the classification. SVM classifies the PQDs from Non-PQDs and KELM classifies PQDs into different classes. The major contribution of this work as follows;

- S-Transform explores the time-frequency characteristics PQDs effectively without any necessity for the selection of mother wavelet. S-transform also ensures least data loss during the time-frequency transformation. Further, S-Transform ensures a better signal clarity and it won't have a cross term problem. Additionally, the statistical features ensure better discrimination between PQDs especially signals with larger deviations.

- Ensemble Classification ensures better classification for both stationary and non-stationary signals. Mixed PQDs are non-stationary in nature and a single classifier won't support for better accuracy. Hence we employed an ensemble classification at which the SVM classifies the input signal into two classes and the KELM classifies the further sub-class.

Rest of the paper is organized as follows; the details of literature survey are explored in 2nd section. The details of proposed method are explored in 3rd section. 4th section provides the details of experimental investigations and last section concludes the paper.

II. Literature Survey

In the past, so many works have been proposed for the detection and classification of PQDs. Broadly the entire methodology is accomplished in three phases; they are feature extraction, feature selection and classification. For feature extraction, the methods like "Wavelet Transform (WT)", "Fast Fourier Transform (FFT)", "Short Time Fourier Transform (STFT)", "Hilbert Huang Transform (HHT)" etc. are used. Next, for feature selection, the method like "Genetic Algorithm (GA)", "Artificial Bee Colony (ABC)", and "Maximum Redundancy Relevancy (MRR)", "Principal Component Analysis (PCA)", Fisher Criterion etc. are used. Finally at classification, the artificial intelligence algorithm such as SVM, "Artificial Neural Networks (ANN)", "Decision Tree (DT)", "Multi-Layer Perceptron (MLP)" etc. are used. Some the approaches employed even deep learning algorithm also for PQDs classification. This section explores the recent methods focused on the detection and classification of PQDs.

Jamali S *et al.* [12] proposed a PQDs classification method to classify totally 16 classes considered based on the IEEE 1159 standard. Each PQD signal is sampled at 6.4 kHz with 10 cycles is processed for feature extraction through different methods. Maximum Relevancy Minimum Redundancy (MRMR), Sequential Forward Selection (SFS) and genetic algorithm are used for the selection of precise features. Then the selected features are fed to several classifiers and the best classifier is found based on the comparison of results. Six classifiers namely K-NN, ANN, SVM, Random Forest, "Linear Discriminant Analysis (LDA)" and DT are employed for classification.

Liu H. *et al.* [13] considered "Fast Discrete Curvelet Transform (FDCT)" [14] and "Singular Spectrum Analysis (SSA)" and deep "Convolutional Neural Networks (CNNs)" to detect and classify the PQDs. Initially the PQD signal is decomposed through FDCT and SSA up to three and six levels respectively. Then they employed CNNs based classifier and Multi-class SVM classifier for the classification of both single and complex PQDs. Totally they considered 31 classes of both real and synthetic PQDs for classification. Similarly, Ma J *et al.* [15] also employed deep learning algorithm for PQDs classification. They used Stacked Auto encoder as a deep learning method to extract high level features and used Particle Swarm Optimization (PSO) and Variances to help the model for PQDs classification.

Wilson L. Rodrigues *et al.* [16] proposed a customized deep learning algorithm for the classification of PQDs from voltage signals. Their CNN model composed of convolutional layers, batch normalization unit, short term memory layer and pooling layer. They used the voltage signals with overlapping windows with different Signal to Noise Ratios (SNRs) and Different sampling rates. Ucar F *et al.* [17] proposed a PQDs classification method in two stages to classify totally six classes. In the first stage, they employed Histogram based method and DWT to detect the majority Power Quality events (PQEs). In the second stage, they employed Extreme Learning Machine (ELM) to classify the PQ events.

E. Sami *et al.* [18] classified totally four single PQDs such as Swell, Sag, Harmonics and Interruption. For such purpose, they employed a Colorized Continuous Wavelet Transform on voltage signal and converted them into images files, i.e., 2D files. Then they are fed to optimized Bayesian CNNs for classification. Their CNN model composed of three convolutional layers, two pooling layers, one fully connected layer and one soft max layer. Amin Akbarpour et al. [19] proposed a direct and straightforward method for PQDs classification. Totally eight PQDs are targeted at classification, they are Swell, Sag, Interruption, Flicker, Harmonics, Sag with harmonics, well with harmonics and interruption. Three machine learning algorithms are employed for classification; they are namely SVM, Decision tree and KNN and conclude that the DT had shown better classification performance.

Singh U and Singh S. N. [20] proposed a multi-objective feature selection mechanism through ACO to minimize the size of feature set and classification error in the classification of PQDs. For feature extraction, they employed the combination of Time-time transform and S-Transform which gives a wide range of features. Three machine learning algorithms are employed for classification; they are namely SVM, Decision tree and KNN. S. Dash, U. Subudhi [21] applied Modified Stockwell Transform with 2nd order Gaussian window for features extraction from PQDs. Three statistical features namely Energy and Standard Deviation of magnitude and phase contour are extracted from ST matrix and fed o SVM for classification. A Meta-heuristic algorithm called "Whale Optimization Algorithm (WOA)" [22] is employed to tune the SVM's hyper parameters. Similarly, Alqam S. J. *et al*. [23] also employed S-transform for time-frequency analysis of PQDs. They employed a rue based DT for classification.

Bravo-Rodríguez *et al.* [24] proposed a S-Transform based hybrid method for PQDs classification. After extracting the features from PQDs through ST, they are fed to three machine learning algorithms namely DT, K-NN and SVM. Further, the process was optimized through GA [25] and "Competitive Swarm Optimization (CSO)" [26]. Belkis Eristi, Huseyin Eristi [27] proposed "ST and Bayesian optimization-based CNN (STBOACNN)" which employs ST and CNN for PQDs classification in hybrid power system. Initially, a contour image is constructed after applying ST over PQD signal. Then the resultant image is fed to CNN for classification. In addition, they employed "Bayesian Optimization Algorithm (BOA)" [28] to tune the hyper parameters of CNN.

W. Zhao et al. [29] used two transform techniques namely Wavelet Transform and S-Transform for feature extraction and two machine learning algorithms namely Decision tree and classification rules for classification of PQDs. The rules are formulated based on the energy spectrum obtained through Wavelet Transform and other seven time frequency features obtained by S-transform. Ismail Topaloglu [30] developed a CNN based classifier to classify the PQDs. They employed an attention model in CNN in which the data is multiplied by the number of elements by the number of epoch time. Ezgi Güney et al. [35] applied S-transform for feature extraction from PQDS and employed two machine learning algorithms namely VM and ANN for classification.

