

Recognition and Classification of Power Quality Disturbances by DWT-MRA and SVM Classifier

Fayyaz Jandan¹

Faculty of Electrical Engineering Department,
Quaid-e-Awam University of Engineering Science and
Technology Campus, Larkana, Pakistan

Syed Abid Ali Shaha³

PhD School of Engineering and Applied Sciences,
Aston University Birmingham UK

Suhail Khokhar²

Faculty of Electrical Engineering Department,
Quaid-e-Awam University of Engineering Science and
Technology, Nawabshah, Pakistan

Farhan Abbasi⁴

Faculty of Electrical Engineering Department,
Quaid-e-Awam University Campus of Engineering Science
and Technology, Larkana, Pakistan

Abstract—Electrical power system is a large and complex network, where power quality disturbances (PQDs) must be monitored, analyzed and mitigated continuously in order to preserve and to re-establish the normal power supply without even slight interruption. Practically huge disturbance data is difficult to manage and requires the higher level of accuracy and time for the analysis and monitoring. Thus automatic and intelligent algorithm based methodologies are in practice for the detection, recognition and classification of power quality events. This approach may help to take preventive measures against abnormal operations and moreover, sudden fluctuations in supply can be handled accordingly. Disturbance types, causes, proper and appropriate extraction of features in single and multiple disturbances, classification model type and classifier performance, are still the main concerns and challenges. In this paper, an attempt has been made to present a different approach for recognition of PQDs with the synthetic model based generated disturbances, which are frequent in power system operations, and the proposed unique feature vector. Disturbances are generated in Matlab workspace environment whereas distinctive features of events are extracted through discrete wavelet transform (DWT) technique. Machine learning based Support vector machine classifier tool is implemented for the classification and recognition of disturbances. In relation to the results, the proposed methodology recognizes the PQDs with high accuracy, sensitivity and specificity. This study illustrates that the proposed approach is valid, efficient and applicable.

Keywords—Power quality disturbances; discrete wavelet transform; multi resolution analysis; support vector machine

I. INTRODUCTION

Today, an industrial progress and trade perform a key role in the economic stability and growth of any country. Not only mechanical but electrical faults and disturbances are largely responsible for the consistent and quality supply of electricity. Rising use of electronic/ solid state converters and fast control equipment has become the requirement for industries as well as for utility companies to get uninterrupted production and supply facilities [1, 2]. Short circuit faults, lightning, switching of induction machines, motors, power transmission lines, capacitor banks, non-conventional power plant integrations etc.

are also the major causes of Power Quality Disturbances (PQD's). The term PQD is by and large outlined as the variation or distortion in normal voltage/ current amplitude, phase angle and frequency. These distortions/disturbances remarkably lessen the reliability, performance and life of electrical equipment. If such disturbances persist and not mitigated instantly, may cause failures of high voltage equipment (i.e. insulation degradation, failures in high-voltage and low-voltage protection systems, incorrect tripping by relaying systems etc.) leading to complete discontinuity of power transmission or distribution network supply systems [3]. PQDs are broadly categorized as, magnitude variation, transients and steady state variations or harmonics. Magnitude variations are classified as the voltage/current sag (0.1-0.9 P.U), swell (1.1-1.8 P.U), interruption (<0.1 P.U) [4], while transients can be further classified as impulsive and oscillatory transients. Consequently, steady-state PQDs are typically classed into harmonics, flicker and notch categories. Due to coinciding abnormal switching and other small operations in power system, variety of single and multiple PQDs (i.e. the combination of different disturbances) may rise [5].

Power quality is a growing concern since 1980s. Increasing researches in this area can be broadly categorized and congregated into six broad aspects. These aspects include PQ fundamental concepts, standards, effects, solutions, sources, instrumentation, mathematical modeling and analysis. For the research in automatic classification and recognition of power quality events, the power quality standards, time domain, transform domain wave analysis and machine learning algorithms are considered. PQDs can be mitigated to some extent by the application of several types and combinations of active/ passive filters and compensators. In order to avoid disturbances and their specific causes, continuous monitoring of voltage / current waveform patterns and instantaneous constraints must be analyzed. Although, the waveform event data collection and manual monitoring's by PQ analyzers are not much reliable and so also requires substantial computational time and resources. In consequence, automatic and intelligent algorithm based measures are adopted essentially for the detection, recognition and classification of

PQDs. This intelligent monitoring and analysis may also lead to help with the purpose of re-establishing and preserving the abnormal power supply in fraction of seconds without even slight interruption. Likewise, at troubleshooting from power generation to utilization systems, fault forecasting and moreover issues with integration of distributed generation (DG's) systems in power system can also be managed [6].

A lot of research work has been carried out in the area of recognition and classification of PQDs. Selection of PQD types, number and types of single and multiple disturbances under observation, choice of signal processing techniques according to their feature extraction capability, selection of features and choice of artificial intelligent technique conferring to classification performance characteristics are the main concerns and challenges. In literature, numerous signal processing techniques are used for waveform analysis. Some frequently used techniques can be classified as Fourier transforms (FT), Kalman filtering, wavelet transforms (WT), Stockwell-transform (S-Transform), Hilbert-Huang transform (HHT), Gabor transform (GT), etc. [3].

In power system PQDs inception is sometimes dynamically slow and sometimes abrupt; for that reason, PQD event detection in the time frame is important for analysis and monitoring. This brings to the application of Wavelet transform. However other transforms can also be implemented. But the limitations of erroneous detection in dynamic and sudden disturbances, complexity, non-localization in time or space and non-suitability with dynamic signal like PQDs, makes wavelet transform a robust and suitable tool for detection of the event in the time-frequency domain. For the classification, features are extracted, based on statistical parameters, from the decomposed and filtered levels of Discrete Wavelet Transform (DWT). These features are then used for the training and testing the artificial intelligence based classifiers [7, 8]. The classification can be applied by the number of techniques like Artificial Neural Network (ANN), Support Vector Machine (SVM), Fuzzy Expert System (FES), Neuro Fuzzy System (NFS) etc. [3, 9, 10]. Although support vector machine classifier has a lot of tuning parameters but it is suitable because of its high classification accuracy, robustness, and less computational time requirements [4, 11].

