Scientific Research Publishing

Optimized Complex Power Quality Classifier Using One vs. Rest Support Vector Machines

David De Yong[1], Sudipto Bhowmik[2], Fernando Magnago[1,2]

[1]Universidad Nacional de Río Cuarto, Río Cuarto, Argentina
[2]Nexant INC, Chandler, AZ, USA
Email: deyongdm@gmail.com, sbhowmik@gmail.com, fernando.magnago@gmail.com

Abstract

Nowadays, power quality issues are becoming a significant research topic because of the increasing inclusion of very sensitive devices and considerable renewable energy sources. In general, most of the previous power quality classification techniques focused on single power quality events and did not include an optimal feature selection process. This paper presents a classification system that employs Wavelet Transform and the RMS profile to extract the main features of the measured waveforms containing either single or complex disturbances. A data mining process is designed to select the optimal set of features that better describes each disturbance present in the waveform. Support Vector Machine binary classifiers organized in a "One Vs Rest" architecture are individually optimized to classify single and complex disturbances. The parameters that rule the performance of each binary classifier are also individually adjusted using a grid search algorithm that helps them achieve optimal performance. This specialized process significantly improves the total classification accuracy. Several single and complex disturbances were simulated in order to train and test the algorithm. The results show that the classifier is capable of identifying >99% of single disturbances and >97% of complex disturbances.

1. Introduction

The significant increase in the inclusion of devices sensitive to current and voltage

fluctuations causes a growing interest in the study of power quality (PQ). A PQ event can be defined as a variation in the regular voltage or current waveform. Some of them can be classified as sags, swells, harmonics, fluctuations, interruptions, and over voltages. IEEE-1159 [1] specifies the characteristics that a waveform must have to be defined as a typical waveform, and classifies different types of disturbances.

The sources of disturbances are very broad, and cause economic losses as well as equipment degradation, for both consumers as well as utilities [2]. Therefore, it is imperative to employ tools to detect, classify and identify PQ events in order to mitigate these effects. Historically, PQ disturbances were analyzed and classified by visual inspection. Hence, the specialist's knowledge played a critical role in the classification and mitigation process. The development of digital measuring devices allowed one to have samples of the waveforms of voltage and current in selected measurement locations, however not all acquired data were useful and required large investment of time for proper root cause analysis. Therefore, it became important to have a tool to help in the process of continuous and automatic disturbance detection. Historically several techniques are used for detection and feature extraction. The more prevalent and effective techniques used are Fourier Transform (FT), Fast Fourier Transform (FFT) [2] [3], Gabor Wigner Transform (GWT) [4], S-Transform (ST) [5], Wavelet Transform (WT) [6], Wavelet Packet Transform (WPT) [7], Sinusoidal Filter method [8] and Kalman Filter (KF) [9].

In the online version of the tool once a PQ event is detected, a set of features are extracted from that waveform in order to reduce the size of the data. This is followed by a classification step in which the classification algorithm links a set of features with appropriate labels that represent the type of disturbance.

Learning techniques based on artificial intelligent (AI) methods are ideal for this kind of task due to their pattern recognition strength. Several classification algorithms that are appropriate for this are Artificial Neural Networks (ANN) [10], Markov Models [11], Fuzzy Logic (FL) [9] and Support Vector Machines (SVM) [12].

Due to the varying causes of power disturbances, it is not uncommon to have the two or more types of disturbances within a measured signal window. A disturbance that consists of a combination of two or more individual disturbances is usually called a complex power quality disturbance. Historically these complex disturbances have not been adequately addressed in previous research. Most of the previous work addressed the problem as an addition to single disturbances analysis but not as a particular problem [13]. Thus, the efficiency of properly classifying these types of disturbances varied widely and required systems that do not lend well to practical application.

For instance, the proposed approach is a multiclass SVM classifier arranged to operate in a One vs Rest architecture, designed to process information in parallel where each classifier defines one class. The main advantages of the proposed methods are: 1) Optimal feature selection, 2) independent parameter configuration

for each stage and each class present in the training set vector, 3) parallel data processing, 4) due the binary classifiers work independently of each other, there is no need to incorporate additional stages to classify complex power quality events.

This paper is organized as follows: Section 2 explains the concept of complex power quality disturbances and presents some previous works that focused on them. In the third Section, a general methodology to design, train and test an SVM classifier is presented. Section 4 explains the experimental test and their results. Finally, Section 5 presents the most important conclusions from the research.

2. Complex Power Quality Disturbances

A complex power quality event is a particular disturbance that comprises of a combination of two or more single disturbances. The most common complex disturbance is a combination of stationary disturbances such as harmonics or fluctuations with a short duration disturbance such as transient surges or sags. Figure 1 shows an example of this class of complex disturbance.

In addition, it is also possible to find a combination of short-duration disturbances, for example, transient surges combined with oscillating voltage dips. Figure 2 illustrates an example of this kind of power quality event.

It is also possible to find complex power quality disturbances as a combination of three or even more single disturbances.

Complex disturbances increase the difficulty during the identification stage due to the co-existences and overlapping of different disturbance characteristics. This complication may result in an incorrect characteristic determination.
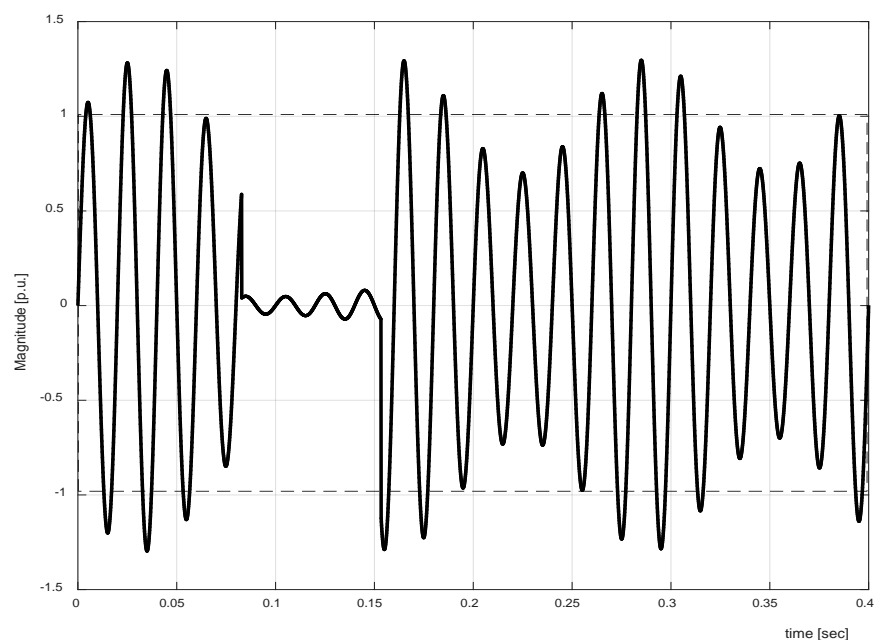