Khokar *et al.* [39] focused on the classification of bot single and multiple PQDs with the help of DWT and "Probabilistic Neural Networks (PNNs)" and ABC. Here, DWT is used for feature extraction and PNN is used for classification. ABC is used for the optimization of two parameters of PNN such as number of features and spread constant. Huang *et al.* [40] proposed an "Optimal Multi-Resolution Fast S-Transform (OMFST)" for feature extraction and "Classification and Regression Tree (CART)" for classification of PQDs in microgrid. OMFST extracts frequency domain features from PQD signal and totally 67 features are extracted from time-frequency analysis. Further, the feature section is accomplished based on the Gini index and they are fed to CART for classification. Li J *et al.* [41] proposed a method for the detection and classification of PQDs based on "Directed Acyclic Graph SVMs (DAG-SVMs)" and "Double Resolution S-Transform (DRST)".

III. Proposed Approach

3.1 Overview

This section illustrates the complete details of proposed approach for the detection and classification of PQDs from voltage signals. Figure.2 shows the overall block diagram of proposed approach which is executed in two phases they are feature extraction and classification. At feature extraction, two different set of features are extracted hey are time-frequency features and statistical features. For the extraction of time-frequency features, we employed the Stock well transform. Under statistical features, we extract totally six features; they are Root Sum of Squared Level (RSS), Mean Absolute Deviation (MAD), Standard Deviation, Minimum, maximum and mean. Then they are fed to classification at where we employed a ensemble classifier, a combination of SVM and KELM.

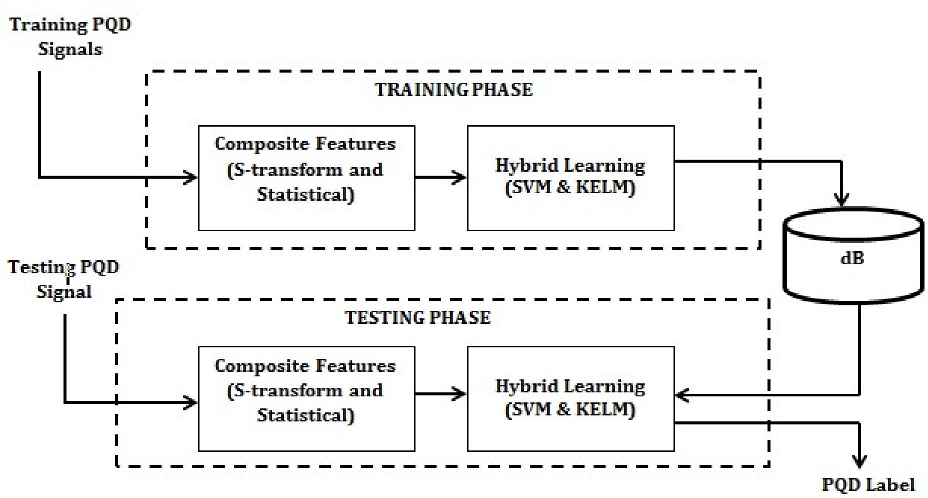


Figure.2 Block Diagram of proposed method

3.2 Feature Extraction

At this phase, we employed two methods namely S-transform and Statistical methods for extracting features from each PQD signal. S-transform intended to provide the time-frequency related features and statistical methods intended to provide statistical characteristics. The complete details of these two methods are explored in the following sub-sections;

A. Time-Frequency Features

In general, the PQD detection and classification is carried out through the voltage signals processing through the electronic and electric equipment. The voltage signals are time base signals which has either continuous or discrete amplitude at every time instance. However, the basic feature such as time and amplitude are not sufficient to detect the PQDs. Hence the researchers extracted different set of features from PQD signals to identify them. Some methods applied Fourier transform which reveals the frequency information of signal, but the FT is applicable for only stationary signals. Next, STFT [31] is a one more time-frequency transformation technique which solves the problem of FT. However, the STFT adapts a fixed width window which cannot reveal the time and frequency resolutions simultaneously. Wavelet Transform [32] is found one of the mostly used multi-resolution analysis technique which can ensure a proper time-frequency resolution of a signal. Wavelet Transform decomposes the signal into different scales and every scale explores specific resolution of the signal. Wavelet Transform applies a pinging window with varying widths and provides both high and low frequency parts of the signal. Finally, the Wavelet transform transforms the signal into a series of wavelet functions. However, the major drawback of Wavelet Transform is the selection of mother wavelet and also the wavelet coefficients lost the data belongs to phase.

To overcome these problems, we use S-Transform, a time-frequency distribution model developed by Stockwell in 1996 [33]. S-Transform is regarded as generalized version of STFT and the extended model of "Continuous Wavelet transform (CWT)". S-Transform is more advantageous than FT and WT. At first, it fixes the modulation sinusoids with respect to time axis; it localizes the scalable Gaussian window translations and dilations. Further, S-Transform ensures a better signal clarity and it won't have a cross term problem. Finally, the computational complexity of S-Transform is very less, i.e., $O(NlogN)$. Since S-transform is regarded as an effective transform, it gives high-frequency resolution at low frequencies and high time resolution at high frequencies. Hence, we applied S-Transform to represent each PQD signal in time-frequency format. The S-Transform of a continuous signal $x(t)$ is expressed as

$$S_x(t,f) = \exp(j2\pi ft)W_x(t,d) \quad (1)$$

Where $W_x(t,d)$ denotes the CWT of a signal $x(t)$. The S-transform is derived as the phase correction of CWT with window being the Gaussian function, Eq.(1) is formulated as a function of CWT of signal. The mathematical expression for $W_x(t)$ is given as

