




Article

A Novel Deep Learning-Based Diagnosis Method Applied to Power Quality Disturbances

Artvin-Darien Gonzalez-Abreu ¹, Miguel Delgado-Prieto ², Roque-Alfredo Osornio-Rios ^{1,*},
Juan-Jose Saucedo-Dorantes ¹ and Rene-de-Jesus Romero-Troncoso ¹

¹ HSPdigital CA-Mecatronica, Engineering Faculty, Autonomous University of Queretaro, San Juan del Rio 76807, Queretaro, Mexico; agonzalez63@alumnos.uaq.mx (A.-D.G.-A.); jsaucedo@hspdigital.org (J.-J.S.-D.); troncoso@hspdigital.org (R.-d.-J.R.-T.)

² MCIA Research Center, Department of Electronic Engineering, Technical University of Catalonia (UPC), 08034 Barcelona, Spain; miguel.delgado@mcia.upc.edu

* Correspondence: raosornio@hspdigital.org

Abstract: Monitoring electrical power quality has become a priority in the industrial sector background: avoiding unwanted effects that affect the whole performance at industrial facilities is an aim. The lack of commercial equipment capable of detecting them is a proven fact. Studies and research related to these types of grid behaviors are still a subject for which contributions are required. Although research has been conducted for disturbance detection, most methodologies consider only a few standardized disturbance combinations. This paper proposes an innovative deep learning-based diagnosis method to be applied on power quality disturbances, and it is based on three stages. Firstly, a domain fusion approach is considered in a feature extraction stage to characterize the electrical power grid. Secondly, an adaptive pattern characterization is carried out by considering a stacked autoencoder. Finally, a neural network structure is applied to identify disturbances. The proposed approach relies on the training and validation of the diagnosis system with synthetic data: single, double and triple disturbances combinations and different noise levels, also validated with available experimental measurements provided by IEEE 1159.2 Working Group. The proposed method achieves nearly a 100% hit rate allowing a far more practical application due to its capability of pattern characterization.

Keywords: autoencoder; deep learning; power quality disturbances; power quality monitoring



Citation: Gonzalez-Abreu, A.-D.; Delgado-Prieto, M.; Osornio-Rios, R.-A.; Saucedo-Dorantes, J.-J.; Romero-Troncoso, R.-d.-J. A Novel Deep Learning-Based Diagnosis Method Applied to Power Quality Disturbances. *Energies* **2021**, *14*, 2839. <https://doi.org/10.3390/en14102839>

Academic Editors: Ahmed Abu-Siada and Andrea Mariscotti

Received: 20 March 2021

Accepted: 12 May 2021

Published: 14 May 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The implementation of recent advanced technologies following the trends of Energy 4.0 and Industry 4.0 allows us to gather, analyze, process, and dispose of a huge amount of data about the diverse processes involved in the different sectors as the industrial and energy [1]. The management of energy systems in Industry 4.0 will enable monitoring, aggregation and control of a large number of individual power generation and consumption units. Energy and electrical systems become more complex with the increase of multiple load profiles connected, and unexpected electrical events can occur causing the appearance of power quality disturbances (PQD) [2]. There is still a need for emerging technologies related to PQD detection and identification to be further developed into cyber-physical systems to implement smart algorithms and new methodologies for Energy 4.0 condition monitoring [3]. The future of the monitoring of power quality (PQ) is moving toward intelligent systems based in the cloud where cyber-physical architectures are implemented. Thereby, techniques of artificial intelligence, as based in deep learning, suppose an improvement in developing structures with higher capabilities of data management.

Power quality on the electrical grid is relevant to industry in light of the harmful, unwanted effects that an uneven electrical supply network is highly likely to cause upon the machinery and equipment attached to the grid. PQD represent random deviation

from an established amplitude and frequency of a sinusoidal waveform for electrical devices and related systems to make proper, adequate utilization. The electric supply chain includes generation, transmission, and consumption. Thus, on the consumption level, commonly ranging between 120 Vac and 600 Vac [4], is where most of the loads are attached to the grid. Consequently, it is at this level in which a major concern in applying PQ monitoring procedures can be spotted [5]. The AC power systems are designed to operate at a sinusoidal voltage at a given magnitude and frequency, typically 50 or 60 Hz. Based upon the previously stated, the ideal waveform of the voltage is defined formally by (1):

$$V(t) = V_m \sin(\omega t + \varphi), \quad (1)$$

where $V(t)$ is the sinusoidal voltage, V_m is the peak amplitude of the signal, ω is the angular frequency, and φ is the phase of the signal. Deviations from such an ideal voltage or current waveform in electrical distribution systems are considered a PQD [6]. Accordingly, the IEEE Std 1159 norm describes the set of possible single PQD, likely to take place in a power supply network [7]. Nonetheless, owing to actual industrial scenarios, where different electrical loads coexist (non-linear loads, lighting devices with electronic ballasts, controlled heating elements, magnet power supplies, battery chargers, furnaces, adjustable speed drives, air-conditioning systems, pumps, or elevators, among others), disturbances can occur within multiple combination forms. Therefore, this leads to complex patterns not directly represented by classical single disturbance descriptions [8].

The nature of the PQD at the industrial power grid leads to complex patterns—a non-periodic and uneven combination of disturbances alongside the work cycles of the related electrical loads connected to the power supply network under monitoring process. This factful event represents a current challenge in PQ diagnosis, since the number of possible disturbance patterns and related variants overcome the capabilities of available methods, usually restricted to a reduced number of single, isolated disturbances [9]. Most of the related literature sources that scaffold the quoted field, point out that the data-driven fault detection and identification approach represents the most promising PQ monitoring strategy [10]. As a whole, the data-driven procedure consists of three steps: feature extraction, feature reduction, and classification. As for the latter, time-domain, frequency domain, and time-frequency domain approaches have been applied to compute significant numerical features characterizing the power line signal. Ref. [11] considers the computation of statistical features using S-transform based multi-resolution analysis of signals to characterize several disturbances for further recognition. Furthermore, Ref. [12] presents a method based on variational mode decomposition of the original signal and the recurrence quantification analysis which reaches a proper characterization of the electrical signals prior to their identification; it follows a data-driven approach. In [13], a study focused on feature selection explores the performance achieved by different subset through features extracted from commonly signal processing techniques. In spite of the good performance achievement, the drawback of this approach is the need to characterize the signal through several signal processing techniques, the outcome of which is a much higher computational complexity. In addition, the work depicted in [14] proposes a novel method to extract the features from the signal by first transforming the 1-dimensional signal into a 2-dimensional signal. Then, for the classification stage, it assesses different machine learning models, such as k-nearest neighbor, multilayer perceptron, and support vector machines, to determine which of these models performs at its best. This work considers the combination of two and three PQD. Yet, all related works sum up that no upper domain exists while dealing with signal characterization, but the fusion of information represents the most promising approach [15]. The reduction of such sets of numerical indicators, regarding their relevance for characterizing patterns, represents a critical data-driven step [16]. Accordingly, non-significant and redundant information must be discarded or attenuated in order to optimize further pattern recognition tasks. Some works as [13,17,18], consider classical algorithms as Principal Component Analysis (PCA), k-Nearest Neighbor (kNN), or sequential forward selection methods. However, on

dealing with an increased variety of patterns—as required by current PQD monitoring applications—, these approaches bring about a limited performance [19].