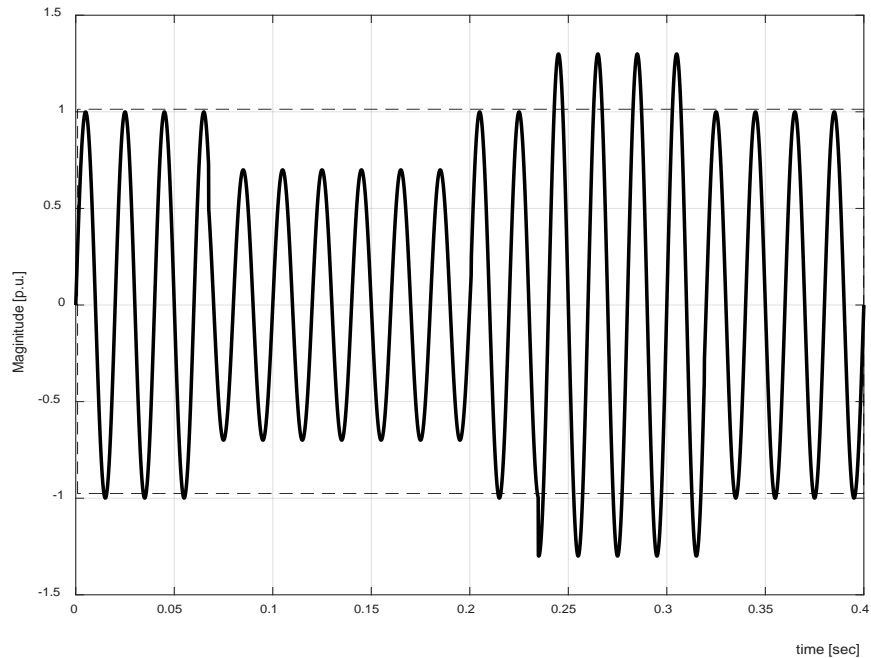**Figure 1.** Momentary interruption and flicker.

Figure 2. Sag and swell.

Some authors addressed these topics by mean different algorithms. For example, authors in [14], presents a comparison between a back-propagation based classifier with a multi-class One vs. One Support Vector Machine classifier. In this article, the SVM classifier rather than the back-propagation achieves better results for the same scenario with complex disturbances. This however depends on an accurate measurement algorithm on multiple nodes of the grid that is not always feasible due to limitations on the deployment of measuring devices and communications infrastructure.

Another alternative method is presented in [15], which proposes the analysis of the signals root mean square (rms) profile to distinguish between different types of PQ events. The identification of transient events is done using WT with four levels of decomposition and the method uses a dynamic ANN to classify harmonics and fluctuations. This method achieves a high percent of correct answers. Although the results are outstanding, the proposed architecture is troublesome. It uses a combination of multiple signal processing techniques and algorithms based on AI. This is typically hard to implement, coordinate and is computationally expensive. Additionally, since the algorithm is based on the WT first coefficient (D_1), it is highly affected by noise present.

Biswal, & Dash, [16], propose a methodology to extract the features based on the ST and a classification technique based on a decision tree. This approach uses seven decision steps to obtain the results and seems to achieve a very high accuracy level for a decision tree based classifier.

Reference [17] uses Tsallis singular entropy, energy entropy, and a modified incomplete ST to extract features, and a decision tree rule to classify single and

complex disturbances. This method achieves goods results. However, the classifier is implemented by a rigid programming structure and involves the calculation of a threshold for each node of the decision tree.

Contemporary research is increasingly using classifiers based on SVM due to their simplicity. L. Gang, & L. Fanguang present a method [18] that uses WT coefficients' energy combined with Principal Component Analysis (PCA) and Independent Component Analysis (ICA) in order to extract the main signal features. The classification is performed using SVM. This methodology becomes complicated mainly at the training stage. The method demonstrates that the PCA reduces the matrix dimension, hence, improves the classification stage performance.

Sovan Dalai [19] proposed a method base on Cross Hilbert-Huang Transform for parameter calculation, PCA to reduce the parameter set, and then a classifier based on a multiclass One vs. Rest SVM. This SVM classifier has a disadvantage of being difficult to train.

Reference [20] suggests a method based on the Ensemble Empirical Mode Decomposition (EEMD) technique to extract the signals features and a multi-label classification technique named Rank Wavelet Support Vector Machine. It preserves the correlation between different event types, improving the accuracy. However, to cover all characteristic of complex disturbances, the maxim decomposition level number of EEMD is set to 11, which increases the computational cost.

A strategy that is quite common, but not always the most appropriate when designing the classification stage, is to treat a complex power quality disturbances as a new type of event, assigning in consequence a new class to each type of complex disturbance [14] [16] [19]. The main disadvantage of this method is that it is necessary to pre-identify all the complex disturbances that may occur and then build the training and testing dataset. Any need to incorporate new disturbances (single or complex) requires the classifier to be re-designed and re-trained.

In addition, most of the previous proposals, the multiclass classifiers are implemented in a one-process unit. This architecture does not allow optimizing the feature extraction based on a particular class. Therefore, they need to use all features to describe all the classes. Furthermore, when it is necessary to add a new class for each additional complex event that wants to be identified, if the classifier is implemented in a one-process unit, the optimization problem could become even more complex.

Based on the current needs and an evaluation of the different methodologies previously presented it can be inferred that development of new algorithms that can handle both single and complex events, easy to implement as well as train, that allow a class-based optimization, and requires low computation cost is needed [13]. Consequently, the main contribution of this work is the development of a system that addresses the aforementioned needs.

3. Proposed Method

The proposed method can be explained in two separate stages:

- The design and training algorithm.
- The classification algorithm.

The summary of the different steps is presented in **Figure 3**. The next few sub-sections briefly explain the objectives of each sub-process that make up the two major stages.