A brief literature review related to proposed algorithm is discussed as follows. Author in [26] proposed Gabor Transform (GT) and feed forward Neural Network to identify the arcing faults in power system. Author in [27] proposed HHT-PNN classification algorithm to detect and classify the single and multiple hybrid PQ disturbances. Author in [28] proposed S-Transform-Based ANN Classifier and Rule-Based Decision Tree. Author in [23] proposed a procedure to detect and classify PQ events having complex perturbations caught in a real time power distribution network using WT-SVM. Author in [4] proposed automatic pattern recognition of Single and Multiple Power Quality Disturbances based on wavelet norm entropy feature and PNN. Author in [25] proposed Variational Mode Decomposition (VMD) and S-transform for feature extraction along with SVM classifier. He selected features on filter and wrapper method. Author in [18] proposed WT-SVM for classification of single and hybrid PQDs and with PNN and feed forward neural network. Author in [9] proposed a

technique for optimal selections of features based in WT-PNN and artificial bee colony based algorithm. Author in [1] proposed tunable-Q wavelet transform and dual multi-class SVM for online detection of PQDs. In [29] a classification method for multiple power quality disturbances using empirical WT adaptive filtering and SVM classifier is proposed

In this study, a new power system disturbance waveform pattern recognition and classification algorithm is proposed with unique feature vector which identifies and classifies the single and hybrid power quality disturbances into classes. The selected PQDs for analysis are frequent in power system operations. These disturbances are generated through parametric model equations in MATLAB workspace/editor environment. Such mathematical model based generated disturbances are then passed through DWT filters, signal processing tool, where not only event detection in time frequency localization is achieved but also feature extraction is performed from approximation and detail levels. Extracted features are used for training and testing of artificial intelligence based support vector machine classifier for the automatic pattern recognition. Real time power system model in Matlab/Simulink environment can also be developed for the physical description of power system disturbances and can be implemented for validation of results [12].

This paper has been organized into six sections. Section I, discussed the detail background of the power quality, including general types of power quality disturbances, issues, causes and general aspects of power quality research and related work. Section II describes wavelet transform in detail. In Section III, support vector machine is discussed. Section IV deals with the methodology of proposed approach in detail. Section V presents results and discussions. Lastly conclusion is presented in Section VI.

II. WAVELET TRANSFORM

With wavelet transform (WT) a waveform is decomposed into the set of basis functions; such basis functions are termed as mother wavelets. A wavelet is fundamentally a wave alternation, having zero mean unlike sinusoid which extends to infinity; it extends and exists for a finite duration. Analysis with WT is well suited to non-periodic signals containing both stationary components and transients, such as the ones that can be found on PQDs. Some well-known mother wavelet types are shown in Fig. 1. The availability of wide range of wavelet derivatives is a key strength of WT analysis. Daubechies 4 (db4) wavelet class is mostly adopted in literature for the analysis of PQ disturbances because of its comparably similar characteristics with PQ events [13]. WT has baseband characteristics in frequency domain. The major advantage of analysis with wavelets is of varying window size, which is wide for slowly varying changes i.e. low frequencies, and narrow for abrupt changes. As a result, there is optimal time-frequency localization in all frequency ranges [14, 15].

WT technique has a significant role in discontinuity/event detection and it has also been found a powerful tool for the feature extraction from waveforms. WT can be achieved by two means i.e. by continuous wavelet transform (CWT) and discrete wavelet transform (DWT). With CWT, scaling and translation of mother wavelet $\phi(t)$ provide the information of

time frequency resolution of the original distorted waveform [16,17]. The mathematical equation of CWT for a given disturbance signal $x(t)$ with respect to $\varphi(t)$ is given as.

$$CWT(c, d) = \frac{1}{\sqrt{c}} \int_{-\infty}^{\infty} x(t) \varphi\left(\frac{t-d}{c}\right) dt \quad (1)$$

In equation (1) c and d are real positive numbers, where c is the scaling factor, inversely proportional to frequency, corresponding to signal dilation or shrinking in time domain. While d is the translation factor corresponds to wavelet shifting.

Although CWT is upright with time frequency analysis but it has some limitations. Computations with CWT by computer simulation is discretized CWT, which is not a true discrete transform. CWT require infinite inputs and the information provided is highly redundant therefore not convenient for computer analysis [18]. Waveform decomposition or reconstruction by CWT requires a significant amount of computational time. Furthermore, CWT is considered substantially sluggish to implement as compared to DWT. The general equation of DWT for signal $x(k)$ is given in equation (2).

$$DWT(m, n) = \frac{1}{\sqrt{c_0^m}} \sum_k x(k) \varphi\left(\frac{n - kd_0 c_0^m}{c_0^m}\right) \quad (2)$$

In the above equation, c_0 and d_0 are discrete scaling and discrete translation factors respectively. Having fixed constant values generally, $c_0 = 2$ and $d_0 = 1$, these parameters, c_0^m and $kd_0 c_0^m$, are taken as constants. Where m and n are the integers, representing frequency localization and time localization, correspondingly. The parameter c_0^m produces oscillatory frequency and length of the wavelet, whereas parameter $kd_0 c_0^m$, credits shifting (translation) position [5].

The idea for DWT computation is same as it is in CWT. In CWT, correlation among a wavelet at different scale and the given signal is calculated by varying the analysis window scale, shifting the window in time, multiplying by the given signal and then integrating over all times. Whereas in Discrete WT, signal x is passed through series of digital high pass (H.P) filters, to analyze high frequencies, and digital low pass (L.P) filters, to analyze low frequencies, where filters are at different cut-off frequencies to evaluate signals at different scales. In this whole process, the resolution of the signal is changed by filtering and the scale is changed by sampling operations [5, 19, 20].

1) *Multi-Resolution analysis:* Mallat and Meyer established the basic framework of Multi resolution analysis (MRA) algorithm. With MRA, decomposition of a waveform can be obtained at various resolution levels and scales of short waveforms i.e. mother wavelets. In literature, MRA is also termed as pyramidal coding which is similar to sub band coding method. Using MRA, multi-level resolution analysis can be performed, where decomposition is repeated up to more than a few levels for increasing frequency resolutions to get detail and approximation coefficient waveforms. MRA decomposition can be mathematically modeled as.

$$(x * g)(n) = \sum_{k=-\infty}^{\infty} x(k)g(n - k) \quad (3)$$

In MRA based wave decomposition, the sample waveform being investigated is passed through half band LP filter $g(k)$, having impulse response g (equation 3). This causes the convolution in discrete time. Similarly, the waveform is also passed concurrently through half band HP filter $h(k)$. For the first level decomposition i.e. down sampling by factor 2, the outputs of the HP and LP filters are referred to as detail level D1 and approximation level A1 respectively. In level 2 decomposition the obtained approximation A1 coefficients are passed through the same HP and LP filters to produce coefficients A2 and D2 respectively. Likewise, A2 coefficients are again passed through filters of the same cut off frequency limits and so on. In this way down sampling is applied for further levels of wave decomposition [19, 21]. This process is largely termed as multi-level decomposition [22]. The filter output relations are mathematically expressed in equation 4 and 5, where k represents number of samples.