Viewed in this way, deep learning techniques are to be considered in multiple industrial fields of applications that deal with high-dimensional sets of data and multiple patterns [20]. Application of deep learning has been presenting good performance in areas like image classification, speech recognition, natural language processing, video processing, and, recently, in areas related to energy management [21]. Autoencoder, convolutional neural networks, or recurrent neural networks are the most common techniques to be used for dealing with complex data involved. There is merely a few works exploring the suitability of such techniques in electrical network monitoring, but even less applied to PQD classification [22]. The work presented in [23] submits data-driven approaches on the grounds of such deep learning techniques. Although the outcoming performances keep being promising, the lack of a common procedure to configure and tune the algorithms still represents a shortcoming that avoids its consideration over real industrial applications [24]. The work presented in [23] looks into the capabilities of deep learning in PQD classification using a convolutional neural network alongside eleven statistical indicators calculated over the four principal components of the electrical signals by the use of a modified version of PCA. Likewise, they use the neural network structure to perform the classification stage. Two classes of multiple, analyzed PQD are regarded by this work as the appearance of the disturbances because some of those disturbances turn out to be opposite in definition. Nevertheless, the method presents good capabilities with simulated data. Their introduction as part of the data-driven procedure represents the necessary step forward to reach the required characterization and management of multiple patterns to be recognized later in combination with classification algorithms, as the well-known neural networks [25].

Thus, the contributions of this paper rely on the proposal of a novel deep learning-based diagnosis method applied to PQD. The originality of the work includes the following key features: (i) a common framework for diagnosis method configuration and tuning of algorithms based on quantitative metrics, (ii) the validation of the autoencoder as a deep learning technique capable of characterizing complex power disturbance patterns, (iii) the consideration of single disturbances, as well as the combination of double and triple disturbances under different signal to noise ratios, and (iv) the validation of the proposed method following both, the classical synthetic signal-based approach, and real measurements available from an open-access database. Moreover, it must be emphasized that the proposed diagnosis method has been validated with the highest number of disturbance patterns compared with the identified state of the art regarding real measurements considered. The latter contributes to the validation of the feasibility of the proposed method. In summary, the contribution of the paper is established on a methodology based on a deep learning technique seeking to contribute to the process of new algorithms to PQ detection, in order to attend the challenges paradigm of Energy 4.0.

This paper is organized as stated next: in Section 2, theoretical background related to the autoencoder as feature reduction approach is presented. Section 3 describes the proposed method. The considered data sets are introduced in Section 4. Section 5 presents and discusses the obtained results. Finally, Section 6 shows the conclusions of this work.

2. Autoencoder as Feature Reduction

Traditional Neural Networks (NN) applications represent shallow approaches, being the three layers feed-forward network (i.e., input, hidden, and output layers), the most common architecture. In such shallow networks, the inputs must be a carefully curated set of parameters obtained by means of feature reduction and/or feature engineering because NN are highly affected by the presence of non-significant and redundant information making the learning process of complex features and relationship from the data difficult [26]. In contrast, deep learning takes a feature learning approach, i.e., features relationships are discovered rather than given. This occurs by taking advantage of the deep networks properties, in which the initial layers extract meaningful features in an unsupervised

manner and the final layers, map these features to the target [27]. By following this approach, the resulting network is apt to working with a wider set of inputs, where the training stage is ideal for the identification of significant features for pattern characterization and recognition, assigning weights autonomously.

Upon the previously stated, the autoencoder is a deep learning technique inspired by NN structures, trained to replicate its input at its output. Thus, the autoencoder structure is divided into two main parts: encoder and decoder. Then, the encoder deals with the representation of the information in the previous layer onto a reduced dimension. Otherwise, the decoder takes the compressed information resulting from the previous layer and returns it to its original dimension. The training stage of an autoencoder is unsupervised and based on the optimization of a cost function to minimize such error between the input and its resulting reconstruction at the output. However, the consideration of multiple layers in deep networks rests upon inaccuracies of the classical backpropagation algorithm failing to update the weights through the layers during the training process as the gradient becomes too small to influence a change and prevents further learning, a problem known as the vanishing gradient [26]. One of the best performing solutions is the implementation of simple three-layers autoencoders trained individually and stacked later, where each subsequent layer uses the hidden layer of the previous autoencoder as the input. After the set of autoencoders is stacked, the deep autoencoder is fine-tuned. Thus, the stacking of different autoencoders is usually considered to reduce or compress the input data to a low dimensionality without losing relevant information to reconstruct the original signal at the output of the structure. An example of such a stacked autoencoder is shown in Figure 1, where the hidden layer with the lowest dimension (i.e., two in this example) represents the reduced set of information resulting from the feature learning reduction process. Such a layer is considered to be used as the input of a posterior classification algorithm as a simple neural network [28].

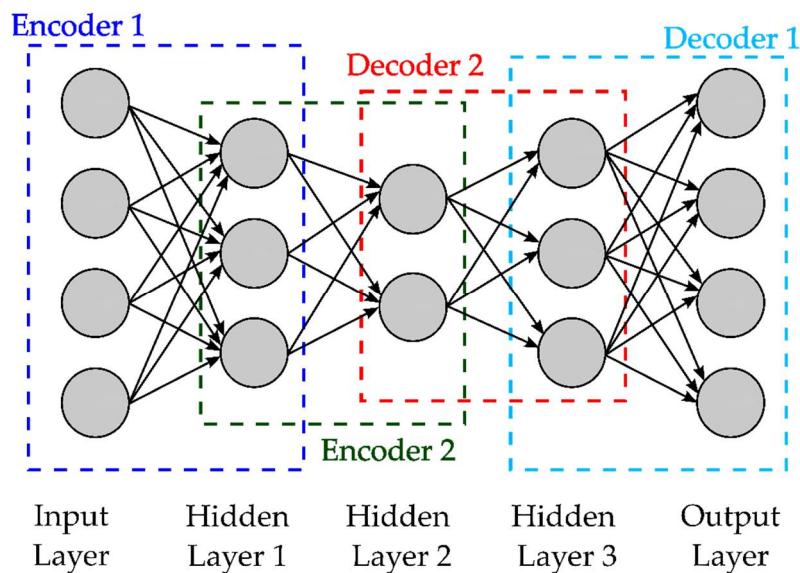


Figure 1. Stacked autoencoder composed of two autoencoders.

Regularized autoencoders use a loss function that encourages the model to have other properties besides the ability to copy its input to its output: Adding penalties to the cost function, an autoencoder could include sparsity of the representation, smallness into the derivative of the representation, and robustness to noise or missing inputs. These coefficients are the L2W regularization term, the sparsity regularization term, and the sparsity proportion term. Thus, the cost function, $J_{(x)}$, related to a regularized autoencoder is presented in (2).

$$J_{(x)} = L(x, \hat{x}) + \lambda * \Omega_{weights} + \beta * \Omega_{sparsity} \tag{2}$$

where $L(x, \hat{x})$ is a loss function (e.g., Mean Squared Error, MSE), that measures the error between the input x and the reconstruction \hat{x} . λ is the coefficient for the L2W regularization term $\Omega_{weights}$, and β is the coefficient for the sparsity regularization term $\Omega_{sparsity}$. The L2W regularization term is shown in (3):

$$\Omega_{weights} = \sum_l^L \sum_j^N \sum_i^K (w_{ji}^{(l)})^2, \quad (3)$$

where L is the number of hidden layers (i.e., one for a basic autoencoder), N is the number of samples, K is the number of variables in the training data set, and $w_{ji}^{(l)}$ represents the value of the weights indexed by $l, j, i = 1, 2, 3 \dots$. In addition, following the sparsity regularization term, the Kullback-Leibler divergence is shown in (4):

$$\Omega_{sparsity} = \sum_{i=1}^{D^{(l)}} KL(\rho \parallel \hat{\rho}_i) = \sum_{i=1}^{D^{(l)}} \rho \log\left(\frac{\rho}{\hat{\rho}_i}\right) + (1 - \rho) \log\left(\frac{1 - \rho}{1 - \hat{\rho}_i}\right); \quad (4)$$

this Equation (4) takes a large value when the average activation probability value, $\hat{\rho}_i$, of a neuron i and its desired value, the sparsity proportion term, ρ , are not close [29]. Thus, such $J_{(x)}$ cost function is approached as an optimization problem during the autoencoder training process.

3. Proposed Methodology

The proposed data-driven fault detection and identification methodology applied to PQ monitoring is divided in four stages. The flowchart of Figure 2 depicts the methodology proposed.

