

Article

Events Classification in Power Systems with Distributed Generation Sources Using an LSTM-Based Method with Multi-Input Tensor Approach

Oswaldo Cortes-Robles , Emilio Barocio * , Ernesto Beltran  and Ramon Daniel Rodríguez-Soto 

Electrical Engineering Department, Universidad de Guadalajara, Guadalajara 44430, Mexico

* Correspondence: emilio.barocio@academicos.udg.mx

Abstract: In this paper, a long short-term memory (LSTM)-based method with a multi-input tensor approach is used for the classification of events that affect the power quality (PQ) in power systems with distributed generation sources. The considered events are line faults (one line, two lines, and three lines faulted), islanding events, sudden load variations, and generation tripping. The proposed LSTM-based method was trained and tested using the signals produced by the events simulated in a study system with distributed generation sources via PSCAD[®]. Then, noise with different levels was added to the testing set for a thorough assessment, and the results were compared with other well-known methods such as convolutional and simple recurrent neuronal networks. The LSTM-based method with multi-input proved to be effective for event classification, achieving remarkable classification performance even in noisy conditions.

Keywords: distributed generation sources; events classification; deep learning; long short-term memory networks; multi-input



Citation: Cortes-Robles, O.; Barocio, E.; Beltran, E.; Rodríguez-Soto, R.D. Events Classification in Power Systems with Distributed Generation Sources Using an LSTM-Based Method with Multi-Input Tensor Approach. *Electricity* **2023**, *4*, 410–426. <https://doi.org/10.3390/electricity4040022>

Academic Editor: Pavlos S. Georgilakis

Received: 27 September 2023

Revised: 1 November 2023

Accepted: 20 November 2023

Published: 5 December 2023



Copyright: © 2023 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

In the last decade, the incorporation of distributed generation units into electrical power systems, producing distributed generation systems (DGSs), has exponentially increased, inspiring a search for new solutions to current concerns such as growing electrical power demand, increasing generation costs, and the reliability of DGSs, among other environmental and social factors [1]. However, distributed generation systems based on renewable sources, which are the most common, have some drawbacks, mainly regarding their intermittent stochastic behavior and the difficulty of controlling the power generation, which directly affect the power quality (PQ) of the DGS [2,3]. Therefore, these facts have brought new challenges for the protection, control, and analysis of DGSs. Accordingly, monitoring of the power quality (PQ) in DGSs has become a main concern for the power electrical industry.

Regarding PQ monitoring of DGSs, there exist different events producing numerous disturbances such as sags, swells, power interruptions, flickers, oscillatory transients, and harmonics, among others. The classification of these PQ disturbances is crucial since they produce equipment failure and overheating, and severely affect the PQ of a DGS, causing an economic impact in losses of more than USD 20 billion in US and EU-25 countries [4–6]. On the other hand, the classification of an event that produces PQ disturbances is even more relevant to acquire a better understanding of the system's behavior. Event classification provides a record of the failures, allowing a deeper analysis of the system's common issues with the aim of avoiding them. Nevertheless, the classification of the event is more complicated than the PQ disturbance classification due to the intrinsic features of each event which produce similar PQ disturbances, which hinders the data generalization process for their classification.

1.1. Definition of the Problem

PQ events produce different time-scaled disturbances that modify the signal wave, affecting the PQ of the DGS [7,8]. Event classification strategies in the literature can be categorized into domain data-driven and model-based approaches [8]. The model-based method is inherently limited by factors such as the unavailability of a physical model and the potential changes in grid topology during reconfiguration. Therefore, there is a significant emphasis on developing data-driven methodologies for event classification. In this sense, within the general framework of disturbance classification, the signals of an electrical measure of the DGS (voltage) are decomposed, and the time-scaled disturbances are classified based on specific extracted features that enhance the disturbance generalization for its classification. This is a well-stated multi-step process where different techniques have been combined to obtain several disturbance classification methods [7,9–13].

On the other hand, different events produce similar PQ disturbances in the signals. This fact is shown in numerous disturbance classification papers by considering different events for producing a certain PQ disturbance for the classification, as in [14–17]. Here, events such as switching loads and generation tripping are produced to generate voltage variations, while faults and switching capacitor bank events produce oscillatory transients. In this regard, in the case of the fault, the principal distinction for the type of fault recognition mainly lies in the number of faulted lines, as can be appreciated in [18–20]. Other examples are the voltage variations caused by the generation/demand relation changes produced in the islanding events, generation tripping, or load variations. Related to this, the similarities in the voltage variations caused by load variations and generation tripping have been proven in [21], while in [22], a major concern for islanding detection is the load/generation switching or faults under generation/demand equality conditions of the system, named the non-detection zone.

Therefore, the main challenge related to the classification of events is the existence of features shared with other events that hinder their discrimination, which makes it necessary to use a different approach to highlight the particular features of each single event. In consequence, event classification is a more challenging assignment than disturbance classification.

1.2. Related Works

The studies proposed in the literature are based on the Discrete Wavelet Transform (DWT) [23–25] for feature extraction and classification, which are efficient mathematical functions to represent the characteristics of a signal and extract information in the time and frequency domains. This technique is especially suitable for dynamic signals. However, its performance is limited in the presence of noisy data, since its accuracy decreases considerably. Similarly, the Discrete Fourier Transform (DFT) and Short Time Fourier Transform (STFT) [8,26] are computationally efficient techniques. Nevertheless, they face challenges when it comes to analyzing and detecting frequency events.

Related to event classification, the most common events considered for this task are the different fault types. According to this, in [18], different types of faults such as line-ground, line-line, line-line-ground, three-line, and three-line-ground, are classified based on the waveforms of the generated signals, using a convolutional neuronal network. The paper shows that the current's waveforms are suitable for fault classification. The method achieves an accuracy of over 99% for the classification of the fault types. Nevertheless, the classification of different events requires a greater effort to carry out.