$$\varphi(n) = \sqrt{2} \sum_k g(n) \varphi(2n - k) \quad (4)$$

$$\varphi(n) = \sqrt{2} \sum_k h(n) \varphi(2n - k) \quad (5)$$

DWT-MRA based decomposition is according to the Nyquist rule, where half the frequencies of the signal have now been removed and half the samples are discarded, schematically shown in Fig. 2. Similarly with this decomposition process the frequency band i.e. half of each filter output characterizes the signal. Therefore time resolution is reduced by factor 2. The frequency resolution is doubled for each next level of decomposition because of half the frequency band of the input of the previous level and so on [12, 14, 19, 23].

Expressions for the approximation A_j and detail level D_j coefficients are:

$$A_{j+1}(n) = \sum_k h(k - 2n)A_j(k) \quad (6)$$

$$D_{j+1}(n) = \sum_k g(k - 2n)A_j(k) \quad (7)$$

The relation for waveform $f(n)$ expanded related to its orthogonal basis of scaling and wavelet function is shown in equation 8. The equation is basically characterized by one set of scaling coefficient and one or several sets of wavelet coefficients

$$f(n) = \sum_k A_1(k)\varphi(n - k) + \sum_k \sum_{j=1} D_j(n)2^{-\frac{j}{2}}\varphi(2^j n - k) \quad (8)$$

In equation 6, 7 and 8, $j=1,2,3,\dots$, represents level of decompositions.

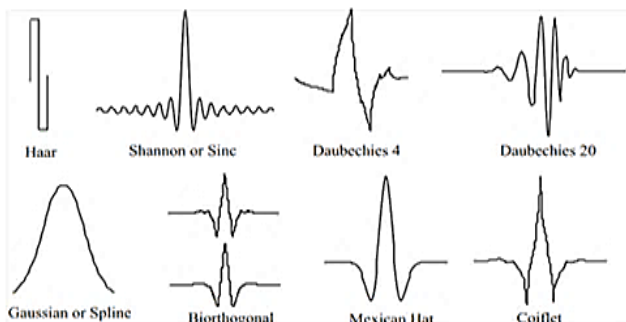


Fig 1. Some Common Types of Mother Wavelets.

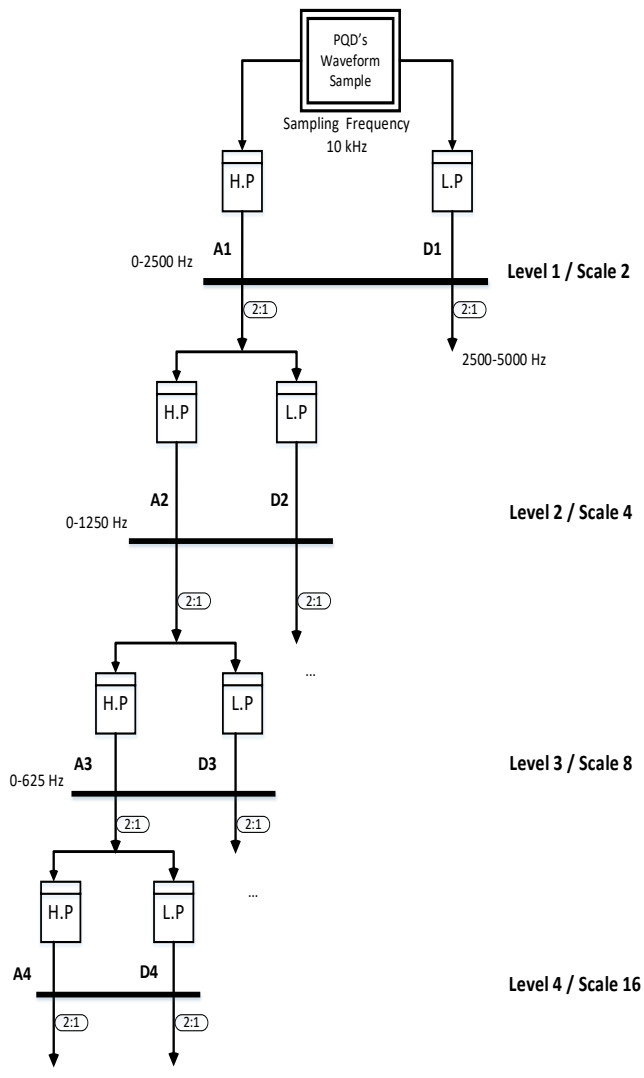


Fig 2. DWT-MRA Decomposition Process Schematic Diagram.

III. SUPPORT VECTOR MACHINE CLASSIFIER

The Support vector machine (SVM) tool, first presented by Vapnik (Vapnik, 1995), is a very powerful, high performance and computationally efficient family of supervised machine learning algorithm. It has wide application in classification and regression (time series prediction like estimation, forecasting, etc.) problems [8, 11]. For PQD waveform pattern recognition,

classification can be performed by utilizing various parameters. In literature, PQD classification is mostly based on statistical learning theory results [5, 8]. For two or more categorized classes of disturbance data, it acts as discriminative classifier typically defined by an optimal hyper plane, separates all the categorized classes by the decision boundary, as shown in Fig. 3. The hyper plane can be defined mathematically as:

$$g(x) = x'\theta + b = 0 \quad (9)$$

$\theta \in R$, for dimension d , comprises the coefficients expressing orthogonal vector to hyper plane.