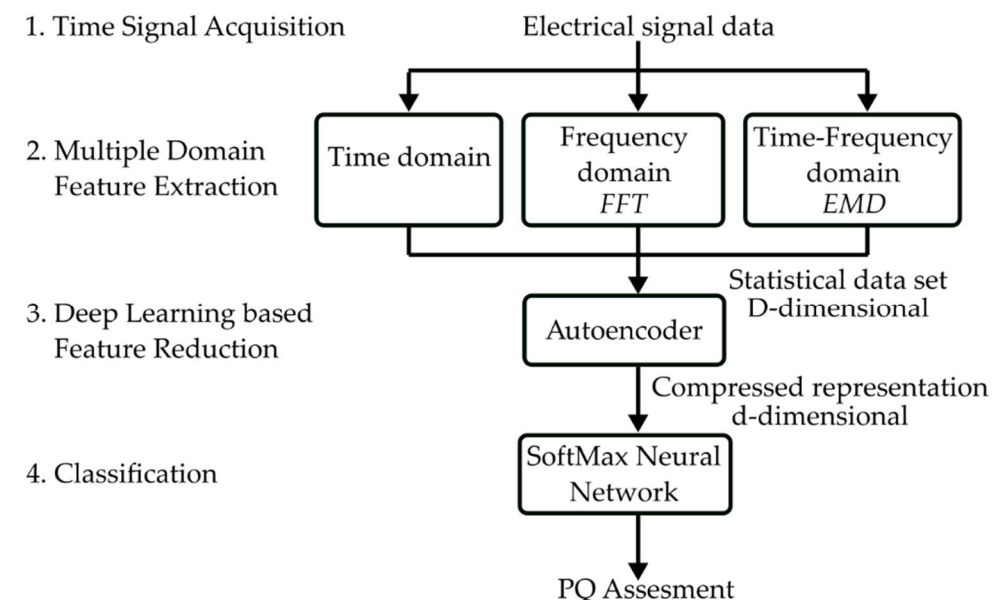


Figure 2. Flowchart of the proposed deep learning-based methodology applied to PQD monitoring.

The first stage of the proposed methodology is related to the definition of the representative synthetic electrical signals for each considered condition, that is: the reference condition, which is the normal operation of the electrical system, and signals including disturbances in a single or combined mode. As stated in the related literature, the generation of such synthetic signals, following the corresponding international standards [14], represents the optimal approach to have a balanced and representative database for training purposes. It must be noted that this work, however, will be onwards extending the

validation of the resulting diagnosis structure with both synthetic and experimental signals provided by a scientific usage database.

The second stage is feature extraction. At this stage, characteristic numerical features are proposed to be extracted from the three possible domains, that is, time, frequency, and time-frequency domains. Taking advantage of the feature learning capabilities of the forthcoming autoencoder structure, the proposed methodology includes such a variety of domains, as all of them have proved to be relevant in different studies [19]. All in all, the fast Fourier transform is computed in the frequency-domain. Additionally, the empirical mode decomposition (EMD) is considered as a signal decomposition approach in the time-frequency domain. EMD has been selected because it offers the ability to self-adapt, it does not require predefining some parameters before its application. This property is a distinguished detail because, unlike other techniques, such as STFT or Wavelet decomposition, where a window size or resolution has to be defined beforehand, likely to lead to the risk of attenuating or fading some effects from the signal. Furthermore, comparisons in the use of these techniques have already been reported showing no general agreement to which specific technique applies [13]. The drawback of EMD is that the number of resulting signals is not controlled, and therefore, multiple signals can emerge. Then, it depends not only on the application but also on the signal to be analyzed, if it happens to contain any noise level. In this particular case, this issue has no impact because, following previous research, it has been identified that the consideration of two to five Intrinsic Mode Functions (IMF) may include significant information for diagnosis purposes [30], being the first two IMF the most significant ones in most of the cases. Thus, from each of the four considered signals (i.e., time-based, FFT-frequency spectrum, and the first two IMF), a set of twenty statistical features are estimated for characterization. Specifically, a number of statistical features is considered: mean, maximum value, root mean square, square root mean, deviation standard, variance, rms shape factor, srm shape factor, crest factor, latitude factor, impulse factor, skewness, kurtosis, 5° moment, 6° moment, energy, entropy, range, form factor, and log energy entropy. The consideration of such indicators represents a powerful strategy to characterize multiple patterns at the time that offer enough generalization capabilities to avoid the risk of overfitting. The detailed equations of these statistical indicators are reported in several studies, as stated in [31] are the first fifteen above listed indicators, and in [23] are the last five.

The third stage corresponds to feature learning and reduction. In this stage, the autoencoder is considered to reduce the original 80-dimensional set of numerical features into a lower-dimensional representation.

In order to reach a numerical feature reduction (i.e., codification), the autoencoder configuration must be carried out. First, the depth, L , and size, l , of each layer for the autoencoder structure must be defined. These two values involve how the input feature vector is codified from an initial D -dimensional set of features to a d -dimensional representation. The selection of various layers intends to preserve the most relevant information and relationships between features, since reducing the information in a single layer could remove important affinities in the input feature vector. The values for the hyperparameters mentioned at the end of section two for the autoencoder configuration, that is, the coefficient λ for the L2W regularization term $\Omega_{weights}$, the coefficient β for the sparsity regularization term $\Omega_{sparsity}$, and ρ , the sparsity proportion term, also need to be selected. The coefficients for the L2W regularization term and the sparsity regularization term, as well as the sparsity proportion term, are involved with the cost function presented in Equation (2) to be minimized. These hyperparameters allow yielding a model that could learn useful features of the input furthermore to perform the encoding task. A proper selection of values for these parameters will be reflected at the end of the autoencoder structure. The cost function to be minimized is the error resulting from the comparison between the original signal and the result of the decoding process of the codification stage. Therefore, low error values means that the codification process (i.e., pattern characterization), is being carried out properly.

As aforementioned, the selection of optimal values to each one of the main autoencoder configuration parameters represents a limitation to its application that is being tackled following empirical approaches. Yet, this option engages the feasibility for practical implementation, since the tuning process becomes unaffordable by non-specialists about the use of autoencoder algorithms. The proposed scheme in this research work, though, includes an autoencoder configuration procedure supported by updated research experience in the field, which provides an accessible and reproducible configuration methodology to promote its application.

Accordingly, in order to obtain a proper configuration, three steps to be followed are proposed. These steps include the definition of the number of hidden layers, the size of each hidden layer, the coefficients for the L2W regularization term and the sparsity regularization term, and the sparsity proportion term. All these parameters are proposed to be determined according to the resulting reconstruction performance optimizing so, the characterization capabilities of the signal. Specifically, the three steps proposed are:

1. Firstly, the selection of the number of hidden layers. A deep network is proposed based on the reduction of the features through the hidden layers' size. The number of layers has to be selected considering the number of inputs. Taking into account the field of the application and following previous research, it is recommended a number between three and five hidden layers [26].
2. Secondly, the size of the hidden layers. This parameter represents the reduction ratio between layers and is also related to the number of features as input. Typical values of proportion between consecutive layers are between one-third and one-tenth of neurons [32]. This range of reduction ratio is supported by [33], which points out that lower values could lead to a poor generalization of the autoencoder.
3. Thirdly, the selection of the values for the coefficients to the regularization terms in the cost function. A search strategy based on coarse fine grid search is considered. With this procedure, a local minimum from the error of reconstruction is certainly expected to be achieved, which means the obtention of a valid representation in the autoencoder structure. Although the set up process depends on several application features, the different patterns related or other hyperparameters considered it could be started by firstly searching any of the terms. In this study, it is proposed as a first step, the search for the values from L2W and sparsity regularization in a range of values between 10 to 1×10^{-6} in a log scale. Then, the selection of the sparsity proportion term in a range between 0.5 and 0.05 with steps of 0.05.