On the other hand, the application of the Stockwell transform (ST) for PQ analysis and classification has acquired significant popularity in recent years [27–30]. It is important to mention that in [30], an islanding identification method using the current signals in a distribution grid with renewable distributed sources is proposed. In the research work, an islanding recognition factor (IRF) is computed by processing the current signal through the Stockwell transform (ST) to recognize the islanding events from the non-islanding events. Then, the discrimination of the islanding events is carried out using decision

rules. The decision rules focus on the islanding identification, leaving aside the individual classification of the rest of the events, globally considering them as faulty and operational events. In this sense, a broader classification of the different events is beneficial for a better understanding of the system through the event record.

In [31], six types of events are produced in a test system with nine buses and three generation sources. The events include line–ground, line–line, and three-line faults, load switching, capacitor switching, and transformer energizing. Here, only the voltage is used for the classification of the six types of events based on wavelet transforms and support vector machines. Although the method obtains high-accuracy results, the success of the method depends on the proposed classification of events, which does not include a variety of events with similar features.

Upon review of the related works, it is clear based on the lack of variety in the classified events that event classification is a challenging assignment. This fact supports the need for more specialized methods capable of classifying different events, even when the events present similar behavior in their features. For this aim, the development of new technologies for sensing and processing the data from a DGS has opened new opportunities for event classification. In this sense, measurement devices such as μ PMUs are being placed at strategic sensing points along DGSs [32], increasing the availability of relevant information necessary to determine the type of event produced in a DGS.

Accordingly, a multi-input approach is proposed in this paper for the classification of varied events with high accuracy.

1.3. Paper Contributions

The aim of this paper is to classify the variety of events produced in power systems with distributed generation sources. The drawback of most of the aforementioned methods for event classification is the lack of variety in the measured variables to more accurately describe each of the events. Therefore, in this paper, different variables obtained from the measurement devices are used to improve the generalization of disturbances and facilitate the classification of an event by using a deep learning (DL)-based technique capable of processing the information from the measurement devices in a closed-loop process.

In this regard, the contributions of this paper to the existing science field under the event classification framework can be summarized as follows:

- A proposed multi-input approach based on a multi-variable tensor;
- An event classification method based on LSTM for event records;
- An analysis of the influence of events on different measured variables of the DGS.

The present paper is structured as follows. The proposed method is described in Section 2.1. Then, the dataset for training the classification model is detailed in Section 2.2. The model training is carried out in Section 2.3. The results are depicted in Section 3. Finally, the conclusions of the research work are drawn in Section 4.

2. Materials and Methods

2.1. Proposed Method

The multi-input approach used in this paper is based on the use of a training/testing multi-variable tensor that contains the signals produced by different disturbance events from the perspective of different variables, such as voltage, current, phase angle, and the active and reactive powers. The multi-variable tensor with the information of the disturbance events allows their classification; otherwise, if it were only focused on the voltage, classification would be a hard task.

2.1.1. Multi-Input Tensor for the Classification of Disturbance Events

Tensors, as multi-dimensional arrays, describe the relationship (multi-linear) between sets of algebraic objects related to a vector space. Therefore, tensors may map objects as scalars, vectors, or even other tensors. For the presented approach, this tensor property

allows the visualization of each disturbance event in a multi-dimensional vector space formed by different vector variables, where its classification is facilitated.

Thus, starting from the sampled data, the signals associated with the system variables are defined as

$$\mathbf{Y}_d^n(t) = [x_{d,1}^n, \dots, x_{d,t}^n, \dots, x_{d,T}^n] \tag{1}$$

where T is the number of time steps for $d = 1, 2, \dots, D$ variables are grouped to represent the n th disturbance event data as:

$$\mathbf{X}_n = \begin{bmatrix} \mathbf{Y}_1^n(t) \\ \vdots \\ \mathbf{Y}_d^n(t) \\ \vdots \\ \mathbf{Y}_D^n(t) \end{bmatrix} \tag{2}$$

Consequently, the data of the disturbance event (\mathbf{X}_n) are equal to a slice of the multi-input dataset tensor, such as:

$$\mathcal{T} = [\mathbf{X}_n(:, :, N)] \tag{3}$$

where N is the number of disturbance events regarded as the batch size of the multi-input tensor, as can be seen in Figure 1.

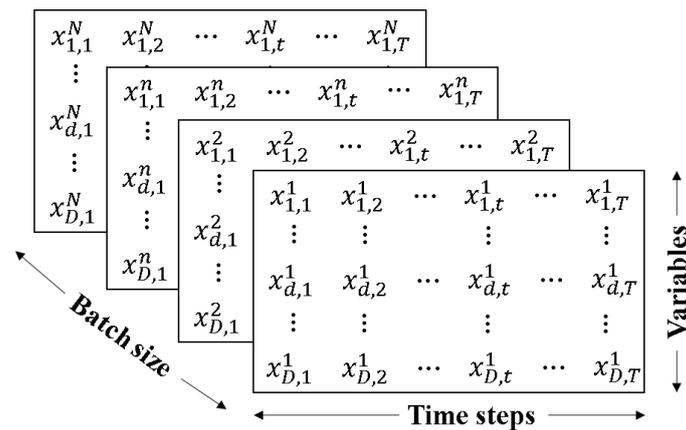


Figure 1. Structure of the multi-variable tensor \mathcal{T} .