3.1. Design and Training Algorithm

The design and training algorithm's main objective is to find the configuration that maximizes the classification's accuracy by optimizing the parameters that rule the behaviour of a classifier based on SVM algorithms. The algorithm's input corresponds to a training arrange that consists of the entire set of N disturbance classes that needs to be classified, for example swell, harmonics, sag, etc. The training set is represented by an [m,s] matrix where m is the number of waveforms and s the amount of samples that represent each waveform, parameter that depends on the selected sample rate and the configured length of the analysis windows.

The algorithm performs a series of calculations in order to extract the optimum set of features that better describes each class of disturbances and to obtain the best configuration of the parameters that govern the accuracy of the learning algorithm.

For more detail, **Figure 4** illustrates the Design and Training algorithm's flowchart.

These calculations are explained next:

3.1.1. Signal Processing

The objective of this process is to transform the training set waveform vector into equivalent representations in order to simplify the process of detecting the presence of a disturbance. Since the proposed method is focused on real, noisy signals, it is necessary to apply a de-noising technique to mitigate the effect of the noise in the sampled waveforms. Ref [21] demonstrates how a de-noising
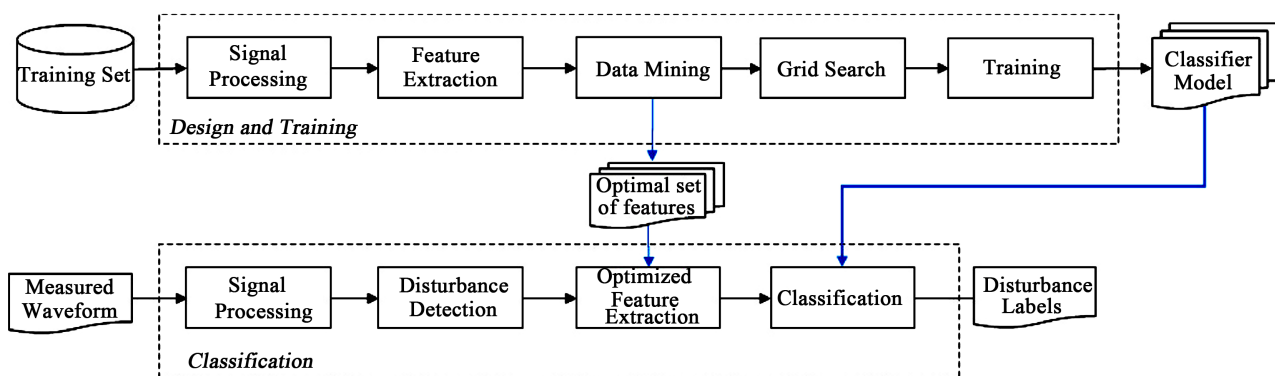

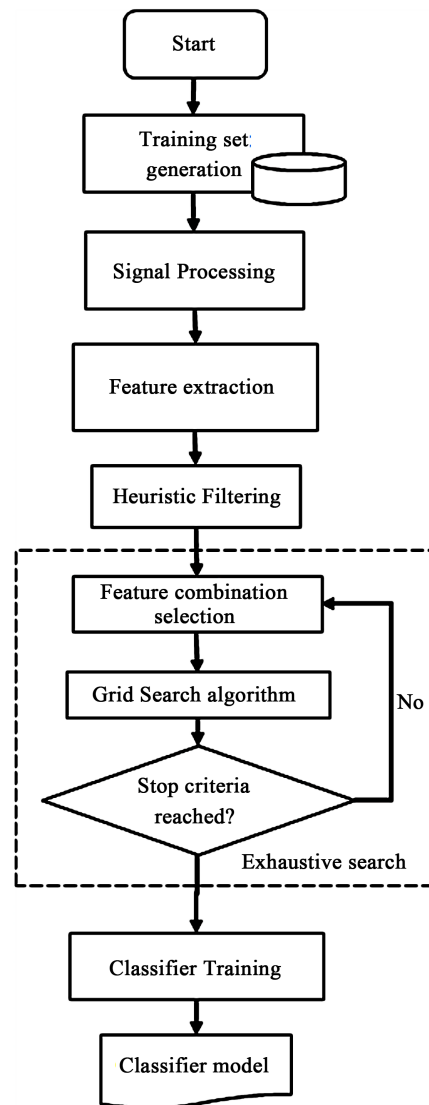**Figure 3.** Design, training and classification process.

Figure 4. Design and training algorithm's flowchart.

scheme improves the classifier ability. This method also narrows the signal duration to a fixed numbers of fundamental cycles and additionally establishes the best sample rate.

3.1.2. Feature Extraction

The information obtained from the sampling process of a representative waveform contains a high percentage of redundant, noisy, and inconsistent information. In order to reduce the data and yet maintain most of the information present in the waveform, feature extraction is typically performed on all the waveforms. A feature is a numeric value obtained from a transformation performed on either the waveform samples or the coefficients obtained from the selected signal processing technique. All the parameters are obtained with the objective of representing some particular characteristic of the original waveform [12] [13] [14] [36]. This is done in two stages.

The first stage's objective is to obtain the minimum set of features that characterizes each particular disturbance. It is important to remark that, at this phase of the algorithm, no information about a given class is used to calculate the features. They are selected to represent all classes present in the training set. The original training set, which dimension is [m, s], is reduced to an [m, n] arrangement, called the feature matrix, where $s >> n$. Since the features are obtained from diverse types of calculations, their dynamic ranges of the finalized parameters are very large. For this reason, the values of every column of the feature matrix are normalized to the region [−1; 1].

3.1.3. Data Mining

The feature extraction reduction procedure, presented in Section 3.1.2, is a process that generates an [m, n] matrix obtained from a set of signal waveforms, where n represents every extracted feature from each one of the m waveforms and is not dependent on or attuned to any particular class of disturbances.

The data mining process is the second stage used to reduce to the dimension of the training set [22]. In this step, the evaluation criteria used to reduce the feature selection, are closely related with every one of the N class included at the original feature set. A subset of j ($j < n$) features from the original n-dimensional set is obtained for each one of the N classes. This is explained in more detail in Section 4.3.3.

Two different techniques are sequentially applied to the training set in order to select an optimal feature subset: *The heuristic filtering* and the *exhaustive search algorithm*.