The proposed autoencoder configuration procedure begins with a set of default values for each of the five parameters. These values are typically considered in different applications [26]. Thus, following the proposed three-step procedure, just one parameter is modified each time, seeking the value as a result of the best MSE value. Once identified, the selection of the next parameter proceeds. It must be of full awareness that the proposed procedure can be iteratively repeated in order to fine-tune the autoencoder.

Finally, the fourth stage of the methodology is related to the classification task. For this purpose, a three softmax layers NN is proposed so as to be trained by the use of the resulting encoded representation from the autoencoder as inputs, and the number of considered classes as output [27].

4. Power Quality Disturbance Data Sets

In order to validate the proposed methodology, a challenging scenario has been considered, including multiple classes, combinatorial disturbances patterns, and different levels of noise. It must be pointed out that most of the available studies related to PQ monitoring deal with single disturbances as a basic framework for validating their methodology. Upon these grounds, just a few of them consider PQ scenarios, including the combination of two disturbances—or even fewer—considering the combination of three disturbance: they represent a challenge and a current industrial requirement. Thus, the set of considered conditions in this study are: C1: Normal, C2: Sag, C3: Swell, C4: Interruption, C5: Flicker,

C6: Harmonics, C7: Oscillatory Transient, C8: Notching, C9: Sag with Harmonics, C10: Sag with Oscillatory Transient, C11: Swell with Harmonics, C12: Swell with Oscillatory Transient, C13: Harmonics with Oscillatory Transient, C14: Harmonics with Flicker, C15: Harmonics with Notching, C16: Sag with Harmonics and with Oscillatory Transient, and C17: Swell with Harmonics and with Oscillatory Transient. These signals are based on the equations and parameters to model them reported in [19,34].

Following such definition and according to the IEEE Std 1159 [7] and related to this work, seven main disturbances patterns are described. Therefore, in the category of short-duration root-mean-square variations, the sag and swell disturbances are firstly defined. Specifically, a sag is defined as a decrease in the RMS voltage between 0.1 per unit (pu), and 0.9 pu, and a swell as the increase in the RMS voltage above 1.1 pu. In both cases, the duration of these disturbances range from 0.5 cycles to 1 min. The third disturbance considered is the interruption: this disturbance is also related to the amplitude of the signal, specifically when the signal decrease less than 0.1 pu for a period not exceeding 1 min. The next disturbance is related to sinusoidal voltages or currents waveforms containing harmonics of the main frequency. These disturbances are referred to as harmonics and are included in the category of waveform distortion. The fifth disturbance is described as voltage fluctuations, which consists of systematic variations of the voltage envelope (or a series of random voltage changes), in which the magnitude does not exceed the voltage in the range between 0.95 pu to 1.05 pu according to [7]. It is important to highlight that voltage fluctuations and term flicker are linked in ANSI/IEEE standards, and, due to this, the literature found in the detection and identification of disturbances regards flicker as a disturbance and continuing in that line also this work. However, it is important to clarify that the flicker is the effect on lightning from these voltage fluctuations phenomena. The sixth disturbance described is the oscillatory transient, which is a sudden, nonpower frequency change in the steady-state condition of voltage, current, or both, which includes positive and negative polarity. Finally, the disturbance named notching is described as a periodic voltage disturbance caused by the normal operation of power electronics devices when current is commutated from one phase to another.

Such set of disturbances has been defined considering multiple studies available in literature sources, as well as excluding those nonsensical conditions in terms of timely, simultaneous appearance over the power line signal (e.g., such as sag with swell, flicker with interruption, interruption with swell, or flicker with interruption with sag, among others).

4.1. Synthetic Data Set

The training stage of the proposed methodology is supported by a set of 1000 synthetic signals for each of the considered conditions, that is, a total training data set of 17,000 time signals. The common parameters of the signals are 60 Hz to its fundamental frequency and an amplitude per unit (pu) aiming at allowing the latter application to different datasets. Each and every one of the different classes is represented for a random generation signal between the ranges of each one of the boundaries that represent the condition related. The signal generation process, as well as the proposed methodology implementation, have been supported by MATLAB 2019b. The signals are created with a time window, which is equivalent to ten periods of the main one. This window size is chosen because the voltage is usually measured on a cycle-by-cycle basis according to [7]. The sampling frequency is 15.36 kHz. Noiseless and three signal-to-noise ratios (i.e., SNR of 50 dB, 40 dB, and 30 dB) are also considered. Figure 3 depicts some of the resulting synthetic disturbances. It must be pointed out that the start and duration length of the disturbance are established randomly within the time window size which provides better generalization capabilities to the final model.

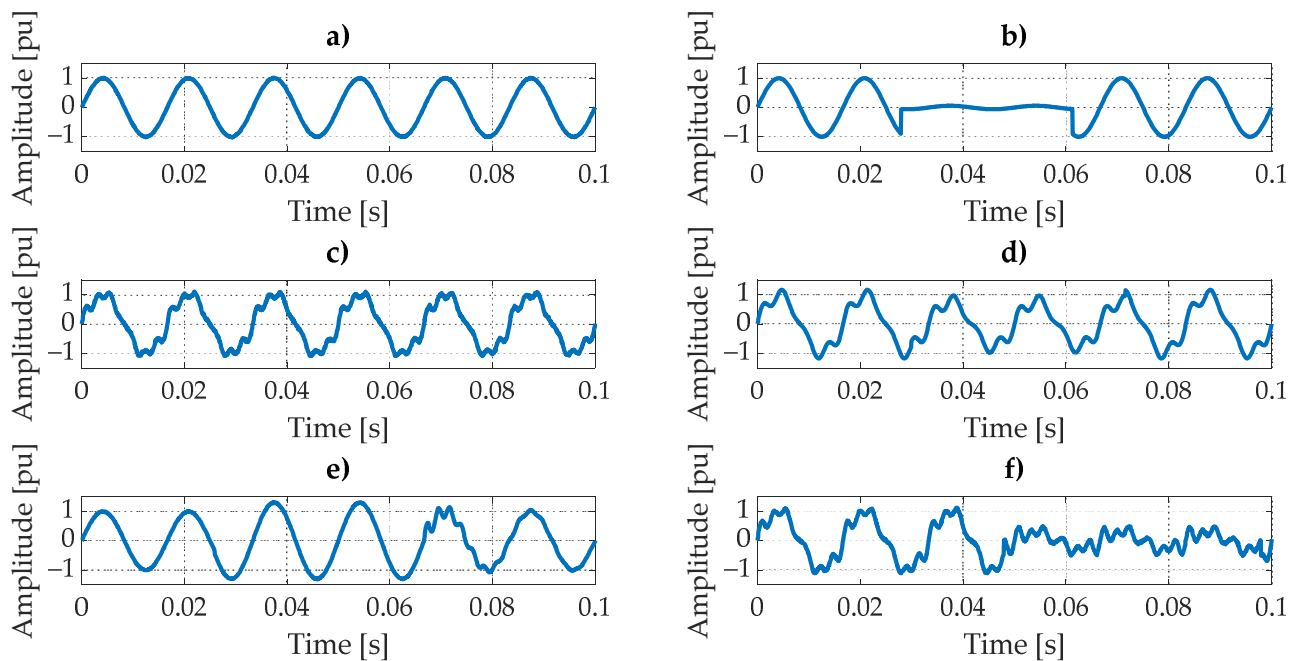


Figure 3. Detail of electrical signals synthetically generated. (a) Normal with SNR of 40 dB; (b) Interruption noiseless; (c) Harmonics with SNR of 30 dB; (d) Sag with Harmonics with SNR of 50 dB; (e) Swell with Oscillatory Transient with SNR of 40 dB; (f) Sag with Harmonics and Oscillatory Transient with SNR of 30 dB.