Figure 1 shows the structure of the multi-input dataset tensor described in the equations from (1) to (3). Here, the slices of the multi-input tensor correspond to the multi-variable data representing the disturbance events for the training and testing of the classification model. The above multi-input dataset tensor \mathcal{T} is afterwards normalized by applying min-max normalization to each variable $\mathbf{Y}_d^n(t)$, as follows:

$$\hat{\mathcal{T}} = [\hat{\mathbf{X}}_n(:, :, N)]$$

$$\hat{\mathbf{X}}_n = \begin{bmatrix} \hat{\mathbf{Y}}_1^n(t) \\ \vdots \\ \hat{\mathbf{Y}}_d^n(t) \\ \vdots \\ \hat{\mathbf{Y}}_D^n(t) \end{bmatrix} \tag{4}$$

$$\hat{\mathbf{Y}}_d^n(t) = \frac{\mathbf{Y}_d^n(t) - \mathbf{Y}_d^n(t)_{min}}{\mathbf{Y}_d^n(t)_{max} - \mathbf{Y}_d^n(t)_{min}}$$

Thereby, the min-max normalization yields a normalized data tensor $\hat{\mathcal{T}}$ with individually scaled variables from 0 to 1, which is used as the multi-input dataset for training and testing of the classification model.

2.1.2. Classification Model Based on LSTM Networks

The generation of the classification model is carried out based on the training data obtained from the 20-fold data split of the normalized multi-input dataset tensor $\hat{\mathcal{T}}$. These training data along with their labels are processed by the long short-term memory (LSTM)-based deep neuronal network (DNN), in order to generate a classification model that fits the expected disturbance events.

2.1.3. Architecture of the LSTM-Based DNN

The deep neuronal network (DNN) is principally composed of recurrent layers based on long short-term memory cells and dense layers. The architecture of the LSTM-based DNN for the classification of disturbance events is shown in Figure 2. In this regard, the normalized training data $\hat{\mathcal{T}}$ are processed by the LSTM-based recurrent layers that mainly deal with the sequential data to obtain crucial features about the disturbance, allowing the dense layers to classify the disturbance event. Furthermore, a batch normalization and a dropout layer were added to encourage the generalization capability of the classification model.

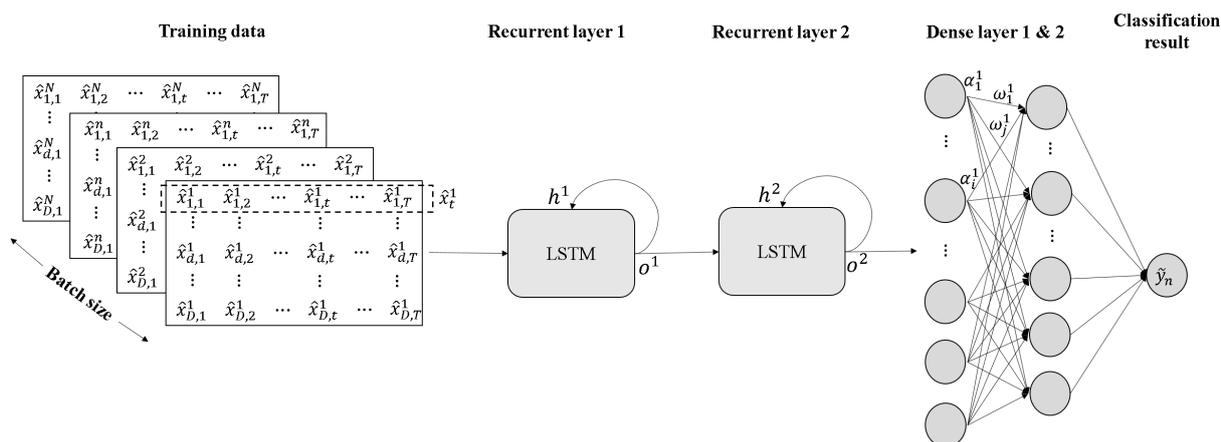


Figure 2. Architecture of the LSTM-based deep neuronal network.

2.1.4. LSTM-Based Recurrent Layers

The present recurrent layers are formed by long short-term memory (LSTM) units. Recently, the LSTM-based DNN has been proven suitable for classifying, processing, and making predictions based on time series data, since there can be lags of unknown duration between important events in a time series. Furthermore, the LSTM units were mainly developed to deal with the vanishing gradient problem [X] that can be encountered when training traditional recurrent neuronal networks (RNNs).

A common LSTM unit is composed of a memory cell, an input gate, an output gate, and a forget gate. The memory cell remembers values over arbitrary time intervals and the three gates regulate the flow of information into and out of the cell.

Figure 3 shows a common LSTM unit structure. In this figure, the variable \hat{x}_t^n represents the multi-variable input data of the batch n at the time t ; the variables o_{t-1}^l and o_t^l refer to the last cell output, and the current output of the layer l (being $l = 1, 2, \dots$, until the LSTM-based layer number), respectively; the variables h_{t-1}^l and h_t^l correspond to the hidden state or memory information of the last and current LSTM unit; lastly, the variables r_t^l , s_t^l , q_t^l , and u_t^l are explained below. The above variables are carried through different operations such as tensor addition (+), concatenation ([|]), and activation functions (ϕ), in order to update the output and hidden state of the unit. Furthermore, the modules σ denote sigmoid functions approximating a binary response 0, 1. Its operation in combination with

the product (\otimes) can be understood as switches that allow, or not, the flow of information giving rise to the input, output, and forgetting information gates.