The exhaustive search algorithm involves the training of the classifier employing different feature set combination. The exhaustive search computational cost increases as the amount of features to be processed grows. To reduce the processing load of the exhaustive search algorithm a heuristic filtering stage is previously applied with the objective to separate the most relevant feature set from the original training set. The results of the filtering process serve as input to the exhaustive search algorithm. Next, both stages are briefly explained:

- Heuristic Filtering: The proposed filtering stage uses a label vector that maps each class to each row of the original feature matrix in order to calculate a feature ranking for each type of disturbance. The implemented methodology is based on Chi-square attributes feature selection [23], Relief-F attributes feature selection [24] and Symmetrical Uncertainty feature selection [25].

- These three heuristic techniques build a sorted list that ranks (from the highest to the lowest) which parameter describes a particular class the best. The algorithm combines the results and generates a unique ranking. Then, according to the user criteria, the j most relevant features are selected for each class. Because of this process, N different feature matrices (whose dimensions are [m, j] are generated from the [m, n] original feature matrix.

- Exhaustive Search Algorithm: In order to find the optimal combination of the features, an exhaustive search strategy is implemented. It consists of testing

the performance of the classifier for all 2^j possible feature combinations. The j features explored by the algorithm are the ones generated at the heuristic filtering stage. This method is known as a wrapper algorithm [26] because it uses the classification algorithm as part of the feature selection process. The algorithm chooses a combination of the *j* features and invokes the grid search algorithm (described in the next section). The grid search algorithm returns the best classification accuracy obtained for this feature set and the combination of the classifier parameters that produce the best performance. Then, the exhaustive search algorithm selects a new feature combination and repeats the calculations. The process is repeated until all combinations are tested. The final output of this stage is a table that contains the accuracy and the classifier parameters for each one of all 2*j* possible feature combinations.

3.1.4. Grid Search Algorithm

A grid search algorithm [27] is a well-known technique that employs a cross-validation methodology to find the best combination of the parameters that govern the classifier. For example, the SVM classifier's behavior is ruled by a combination of two (rarely more) parameters: the box constraint parameter C, and some parameter related with the selected kernel.

3.1.5. Training

Finally, the classifier is configured with the results of the grid search algorithm and trained using the features selected by the data mining process.

A trained and optimized classifier model is the outcome of the of the design and training algorithm.

In the next section of this paper, we explain how the classification algorithm to identify a disturbance in a measured waveform will use this conceptual model.

3.2. Classification Algorithm

The objective of the classification algorithm is to process a waveform, detect the presence of a disturbance, indicate when the disturbance starts (only for short time disturbances) and classify them into a predefined group.

Figure 5 shows the Classification algorithm's flowchart that consists of a series of processes that are explained below.

3.2.1. Signal Processing

The signal processing techniques applied to train the classifier must be the same as that implemented in the Design and training algorithm (Section 3.1.1).

3.2.2. Disturbance Detection

The purpose of this module is to detect the presence of an abnormality in the sampled signal and identify the instant when the power quality disturbance event begins or ends. If no disturbances are detected the classification algorithm discards all samples obtained from the measured waveforms.

Several methods have been developed to detect a disturbance in a waveform. Methods in reference [26] were used in our classifier.
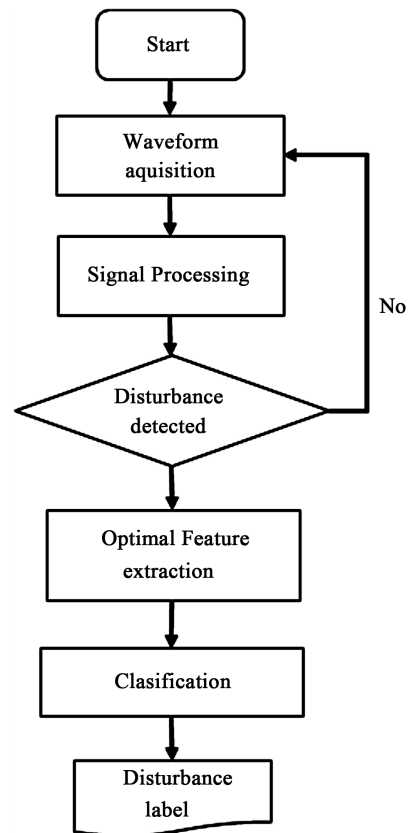
Figure 5. Classification algorithm's flowchart.

3.2.3. Optimized Feature Extraction

The objective of this step is similar to the method presented in Section 3.1.2. However, this process only extracts the optimum feature set according to the Data Mining process results calculated in the *Design and training stage*. This reduced subset of features allows faster computation and thus ideal for real-time implementation.

3.2.4. Classification

This process uses the trained classifier model obtained from the design and training algorithm, to categorize the set of features extracted in the previous classification algorithm's stages. This results in a label that indicates which class the measured disturbed waveform belongs to.

4. Experimental Results

To test the proposed algorithm with Complex Power Quality disturbances a One vs. Rest of five binary SVM classifiers is developed. This section is organized in the following way: The first subsection presents the classifier architecture. Then, a description of the training set used to train and test the classifier is provided. The third subsection presents the techniques employed in the Design and Training algorithm and the respectively obtained results. Finally, the fourth subsection presents the classification results.

4.1. Classifier Architecture

A kernel-based methodology called Support Vector Machine (SVM) is selected to build the classifier. Support Vector Machine mathematical theory can be found in [28] and [39].

Different machine learning methods were considered for classification stage: Support Vector Machine (SVM), Probabilistic Neural Network (PNN) and Extreme Learning Machine (ELM).

SVM method is selected mainly because: It has a strong founding theory; In general, the optimization problem involved in the training reaches the global optimum due to convex quadratic programming; It has no issue for choosing a proper number of parameters; It is less prone to over fitting; Yields more clear results and a geometrical interpretation; Since SVM is trained using dual representations and sparse arrays it is very efficient.

According to [29] [30] [31] SVM performs better than PNN and algorithms based on k-nearest neighbor.

In [32] a comparative study between SVM and ELM is performed. According to the author both methods have an outstanding generalization ability but SVM performs better when the training set is small. That is an important attribute in Power Quality problems where it is not easy to have a big database of measured disturbances to configure a training set.

Another comparative study concludes that ELM and SVM have similar accuracy performance for the most classification problems [33]. According to the author, running times on small datasets show that SVM is the fastest method.

In [34] a comparison between ELM and SVM over a particular area of classification, *i.e.* text classification, is conducted. The results of benchmarking experiments with SVM show that for many categories SVM still outperform ELM.