Following the proposed methodology, a set of 80 statistical features are estimated for each signal, resulting from the time, frequency, and time-frequency domains processing. Once the numerical feature database is generated, a training set and test set are defined. The training set includes 90% of the samples and the test set the remaining 10%. Both sets will keep a balanced representation of each class and a 10-fold-cross validation approach is taken into account throughout the training process. Thus, the final data sets are defined by a matrix of $(c * m) * n$, where c is the number of classes considered (i.e., 17), m is the number of signals, 900 for each class in the training set and 100 for each class in the test set, and finally, and n is the number of estimated features, that is, the 80 statistical features considered in this methodology.

4.2. Experimental Data Set

With the purpose to extend the validation of the submitted methodology, a real data set has additionally been considered for test purposes over the trained diagnosis model. Such experimental database contains representative signals of some of the PQ scenarios considered. The experimental database accounted in this study is supplied by the IEEE P1159.2 working group and referred to by some studies in the field: in [24], in order to verify the effectiveness of his proposed for the classification of PQD signals waveforms; in [34], eleven waveforms signals from this database have been utilized to demonstrate the successful classification with his method; and, in [35], five real signals have been used to demonstrate the ability of the proposed approach to identify the disturbances signals. All in all, these real signals have been used to validate the approaches presented in PQ studies.

The signals of such database are sampled at 15,360 Hz and contain six cycles at 60 Hz operation frequency, an amplitude in pu, and an SNR estimated at 45 dB [24]. Although most of the related studies consider less than 10 samples of such database, in this particular endeavor, which pursues to validate the generalization capabilities and performance of the proposed methodology, a total of 48 signals have been considered. It must be remarked on that the signals in this database are not originally labeled. Thus, the labeling process is achieved by the scientific community who observed the signal features. Consequently, the set of 48 measurements considered in this study, is representative of eight different

conditions related to the PQD considered, specifically: C2, C4, C9, C10, C12, C13, C16, and C17. Figure 4 depicts two signals from this real database.

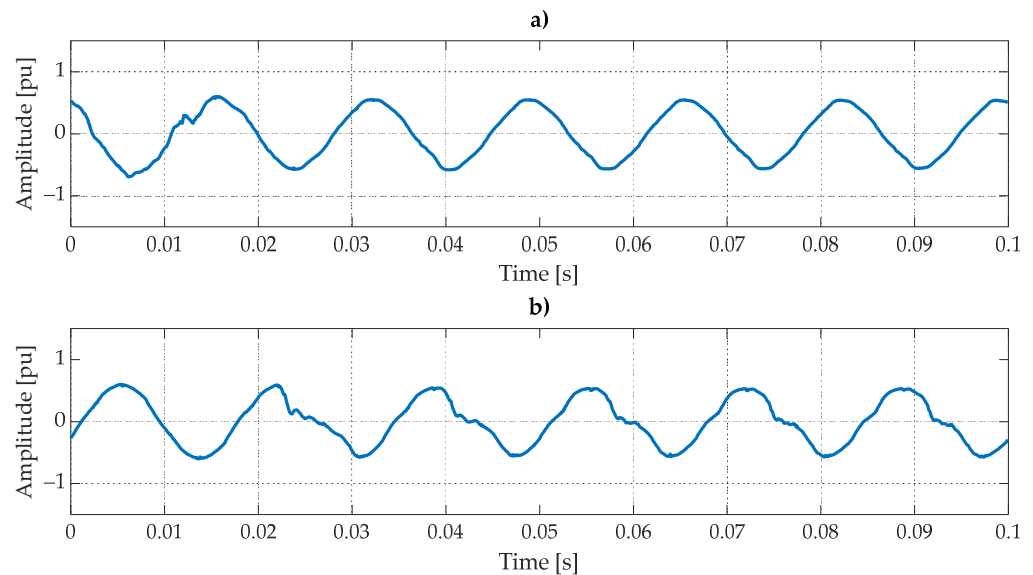


Figure 4. Real database signal. (a) Sag disturbance; (b) Sag with harmonics.

5. Results and Discussion

Applying the proposed method previously described, the selection of optimal values for the autoencoder configuration is carried out to begin with. Table 1 summarizes the resulting hyperparameters.

Table 1. Resulting hyperparameters for the implementation of the stacked autoencoder.

Hyperparameter	Value Selected	
Number of autoencoders	3	
Hidden layer sizes	Layer	Size
	1	70
	2	50
	3	30
L2W Regularization	1×10^{-6}	
Sparsity Regularization	1	
Sparsity Proportion	0.2	

5.1. Feature Learning Process

Broadly speaking, a number of statistical indicators taken from the literature are extracted from the time, frequency and time-frequency domains of an electrical signal. These numerical indicators (80 actually) are the input to the autoencoder model to be encoded. The autoencoder aims to extract, automatically, useful characteristics to properly encode and decode input signals by minimizing reconstruction error. In this regard, a qualitative and quantitative analysis of such reconstruction task is proposed, in order to validate the effectiveness of the feature learning process. For this purpose, some representative measurements have been selected to analyze the reconstruction performance from a qualitative point of view. Two examples of such reconstruction can be observed in Figure 5, where the input vector containing the 80 statistical features, its resulting reconstruction, and the corresponding error are depicted.

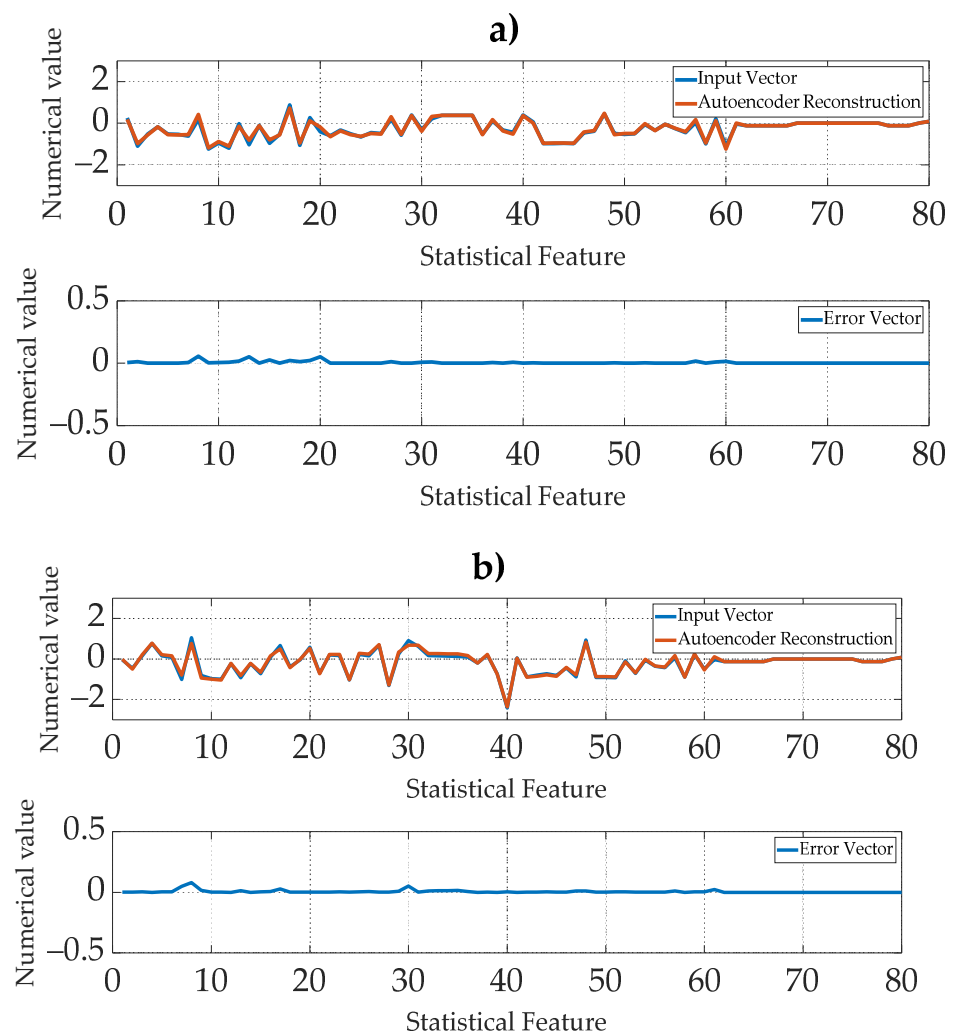


Figure 5. Original feature signals, its reconstruction and the resulting errors. (a) Signal including flicker disturbance; (b) Signal including harmonic and oscillatory transient disturbances. In x -axis, 1–20: Time features; 21–40: Frequency features; 41:80 Time-frequency features. In y -axis, numerical value of statistical features.