Hyper plane is a linear decision boundary that splits the space for classification into two parts. Similarly, it can be a nonlinear decision surface boundary for classifying multiple classes data [24, 25]. For inseparable and complex data kernel functions are adopted. These functions transform data to large dimensional feature space where input data becomes more separable i.e. maximum margin hyper planes are established, related to original input feature space. Gaussian or radial basis function, sigmoid, polynomials, exponential radial basis functions, splines Etc. are the generally used types of Kernel functions in the literature [24]. In this work, Gaussian kernel is adopted for binary classification and feature mapping. The mathematical relation for Gaussian kernel is given as:

$$f(x_i, l_j) = \frac{\left(\exp - \left(\|x_i - l_j\|^2\right)\right)}{2\sigma^2} \quad (10)$$

In equation 10 x_i represents feature and l_j , land mark point whereas σ is a Gaussian kernel parameter, features $f(x_i, l_j)$ to vary more smoothly. For the classification let the training vector is $x_j \in R$, along with their categories $y_j = (-, +)$ where algorithm searches maximum margin length i.e. the region which contains no observations, for an optimal hyper plane and places the observations in the positive and negative class categories. From equation 9, for Separable classes classification, an objective is to minimize $\|\theta\|$ with respect to b and θ . So that for all feature vectors (x_j, y_j) , $y_j g(x_j) \geq 1$. When support vectors x_j are on the boundary, $y_j g(x_j) = 1$. To optimize and minimize the objective, SVM algorithm uses Lagrange multiplier, as shown in equation 11. To minimize the equation 11, subject to $\sum a_j y_j = 0$.

$$0.5 \sum_{j=1}^a \sum_{k=1}^n a_j a_k y_j y_k x_j x_k - \sum_{j=1}^n a_j \quad (11)$$

where $a_j \geq 0$ and $j=1,2,\dots,n$.

$$a_j [y_i g(x_j) - 1 + \delta_j] = 0 \quad (12)$$

$$\delta_j (C - a_j) = 0 \quad (13)$$

Whereas,

$$g(x_j) = \varphi(x_j)' \theta + b \quad (14)$$

In above equation φ is a kernel function. δ_j is entitled slack parameter and C is regularization parameter. For perfectly separable classes slack parameter, $\delta_j = 0$.

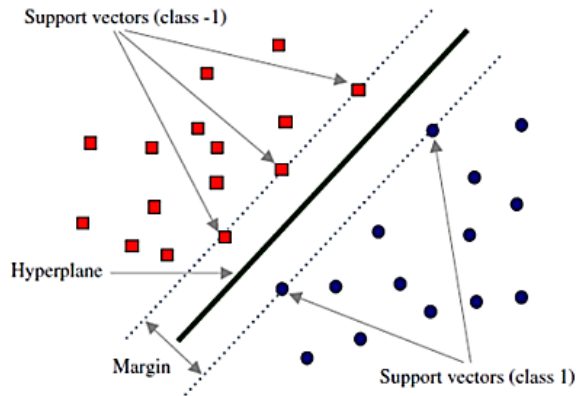


Fig 3. Support Vector Machine Classification Schematic.

In case of inseparable classes, minimizing equation 15 with respect to θ, b and δ_j , subject to equation 16.

$$0.5\|\theta\|^2 + C \sum \delta_j \quad (15)$$

$$y_j g(x_j) \geq 1 - \delta_j \quad (16)$$

Where $\delta_j \geq 0$, except $0 \leq a_j \leq C$, for all $j = 1, 2, 3, \dots, n$. SVM Score function is shown in equation 17 as:

$$\hat{g}(x) = \sum_{j=1}^n \hat{a}_j y_j x'_j + \hat{b} \quad (17)$$

Where \hat{b} is the bias estimate and \hat{a}_j is j th vector estimate. The SVM classifies new observation z using $\text{sign}(\hat{g}(z))$. Non-linear boundary in SVM works in transformed predictor space to get optimal hyper plane. The dual formalization for nonlinear SVM is represented in equation 18 with respect to a_1, a_2, \dots, a_n , subject to $\sum a_j y_j = 0$, where $0 \leq a_j \leq C$.

$$0.5 \sum_{j=1}^n \sum_{k=1}^n a_j a_k y_j y_k G(x_j, x_k) - \sum_{j=1}^n a_j \quad (18)$$

For all $j=1, 2, \dots, n$. and the KKT complementarity conditions. $G(x_j, x_k)$ are the elements of the Gram matrix. The resulting score function for SVM is given in equation 19.

$$\hat{g}(x) = \sum_{j=1}^n \hat{a}_j y_j G(x_j, x_k) + \hat{b} \quad (19)$$

SVM has better generalization ability (i.e. ability to learn a rule for classifying training data to ability of resulting rule to classify testing data.). It may handle large feature vector dimension space and also has no over fitting issue for large classification problems as compared to logistic regression and neural network or other conventional classifiers. SVM training is much easier than training artificial neural networks [1, 2].

IV. METHODOLOGY

The proposed algorithm for PQD recognition comprises of four stages, namely, PQ disturbances sample data generation, Feature extraction, classification and decision stage, as shown in flow chart, Fig. 4.

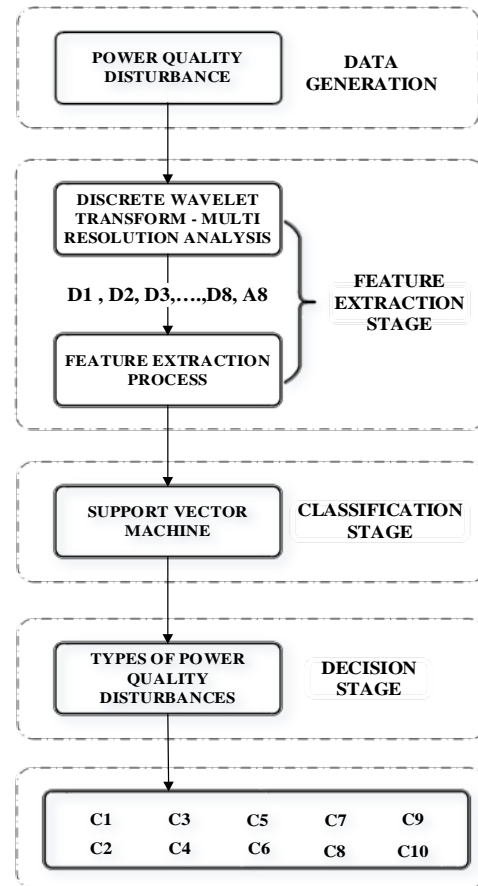


Fig 4. Flow Chart of Proposed Methodology.

A. PQ Disturbances Pattern Generation

Real Power quality waveforms may frequently exhibit slowly or abrupt changing trends, oscillations punctuated with harmonics, transients or other disturbances. In contrast, these changes are the important part of the data both perceptually and in terms of information of abnormality they provide. For the classification of disturbances, PQDs data can be obtained through real time PQ loggers or can be generated using parametric equations [25], where equation parameters are based on Categories & Characteristics of power system electromagnetic phenomenon, IEEE STD 1159-2009. PQD generation through parametric equations is expedient, variety of samples for any type of disturbance either single or multiple signals can be simulated. Disturbance magnitude and duration over cycles can be changed in a wide range and controlled manner according to disturbance type and IEEE standards. In this work events of pure voltage sine wave, sags, swells, interruptions, harmonics, impulsive transients, oscillatory transients, sag with harmonics and swell with harmonics are considered for generation. 100 random sample cases with 10 cycles at 50Hz (0.2 seconds) for each disturbance were produced. The sampling frequency was 10 kHz i.e. 200 points per cycle. Both magnitudes, as well as time of event occurrence, were diverse in accordance with the aforementioned standard. Fig. 5 shows only one random sample waveform of each category of generated disturbance.