$$W_x(\tau,d) = \int_{-\infty}^{\infty} x(t)\omega(t-\tau,d)dt \quad (2)$$

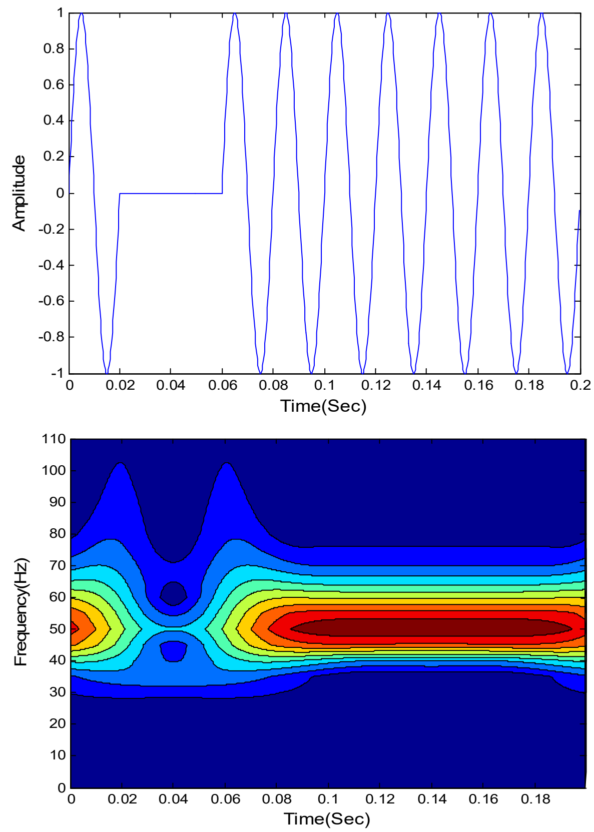
Where $\omega(t,f)$ is a Gaussian Mother wavelet expressed as

$$\omega(t,f) = \frac{|f|}{\sqrt{2\pi}}\exp\left(\frac{-f^2t^2}{2}\right)\exp(-j2\pi ft) \quad (3)$$

Where d signifies the inverse of frequency f, i.e., $d = 1/f$. Thus the mathematical expression of S-Transform is changed as

$$S_x(\tau, d) = \frac{|f|}{\sqrt{2\pi}} \exp\left(\frac{-f^2(t-\tau)^2}{2}\right) \exp\left(-j2\pi f t\right) dt \quad (4)$$

From Eq.(3), it can denoted that the S-transform's width is totally dependent on the frequency f, hence, the width becomes wider for the decrease in frequency and becomes narrower for an increase in the frequency [34]. An example S-transform response of Interruption is shown in Figure.3.



(a) (b)

Figure.3 (a) Interruption and (b) S-transform response of Interruption

B. Statistical features

Since the statistics of 2-D signal are different from PQD to PQD, they can explore significant knowledge about their statistical nature. For example the Sags mean is different from Swell's Mean because they are much deviated in their magnitudes. Hence, we employed to describe the PQD through their statistical features. For the computation of statistical features, each PQD is windowed through a window of time span 0.5 seconds. Totally six features are computed under this category, they are namely RSS, MA), Standard Deviation (SD), Minimum, maximum and mean. They are explained as follows;

Mean: Mean explores the average amplitude variations in the signal. For a given w, the mean is computed as a summation of amplitudes of all samples followed by division of sum with total number of samples present in the window. Mathematically, it is expressed as

$$\mu = \frac{1}{size(w)} \sum_{i=1}^{size(w)} s_i \quad (5)$$

Where s_i denotes the i^{th} sample's amplitude.

Minimum and Maximum: Minimum and Maximum value explores the least and most values among the given input samples. These features help in the discrimination between sag and swell signals and also between sag with harmonics and swell with harmonics. For a given window *w*, the minimum and maximum are computed as;

$$Mx = Max(w) \quad (6)$$

$$Mn = Min(w) \quad (7)$$

SD: SD explores the PQD signal's statistical distribution with respect to mean. For a given window w, initially mean is computed and then each sample's amplitude is subtracted from mean followed by summation, normalization and square root computation. Mathematically, MAD is computed as

$$\sigma = \sqrt{\frac{1}{size(w)} \sum_{i=1}^{size(w)} (s_i - \mu)^2} \quad (8)$$

MAD: It reveals the signal's variability. For a given window *w*, mean is calculated at first and then each sample's amplitude is subtracted from mean followed by summation and normalization is performed. Since the means of PQDs are different in nature, the MAD ensure better discrimination between PQDs. Mathematically, MAD is computed as

$$MAD = \frac{1}{size(w)} \sum_{i=1}^{size(w)} (p_i - \mu) \quad (9)$$

RSS: It is measured as the square root of mean of summation of squared amplitudes of each sample in the window. For a given window, initially each sample is squared and then all the values are subjected to summation followed by normalization and square root computation. RSS alleviate the difference between PQD signals and noises perfectly since the squared amplitude clears the ambiguity. For a given window *w*, the RSS is computed as;

$$RSS = \sqrt{\frac{1}{size(w)} \sum_{i=1}^{size(w)} (p_i)^2} \quad (10)$$

3.3 Ensemble Classification

Under the Ensemble classifier, we considered two different machine learning algorithms namely KELM and SVM. SVM is employed at the first stage of the classification while at the second stage classification, KELM is employed. During the training phase, the system learns only two kinds of features they are PQD features and normal features. The SVM algorithm assigns these two features with two labels, one label for PQD set and another label for normal set. Next for the PQDs classification, the system uses KELM algorithm and the input data is only PQDs feature set. During the testing phase, initially the features are computed through S-transform and Statistical measures and then processed through the SVM algorithm to find its class. If it is classified as PQD then it is again processed through KELM algorithm to find its further class. The complete details about the two machine learning algorithms are demonstrated here.

A. SVM

SVM mainly determines a hyperplane which can place the samples from inside class. Different types of kernel functions are employed in the SVM which can hypothesize the nonlinear and linear features thereby it can model a hyperplane strictly according to the functionality of the kernel function. At this case, we used a tool called as LIBSVM to test the model. Consider $x_i \in R^n, i = 1,2, \dots . l$ as vector of input features of two classes, and then SVM determines a vector of output features $y_i \in R^l$ in such a way $y_i \in \{-1,1\}$. Based on this theory, the mathematical expression of SVM is given as

$$f(t) = sgn\left(\sum_{i=1}^{l} y_i \alpha_i K(x_i, x) + b\right) \quad (11)$$

Where α_i = the coefficients of Lagrange multiplier of the i^{th} feature, $K(x_i, x)$= kernel function and b = subjective constant. Mathematically, the function of kernel is expressed as

$$K\big((x_i, x)\big) = exp\left(\frac{-\|(x_i - x)\|^2}{\sigma^2}\right), \sigma \in R \quad (12)$$

In accordance to the functionality theory, a kernel $K(x_i, x)$ is determined as positive definite kernel if it satisfies the mercer's condition. For a given input signal, the SVM assigns two labels such as -1 for Normal Class and +1 for PQD class.