The qualitative inspection of such results shows that the autoencoder properly performs the characterization of the original signals since neither significant deviation nor averaged patterns are delivered. This fact is especially positive considering the differences among considered signals. From a quantitative point of view, the feature learning performance is also validated, as the resulting reconstruction error estimated through MSE is 0.0143 for Figure 5a and 0.002 for Figure 5b. Such values are representative, owing to the fact that the average MSE of the whole data set is 0.0327, in all cases with a very low resulting error [24].

5.2. Classification Performance

As aforementioned, pattern recognition is carried out by adding one single neural layer to the end of the stacked autoencoder structure. The number of neurons of this last layer must be equal to the number of considered conditions (i.e., seventeen). Accordingly, once trained, the result of the test stage is shown in Table 2, in which the classification accuracy for each one of the considered conditions is detailed.

Table 2. Classification accuracy of PQ disturbance conditions applying the stacked autoencoder-neural network method.

PQ Disturbance Condition	Accuracy (%)
C1: Normal	100
C2: Sag	99
C3: Swell	100
C4: Interruption	100
C5: Flicker	100
C6: Harmonics	100
C7: Oscillatory Transient	100
C8: Notching	99
C9: Sag with Harmonics	98
C10: Sag with Oscillatory Transient	100
C11: Swell with Harmonics	99
C12: Swell with Oscillatory Transient	100
C13: Harmonics with Oscillatory Transient	99
C14: Harmonics with Flicker	99
C15: Harmonics with Notching	100
C16: Sag with Harmonics and with Oscillatory Transient	98
C17: Swell with Harmonics and with Oscillatory Transient	100

It has been obtained as an outcome a 99.47%, considering the proposed diagnosis method for the seventeen conditions. Related to this, it can be observed that the classification accuracy reaches its high in the vast majority of the conditions. However, some classes exhibit misclassification between one and two percent. In the case of single disturbance conditions, that is, from C1 to C8, the observed misclassifications are related to low-intensity severities of the corresponding disturbances. Thus, the single misclassification observed in the sag condition (i.e., C2), emerges from misclassification as interruption condition (i.e., C4). This response, however, is coherent with the underlying physical effect of the disturbance characterized by the features. So, the misclassified sag disturbance sample is due to the extremely low value of the corresponding severity (i.e., less than 0.1 pu), which is recognized as an interruption. As for two combined disturbances, that is, from C9 to C15, the condition presenting the highest misclassification error (i.e., 2%), is over the C11 disturbance corresponding to sag with harmonics. For the complex cases referred to as the combination of three disturbances—C16 and C17—, the observed misclassification (i.e., 2% in C16 condition) results from the assignment to simpler disturbance conditions. This means that the misclassified samples are recognized as a single disturbance condition (i.e., C9), or two combined disturbance conditions (i.e., C10), but, in all cases, conditions that include some of the disturbances present in the original samples. In addition to all this discussion, it is important to remark on the fact that all samples corresponding to the normal condition are properly classified.

5.3. Validation of the Method at Different SNR

The electrical noise caused by measurement equipment, instrumentation devices, or electrical loads connected to the power electrical system is typical and, in consequence, represents an important aspect to consider in the assessment of a power disturbance monitoring solution. Owing to the previously stated, different signal-to-noise ratios have been considered. Following other studies in the field for comparison purposes, low, medium, and high ratios have been considered over the test sets of the proposed method, that is, 50 dB, 40 dB, and 30 dB, respectively [36]. Table 3 shows the resulting overall classification accuracy at these three signals-to-noise ratios.

Table 3. Resulting classification accuracy at different SNR.

	SNR			Average
	50 dB	40 dB	30 dB	
Performance	98.35%	98.37%	97.48%	98.06%

It should be noted that the obtained results indicate that the performance of the proposed stacked autoencoder-neural network structure is almost unaffected (i.e., less than 2%), by the SNR appearance over the measurements. The performance decay throughout the three SNR occurs because of a slight increase in the misclassification of the complex signals with three combined disturbances. Yet, like in the previous noiseless scenario, the misclassifications meet connection with simpler conditions including, at least, one of the disturbances in the considered measurement.

5.4. Validation of the Method with Real Signals

In order to validate the proposed method, a challenging set of 48 signals from a real database have been considered. Table 4 shows the resulting confusion matrix of the classification.

Table 4. Confusion matrix for the real signals analyzed.

		Actual Condition									
		2	4	5	9	10	12	13	16	17	
Predicted Condition	2	17									
	4		1								
	5				2	1					
	9				11			1			
	10					1					
	12						1				
	13							2			
	16								5		
	17										6

The confusion matrix displays a classification accuracy of 91.66% in regard to the 48 signals with disturbances in the real database. An example of the correct classification is the signal shown in Section IV, Figure 4 contains a sag disturbance as it is observed due to the amplitude of 0.5 pu. Next, as expected, the resulting classification also indicates that it corresponds to a sag disturbance. Most of the misclassifications (i.e., three of the four) are related to a flicker disturbance due to a low level of the correct disturbance that is misclassified with the flicker condition. Besides, there is another misclassified signal: a sample of C13 is assigned as C9. In this case, the condition misclassified is the combination of the condition predicted and another single disturbance, the sample contains more characteristics from C9 than C13. In this sense, the misclassification is expected as seen in Figure 5, where the disturbance recognized in the signal is the combination of sag with harmonics and even it could be considered an interruption in the part of the third cycle. Nevertheless, it is classified as a flicker disturbance.

5.5. Comparative with Other Recent Publications in the Field

Finally, so as to set the potentiality of the proposed methodology in context, some of the most significant works previously published in the related literature sources have been compiled and organized, as shown in Table 5.

Table 5. Performance comparison with recent literature.

Ref.	Year	Feature Extraction Technique	Classification Technique	Data Type	Type of Multiple PQD	Number of Disturbances Considered	Accuracy Synthetic/Real
[33]	2019	MST	PSSAE	Synthetic/Emulated	double	12/8	99.46/94.8
[36]	2019	1-D ST+TT	DT	Synthetic/Real	double	14/7	99.93/98.57
[35]	2019	EWT	SVM	Synthetic/Emulated and Real	double	15/5	95.56/-
[37]	2019	ST	DT	Synthetic/Real	double *	16/3	99.47/-
[23]	2019	IPCA	1-D CNN	Synthetic	double *, triple *	11	99.92
[14]	2020	2D-RT	KNN	Synthetic/Emulated	double, triple	17/17	99.26/97.62
[38]	2018	High-Order Statistics	Neuro-Tree	Synthetic/Emulated and Real	double, triple	19/19	97.8/100 **
Proposed	2020	FFT+EMD	NN	Synthetic/Real	double, triple	16/8	99.47/91.66

-: Not reported; *: Some of the reported disturbances are not in a simultaneous apparition; **: Not detail of the proportion from the real conditions.

It can be observed that there is a clear difference between works limited to combinations of two disturbances and those that include combinations of three disturbances. The complexity of the electrical systems is being increased and, as a consequence, most recent studies tend to consider the appearance of three combined disturbances at the very same moment as this proposed method does. From such studies that consider three combined disturbances, there is also a clear difference between those using emulated signals and those considering real ones. In this regard, although emulated signals represent a proper approach to validate proposed methods, the challenge facing real signals provides a greater validation of the method generalization and its eventful impact on real applications. An ultimate consideration is the fact that the methods and techniques applied in the related literature sources require higher specialization and proficiency. Furthermore, they do not provide a common procedure to configure or adjust the parameters involves restricting its application to technical and industrial sectors.