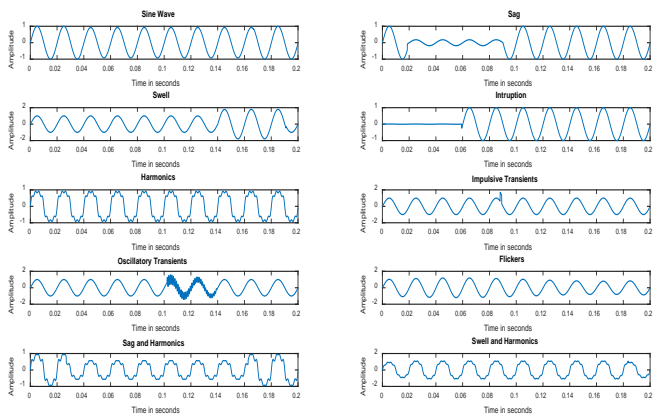


Fig 5. Generated Sample of Each Selected PQD Type through Parametric Equations.

B. Feature Extraction

Feature extraction stage is the most important step for the machine learning based pattern recognition and classification of PQD problems. Extracted features are the measured data, obtained from the waveform samples to develop a feature vector. This feature vector should be dimensionally concise so that learning and generalization process in classifier algorithms for classification can be implemented effectively. Feature extraction stage consists of two sub stages. In the first stage, all the generated samples for each disturbance class are decomposed up to 8 levels to get wavelet coefficients using DWT-MRA, where the wavelet coefficients are A_j approximation and D_j detail levels. Thus for each class of disturbance D_1 - D_8 and A_8 coefficients vectors are extracted. In order to reduce the obtained data and to enhance the classifiers performance effective and suitable statistical parameters are proposed for feature vector development in the second stage of feature extraction.

1) *Energy feature*: The energy feature mining is according to the property which states that energy of the signal in time domain and in the frequency domain are equal, as frequency domain signal $X[n]$, i.e. Fourier transformed signal contains all the information about $X[t]$. This property is termed as Parseval's Theorem.

$$E_{sig}(t) = \frac{1}{T} \int_0^T |x(t)|^2 dt \quad (20)$$

$$E_{sig}(n) = \frac{1}{N} \sum_{j=1}^N |X[n]|^2 \quad (21)$$

Where t denotes the time period and n is the sampling period of the signal waveform. The energy features of the PQDs are obtained from wavelet coefficients A_j and D_j , which are obtained at various frequency bands for each of the disturbance types. The energy feature vector consists of energy percentage corresponding to the respective wavelet coefficients, which are calculated by the relations shown in equation 22 and 23.

$$E_{D_i} = \sum_{j=1}^N |D_{i,j}|^2 \quad (22)$$

$$i = 1, 2, 3, \dots, l$$

$$E_{A_i} = \sum_{j=1}^N |A_{i,j}|^2 \quad (23)$$

$$E_i = [E_{A_1} E_{A_2} \dots E_{A_l} E_{D_1} E_{D_2} \dots E_{D_l} E_{D_l}] \quad (24)$$

E_{A_j} and E_{D_j} are the energies of wavelet-approximation and detail coefficients up to level j and E_i is the energy feature vector.

2) *Entropy feature*: Entropy parameter has been found suitable as a feature for PQD classification and recognition. In information theory entropy is generally regarded as the precise indicator of disorder ness, imbalance and uncertainties relating to random variables that may be gained by the observations (in this case D_i and A_8), whereas entropy is always greater than or equal to zero. Its outcomes can be generalized to provide information about specific events and outliers. Shannon entropy, a decreasing functions of a scattering of random variables and is maximum when all outcomes are equally to be expected. Following are the relations for Shannon entropy for detail and approximation level coefficients.

$$Ent_{D_i} = - \sum_{j=1}^N D_{i,j}^2 \log(D_{i,j}^2) \quad (25)$$

$$Ent_{A_i} = - \sum_{j=1}^N A_{i,j}^2 \log(A_{i,j}^2) \quad (26)$$

Where $A_{i,j}$ and $D_{i,j}$ is the probability of the occurrence of feature values $\{D_1, \dots, D_8, A_8\}$ and $i = 1, \dots, 10$. Over all entropy features obtained from the MRA based DWT for any PQ Signals are given by

$$Ent_i = [Ent_{D_1} Ent_{D_2} \dots Ent_{D_l} Ent_{D_l}, Ent_{A_8}] \quad (27)$$

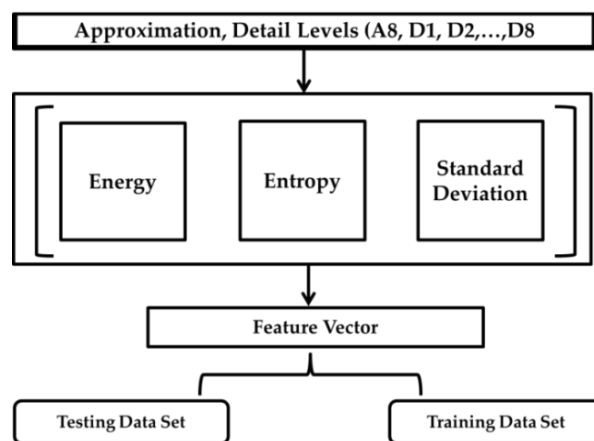


Fig 6. Feature Vector Development Flow Chart.

3) *Standard deviation*: Standard deviation feature measures the dispersion of an event frequency distribution or it can also be defined as a parameter that shows the way in which a probability function or probability density function is centered around its mean which is equal to the root of moment in which the deviation from mean is squared.

$$S.D_{D_j} = \sqrt{\frac{\sum_{j=1}^N (D_j - mean)^2}{N - 1}} \quad (28)$$

$$S.D_{A_j} = \sqrt{\frac{\sum_{j=1}^N (A_j - mean)^2}{N - 1}} \quad (29)$$

$$SD_i = E_i = [SD_{D_1} SD_{D_2} \dots SD_{D_i} SD_{D_i}] \quad (30)$$

$$Feature_i = [E_i Ent_i SD_i] \quad (31)$$

The feature vector obtained for i samples is used for the classification purpose in the classification stage for each of the PQDs type. Where half of the data is utilized for training the classifier and other half is used for testing the classifier performance. Schematic diagram for the feature vector development process is shown in Fig. 6.