B. KELM

ELM is found as one of the popular approach which was developed by Huang et al. [36] at first. It is regarded as a feed formal neural network which consists of a single hidden layer. Since the conventional machine learning algorithms require more parameters tuning, this is regard as a local optimal solution. However in ELM, there is no such requirement and it requires only to tune the hidden nodes count in the network. Furthermore ELM also won't expect the weight adjustments in the input layer and the Hidden layers bias. It can be regarded as an optimal solution in global fashion [37]. Due to these reasons the ELM is considered as a fastest algorithm which has fastest convergence rate. Consider the training dataset is represented as $T_r = \{(x_i, \boldsymbol{t_i}), i = 1,2, \dots, N\}$, where $\boldsymbol{x_i} = [x_i^1, x_i^2, \dots, x_i^n]$ is the input feature vector and $\boldsymbol{t_i} = [t_i^1, t_i^2, \dots, t_i^n]$ is the corresponding target vector, the major requirement is to find an optimal model under testing process. Accordingly $\boldsymbol{y_i} = [y_i^1, y_i^2, \dots, y_i^n]$ is an output vector needs to be determined with the help of ELM, hence, it can be described as

$$\boldsymbol{y_i} = \sum_{j=1}^{l} \beta_j g_j(\boldsymbol{x_i}) = \sum_{j=1}^{l} \beta_j g\big(\boldsymbol{\alpha_j} . \boldsymbol{x_i} + \boldsymbol{c_j}\big), i = 1, \dots, N \quad (13)$$

Where $\boldsymbol{\alpha_j}$ is called as Weighted vector which illustrates the weight between hidden layer and input layer, β_j is called as Weighted vector which illustrates the weight between hidden layer and output layer, $g\big(\boldsymbol{\alpha_j} . \boldsymbol{x_i} + \boldsymbol{c_j}\big)$ is the hidden layer's activation function and bias function is denoted by c_j. As c_j and α_j are random processes in the hidden layer, they needs to be determined at hidden layers nodes. Towards the weights and hidden layer nodes computation, the error obtained between output vector and target vector are presumed to be zero as shown below.

$$\sum_{i=1}^{N}\left\|\boldsymbol{t_j}-\boldsymbol{y_j}\right\|=0 \quad (14)$$

Substitute Eq.(13) in Eq.(14), then

$$\boldsymbol{t_j}=\sum_{j=1}^{l}\beta_j g\left(\boldsymbol{\alpha_j}.\boldsymbol{x_i}+c_j\right), i=1,\dots,N \quad (15)$$

Upon expanding the Eq.(15), the resultant of expansion is shown in the Eq.(16), as follows;

$$\begin{bmatrix} g(\boldsymbol{\alpha_1}.\boldsymbol{x_1}+c_1) & \dots & g(\boldsymbol{\alpha_l}.\boldsymbol{x_1}+c_l) \\ \vdots & \ddots & \vdots \\ g(\boldsymbol{\alpha_1}.\boldsymbol{x_N}+c_1) & \dots & g(\boldsymbol{\alpha_l}.\boldsymbol{x_N}+c_l) \end{bmatrix}_{N\times l} . \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_l^T \end{bmatrix}_{l\times m} = \begin{bmatrix} t_1^T \\ \vdots \\ t_l^T \end{bmatrix}_{N\times m} \quad (16)$$

Eq.(16) can be written as

$$\boldsymbol{H}\beta=\boldsymbol{T} \text{ and } \beta=\boldsymbol{H}^{+}\boldsymbol{T} \quad (17)$$

Where **H** is the Hidden layer's output matrix**,** β is the Hidden layer's weight and T is target vector's the output matrix. $\boldsymbol{H}^{+}$ is the **"**Moore-Penrose (MP)" generalized inverse of H matrix**,** is obtained as follows;

$$\boldsymbol{H}^{+}=\boldsymbol{H}^{\boldsymbol{T}}(\boldsymbol{H}\boldsymbol{H}^{\boldsymbol{T}})^{-\boldsymbol{1}} \quad (18)$$

However, ELM experienced poor classification performance under its accomplishment over some unknown datasets. Thus, we developed a new version of ELM by introducing a kernel function in it and named as KELM. According to the modified expression, the KELM is formulated as

$$f(\boldsymbol{x})=h(\boldsymbol{x})\beta=h(\boldsymbol{x})\left(\boldsymbol{H}^{\boldsymbol{T}}\left(\frac{\boldsymbol{I}}{\boldsymbol{C}}+\boldsymbol{H}\boldsymbol{H}^{\boldsymbol{T}}\right)^{-\boldsymbol{1}}\right) \quad (19)$$

Where C denotes a penalty parameter and *I* is an identity matrix. The Kernel function of KELM is defined as

$$KELM_{i,j}=h(\boldsymbol{x_i})h\left(\boldsymbol{x_j}\right)=K(\boldsymbol{x_i},\boldsymbol{x_j}) \quad (20)$$

Then Eq.(19) is changed as

$$f(\boldsymbol{x})=\begin{bmatrix} K(\boldsymbol{x},\boldsymbol{x_1}) \\ \vdots \\ K(\boldsymbol{x},\boldsymbol{x_N}) \end{bmatrix} . \left(\boldsymbol{H}^{\boldsymbol{T}}\left(\frac{\boldsymbol{I}}{\boldsymbol{C}}+\boldsymbol{H}\boldsymbol{H}^{\boldsymbol{T}}\right)^{-\boldsymbol{1}}\right) \quad (21)$$

The kernel function shows significant impact on the performance of KELM. Furthermore, the exact kernel function insertion is challenging task in KELM. Here, we consider two different kernels such as RBF and polynomial to insert in the KELM. The polynomial kernel has a great generalization capability and hence regarded as generalized function which was employed in so many approaches to deal with generalized learning capability [38]. Moreover, it is also considered as global kernel and its mathematical expression is given as follows;

$$K_p\left(\boldsymbol{x},\boldsymbol{x_j}\right)=(\boldsymbol{x}.\boldsymbol{x_i}+b)^p \quad (22)$$

On the other hand, the RBF kernel is regarded as local kernel function which has strong ability of learning and weak ability at generalization of solution. Based on the local functionality, the RBF kernel is mathematically articulated as

$$K_R(\boldsymbol{x},\boldsymbol{x_j}) = exp\left(-\frac{\|x-x_i\|^2}{2\sigma^2}\right) \quad (23)$$

By integrating these two kernels, a composite kernel function is designed by assigning individual weights to each kernel. Mathematically, the composite kernel function is represented as

$$K_C(\boldsymbol{x},\boldsymbol{x_j}) = w_1.K_p(\boldsymbol{x},\boldsymbol{x_j}) + (1-w_1).K_R(\boldsymbol{x},\boldsymbol{x_j}), w_1 \in [0,1] \quad (24)$$

The proposed composite kernel is more advantageous than the individual kernels. As the polynomial kernel suffers from weak ability at learning, it is solved through RBF kernel. Next, the weak ability at generalization of solution with RBF kernel is solved through polynomial kernel. Thus, the proposed approach can ensure better IDS with strong learning ability along with strong generalization capability. Such kind of classifiers produces better classification results.