Two main aspects must be discussed related to the advantages and limitations of the proposed methodology. Firstly, those related to the proposed feature-fusion scheme included in the methodology. Although related literature does not prioritize the use of one domain over the others (i.e., time, frequency, or time-frequency domains), it is clear that the consideration of the three domains increases the electrical signal characterization possibilities and, then, the forthcoming pattern recognition accuracy. In this regard, the proposed methodology provides an adaptive fault detection and identification scheme in which the feature calculation can be tackled fusing multiple domains, even, if required, by introducing multiple domain's techniques. Secondly, those related with the use of an autoencoder as deep learning feature reduction. The main limitation related to the consideration of such technique is the tuning of the corresponding hyperparameters. This fact implies that an optimization procedure must be carried out during the off-line stage related to the model training. The proposed methodology considers a classical coarse-fine grid search taking advantage of the most significant value ranges identified for each hyperparameter in the literature. In spite of the fact that the proposed methodology exhibits good results in terms of autoencoder performances, the obtained values correspond to local minimum in the reconstruction error cost function. Related to the computational burden, the proposed algorithm has a computational time required for the execution, on average, 16.5 milliseconds, considering a personal computer with the following characteristics: Intel (R) Core (TM) i5-6500 CPU @3.2 GHz, 16 GB RAM, and NVIDIA Quadro P620 GPU. This computation time for monitoring is sufficient, perhaps not for control, as this work is not intended to cover the control area. This detection time is much shorter than the time needed to make decisions regarding the production and network's distribution actions owing to their slower dynamics. To sum up, these types of algorithms have not been designed in

terms of real-time. Nevertheless, they are helpful for the controller of the grid or power system to make a decision as soon as possible.

6. Conclusions

This work contributes to the development of an algorithm as a proposal in the towards the management of PQD that may arise in the context of Industry 4.0 through a novel deep learning-based diagnosis method applied to PQD. There are three important aspects of this approach. The first one being the use of a stacked autoencoder-neural network based structure in the methodology for PQD characterization, detection and identification. In this regard, the capabilities of the stacked autoencoder are validated as deep learning tools, the purpose of which is to learn significant features from the electrical signals considering different disturbances. The feature learning process is supported by a domain-fusion approach, including time, frequency, and time-frequency domains through the consideration of two signal processing techniques: the FFT and the EMD. The second important aspect is the consideration of multiple conditions of disturbance, above all, the complexity caused by the combination of three different disturbances. There must be a strong emphasis on the favorable performance of all sorts of considered SNR. Finally, the achievement of high performances. In spite of the fact that some of the measurements corresponding to complex scenarios with three combined disturbances are misclassified, it ought to be made of awareness that the assigned condition always includes single or combined disturbances that are present in the original signal. Another important aspect about the proposed method to bear in mind is its flexibility. It means that other features different from those considered in the proposal could be added since the method shows good feature learning capabilities during the characterization stage. The contribution of the work is the fact that it presents a different approach to PQ monitoring for the identification of disturbances. The validation of this work was carried out in two ways: with synthetic signals and with real signals. The synthetic signals have been generated using a set of definitions provided by the standard and developed by software. The second part of the validation has been performed considering real signals from an IEEE database. All in all, the performance of the proposed method is validated through the consideration of 17 conditions, 4 noise levels, and synthetic and real signals, achieving a performance of 99.47% and 91.66%, accordingly. Such proposed method involves the validation of a solution for the detection and classification of PQD that allows the consideration of multiple patterns, including a configuration procedure for its application. Therefore, future works consider the evolution towards incremental learning systems: they also embrace far more complex data and new functionalities, leading the research towards the integration of Novelty Detection and Incremental Learning that have been proposed in other fields of application. Under this framework, it is intended to extend the research towards the performance of this methodology in different industrial applications, in which different disturbance patterns could be found resulting from specific industrial processes. Thus, facing one of the future challenges in the field, that is, the detection of novel disturbances and their learning for latter recognition. For this purpose, the methodology needs to be implemented on digital platforms to optimize computational performance, with online and real-time applications to have in mind.

Author Contributions: Conceptualization, A.-D.G.-A., M.D.-P. and R.-A.O.-R.; Methodology, A.-D.G.-A., M.D.-P. and R.-A.O.-R.; Software, A.-D.G.-A. and M.D.-P.; validation, R.-A.O.-R. and J.-J.S.-D.; Writing—original draft, A.-D.G.-A. and M.D.-P.; Writing—reviewing and editing, R.-A.O.-R. and R.-d.-J.R.-T. All authors have read and agreed to the published version of the manuscript.

Funding: This research work has been partially supported by FOIUQAQ-2018 FIN 201812 and CONACyT doctoral scholarship number 735042. This work has been co-financed by the European Regional Development Fund of the European Union in the framework of the ERDF Operational Program of Catalonia 2014–2020, grant number 001-P-001643.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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