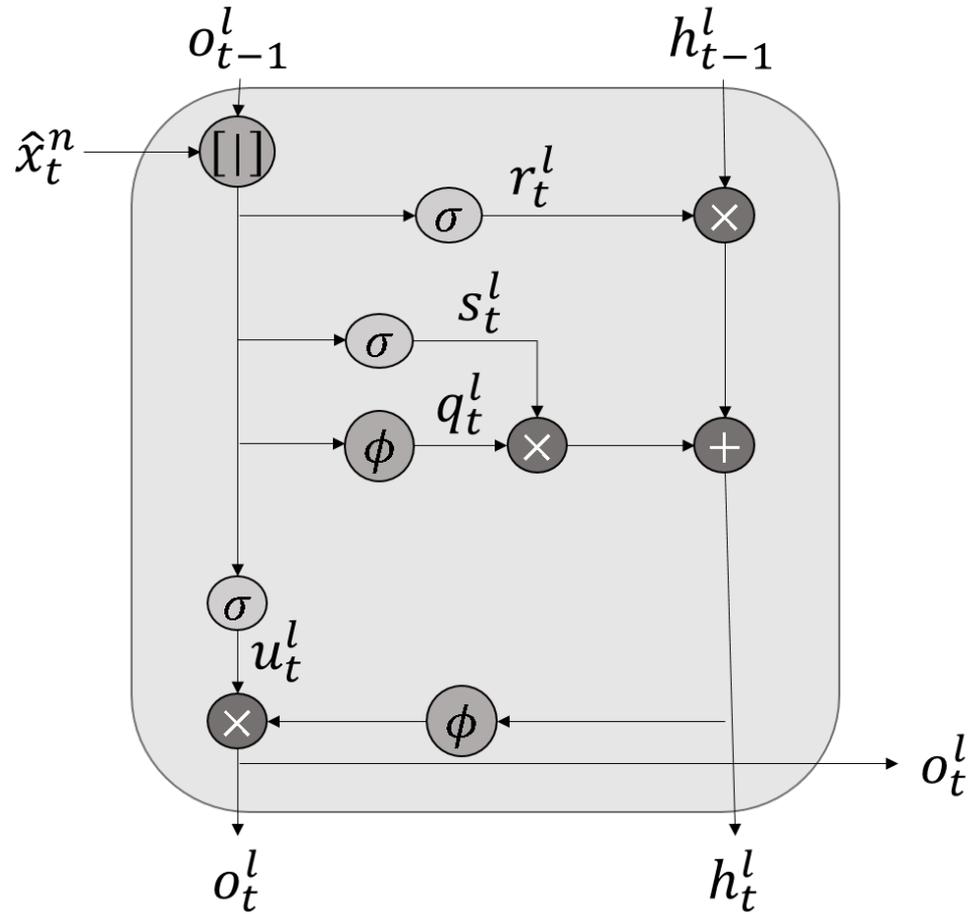


Figure 3. Long short-term memory unit structure.

In this regard, the long-term memory allows crucial information to pass from h_{t-1}^l to h_t^l . Then, the current hidden state h_t^l is obtained as the sum of information in the last memory cell h_{t-1}^l that we want to preserve, and the new information q_t^l that we want to add to the memory cell, as follows:

$$h_t^l = r_t^l * h_{t-1}^l + s_t^l * q_t^l \tag{5}$$

where the information in q_t^l to add is obtained from processing the input data and the previous output:

$$q_t^l = \phi(w_q^l [\hat{x}_t^n | o_{t-1}^l] + b_q^l) \tag{6}$$

On the other hand, r_t^l and s_t^l are switches that control the forgetting or remembering of the information in the past memory and the new information, depending on the x_t^n input data and the previous output o_{t-1}^l . Therefore, r_t^l and s_t^l are computed as:

$$r_t^l = \sigma(w_r^l [\hat{x}_t^n | o_{t-1}^l] + b_r^l) \tag{7}$$

and

$$s_t^l = \sigma(w_s^l [\hat{x}_t^n | o_{t-1}^l] + b_s^l) \tag{8}$$

where w^l are the weights, and b^l are the biases associated to the switches q_t^l, r_t^l, s_t^l .

Moreover, the current output is obtained from the resulting operation (\otimes) between the previous output and the input data by the selection switch u_t , calculated using:

$$u_t^l = \sigma\left(w_u^l \left[\hat{x}_t^n \mid o_{t-1}^l\right] + b_u^l\right) \quad (9)$$

and the updated hidden state h_t^l , as:

$$o_t^l = u_t^l * \phi\left(w_o^l h_t^l + b_o^l\right) \quad (10)$$

Then, the output of the first LSTM-based layer (o^1) is thereupon used as the input of the following layer, repeating the process from Equations (5)–(10).

2.1.5. Dense Layers

On the other hand, the dense layers process the output of the final recurrent layer (o^2) in order to obtain a predicted label of the disturbance event classified. Thus, the dense unit processes the input data as a result of the activation of the unit by its activation function:

$$\alpha_i^l = \phi(z_i^l) = \max(0, z_i^l) \quad (11)$$

where α_i^l is the result of the activation of the i th unit in the l th dense layer. For the present classification model, the activation Relu is selected as the activation function of the first dense layer. The Relu activation function yields that α_i^l be zero when the value of z_i^l is less than zero; otherwise, α_i^l is equal to z_i^l . In this regard, z_i^l depends on the weights and bias associated with the present unit with respect to the last layer:

$$z_i^l = \sum_j w_j^l \alpha_j^{l-1} + b_i^l \quad (12)$$

where w_j^l are the weights associated with the results of the units in the previous layer α_j^{l-1} , and b_i^l is the bias of the unit. In this sense, for the first dense layer, the previous results would be the results from the last recurrent layer (i.e., $\alpha_j^{l-1} = o_t^2$).

Finally, the predicted label of the classified disturbance event is obtained at the last dense layer, computing the output of the previous dense layer by its Softmax activation function, as follows:

$$\tilde{y}_n = \phi(z_i^l) = \frac{e^{z_i^l}}{\sum_j e^{z_j^l}} \quad (13)$$

where \tilde{y}_n is the predicted label for the n th disturbance event classified.