IV. Experimental Analysis

4.1 Simulation setup

To validate the proposed method experimentally, here we generate synthetic PQD signals in different number. Table.1 shows the generalized mathematical expressions through several control parameters used to generate PQDs. Here we consider totally 11 different signals among which one is normal signal which is represented with a pure sinusoidal wave form of frequency 50 Hz and amplitude of 1.0 p.u. Next, the PQDs are represented through the deviated Sinusoid signals and the deviations are attained through different control parameters. The entire signals are generated through MATLAB tool and the total recording time of each signal is kept as 0.4 seconds. Next, table.2 shows the number of signal generated, trained and tested. Out of total generated signals, 70% of signals are used for training and the remaining 30% are used for testing.

Table.1 controlling parameters of various PQDs

Class	Name	Expression	Controlling Parameters
C1	Sag	$s(t) = [1-\alpha(u(t-t_1) - u(t - t_2))]\sin(\omega_n t)$	$0.9 \geq \alpha \geq 0.1$ & $9T \geq t_2 - t_1 \geq T$
C2	Swell	$s(t) = [1+\alpha(u(t-t_1) - u(t - t_2))]\sin(\omega_n t)$	$0.8 \geq \alpha \geq 0.1$ & $9T \geq t_2 - t_1 \geq T$

C3	Interruption	$s(t) = [1 - \alpha(u(t - t_1) - u(t - t_2))] \sin(\omega_n t)$	$1 \geq \alpha \geq 0.9$ & $9T \geq t_2 - t_1 \geq T$
C4	Flicker	$s(t) = [1 + \alpha \sin(2\pi\beta t)] \sin(\omega_n t)$	$0.1 \leq \alpha \leq 0.2,\ 5Hz \leq \beta \leq 20Hz$
C5	Harmonics	$s(t) = \alpha_1 \sin(\omega_n t) + \alpha_3 \sin(3\omega_n t) + \alpha_5 \sin(5\omega_n t) + \alpha_7 \sin(7\omega_n t)$	$0.05 \leq \alpha_3, \alpha_5, \alpha_7 \leq 0.15, \sum \alpha_i^2 = 1$
C6	Oscillatory Transients	$s(t) = \sin(\omega_n t) + \alpha \exp\left(-(t - t_1)/\tau\right)(u(t - t_2))\ \sin(2\pi f_n t)$	$0.1 \leq \alpha \leq 0.8,\ 0.5T \leq t_2 - t_1 \leq 3T,\ 300Hz \leq f_n \leq 900Hz,\ 8ms \leq \tau \leq 4m$
C7	Spike	$s(t) = \sin(\omega_b t) + sign(\sin(\omega_b t)) \times \sum_{n=0}^{9} k\big[u\big(t - (t_1 + 0.2n)\big) - u(t - (t_2 + 0{,}02n))\big]$	$0.1 \leq k \leq 0.4,\ 0 \leq t_1, t_2 \leq 0.5T,$ $0.01T \leq t_2 - t_1 \leq 0.05T$
C8	Notch	$s(t) = \sin(\omega_b t) - sign(\sin(\omega_b t)) \times \sum_{n=0}^{9} k\big[u\big(t - (t_1 + 0.2n)\big) - u(t - (t_2 + 0{,}02n))\big]$	$0.1 \leq k \leq 0.4,\ 0 \leq t_1, t_2 \leq 0.5T,$ $0.01T \leq t_2 - t_1 \leq 0.05T$
C9	Sag + Harmonics	$s(t) = \big[1 - \alpha\big(u(t - t_1) - u(t - t_2)\big)\big] \times [\alpha_1 \sin(\omega_n t) + \alpha_3 \sin(3\omega_n t) + \alpha_5 \sin(5\omega_n t) + \alpha_7 \sin(7\omega_n t)]$	$\sum \alpha_i^2 = 1,\ 0.15 \geq \alpha_5, \alpha_7, \alpha_3 \geq 0.05,\ 0.9 \geq \alpha \geq 0.1$ & $9T \geq t_2 - t_1 \geq T$

C10	Swell + Harmonics	$s(t) = \left[1 + \alpha\left(u(t - t_1) - u(t - t_2)\right)\right] \times [\alpha_1 \sin(\omega_n t) + \alpha_3 \sin(3\omega_n t) + \alpha_5 \sin(5\omega_n t) + \alpha_7 \sin(7\omega_n t)]$	$\sum \alpha_i^2 = 1,\ 0.15 \geq \alpha_5, \alpha_7, \alpha_3 \geq 0.05,\ 0.8 \geq \alpha \geq 0.1\ \&\ 9T \geq t_2 - t_1 \geq T$
C11	Interruption + harmonics	$s(t) = \left[1 - \alpha\left(u(t - t_1) - u(t - t_2)\right)\right] \times [\alpha_1 \sin(\omega_n t) + \alpha_3 \sin(3\omega_n t) + \alpha_5 \sin(5\omega_n t) + \alpha_7 \sin(7\omega_n t)]$	$\sum \alpha_i^2 = 1,\ 9T \geq t_2 - t_1 \geq T,\ 0.15 \geq \alpha_5, \alpha_7, \alpha_3 \geq 0.05,\ 1 \geq \alpha \geq 0.9$
C12	Normal	$s(t) = \sin(\omega_n t)$	$\omega_n = 2\pi \times 50$ rad/sec

Table.2 Simulation Setup

Class	Total Signals Generated	Trained	Tested
Sag – C1	297	208	89
Swell – C2	264	185	79
Interruption – C3	66	46	20
Flicker - C4	32	22	10
Harmonics – C5	27	19	8
Oscillatory Transients – C6	288	202	86
Spike – C7	144	101	43
Notch – C8	144	101	43
Sag + Harmonics – C9	243	170	73
Swell + Harmonics – C10	216	151	65
Interruption + harmonics – C11	54	37	17
Normal - C12	10	7	3
Total	**1785**	**1249**	**536**