To test the proposed method, five binary Support Vector Machine classifiers configured in a One vs. Rest architecture is set up as shown in **Figure 6**.
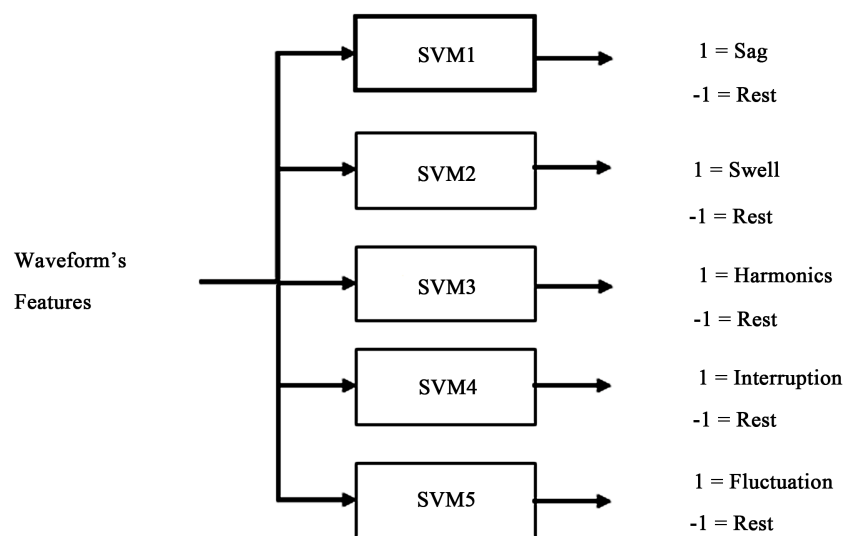

Figure 6. SVM one vs rest classifier.

Due to their unique architecture, the classifier can classify single as well as complex disturbances without the need of adding new binary classifiers.

The selected kernel function for each of the five SVM binary classifiers is the Radial Basis Function (RFB) because it proves to be the most appropriate function for pattern recognition [35]. The parameters that rule the SVM training are *C*, also called *box constraint*, and *Sigma*, which governs the kernel function mapping behavior. The simultaneous configuration of both parameters rules the classifier's accuracy rate.

4.2. Training Set Configuration

To train the classifier, 2600 disturbances were generated using a MATLAB tool developed by the authors [36].

Table 1 and Table 2 summarize the distribution of the training set and the labels assigned to each one of the binary classifiers presented in Figure 6.

Table 1. Single disturbances training set.

Single Quality Event	Total	Class	Sag	Swell	Harmonics	Interruption	Fluctuations
Sag	400	1	1	−1	−1	−1	−1
Swell	400	2	−1	1	−1	−1	−1
Harmonics	400	3	−1	−1	1	−1	−1
Interruption	400	4	−1	−1	−1	−1	−1
Fluctuations	400	5	−1	−1	−1	1	1

Table 2. Complex disturbances training set.

Complex Power Quality Event	Total	Class	Sag	Swell	Harmonics	Interruption	Fluctuations
Harmonics and Sag	100	3 + 1	1	−1	1	−1	−1
Harmonics and Swell	100	3 + 2	−1	1	1	−1	−1
Harmonics and Interruption	100	3 + 4	−1	−1	1	1	−1
Fluctuations and Sag	100	5 + 1	1	−1	−1	−1	1
Fluctuations and Swell	100	5 + 2	−1	1	−1	−1	1
Fluctuations and Interruption	100	5 + 4	−1	−1	1	−1	1

While there is a wide range of disturbances, to simplify the analysis, only a subset that contains the most common types of disturbances is considered in this paper.

4.3. Results of the Design and Training Stage

The following subsections explain the details of the techniques used and the associated results for each process of the stages shown in Figure 3.

4.3.1. Signal Processing

To process the simulated or measured waveforms, the sample rate is configured to 10 [Kilo sample/sec]. Snapshots of 400 ms are used to for each waveform's length, which is equivalent to 20 cycles of an undisturbed signal (assuming a fundamental frequency of 50 Hz).

Before the selection of Wavelet Transform (WT) as the signal processing methodology, other alternatives were studied. For example Stockwell Transform (ST) and Gabor Transform (GT). Previous work determined that these two signal processing methods perform very well with signals, which include noise. However, WT is better in term of simplicity and computational cost, therefore WT was selected for the signal processing stage [2].

A nine-level Discrete Wavelet Transform (DWT) using Daubechies number four wavelet mother was selected [37].

To complete the set of relevant features, the root mean square profile calculation is also proposed.

4.3.2. Feature Extraction

The feature extraction algorithm calculates the signals rms. profile as well as the nine DWT coefficients of the 2600 waveform of the training set to obtain the parameters presented in Table 3. Subsequent stages are used to reduce the number of features that are needed to represent each type of disturbance.

Table 3. Complete set of features.

Signal Processing Technique	Extracted Feature	Nomenclature
RMS Profile	Minimum RMS Profile	Min RMS
	Maximum RMS Profile	Max RMS
Wavelet Transform	Maximum d_i module coefficient	Max abs d_i
	Standard Deviation d_i coefficient	Std d_i
	Normalized Energy d_i coefficient	d_i Energy
	Maximum a_9 module coefficient	Max abs a_9
	Standard Deviation a_9 coefficient	Std a_9
	Normalized Energy a_9 coefficient	a_9 Energy

Where i represents the ith calculated wavelet level.

As an output of this process, a [2600, 32] matrix is obtained. This matrix contains all features that characterize each type of disturbance.

4.3.3. Data Mining

This section presents the reduced selected features using the techniques elaborated in Section 3.1.3. Table 4 illustrates the results obtained for the heuristic filtering process.

From the original [2600, 32] feature matrix, five matrixes were obtained, one for each class of disturbances, whose dimension are equal or less than [2600, 7].

After the number of features is significantly reduced by the filtering stage, it is important to find which combination of them produces the most accurate percentage in the training and validation stage. Table 5 shows the results of the exhaustive search algorithm presented in Section 3.1.3.

To train and test the algorithm performance, 60% of the 2600 disturbances are used for the supervised training of the classifier, while the remaining 40% are employed for the validation process.

4.3.4. Grid Search Algorithm Results

The results of the grid search algorithm are presented in Table 6. It shows the best parameter combinations that govern each binary SVM stage with the achieved validation accuracy. These parameters combinations are obtained for the feature combination presented in Table 5.