The complete methodology for the classification of different disturbance events is summarized in Figure 4. From this figure, it can be seen that the signals associated with the different variables of the system are sampled from the studied electrical power system (EPS) with distributed generation sources (DGS) for the different events (Equation (2)). Then, the signals collected are grouped by their corresponding event to form slices of the multi-variable tensor (Equation (3)). Afterwards, the signals of the tensor are normalized (Equation (4)), and the normalized tensor is split into the training and testing datasets. Subsequently, the training dataset along with its event labels is used to generate the model for the event classification.

Lastly, the trained model is used to classify the testing data, and the resulting classes are compared to the actual event labels of the testing set in order to evaluate the performance of the classification model.

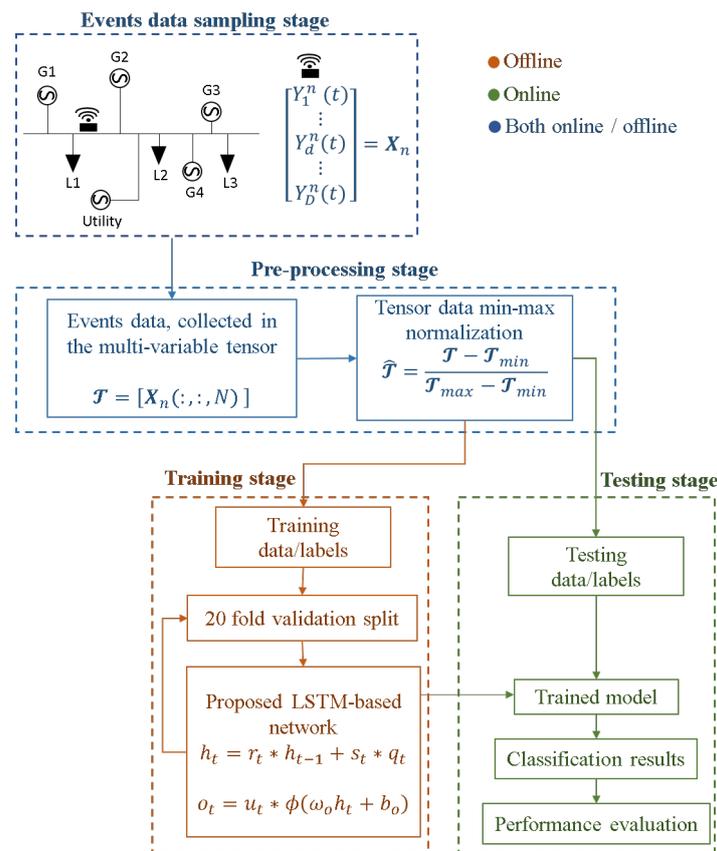


Figure 4. Methodology for the proposed approach.

2.2. Dataset Generation for Disturbance Event Classification

The methodology for the dataset generation is explained in this section. Hence, the distributed generation study system is first described. Next, the considerations for the disturbance event generation, including the type of event and parameter variation, are depicted. Finally, the system's selected variables for building the multi-input dataset tensor are listed.

2.2.1. Distributed Generation Study System

The distributed generation study system consists of a PSCAD-modeled distributed network connected to a high voltage (HV) network rated at 115 kV, via a 0.6/115 kV step up transformer. Here, different types of generation units such as a wind turbine, photovoltaic array, and synchronous generator are connected to the 0.6 kV-rated point of common coupling (PCC). Figure 5 shows the study system.

The attributes of the system's distributed generation units are the following.

1. Wind Generator type 1 (WG 1). Different from the previous generation unit, this is a fixed-speed wind turbine with a squirrel-cage induction generator. The generator operates at a line voltage of 0.6 kV with a frequency of 60 Hz. The wind turbine is represented by the input torque ($T = -0.8$ PU) to the generator. This type of induction machine cannot excite itself. Therefore, in order to reduce the amount of reactive power drawn into the machine during startup (hence, limit inrush currents), it uses a thyristor-based soft starter.
2. Wind Generator type 3 (WG 3). This is a variable-speed wind turbine with a doubly fed wound-rotor induction generator. The generator operates at a line voltage of 0.6 kV with a frequency of 60 Hz. The wind turbine is represented by the input torque ($T = -0.25$ PU) to the generator. Through the use of power electronics, reactive power can be supplied to the machine via the rotor. Hence, no reactive power needs to be

- drawn from the system during start-up. This wind turbine is located 2 km away from the system bus and is connected through a 25 kV transmission line (represented by a pi-section). Note that the voltage is stepped up to 25 kV along the transmission line and stepped back down to 0.6 kV at the wind generator.
3. Photovoltaic generator. A positive and negative DC voltage is outputted from the PV array and sent to a DC/DC converter for the purpose of maximum power point tracking. The DC voltage is then sent through a power electronic inverter, which converts it to an AC voltage with a magnitude of approximately 0.23 kV and a frequency of 60 Hz. The voltage is then stepped up using a 0.23/0.6 kV step up transformer and sent to the system.
 4. Synchronous Generator (Synch. Gen.). A synchronous generator is driven by a small hydro turbine which is initialized to operate at its rated conditions. The amount of power generated by the turbine is controlled by the governor. The synchronous generator is rated at 100 kVA, with a line voltage of 0.6 kV and a frequency of 60 Hz. Its field windings are connected to an exciter, which is used to magnetize the machine. Hence, no reactive power will be drawn from the system.