C. Classification Stage

The $Feature_i$ feature vector, developed in equation 31, comprises of 27 dimensions of feature dataset for 100 samples of each of the PQ disturbance class i.e. 27×100 . From 100 samples of each disturbance class, half of the data set ($27 \times$

50) has been used for training the SVM classifier and rest of data is for testing purpose. For classification training with SVM one vs. one (1Vs.1) approach is adopted as shown in figure. Where in each SVM training node, $i=1$ class is trained against all classes. Similarly for next SVM training node that aforesaid $i=1$ class is replaced with $i=2$ class and training is done with all other classes. This process was iterated until all classes were passed through training. With this training process SVM develop algorithm functions i.e. $F(Cn)$, for binary data classification and outlier detection of n classes. Therefore 1 vs.1 approach may allow the SVM classifier to have a very upright training performance with this multi class classification problem. Testing of classifier, for each class, results the positive scores and negative scores for classified and misclassified class samples respectively. The label for classified disturbance sample was set to 1 and misclassified sample was set to 0. The input to the classifier is the time domain disturbance signal. Fig. 7 shows the one vs. one SVM binary classification schematic diagram. In this work, Lib-SVM Matlab Software toolbox library is used for classification. Lib-SVM consists of the default Sequential Minimal Optimization (SMO) solver; it reduces one-norm problem by a series of two point minimizations and contains bias terms and uses linear constraints in the model. RBF or Gaussian kernel was used for classifying non separable feature vectors i.e. nonlinear, where the kernel scale was set to 5.2 and box constraint to level 1 and extracted feature data was standardized for computational simplification.

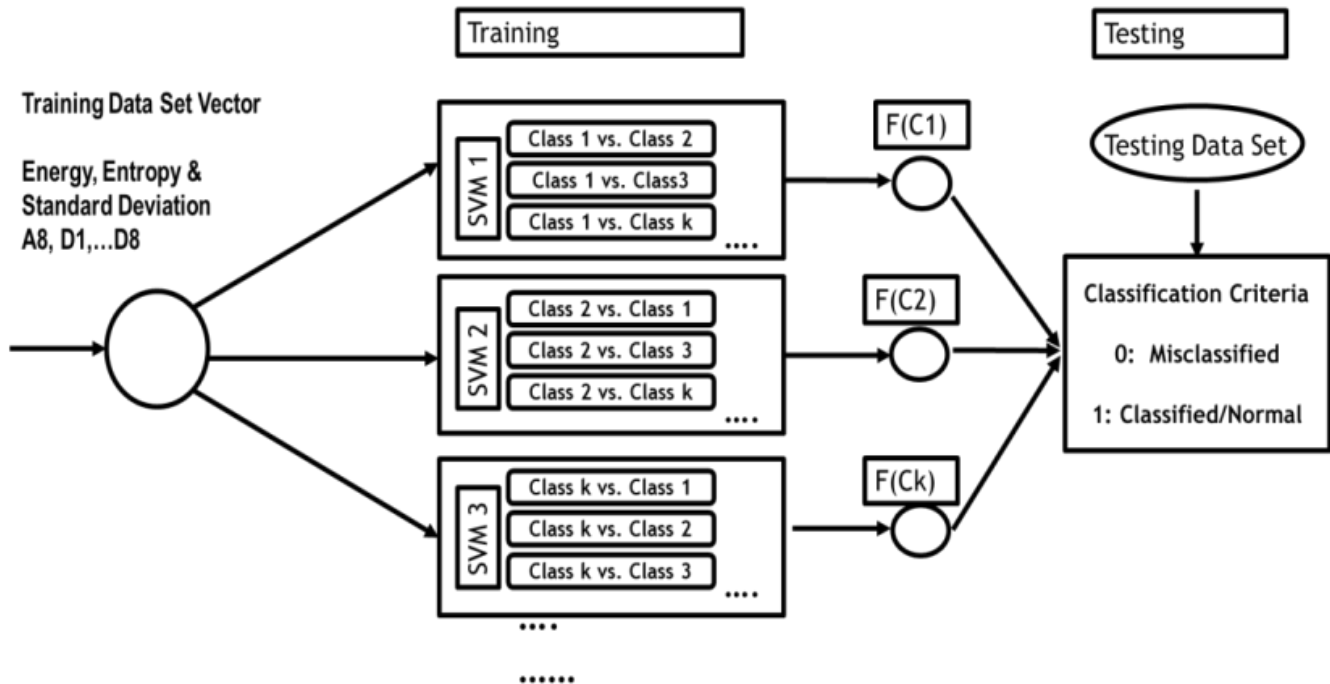


Fig. 7. One vs. One SVM Binary Classification Schematic.

VI. RESULTS AND DISCUSSION

For all that samples of ten generated disturbances, time frequency localization of each disturbance event is achieved and succeeded by the DWT-MRA technique. In order to show the DWT- MRA time frequency event localization characteristic, only one sample for sag with harmonics and its 1-5 detail, having high frequency, and 5th level discretized approximation coefficients, having the fundamental frequency components, are plotted in Fig. 8. These levels are obtained by decomposition of original generated sample via Daubechies fourth order (dB4) wavelet filter. For recognition and classification 1-8 details and 8th approximation coefficients are utilized for coefficients feature vector development. The first level signal has a frequency range of $f/2 - f/4$ i.e. 2500-5000Hz, where f is the sampling frequency. In the same way second level has frequency range of $f/4 - f/8$ i.e. 2500-1250 Hz. The third, fourth, fifth, sixth, seventh and eighth level decomposed disturbance signal have the frequency ranges of $f/8 - f/16, \dots, f/256 - f/512$ correspondingly.

To reduce the data set and to enhance the accuracy of classifier, statistical analysis is performed on coefficient vectors by pulling out energy, Shannon’s entropy and standard deviation information. Out of (27×100) of the samples of approximation and detail coefficient features half of samples i.e. (27×50) dimension of each disturbance are used for training SVM classifier. However, for testing the classifiers classification performance and evaluation of trained model, remaining extracted feature vectors are utilized in SVM predictor model for all the ten types of PQ disturbances. Such disturbance classes are presented in confusion matrix, shown in Table I, where the matrix is constructed with actual classes vs. classifiers predicted classes.