Once the best set of features that represent each disturbance and the optimum parameters *C* and *Sigma* that govern each binary SVM classifier is found, the design stage is concluded. Then, each binary classifier is trained using the LibSVM library [38].

Table 4. Heuristic feature filtering results.

PQ Event	Selected Feature	Dimension
Sag	Min RMS, Max RMS, Max abs a_9, d_7 Energy, Max abs d_5, Std d_4.	2600×6
Swell	Min RMS, Max RMS, Std d_9, d_7 Energy, Max abs d_4, Max abs d_3, Max abs d_2.	2600×7
Harmonics	Min RMS, Max RMS, Std d_9, Max abs d_9, Std d_8, Max abs d_5.	2600×6
Interruption	Min RMS, Max RMS, Max abs a_9, Std d_9, Max Abs d_4.	2600×5
Fluctuations	Min RMS, Max RMS, Max Abs a_9, Std d_9, Max abs d_9, Std d_8, Max abs d_8.	2600×7

Table 5. Exhaustive search algorithm results.

PQ Event	Selected Feature	Validation Accuracy	Dimension
Sag	Min RMS, Max RMS, Max abs a_9, Std d_4.	98.8%	2600×4
Swell	Min RMS, Max RMS, Std d_9, Max abs d_2.	99.6%	2600×4
Harmonics	Std d_8, Max abs d_5.	100%	2600×2
Interruption	Min RMS.	100%	2600×1
Fluctuations	Min RMS, Max RMS, Max Abs a_9, Std D_8.	99.7%	2600×4

Table 6. Grid Search Algorithm results.

SVM Classifier	Sigma	C	Validation Accuracy
1. Sag	2	4000	98.8%
2. Swell	2	1000	99.6%
3. Harmonics	3.5	3000	100%
4. Interruption	$1e^{-5}$	50	100%
5. Fluctuations	4	500	99.7%

4.4. Results of Classifier Algorithm

To test the classifier architecture designed and optimized by the process presented in Section 4.3, two scenarios are used. In the first scenario, the classifier is tested using a set of single disturbances. On the other hand, the second scenario tests the classifier with a set of complex power quality disturbances.

4.4.1. Scenario 1: Single Power Quality Events

Although this paper focuses on complex disturbances analysis, first at all, it is necessary testing the algorithm performance with simple disturbances.

To test the algorithm, 1000 waveforms are generated, 200 for each type of PQ events. All parameters that govern the disturbances, like magnitude, inception angle, duration, among others, are randomly generated considering the ranges established in [1].

The confusion matrix represented in Table 7 shows the calculated results.

Analyzing Table 7, it can be concluded that the designed classifier performs significantly well because it can correctly classify more than 99.7% of the proposed single disturbances.

One dataset from the harmonics and interruption set are partially classified as a complex disturbance containing the respective single disturbance. This may be inferred as a partially correct classification.

4.4.2. Scenario 2: Complex Power Quality Events

To test the algorithm for a complex power quality scenario, a set 1200 waveforms are generated with a combination of simulated waveforms with real waveforms measured in an oil factory [39]. Similar to scenario 1, the parameters that govern the event are randomly selected.

The results are summarized in the matrix presented in Table 8.

The values displayed with parenthesis () refer to the event index described in Table 7.

Considering a total of 1200 complex power quality events used to test the algorithm, only 33 were misclassified giving a success rate of 97.25%.

Analyzing the erroneous classification data set, the classifier was capable of identifying one of the two disturbances that was present in the complex event and thus was partially classified. In other words, 2334 disturbances, from 2400, were correctly classified. Under this consideration the complex power quality accuracy rate reach the 98.583%.

Table 7. Single power quality classification results.

	Sag	Swell	Harmonics	Interruption	Fluctuations	Misclassified	Double
1. Sag	200	0	0	0	0	0	0
2. Swell	0	200	0	0	0	0	0
3. Harmonics	0	0	198	0	0	1	1 (2 + 3)
4. Interruption	0	0	0	199	0	0	1 (4 + 3)
5. Fluctuations	0	0	0	0	200	0	0

Table 8. Complex power quality classification results.

	Harm. and Sag	Harm. and Swell	Harm. and Int.	Fluct. and Sag	Fluct. and Swell	Fluct. and Int.	Single	Triple
Harmonics and Sag	197	0	0	1	0	0	1 (3)	1 (1 + 3 + 5)
Harmonics and Swell	0	198	0	0	0	0	1 (3)	0
Harmonics and Interruptions	0	0	194	0	0	1	4 (3)	1 (3 + 4 + 5)
Fluctuation and Sag	0	0	0	190	0	0	10 (5)	0
Fluctuations and Swell	0	0	0	0	189	0	11 (5)	0
Fluctuations and Interruptions	0	0	0	0	0	199	1 (5)	0

4.4.3. Comparative Results

The accuracy to identify complex power quality disturbances of different methodologies is compared in Table 9.

5. Conclusions

This paper proposes a simple, efficient, fast and easily trainable method to classify single and complex power quality disturbances. The methodology is based on a combination of the Discrete Wavelet Transform (DWT) and the *rms profile* of each of the measured disturbances for *feature extraction*: a two-stage method to select the optimum set of representative features that reduce the feature set considerably maximizing the accuracy of the classification. A One vs. Rest multiclass SVM classifier was developed as a binary node array, and it was used to classify the extracted features.

The proposed methodology does remarkably well in classifying all single disturbances and outperforms most of the contemporary methodologies. The accuracy achieved exceeds those presented in [15] [16] [19] [20]. In addition, the designed method demonstrates that it is possible to identify a significant amount of complex power quality disturbances using only five binary decision stages (one for each single disturbance). This shows that complex disturbances need not be treated as separate classes like the classifiers presented in [14] [16] [19] but can be accurately classified with the same class as the single disturbance. Each binary classifier can be trained and optimized to distinguish both the single as well as the inclusive complex disturbance. This is one of the major contributions of this paper because it makes the classifier simpler, faster and easier to train.

Table 9. Comparative results.