References

1. Ferrero, R.; Collotta, M.; Bueno-Delgado, M.V.; Chen, H.-C. Smart Management Energy Systems in Industry 4.0. *Energies* **2020**, *13*, 382. [\[CrossRef\]](#)
2. Montoya, F.G.; Baños, R.; Alcayde, A.; Montoya, M.G.; Manzano-Agugliaro, F. Power Quality: Scientific Collaboration Networks and Research Trends. *Energies* **2018**, *11*, 2067. [\[CrossRef\]](#)
3. Bravo-Rodríguez, J.C.; Torres, F.J.; Borrás, M.D. Hybrid Machine Learning Models for Classifying Power Quality Disturbances: A Comparative Study. *Energies* **2020**, *13*, 2761. [\[CrossRef\]](#)
4. ANSI C84.1-2006. *American National Standard for Electric Power Systems and Equipment—Voltage Ratings (60 Hertz)*; National Electrical Manufacturers Association: Rosslyn, VA, USA, 2006.
5. Singh, B.; Chandra, A.; Al-Haddad, K. *Power Quality: Problems and Mitigation Techniques*, 1st ed.; Wiley: Hoboken, NJ, USA, 2015; ISBN 9781118922057.
6. Bollen, M.H.J.; Gu, I.Y.-H. *Signal Processing of Power Quality Disturbances*; John Wiley & Sons: Hoboken, NJ, USA, 2006; ISBN 9780471931317.
7. IEEE Power and Energy Society. *IEEE Recommended Practice for Monitoring Electric Power Quality*; IEEE Standard 1159–2009 (Revision IEEE Std 1159–2009); IEEE Power and Energy Society: New York, NY, USA, 2019; Volume 2019, ISBN 9780738159393.
8. Liu, H.; Hussain, F.; Shen, Y.; Arif, S.; Nazir, A.; Abubakar, M. Complex Power Quality Disturbances Classification via Curvelet Transform and Deep Learning. *Electr. Power Syst. Res.* **2018**, *163*, 1–9. [\[CrossRef\]](#)
9. Wang, J.; Xu, Z.; Che, Y. Power Quality Disturbance Classification Based on DWT and Multilayer Perceptron Extreme Learning Machine. *Appl. Sci.* **2019**, *9*, 2315. [\[CrossRef\]](#)
10. Minh Khoa, N.; van Dai, L. Detection and Classification of Power Quality Disturbances in Power System Using Modified-Combination between the Stockwell Transform and Decision Tree Methods. *Energies* **2020**, *13*, 3623. [\[CrossRef\]](#)
11. Mahela, O.P.; Shaik, A.G. Recognition of Power Quality Disturbances Using S-Transform Based Ruled Decision Tree and Fuzzy C-Means Clustering Classifiers. *Appl. Soft Comput.* **2017**, *59*, 243–257. [\[CrossRef\]](#)
12. Cortes-Robles, O.; Barocio, E.; Segundo, J.; Guillen, D.; Olivares-Galvan, J.C. A Qualitative-Quantitative Hybrid Approach for Power Quality Disturbance Monitoring on Microgrid Systems. *Measurement* **2020**, *154*, 107453. [\[CrossRef\]](#)
13. Jamali, S.; Farsa, A.R.; Ghaffarzadeh, N. Identification of Optimal Features for Fast and Accurate Classification of Power Quality Disturbances. *Measurement* **2018**, *116*, 565–574. [\[CrossRef\]](#)
14. Karasu, S.; Saraç, Z. Classification of Power Quality Disturbances by 2D-Riesz Transform, Multi-Objective Grey Wolf Optimizer and Machine Learning Methods. *Digit. Signal Process.* **2020**, *101*, 102711. [\[CrossRef\]](#)
15. Li, P.; Gao, J.; Xu, D.; Wang, C.; Yang, X. Hilbert-Huang Transform with Adaptive Waveform Matching Extension and Its Application in Power Quality Disturbance Detection for Microgrid. *J. Mod. Power Syst. Clean Energy* **2016**, *4*, 19–27. [\[CrossRef\]](#)
16. Lopez-Ramirez, M.; Cabal-Yepez, E.; Ledesma-Carrillo, L.; Miranda-Vidales, H.; Rodriguez-Donate, C.; Lizarraga-Morales, R. FPGA-Based Online PQD Detection and Classification through DWT, Mathematical Morphology and SVD. *Energies* **2018**, *11*, 769. [\[CrossRef\]](#)
17. Kumar, P.K.A.; Vijayalakshmi, V.J.; Karpagam, J.; Hemapriya, C.K. Classification of Power Quality Events Using Support Vector Machine and S-Transform. In Proceedings of the 2016 2nd International Conference on Contemporary Computing and Informatics (IC3I), Noida, India, 14–17 December 2016; Volume 7, pp. 279–284.
18. Zhu, R.; Gong, X.; Hu, S.; Wang, Y. Power Quality Disturbances Classification via Fully-Convolutional Siamese Network and k-Nearest Neighbor. *Energies* **2019**, *12*, 4732. [\[CrossRef\]](#)
19. Mishra, M. Power Quality Disturbance Detection and Classification Using Signal Processing and Soft Computing Techniques: A Comprehensive Review. *Int. Trans. Electr. Energy Syst.* **2019**, *29*. [\[CrossRef\]](#)
20. Wang, S.; Chen, H. A Novel Deep Learning Method for the Classification of Power Quality Disturbances Using Deep Convolutional Neural Network. *Appl. Energy* **2019**, *235*, 1126–1140. [\[CrossRef\]](#)
21. Fu, L.; Zhu, T.; Pan, G.; Chen, S.; Zhong, Q.; Wei, Y. Power Quality Disturbance Recognition Using VMD-Based Feature Extraction and Heuristic Feature Selection. *Appl. Sci.* **2019**, *9*, 4901. [\[CrossRef\]](#)
22. Cai, K.; Hu, T.; Cao, W.; Li, G. Classifying Power Quality Disturbances Based on Phase Space Reconstruction and a Convolutional Neural Network. *Appl. Sci.* **2019**, *9*, 3681. [\[CrossRef\]](#)
23. Shen, Y.; Abubakar, M.; Liu, H.; Hussain, F. Power Quality Disturbance Monitoring and Classification Based on Improved PCA and Convolution Neural Network for Wind-Grid Distribution Systems. *Energies* **2019**, *12*, 1280. [\[CrossRef\]](#)
24. Liu, H.; Hussain, F.; Yue, S.; Yildirim, O.; Yawar, S.J. Classification of Multiple Power Quality Events via Compressed Deep Learning. *Int. Trans. Electr. Energy Syst.* **2019**, *29*, 1–14. [\[CrossRef\]](#)
25. Thirumala, K.; Prasad, M.S.; Jain, T.; Umarikar, A.C. Tunable-Q Wavelet Transform and Dual Multiclass SVM for Online Automatic Detection of Power Quality Disturbances. *IEEE Trans. Smart Grid* **2018**, *9*, 3018–3028. [\[CrossRef\]](#)
26. Goodfellow, I.; Bengio, Y.; Courville, A. *Deep Learning*; MIT Press: Cambridge, MA, USA, 2016.

27. Buduma, N.; Lacascio, N. *Fundamentals of Deep Learning*, 1st ed.; Loukides, M., Cutt, S., Eds.; O'Reilly: Sebastopol, CA, USA, 2017; ISBN 9789814618038.
28. Cardoso, E.S. Advanced Energy Management Strategies for HVAC Systems in Smart Buildings. Ph.D. Thesis, Univesitat Politecnica de Catalunya, Barcelona, Spain, December 2019.
29. Olshausen, B.A.; Field, D.J. Sparse Coding with an Overcomplete Basis Set: A Strategy Employed by V1? *Vision Res.* **1997**, *37*, 3311–3325. [[CrossRef](#)]
30. Zhan, L.; Ma, F.; Zhang, J.; Li, C.; Li, Z.; Wang, T. Fault Feature Extraction and Diagnosis of Rolling Bearings Based on Enhanced Complementary Empirical Mode Decomposition with Adaptive Noise and Statistical Time-Domain Features. *Sensors* **2019**, *19*, 4047. [[CrossRef](#)]
31. Saucedo-Dorantes, J.J.; Delgado-Prieto, M.; Osornio-Rios, R.A.; de Jesus Romero-Troncoso, R. Multifault Diagnosis Method Applied to an Electric Machine Based on High-Dimensional Feature Reduction. *IEEE Trans. Ind. Appl.* **2017**, *53*, 3086–3097. [[CrossRef](#)]
32. He, J.; Ouyang, M.; Yong, C.; Chen, D.; Guo, J.; Zhou, Y. A Novel Intelligent Fault Diagnosis Method for Rolling Bearing Based on Integrated Weight Strategy Features Learning. *Sensors* **2020**, *20*, 1774. [[CrossRef](#)]
33. Qiu, W.; Tang, Q.; Liu, J.; Teng, Z.; Yao, W. Power Quality Disturbances Recognition Using Modified S Transform and Parallel Stack Sparse Auto-Encoder. *Electr. Power Syst. Res.* **2019**, *174*, 105876. [[CrossRef](#)]
34. Alam, M.R.; Bai, F.; Yan, R.; Saha, T.K. Classification and Visualization of Power Quality Disturbance-Events Using Space Vector Ellipse in Complex Plane. *IEEE Trans. Power Deliv.* **2020**, *8977*, 1. [[CrossRef](#)]
35. Thirumala, K.; Pal, S.; Jain, T.; Umarikar, A.C. A Classification Method for Multiple Power Quality Disturbances Using EWT Based Adaptive Filtering and Multiclass SVM. *Neurocomputing* **2019**, *334*, 265–274. [[CrossRef](#)]
36. Singh, U.; Singh, S.N. A New Optimal Feature Selection Scheme for Classification of Power Quality Disturbances Based on Ant Colony Framework. *Appl. Soft Comput.* **2019**, *74*, 216–225. [[CrossRef](#)]
37. Huang, N.; Wang, D.; Lin, L.; Cai, G.; Huang, G.; Du, J.; Zheng, J. Power Quality Disturbances Classification Using Rotation Forest and Multi-resolution Fast S-transform with Data Compression in Time Domain. *IET Gener. Transm. Distrib.* **2019**, *13*, 5091–5101. [[CrossRef](#)]
38. Ribeiro, E.G.; Mendes, T.M.; Dias, G.L.; Faria, E.R.S.; Viana, F.M.; Barbosa, B.H.G.; Ferreira, D.D. Real-Time System for Automatic Detection and Classification of Single and Multiple Power Quality Disturbances. *Measurement* **2018**, *128*, 276–283. [[CrossRef](#)]