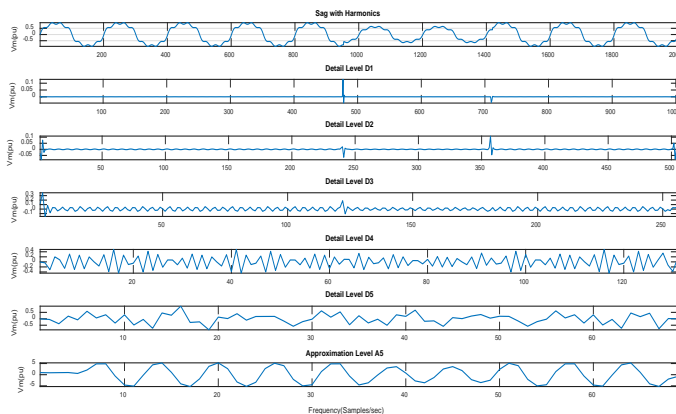


Fig 8. Generated Sag with Harmonics disturbance sample with its detail levels 1– 5 and 5th level Approximation coefficient waveforms.

TABLE I. CONFUSION MATRIX

		PREDICTED CLASSES									
		C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
ACTUAL CLASSES	C1	50	0	0	0	0	0	0	0	0	0
	C2	0	44	0	6	0	0	0	0	0	0
	C3	0	0	49	0	0	0	0	1	0	0
	C4	0	4	0	46	0	0	1	0	0	0
	C5	0	0	0	0	50	0	0	0	0	0
	C6	6	0	0	0	0	43	1	0	0	0
	C7	9	0	0	0	0	0	41	0	0	0
	C8	2	0	0	0	0	0	0	48	0	0
	C9	0	0	0	0	0	0	0	0	50	0
	C10	0	0	0	0	0	0	0	0	1	49

In Tables I, II and III, C represents labels for disturbance types which are specified as follows:

- C1** Sine Wave
- C2** Sag
- C3** Swell
- C4** Interruption
- C5** Harmonics
- C6** Impulsive Transients
- C7** Oscillating Transients
- C8** Flickers
- C9** Sag and Harmonics
- C10** Swell and Harmonics

The diagonal elements in confusion matrix represent true positives (TP) i.e. correctly classified disturbances, whereas off diagonal elements shows misclassified disturbance samples. From matrix, all rows elements except diagonal element are false negatives (FN) i.e. sample is classified as predicted class but actually it is not. Similarly, all the column elements except diagonal elements are false positives (FP). Classifier performance summary is shown in Table II, where positive rate of disturbance class refers sensitivity of classifier or recall rate. The overall accuracy is found 94%.

TABLE II. CLASSIFIER PERFORMANCE SUMMARY

Classes	T.P Rates (%)	FN Rates (%)	Positive Predictive Rate (%)	False Discovery Rate(%)
C1	100%	--	75%	25%
C2	88%	12%	92%	8%
C3	98%	2%	100%	--
C4	92%	8%	88%	12%
C5	100%	--	100%	--
C6	86%	14%	100%	--
C7	82%	18%	98%	2%
C8	96%	4%	98%	2%
C9	100%	--	98%	2%
C10	98%	2%	100%	--
Overall Accuracy 94 %				

TABLE III. CORRECTLY CLASSIFIED VERSUS MISCLASSIFIED SAMPLE RESULTS

Types Of PQDs	Testing Patterns	Correct (Classified)	Misclassified
C1	50	50	00
C2	50	44	06
C3	50	49	01
C4	50	46	04
C5	50	50	00
C6	50	43	07
C7	50	41	09
C8	50	48	02
C9	50	50	00
C10	50	49	01
Average	50	47	3

Tables I, II and III clearly shows that proposed algorithms has effectively classified the eight distinct and two hybrid PQDs. Performance of classifier is found up to mark. As fifty disturbances sample patterns for all ten disturbance classes were employed for testing purpose in SVM classifier. In consequence, an average of 47 samples out of 50 is found correctly classified whereas 3 samples are found as misclassified. However class wise correctly classified and misclassified samples numbers are also tabulated in Table III. The classification results show that proposed algorithm is effective and due to its simplicity and less computational requirements it is suitable and nominally applicable.

VII. CONCLUSION

This paper has presented an automatic machine learning based PQDs pattern classification by adopting statistically unique extracted features of Energy, Entropy and Standard deviation. These features have been calculated from the range

of one to 8th level decomposed wavelet coefficients for each of the randomly generated disturbance samples. Such Disturbances were obtained from IEEE standard limits based parametric equations. With proposed feature combination and distinctive disturbance types, proposed approach of using multi-resolution analysis based discrete wavelet transform and support vector machine algorithm provides the better classification results with distinct and multiple power quality disturbance classes, in spite of small training data set. Classification using optimization algorithms for optimal feature selections are more effective due to non-selection redundant features, but it requires more computational resources, time and complex simulations. Therefore with small data set and limited resources, DWT-SVM with Gaussian kernel provides the best classification in this case. The proposed work exhibits a promising agreement with simulation results.

ACKNOWLEDGMENT

This work is supported by Quaid-e-Awam University Campus of Engineering Science and Technology, Larkana, Sindh Pakistan. And Quaid-e-Awam University of Engineering Science and Technology, Nawabshah, Pakistan

Conflicts of Interest: The authors declare no conflict of interest.