Liu Z. [20]	EEMD	Rank-SVM	-	95.36
Biswall [16]	FDST	DT	-	98.19
Cheng Long Chuang [15]	WT	DS-ANN	ES	98.4875
Soval Dalai [19]	CHT	ESVM-PCA	-	97.6
Proposed Method	*WT*	*OvR SVM*	*ES*	**98.588**

EEMD = Ensemble Empirical Mode Decomposition; FDST = Fast Discrete Stockwell Transform; WT = Wavelet Transform; CHT = Cross Hilbert Transform; DT = Decision Tree; DS ANN = Dynamic Structural Neural Network; ICA = Independent Component Analysis; PCA = Principal Component Analysis; OvR SVM = One versus Rest Support Vector Machine; ES = Exhaustive Search.

This paper also demonstrates that excellent results can be achieved using a small set of features that are appropriately selected. The whole process can be parallelized because each node can be processed independently leading to faster computation times and thus ideal for online real-time implementation.

When a new complex power quality event needs to be included, the method has to be completely retrained to allow each classifier to consider the new event. This fact represents a weakness of the proposed method, which is shared with most of the algorithms based on linear learning. However, the classification remains robust even with increasing complexity of disturbances present in the signal compared to the ones presented in previous works [14] [16] [17] [19] even though for the 400 ms window of measurement it is relatively rare to have a significant number of events within the sampled signal.

Future work will focus on finding an optimum training set size that can be present and still provide acceptable results as well as overcoming the need for a full retraining in cases of newer exotic disturbances.

According to [33], SVM and ELM have similar accuracy results, therefore, the selection of the most appropriate machine learning algorithm is a problem dependent decision. Future works will focus on comparing the accuracy of both classifiers for Power Quality disturbance classification problem.

Acknowledgments

References

[1] IEEE Std 1159-2009—IEEE Recommended Practice for Monitoring Electric Power Quality. 26 June 2009, c1-81. http://ieeexplore.ieee.org/document/5154067/

[2] Roscoe, A.J., Burt, G.M. and McDonald, J.R. (2009) Frequency and Fundamental Signal Measurement Algorithms for Distributed Control and Protection Applications. *Generation, Transmission & Distribution*, **3**, 485-495. https://doi.org/10.1049/iet-gtd.2008.0517

[3] Heydt, G.T., Fjeld, P.S., Liu, C.C., Pierce, D., Tu, L. and Hensley, G. (1999) Applications

of the Windowed FFT to Electric Power Quality Assessment. *IEEE Transactions on Power Delivery*, **14**, 1411-1416. https://doi.org/10.1109/61.796235

[4] Cho, S.-H., Jang, G. and Kwon, S.-H. (2010) Time-Frequency Analysis of Power-Quality Disturbances via the Gabor-Wigner Transform. *IEEE Transactions on Power Delivery*, **25**, 494-499. https://doi.org/10.1109/TPWRD.2009.2034832

[5] Dash, P.K., Panigrahi, B.K. and Panda, G. (2003) Power Quality Analysis Using S-Transform. *IEEE Transactions on Power Delivery*, **18**, 406-411. https://doi.org/10.1109/TPWRD.2003.809616

[6] Santoso, S., Powers, E.J., Grady, W.M. and Hofmann, P. (1996) Power Quality Assessment via Wavelet Transform Analysis. *IEEE Transactions on Power Delivery*, **11**, 924-930. https://doi.org/10.1109/61.489353

[7] Liu, L.Y. and Zeng, Z.Z. (2008) The Detection and Location of Power Quality Disturbances Based on Orthogonal Wavelet Packet Transform. *Third International Conference on Electric Utility Deregulation and Restructuring and Power Technologies*, Nanjing, 6-9 April 2008, 1831-1835.

[8] Radil, T., Ramos, P.M. and Serra, A.C. (2008) Detection and Extraction of Harmonic and Non-Harmonic Power Quality Disturbances Using Sine Fitting Methods. *13th International Conference on Harmonics and Quality of Power*, Wollongong, 28 September-1 October 2008, 1-6.

[9] Dash, P.K. and Chilukuri, M.V. (2004) Hybrid S-Transform and Kalman Filtering Approach for Detection and Measurement of Short Duration Disturbances in Power Networks. *IEEE Transactions on Instrumentation and Measurement*, **53**, 588-596. https://doi.org/10.1109/TIM.2003.820486

[10] das Merces Machado, R.N., Bezerra, U.H., Pelaes, E.G., de Oliveira, R.C.L. and de Lima Tostes, M.E. (2009) Use of Wavelet Transform and Generalized Regression Neural Network (GRNN) to the Characterization of Short-Duration Voltage Variation in Electric Power System. *IEEE Latin America Transactions*, **7**, 217-222. https://doi.org/10.1109/TLA.2009.5256832

[11] Jaehak, C., Powers, E.J., Grady, W.M. and Bhatt, S.C. (2002) Power Disturbance Classifier Using a Rule-Based Method and Wavelet Packet-Based Hidden Markov Model. *IEEE Transactions on Power Delivery*, **17**, 233-241. https://doi.org/10.1109/61.974212

[12] Biswal, B., Biswal, M.K., Dash, P.K. and Mishra, S. (2013) Power Quality Event Characterization Using Support Vector Machine and Optimization Using Advanced Immune Algorithm. *Neurocomputing*, **103**, 75-86. https://doi.org/10.1016/j.neucom.2012.08.031

[13] Khokhar, S., Mohd Zin, A.A.B., Mokhtar, A.S.B. and Pesaran, M. (2015) A Comprehensive Overview on Signal Processing and Artificial Intelligence Techniques Applications in Classification of Power Quality Disturbances. *Renewable and Sustainable Energy Reviews*, **51**, 1650-1663. https://doi.org/10.1016/j.rser.2015.07.068

[14] Lin, W.-M., Wu, C.-H., Lin, C.-H. and Cheng, F.-S. (2006) Classification of Multiple Power Quality Disturbances Using Support Vector Machine and One-versus-One Approach. *International Conference on Power System Technology*, Chongqing, 22-26 October 2006, 1-8. https://doi.org/10.1109/ICPST.2006.321956

[15] Chuang, C.-L., Lu, Y.-L., Huang, T.-L., Hsiao, Y.-T. and Jiang, J.-A. (2005) Recognition of Multiple PQ Disturbances Using Wavelet-Based Neural Networks—Part 2: Implementation and Applications. *Transmission and Distribution Conference and Exhibition: Asia and Pacific*, Dalian, 18 August 2005, 1-6.