B. Results

After the simulation of test signals, the obtained classification results are formulated in the form of a confusion matrix of size $N \times N$ where N denotes the total number of classes intended to classify. Table.3 shows the resultant confusion matrix constructed after the simulation of proposed mechanism with multiple features and two machine learning algorithms. From the results, we can see that the proposed approach had identified more TPs for single PQDs than the mixed PQDs. Next, the performance is evaluated through three performance metrics namely Recall, precision, and F-Score, shown in Table.4. Out of 12 PQDs, four PQDs are recognized more accurately as they have gained recall rate of 100%. On the other hand, along with few single PQDs, the mixed PQDs have gained 100% precision. As the statistical feature ensures more differentiation between mixed PQDs, the False Positive count is less even for mixed PQDs. Hence, they are precisely classified with 100% precision. Further, the maximum misclassification is observed between C1 (Sag) and C3 (Interruption) because our synthetic Sag and Interruption signals are in resemblance with each other. The False negative Rate between interruption and Sag is observed as approximately 20%. Further, we observed that some of the mixed PQDs are classified as the base signals. For example out of 73 Sag with harmonics, 2 are classified as Harmonics and 2 are classified as Sag. Similarly, out of 65 swell with harmonics, 2 are classified as swell and out of 17 interruption with harmonics, 2 are classified as harmonics.

Table.3 Confusion matrix of overall system

	C1	**C2**	**C3**	**C4**	**C5**	**C6**	**C7**	**C8**	**C9**	**C10**	**C11**	**C12**	**Total**
C1	**89**	0	0	0	0	0	0	0	0	0	0	0	**89**
C2	0	**79**	0	0	0	0	0	0	0	0	0	0	**79**
C3	4	0	**16**	0	0	0	0	0	0	0	0	0	**20**
C4	0	0	0	**9**	1	0	0	0	0	0	0	0	**10**
C5	0	0	0	0	**8**	0	0	0	0	0	0	0	**8**
C6	0	2	0	0	1	**83**	0	0	0	0	0	0	**86**
C7	0	0	0	0	0	1	**41**	1	0	0	0	0	**43**
C8	0	0	0	0	0	0	3	**40**	0	0	0	0	**43**
C9	2	0	0	0	2	0	0	0	**69**	0	0	0	**73**
C10	0	1	0	0	2	0	0	0	0	**62**	0	0	**65**
C11	0	0	0	0	2	0	0	0	0	0	**15**	0	**17**
C12	0	0	0	0	0	0	0	0	0	0	0	**3**	**3**

Total	95	82	16	9	16	84	44	41	69	62	15	3	**536**

Table.4 Performance metrics of overall detection and classification system

Class	Recall (%)	Precision (%)	F-score (%)
C1	100.00	93.3800	96.7400
C2	100.00	96.3400	98.1400
C3	80.0000	100.00	88.8900
C4	90.0000	100.00	94.7400
C5	100.00	50.0000	66.6700
C6	96.5100	98.8100	97.6500
C7	95.3500	93.1800	94.2500
C8	93.0200	97.5600	95.2400
C9	94.5200	100.00	97.1800
C10	95.3800	100.00	97.6400
C11	88.2400	100.00	93.7500
C12	100.00	100.00	100.00

Table.5 shows the impact of feature extraction and machine learning methods on the detection and classification of PQDs. For this purpose, we employed different combination like ST+SVM, ST+KELM, SF+SVM and SF+KELM and computed F-score for every simulation. From the results, the best combination is found at ST + KELM since it gained an average F-score of 94.6484%. Compared with statistical features, the multi-resolution features extracted by S-transform are more informative and provide more information about the time-frequency characteristics of PQDs. Further, the KELM employed a composite kernel strategy to train the detection system and hence, the both stationary and non-stationary signals are classified accurately. Further, the least F-score is observed at the combination of SF + SVM because the statistical features provide similar values for common signals like sag & interruption, and Flicker & Oscillatory transient etc.

Table.5 F-score computation at different combinations of proposed approach (ST: S-Transform, SF: Statistical Features)

Class	ST + SVM	ST + KELM	SF+ SVM	SF + KELM
C1	93.5820	94.8210	94.0630	94.2670

C2	97.4120	96.2210	95.4650	95.6680
C3	85.4410	86.9710	86.2120	86.4190
C4	91.4520	92.8210	92.0660	92.2670
C5	68.4420	64.7510	63.9980	64.1980
C6	93.2170	95.7310	94.9790	95.1740
C7	92.1380	92.3310	91.5780	91.7790
C8	93.3360	93.3210	92.5620	92.7660
C9	96.8685	95.2610	94.5080	94.7080
C10	95.3350	95.7210	94.9620	95.1660
C11	89.4780	91.8310	91.0750	91.2750
C12	100.00	100.00	100.00	100.00

Figure.4 and Figure.5 shows the impact of SNR and sampling rates on the detection and classification of PQDs through Precision and Recall respectively. As the SNR increases, the noise present in the PQD signal gets vanished and the signal becomes more qualitative. Such kind of qualitative signal ensure better features and helps the detection system in accurate classification. Hence, we achieved more precision and recall at larger SNR than at the lower SNRs. Next, the sampling rate also plays a vital role in the PQD classification. Towards such analysis, we conduct a simulation study with varying sampling rates, i.e., different samples are kept for each window like 16 samples/window, 32 samples/window and 64 samples/window. As the number of samples is more in each window, the system can gain more knowledge about the features of PQDs. Hence the better performance is achieved at larger sampling rates. The maximum detection rate is observed as 98.3250% and it is achieved at the sampling rate of 64 samples/window and SNR of 40 dB. Similarly, the maximum precision is observed as 99.5620% and it is achieved at the sampling rate of 64 samples/window and SNR of 40 dB.