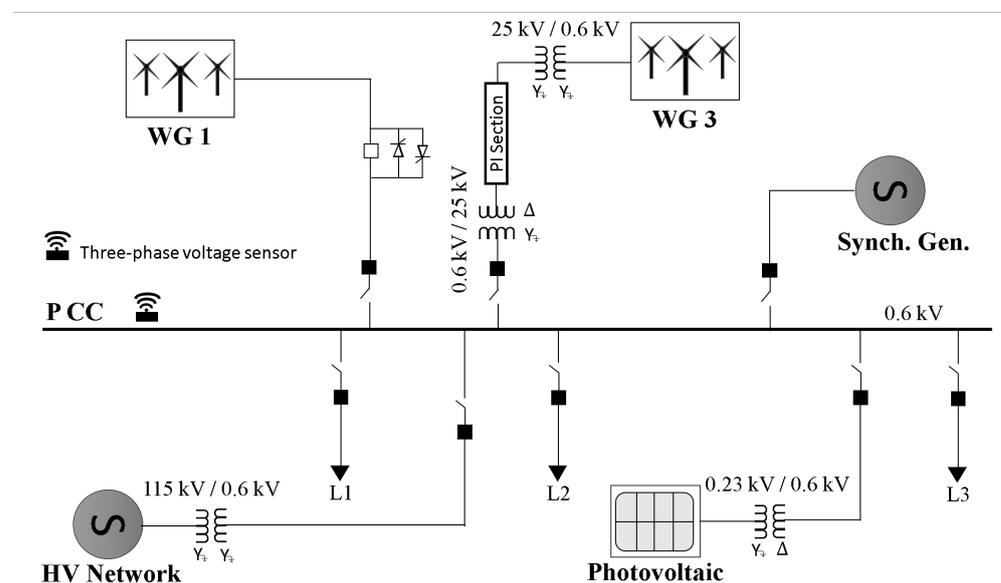


Figure 5. Distributed generation study system.

Besides the generation units, the distributed system has three loads located along the PCC demanding active and reactive power. The active and reactive power rating of the generation units and loads are shown in Table 1.

Table 1. Active and reactive power rating for generation/demand.

Power	WG 1	WG 3	Photovoltaic	Synch. Gen.	L1	L2	L3
P (kW)	19.34	59.82	21.83	90.92	60	60	80
Q (kVAr)	26.17	7.644	0.302	39.54	20	20	40

Thus, the produced disturbance events for the classification are selected based on the system traits and different considerations as suggested below in the dataset generation section.

2.2.2. Produced Disturbance Events

The classes of disturbance events produced in the study system for event classification are listed in Table 2.

Table 2. Classes of disturbance events.

Class	Event Description
C1	Normal working (undisturbed)
C2	One line faulted
C3	Two lines faulted
C4	Three lines faulted
C5	Islanding event
C6	Sudden load variation
C7	PV generation unit tripping
C8	Synchronous generation unit tripping

Eight different classes of events are produced in the study system. The events include the normal working of the system, various types of faults (mono-phase, line-to-line, and three-phase), disconnection from the main network (islanding), load variation, and generation tripping.

For the fault diagnosis, all types of faults are grounded. On the other hand, the faulted line is shifted for the mono-phase and line-to-line analysis cases. Additionally, the following facts were applied to the event cases for dataset generation:

- Power system demand reduction from 100% to 10% (10 cases);
- Disturbance event time-span of 0.2, 0.4, 0.6 s (three cases);
- Disturbance event onset time in 0.1, 0.5, 0.7 s (three cases).

The above considerations generate 90 cases (per phase) for each event class, giving a total number of 2160 event cases in the dataset.

2.2.3. Selection of the System Variables

Voltage and current are the most common system variables used for the characterization of a system's behavior. The voltage was widely used for disturbance classification in [5–8], while the current was used in [18–20] for fault classification. Nevertheless, the features of some disturbance events' behavior require more system variables to classify them. For example, the islanding event (C5), load variation (C6), and different units' generation tripping (C7, C8) produce similar voltage and current variations, due to the imbalance of the demand/generation relation. This fact severely affects their classification by only using the voltage and current variables. In this regard, variables of active power, reactive power, and phase angle allow a better generalization of the events, improving their classification.

For each of the 2160 event cases, the disturbed signals corresponding to the Voltage (V), Current (I), Active power (P), Reactive power (Q), and Phase angle (θ) produced by the event are obtained from the PCC with a sampling frequency of 1kHz. The recorded signals for the dataset are organized in a multi-input tensor with dimensions of $2160 \times 5 \times 1000$, where 2160 corresponds to the total number of generated events, 5 is the number of system variables, and 1000 is the signal's resolution.

2.3. Training of the Classification Model

The training of the model is carried out using the multi-input dataset tensor through the computation of the loss function of the LSTM-based DNN. Furthermore, certain actions are employed in order to prevent the overfitting of the data during the training of the classification model.

2.3.1. Prevention of Overfitting in Training

Overfitting of the classification model is one of the main issues in DNN. Overfitting produces an erroneous generalization of the data, causing classification problems when classifying new data. For this reason, the following actions are applied to prevent data overfitting in the training process of the classification model.

- Twenty-fold validation split of the training data. Here, the training data are separated into training (80%) and validation (20%) data, aiming to validate with new data the accuracy of the classification model during each epoch.
- Balance of the training dataset. For this action, each class in the training dataset has the same number of elements.
- Batch normalization. A batch normalization layer is included after the LSTM layers for the normalization of the output data information.

The classification model is trained through loss function computation using the training data with the above considerations.

2.3.2. Loss Function Computation

In the present method, the sparse categorical cross entropy is used to calculate the losses to train the model. The loss function is defined as:

$$H(y, \tilde{y}) = - \sum_{n=1}^N y_n \log(\tilde{y}_n) \quad (14)$$

where $n \in N$ is the total number of classification events, and y_n and \tilde{y}_n are the actual and predicted labels for the n th disturbance event in the label vector y and \tilde{y} , respectively. The loss function is solved through backpropagation using the adam optimiser, letting a trained model be H .