REFERENCES

- [1] K. Thirumala, M. S. Prasad, T. Jain, and A. C. Umarikar, "Tunable-Q wavelet transform and dual multiclass SVM for online automatic detection of power quality disturbances," IEEE Transactions on Smart Grid, vol. 9, pp. 3018-3028, 2018.
- [2] H. Liu, F. Hussain, Y. Shen, S. Arif, A. Nazir, and M. Abubakar, "Complex power quality disturbances classification via curvelet transform and deep learning," Electric Power Systems Research, vol. 163, pp. 1-9, 2018.
- [3] S. Khokhar, A. A. B. Mohd Zin, A. S. B. Mokhtar, and M. Pesaran, "A comprehensive overview on signal processing and artificial intelligence techniques applications in classification of power quality disturbances," Renewable and Sustainable Energy Reviews, vol. 51, pp. 1650-1663, 2015.
- [4] S. Khokhar, A. A. Mohd Zin, A. S. Mokhtar, and N. Zareen, "Automatic Pattern Recognition of Single and Multiple Power Quality Disturbances," Australian Journal of Electrical and Electronics Engineering, 2015.
- [5] M. ElNozahy, R. El-Shatshat, and M. Salama, "Single-phasing detection and classification in distribution systems with a high penetration of distributed generation," Electric Power Systems Research, vol. 131, pp. 41-48, 2016.
- [6] F. A. Borges, R. A. Fernandes, I. N. Silva, and C. B. Silva, "Feature extraction and power quality disturbances classification using smart meters signals," IEEE Transactions on Industrial Informatics, vol. 12, pp. 824-833, 2016.
- [7] Z. Liu, Q. Hu, Y. Cui, and Q. Zhang, "A new detection approach of transient disturbances combining wavelet packet and Tsallis entropy," Neurocomputing, vol. 142, pp. 393-407, Oct 22 2014.
- [8] M. Biswal and P. K. Dash, "Measurement and Classification of Simultaneous Power Signal Patterns With an S-Transform Variant and Fuzzy Decision Tree," IEEE Transactions on Industrial Informatics, vol. 9, pp. 1819-1827, 2013.
- [9] S. Khokhar, A. A. Mohd Zin, A. P. Memon, and A. S. Mokhtar, "A new optimal feature selection algorithm for classification of power quality disturbances using discrete wavelet transform and probabilistic neural network," Measurement, vol. 95, pp. 246-259, 2017.
- [10] S. Salcedo-Sanz, J. L. Rojo-Álvarez, M. Martínez-Ramón, and G. Camps-Valls, "Support vector machines in engineering: An overview,"

- Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, vol. 4, pp. 234-267, 2014.
- [11] S. Khokhar, A. A. M. Zin, A. P. Memon, and A. S. Mokhtar, "A new optimal feature selection algorithm for classification of power quality disturbances using discrete wavelet transform and probabilistic neural network," *Measurement*, vol. 95, pp. 246-259, 2017.
- [12] J. Barros, R. I. Diego, and M. de Apráziz, "Applications of wavelets in electric power quality: Voltage events," *Electric Power Systems Research*, vol. 88, pp. 130-136, 2012.
- [13] M. Tuljapurkar and A. A. Dharme, "Wavelet based signal processing technique for classification of power quality disturbances," in *Proceedings - 2014 5th International Conference on Signal and Image Processing, ICSIP 2014*, 2014, pp. 337-342.
- [14] A. A. Abdoos, Z. Moravej, and M. Pazoki, "A hybrid method based on time frequency analysis and artificial intelligence for classification of power quality events," *Journal of Intelligent & Fuzzy Systems*, vol. 28, pp. 1183-1193, 2015 2015.
- [15] M. B. Latran and A. Teke, "A novel wavelet transform based voltage sag/swell detection algorithm," *International Journal of Electrical Power & Energy Systems*, vol. 71, pp. 131-139, 10// 2015.
- [16] P. Kanirajan and V. S. Kumar, "Power Quality Disturbance Detection and Classification using Wavelet and RBFNN," *Applied Soft Computing*, vol. 35, pp. 470-481, 2015.
- [17] 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," *International Journal of Electrical Power & Energy Systems*, vol. 49, pp. 95-103, 2013.
- [18] A. A. M. Z. Suhail Khokhar, M. A. Bhayo, A. S. Mokhtar, "Automated recognition of single & hybrid power quality disturbances using wavelet transform based support vector machine," *Jurnal Teknologi*, vol. 79, pp. 97-105, 2017.
- [19] S. Ekici, "Classification of power system disturbances using support vector machines," *Expert Systems with Applications*, vol. 36, pp. 9859-9868, 2009.
- [20] F. M. Arrabal-Campos, F. G. Montoya, R. Baños, J. Martínez-Lao, and A. Alcayde, "Simulation of power quality disturbances through the wavelet transform," in *Harmonics and Quality of Power (ICHQP)*, 2018 18th International Conference on, 2018, pp. 1-5.
- [21] N. Huang, S. Zhang, G. Cai, and D. Xu, "Power Quality Disturbances Recognition Based on a Multiresolution Generalized S-Transform and a PSO-Improved Decision Tree," *Energies*, vol. 8, pp. 549-572, Jan 2015.
- [22] S. Jeevitha and M. C. Mabel, "Analysis of Multi-Disturbance Power Signals on Compression: A Wavelet Based Approach," *Journal of Computational and Theoretical Nanoscience*, vol. 15, pp. 345-350, 2018.
- [23] D. De Yong, S. Bhowmik, and F. Magnago, "An effective Power Quality classifier using Wavelet Transform and Support Vector Machines," *Expert Systems with Applications*, vol. 42, 2015.
- [24] D. D. Ferreira, J. M. de Seixas, and A. S. Cerqueira, "A method based on independent component analysis for single and multiple power quality disturbance classification," *Electric Power Systems Research*, vol. 119, pp. 425-431, Feb 2015.
- [25] A. A. Abdoos, P. Khorshidian Mianaei, and M. Rayatpanah Ghadikolaei, "Combined VMD-SVM based feature selection method for classification of power quality events," *Applied Soft Computing*, vol. 38, pp. 637-646, 1// 2016G. Eason, B. Noble, and I. N. Sneddon, "On certain integrals of Lipschitz-Hankel type involving products of Bessel functions," *Phil. Trans. Roy. Soc. London*, vol. A247, pp. 529-551, April 1955. (*references*)
- [26] T. A. Kawady, N. I. Elkalashy, A. E. Ibrahim, and A.-M. I. Taalab, "Arcing fault identification using combined Gabor Transform-neural network for transmission lines," *International Journal of Electrical Power & Energy Systems*, vol. 61, pp. 248-258, 2014.
- [27] R. Kumar, B. Singh, and D. T. Shahani, "Recognition of Single-stage and Multiple Power Quality Events Using Hilbert-Huang Transform and Probabilistic Neural Network," *Electric Power Components and Systems*, vol. 43, pp. 607-619, 2015/04/03 2015.
- [28] R. Kumar, B. Singh, D. T. Shahani, A. Chandra, and K. Al-Haddad, "Recognition of Power-Quality Disturbances Using S-Transform-Based ANN Classifier and Rule-Based Decision Tree," *IEEE Transactions on Industry Applications*, vol. 51, pp. 1249-1258, Mar-Apr 2015.
- [29] K. Thirumala, S. Pal, T. Jain, and A. C. Umarikar, "A classification method for multiple power quality disturbances using EWT based adaptive filtering and multiclass SVM," *Neurocomputing*, 2019.