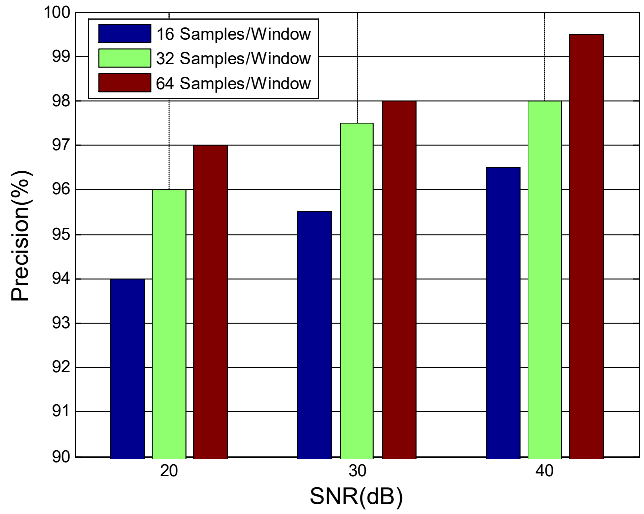


Figure.4 Precision for varying SNRs at different Sampling rates

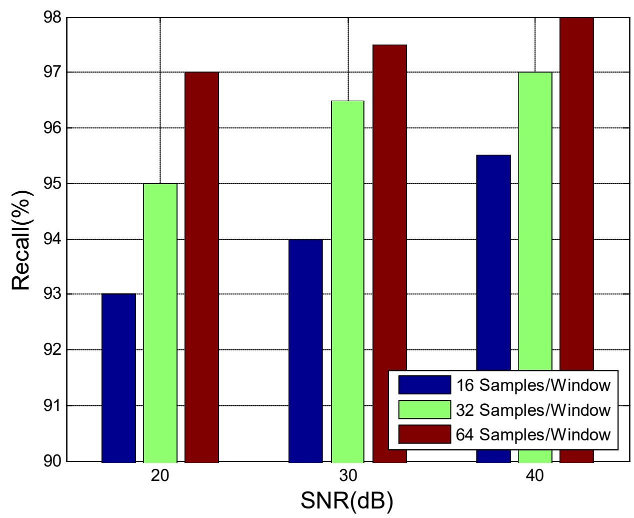


Figure.5 Recall for varying SNRs at different Sampling rates

Table.6 Comparison of accuracies between proposed and existing methods

Reference No.	**Methodology**	**Number of Classes**	**Accuracy (%)**			
			20 DB	**30 DB**	**40 DB**	**50 DB**
[39]	Wavelet Transform and Probabilistic Neural Networks	14	86.86	91.93	93.71	94.57
[29]	Wavelet Features with K-NN	8	96.78	98.00	98.48	98.56
[40]	Fast S-Transform and Embedded Decision Tree	12	91.50	98.58	98.83	98.92

[41]	S-Transform and DAG-SVM	9	97.77	-	-	-
[24]	S-Transform and CSO-SVM	13	97.76	-	-	-
Proposed	S-Transform and Statistics with Ensemble Learning	12	97.96	98.69	99.20	99.45

Table.6 shows the comparison between proposed and several conventional methods through the methodology employed, classes considered and accuracy obtained at different SNRs. Among the existing methods, we considered both single PQDs classification methods and mixed PQDs classification methods. Most of the methods employed Wavelet and S-transform for feature extraction as they are more effective in the provision of time-frequency features. However they were not much concentrated on statistical analysis and classifier. Considering additional features along with time-frequency features, improves the analytical capability of system thereby increasing the classification accuracy. For example, [39] and [29] used simple Wavelet transform and simple machine learning algorithms. Hence [39] experienced less accuracy of 86.86% only. Even at larger SNRs it gained only 94.57% accuracy which shows that it had limited capacity at mixed PQDs. Next, [41] considered only single PQDs and hence it gained 97.77% accuracy. Even though [24] employed an additional CSO algorithm for the optimization SVMs parameters, it gained only 97.76% accuracy which shows that it has limited capability for mixed PQDs classification. Further, we observed that as the SNR progresses, the accuracy increases and reached maximum value at 50 DB. At every SNR, the proposed method gained more accuracy than all the existing methods. The main reason is that the proposed method involved two set of features and two machine learning algorithms. Such kind of accomplishment ensures better discrimination between mixed PQDs and sources of mixed PQDs.

V. Conclusion

PQDs are serious problems in the electrical field which causes serious damages to the appliances and devices connected. Hence, the prior detection and classification of PQDs is a significant research area. Towards such aspect, this paper proposed a new method which describes each PQD with two set of features and employs two machine learning algorithms for classification. S-transform is employed to analyze time-frequency features and statistical features are extracted along with. SVM and KELM machine learning algorithms are employed at classification. Simulation on the develop stem through synthetic PQDs data shows the superiority in the classification of totally 12 types of PQDs which include both single and mixed PQDs. Especially, the proposed method shown its significance at mixed PQDs which are considered complex signals for existing methods. On an average, the proposed gained an improvement of 0.5896% from single PQDs classification methods and 5.5230%f from mixed PQDs classifications methods.

References

[1] Roldán-Blay C, Escrivá-Escrivá G, Roldán-Porta C. Improving the benefits of demand response participation in facilities with distributed energy resources. *Energy.* 2019; 169:710–8

[2] Hirsch A, Parag Y, Guerrero J. Microgrids: a review of technologies, key drivers, and outstanding issues. *Renew Sustain Energy Rev*. 2018;90:402–11

[3] Silva IRS, Rabêlo RAL, Rodrigues JJPC, Solic P, Carvalho A. A preference-based demand response mechanism for energy management in a microgrid. *J Clean Prod* 2020; 255:120034

[4] Atems B, Hotaling C. The effect of renewable and nonrenewable electricity generation on economic growth. *Energy Policy*. 2018; 112:111–8.

[5] Elbasuony GS, Aleem SHA, Ibrahim AM, Sharaf AM. A unified index for power quality evaluation in distributed generation systems. *Energy.* 2018; 149:607–22.

[6] Mahela OP, Shaik AG, Gupta N. A critical review of detection and classification of power quality events. *Renew Sustain Energy Rev.* 2015;41:495–505.

[7] Hooshmand R, Enshaee A. Detection and classification of single and combined power quality disturbances using fuzzy systems oriented by particle swarm optimization algorithm. *Electr Power Syst Res*. 2010; 80(12):1552–611.

[8] Kapoor R, Gupta R, Jha S, Kumar R, et al. Detection of power quality event using histogram of oriented gradients and support vector machine. *Measurement.,* 2018; 120:52–755.

[9] O. J. Singh and Dr. P. Winston, "A Survey on Classification of Power Quality Disturbances in a Power System", *International Journal of Engineering Research and Applications (IJERA),* Vol.4, Issue.8, August 2014, pp.80-84.