[16] Biswal, M. and Dash, P.K. (2013) Detection and Characterization of Multiple Power Quality Disturbances with a Fast S-Transform and Decision Tree Based Classifier. *Digital Signal Processing*, **23**, 1071-1083. https://doi.org/10.1016/j.dsp.2013.02.012

[17] Liu, Z., Cui, Y. and Li, W. (2015) Combined Power Quality Disturbances Recognition Using Wavelet Packet Entropies and S-Transform. *Entropy*, **17**, 5811. https://doi.org/10.3390/e17085811

[18] Liu, G., Li, F.G., Wen, G.L., Ning, S.K. and Zheng, S.G. (2013) Classification of Power Quality Disturbances Based on Independent Component Analysis and Support Vector Machine. 2013 *International Conference on Wavelet Analysis and Pattern Recognition (ICWAPR)*, Tianjin, 14-17 July 2013, 115-123.

[19] Dalai, S., Dey, D., Chatterjee, B., Chakravorti, S. and Bhattacharya, K. (2013) Cross Hilbert-Huang Transform Based Feature Extraction Method for Multiple PQ Disturbance Classification. 2013 *IEEE* 1*st International Conference on Condition Assessment Techniques in Electrical Systems* (*CATCON*), Kolkata, 6-8 December 2013, 314-317. https://doi.org/10.1109/CATCON.2013.6737519

[20] Liu, Z., Cui, Y. and Li, W. (2015) A Classification Method for Complex Power Quality Disturbances Using EEMD and Rank Wavelet SVM. *IEEE Transactions on Smart Grid*, **6**, 1-1. https://doi.org/10.1109/TSG.2015.2397431

[21] Yang, H.-T. and Liao, C.-C. (2001) A De-Noising Scheme for Enhancing Wavelet-Based Power Quality Monitoring System. *IEEE Transactions on Power Delivery*, **16**, 353-360. https://doi.org/10.1109/61.924810

[22] Manimala, K., Selvi, K. and Ahila, R. (2012) Optimization Techniques for Improving Power Quality Data Mining Using Wavelet Packet Based Support Vector Machine. *Neurocomputing*, **77**, 36-47. https://doi.org/10.1016/j.neucom.2011.08.010

[23] Huan, L. and Setiono, R. (1995) Chi2: Feature Selection and Discretization of Numeric Attributes. 2012 *IEEE* 24*th International Conference on Tools with Artificial Intelligence*, Herndon, 5-8 November 1995, 388-391.

[24] Robnik-Šikonja, M. and Kononenko, I. (2003) Theoretical and Empirical Analysis of ReliefF and RReliefF. *Machine Learning*, **53**, 23-69. https://doi.org/10.1023/A:1025667309714

[25] Hall, M.A. and Smith, L.A. (1998) Practical Feature Subset Selection for Machine Learning. Department of Computer Science, University of Waikato, Hamilton, New Zealand.

[26] Hall, M.A. and Smith, L.A. (1999) Feature Selection for Machine Learning: Comparing a Correlation-Based Filter Approach to the Wrapper. Department of Computer Science, University of Waikato, Hamilton, New Zealand.

[27] Lin, C.-J. (2010) A Practical Guide to Support Vector Classification. Department of Computer Science, National Taiwan University, Taipei.

[28] Vapnik, V. (1995) The Nature of Statistical Learning Theory, Data Mining and Knowledge Discovery. Springer, Berlin.

[29] Mahela, O.P., Shaik, A.G. and Gupta, N. (2015) A Critical Review of Detection and Classification of Power Quality Events. *Renewable and Sustainable Energy Reviews*, **41**, 495-505. https://doi.org/10.1016/j.rser.2014.08.070

[30] Kalatzis, I., Piliouras, N., Ventouras, E., Papageorgiou, C.C., Rabavilas, A.D. and Cavouras, D. (2003) Comparative Evaluation of Probabilistic Neural Network versus Support Vector Machines Classifiers in Discriminating ERP Signals of Depressive Patients from Healthy Controls. *Proceedings of the* 3*rd International Symposium on Image and Signal Processing and Analysis*, 18-20 September 2003, 981-985. https://doi.org/10.1109/ISPA.2003.1296422

[31] Modaresi, F. and Araghinejad, S. (2014) A Comparative Assessment of Support Vector Machines, Probabilistic Neural Networks, and K-Nearest Neighbor Algorithms for Water Quality Classification. *Water Resources Management*, **28**, 4095-4111. https://doi.org/10.1007/s11269-014-0730-z

[32] Liu, X., Gao, C. and Li, P. (2012) A Comparative Analysis of Support Vector Machines and Extreme Learning Machines. *Neural Networks*, **33**, 58-66. https://doi.org/10.1016/j.neunet.2012.04.002

[33] Chorowski, J., Wang, J. and Zurada, J.M. (2014) Review and Performance Comparison of SVM- and ELM-Based Classifiers. *Neurocomputing*, **128**, 507-516. https://doi.org/10.1016/j.neucom.2013.08.009

[34] Liu, Y., Loh, H.T. and Tor, S.B. (2005) Comparison of Extreme Learning Machine with Support Vector Machine for Text Classification. *The Proceedings of the 18th International Conference on Innovations in Applied Artificial Intelligence*, Bari, 22-24 June 2005, 390-399. https://doi.org/10.1007/11504894_55

[35] Hsu, C.-W. and Lin, C.-J. (2002) A Comparison of Methods for Multiclass Support Vector Machines. *IEEE Transactions on Neural Networks*, **13**, 415-425. https://doi.org/10.1109/72.991427

[36] De Yong, D., Reineri, C. and Magnago, F. (2013) Educational Software for Power Quality Analysis. *IEEE Latin America Transactions*, **11**, 479-485. https://doi.org/10.1109/TLA.2013.6502849

[37] Erişti, H., Yıldırım, Ö., Erişti, B. and Demir, Y. (2013) 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*, **49**, 95-103. https://doi.org/10.1016/j.ijepes.2012.12.018

[38] Chang, C.-C. and Lin, C.-J. (2011) LIBSVM: A Library for Support Vector Machines. *ACM Transactions on Intelligent Systems and Technology*, **2**, 1-27. https://doi.org/10.1145/1961189.1961199

[39] De Yong, D., Bhowmik, S. and Magnago, F. (2015) An Effective Power Quality classifier using Wavelet Transform and Support Vector Machines. *Expert Systems with Applications*, **42**, 6075-6081. https://doi.org/10.1016/j.eswa.2015.04.002