Then, the trained model H is used to classify different disturbance events in the distributed generation study system.

3. Results and Discussion

In this section, the results of the validation and testing of the classification method are shown. First, the trained model is validated using the 20-fold validation data. The classification results of the validated model are depicted by its classification training/validation accuracy curves. Then, the validated model is thoroughly tested using the testing data with different noise levels. Here, the results of the proposed approach are compared with other well-known methods proven for classification in the literature, such as simple recurrent neuronal networks (RNN) and convolutional neuronal networks (CNN). Finally, laboratory data are classified via the proposed method and the results are discussed.

3.1. Validation of the Classification Model

Validation of the classification model is crucial during the training process and aims to prevent the model from overfitting. For this reason, twenty percent of the training data are separated during the training and used to validate the classification accuracy of the model. In this regard, the classification training/validation accuracy curves are depicted in Figure 6.

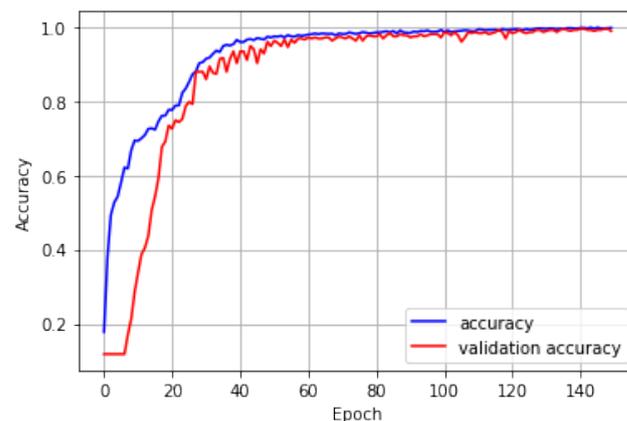


Figure 6. Accuracy curves for the training/validation of the LSTM-based classification model.

From Figure 6, it can be seen that the validation accuracy of the LSTM-based model reaches 99.5% along the 150 epochs of the training, close to 100% of training accuracy. Furthermore, there are no elongated down spikes during the training process. These facts suggest that the model was adequately trained, avoiding overfitting. On the other hand, the final validation accuracy forecasts the accuracy ratio for the classification of new data.

3.2. Testing of the Validated Classification Model Using Disturbance Events Produced in the Study System

For testing the validated model, 800 new disturbance events (100 per class) are produced in the study system considering the parameters stated in Section 2.2.2. The classification results for the new events are shown in Figure 7.

1	100	0	0	0	0	0	0	0
2	0	100	0	0	0	0	0	0
3	0	0	100	0	0	0	0	0
4	0	0	0	100	0	0	0	0
5	0	0	0	0	100	0	0	0
6	0	0	0	0	1	99	0	0
7	0	0	0	0	0	0	100	0
8	0	0	0	0	1	0	0	99
	1	2	3	4	5	6	7	8

Figure 7. Confusion matrix for classification of the disturbance events.

The confusion matrix in Figure 7 allows the identification of the classification success of the new events, with only one misclassification existing for classes 6 and 8. Furthermore, the confusion matrix enables the assessment of the model’s performance through evaluation metrics. In this regard, different evaluation metrics which have been widely used in the literature were selected to assess the LSTM-based model performance for the classification of disturbance events.

For this aim, the selected evaluation metrics, which focus on the quantification of true/false positive and negative classifications, are recall, precision, F1-Measure, and accuracy [17]. The computation of the evaluation metrics recall, precision, and F1-Measure for the assessment of the method’s classification performance is shown in Figure 8.

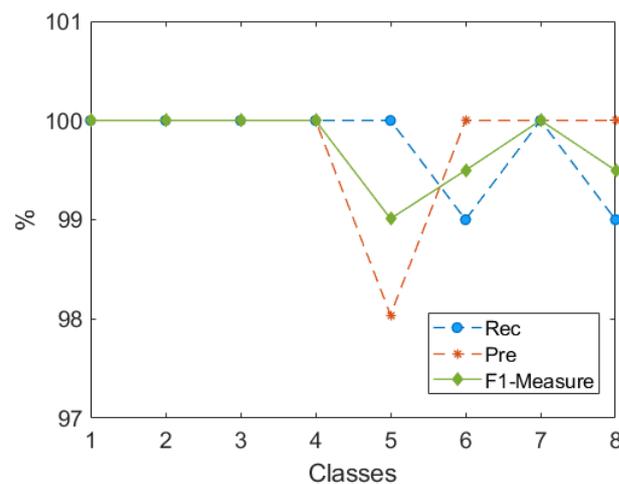


Figure 8. Results of the classification performance assessment by the evaluation metrics.

In Figure 8, the classification performance per class of the PQ events can be seen. This graph easily enables the identification of misclassification issues, aiming to determine any classification problems due to training failure. In this respect, the graph shows a decrease in the F1-measure only for classes 5, 6, and 8, corresponding to the islanding events, the sudden load variations, and the synchronous generation unit tripping, respectively. For the case of the sudden load variations (C6) and the synchronous generation unit tripping (C8), their decreasing responses to a single misclassification as an islanding event (C5) affected their recall metric computation and were captured by the F1-Measure. The worst classification performance, based on the F1-Measure, is for the islanding events (C5) due to the false positives in C6 and C8. This classification issue occurred because the islanding event is reflected in the generation/demand relation, such as load variations and unit tripping. Nevertheless, the method's classification performance is still outstanding, even with the similarities between the intrinsic features of some events, such as the different types of faults (C2–C4), or those related to the generation/demand imbalance (C5–C8).