[10] Mozaffari, Mahsa, Keval Doshi, and Yasin Yilmaz., "Real-Time Detection and Classification of Power Quality Disturbances", *Sensors* 22, no. 20: 7958, 2022.

[11] J. E. Caicedo, D. A. Martinez, E. Rivas Trujillo, "A systematic review of real time detection and classification of power quality disturbances", *Prot. Control mod Power Syst,* 8, 3, 2023.

[12] Jamali S, Farsa AR, Ghaffarzadeh N. Identification of optimal features for fast and accurate classification of power quality disturbances. *Measurement.* 2018, 116:565–74.

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

[14] Emmanuel Cand`es, Laurent Demanet, David Donoho and Lexing Yin, "Fast Discrete Curvelet Transforms", *Multiscale Modeling & Simulation*, Vol. 5, Iss. 3 (2006).

[15] Ma J, Zhang J, Xiao L, Chen K, Wu J. Classification of power quality disturbances via deep learning. *IETE Tech Rev.* 2017;34(4):408–15.

[16] Wilson L. Rodrigues Jr, Fabbio A. S. Borges, Antonio O. de Carvalho Filho, Ricardo de A. L. Rabelo, "A Deep Learning Approach for the Detection and Classification of Power Quality Disturbances with Windowed Signals", *SN Computer Science* (2021) 2:64.

[17] Ucar F, Alcin OF, Dandil B, Ata F (2018) Power quality event detection using a fast extreme learning machine. Energies, 2018, 11, 145.

[18] Sami Ekici, Ferhat Ucar, Besir Dandil, Reza Arghandeh, "Power quality event classification using optimized Bayesian convolutional neural networks", *Electrical Engineering,* (2021) 103:67–77.

[19] Amin Akbarpour, Mehdi Nafar, Mohsen Simab, "Multiple power quality disturbances detection and classification with fluctuations of amplitude and decision tree algorithm", Electrical Engineering (2022) 104:2333–2343.
[20] Singh U, Singh SN (2019) A new optimal feature selection scheme for classification of power quality disturbances based on ant colony framework. Appl Soft Comput 74:216–225.
[21] S. Dash, U. Subudhi (2019) Multiple power quality event detection and classification using modified S transform andWOAtuned SVM classifier. *Electr Eng Syst Sci.*
[22] S. Mirjalili and A. Lewis,. "The Whale Optimization Algorithm." in Advances en Engineering Software,. vol. 95,pp. 51-67,May 2016
[23] Alqam SJ, Zaro FR (2019) Power quality detection and classification using S-transform and rule-based decision tree. *Int J Electr Electron Eng Telecommun*. VOl.8, No.9, 2018.
[24] Bravo-Rodríguez JC,Torres FJ, BorrásMD(2020) Hybrid machine learning models for classifying power quality disturbances: a comparative study. Energies 13:2761.
[25] Siedlecki, W.; Sklansky, J. A note on genetic algorithms for large-scale feature selection. *Pattern Recognit. Lett*. 1989, 10, 335–347.
[26] Cheng, R.; Jin, Y. A competitive swarm optimizer for large scale optimization. *IEEE Trans. Cybern.* 2014, 45, 191–204.
[27] Belkis Eristi, Huseyin Eristi, "A new deep learningmethod for the classification of power quality disturbances in hybrid power system", *Electrical Engineering* (2022) 104:3753–3768.
[28] Sameen MI, Pradhan B, Lee S (2020) Application of convolutional neural networks featuring Bayesian optimization for landslide susceptibility assessment. *CATENA* 186:104249.
[29] Wenjing Zhao1* , Liqun Shang1 and Jinfan Sun, "Power quality disturbance classification based on time-frequency domain multi-feature and decision tree" *Protection and Control of Modern Power Systems* (2019) 4:27.
[30] Ismail Topaloglu, "Deep Learning Based a New Approach for Power Quality Disturbances Classification in Power Transmission System", *Journal of Electrical Engineering & Technology* (2023) 18:77–88.
[31] Daubechies I (1992) Ten lectures of wavelets. *CBMS-NSF Reg Conf Ser Appl Math*.
[32] Kim D I, Chun T Y, Yoon S H et al (2017)Wavelet-based event detection method using PMU data. *IEEETrans Smart Grid* 8:1154–1162.
[33] Stockwell, R. G., Mansinha, L. & Lowe, R. Localization of the complex spectrum: The S transform. *IEEE Trans. Signal Process.* 44, 998–1001 (1996).
[34] Rutkowski, G., Patan, K. & Leśniak, P. In *Intelligent Systems in Technical and Medical Diagnostics* 279–289 (Springer, 2014).
[35] Ezgi Güney, Ozan Çakmak , Çağrı Kocaman, "Classification of Stockwell Transform Based Power Quality Disturbance with Support Vector Machine and Artificial Neural Networks", *Journal of Intelligent Systems: Theory and Applications* 5(1) 2022: 75-84.
[36] G.B. Huang, Q.Y. Zhu, C.K. Siew, Extreme learning machine: Theory and applications, *Neuro-computing* 70 (2006) 489–501.
[37] J. Wu, Y. Zhu, Z.C. Wang, Z.J. Song, X.G. Liu, W.H. Wang, Z.Y. Zhang, Y.S. Yu, Z.P. Xu, T.J. Zhang, J.H. Zhou, A novel ship classification approach for high resolution SAR images based on the BDA-KELM classification model, *Int. J. Remote Sens*. 38 (2017) 6457–6476.

[38] Z.D. Tian, S.J. Li, Y.H. Wang, X.D. Wang, Wind power prediction method based on hybrid kernel function support vector machine, *Wind Eng*. 42 (2018) 252–264.
[39] Khokhar, S.; Zin, A.A.M.; Memon, A.P.; Mokhtar, A.S. A new optimal feature selection algorithm for classification of power quality disturbances using discrete wavelet transform and probabilistic neural network. *Measurement* 2017, 95, 246–259.
[40] Huang, N.; Peng, H.; Cai, G.; Chen, J. Power Quality Disturbances Feature Selection and Recognition Using Optimal Multi-Resolution Fast S-Transform and CART Algorithm. *Energies* 2016, 9, 927.
[41] Li, J.; Teng, Z.; Tang, Q.; Song, J. Detection and Classification of Power Quality Disturbances Using Double Resolution S-Transform and DAG-SVMs. *IEEE Trans. Instrum. Meas*. 2016, 65, 2302–2312.