It is clear that the LSTM-based method has remarkable performance in classifying the produced events in noise-free conditions. However, noise is naturally present in real-world systems. For this reason, the testing events were corrupted by adding different levels of Gaussian noise to mimic real-world signals. Then, the LSTM-based method was tested using the noisy events of the testing set with 50 dB, 40 dB, and 30 dB of Gaussian noise added. The accuracy results of classification of the testing set under different noise levels are shown in Figure 9.

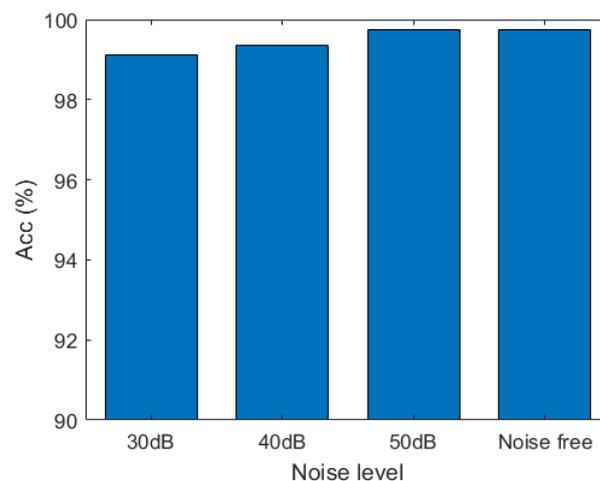


Figure 9. Classification accuracy results under noisy conditions.

The accuracy results of the classification of the noisy testing set depicted in Figure 9 show an outstanding performance of the LSTM-based method even with 30dB of added noise, achieving 99.12% accuracy. Meanwhile, the method achieves an accuracy value of 99.35% for added noise of 40 dB. Finally, for a low noise level of 50 dB, the accuracy reaches 99.75%, comparable to the noise-free testing set. In summary, the LSTM-based method performs remarkably well in event classification even in noisy conditions, maintaining an accuracy over 99% for all cases. In this regard, the exceptional classification performance is due to the multi-input tensor array which accomplishes the correct data generalization to facilitate PQ event classification.

In order to keep assessing the LSTM-based method's performance, it is compared against other well-known methods for data classification reported in the literature. For this purpose, the methods selected are simple recurrent neuronal networks (RNNs), and convolutional neuronal networks (CNNs).

For this case, the RNN has the same architecture as the LSTM depicted in Figure 2, but with RNN neurons instead of LSTM neurons. On the other hand, the CNN architecture is adapted from [33], where it is used for PQ disturbance classification. The CNN has

four 2-D convolutional layers and three dense layers. The first two convolutional layers with 32 filters are followed by a MaxPooling layer and a dropout layer. Then, two more convolutional layers with 64 filters are followed by a GlobalMaxPooling layer. Lastly, there are two dense layers with Relu activation function for their 256 and 128 neurons, respectively, and a final dense layer with eight neurons activated by Softmax function corresponding to the number of PQ events to classify. The input data are the multi-input tensor arrays for all the cases. The training and testing for all the methods were carried out in Google Colaboratory (CO).

The results of the method performance comparison, detailed in Table 3, show a remarkable accuracy rate of over 90% for the compared methods. This is because of the multi-input tensor array used as input for all methods. Nevertheless, it can be seen that the method with the best classification accuracy results is LSTM-based. In this regard, the CNN-based method shows comparable results with respect to the LSTM, but with a much longer training time even when the CNN employs fewer training epochs. Furthermore, the LSTM avoids the vanishing gradient problem present in the RNN, which evidently affects the RNN's classification accuracy.

Table 3. Classification performance comparison.

Type	Noise Levels			Free	ETP ^a	TT ^b (s)
	30 dB	40 dB	50 dB			
LSTM	99.12	99.35	99.75	99.75	150	25
RNN	94.86	95.37	96.12	96.55	150	24
CNN	99.12	99.12	99.25	99.55	10	312

^a (Elapsed Training Epochs). ^b (Training Time).

3.3. Discussion

The outstanding classification performance is achieved due to the successful combination of the system variables selected for building the multi-input dataset tensor. This fact is demonstrated by the comparison of the validation accuracy performance shown in Figure 10.

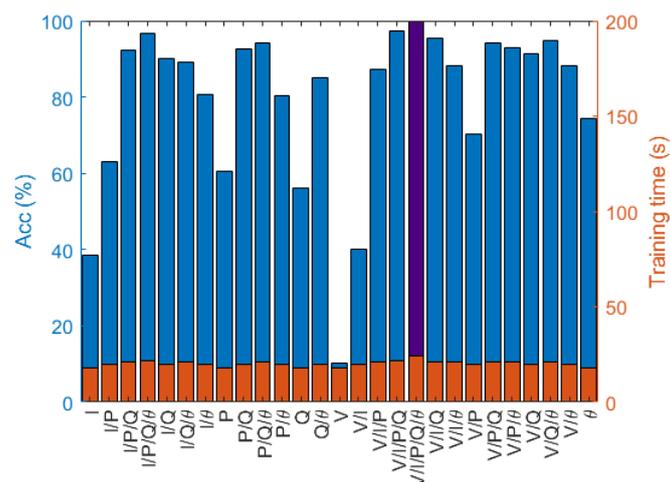


Figure 10. Validation accuracy for different system variable combinations.

From Figure 10, it can be appreciated that the validation accuracy for the single variables voltage (V), current (I), active power (P), reactive power (Q), and phase angle (θ) is low compared with the rest of the combinations except for the combination of V/I, as was expected. In this regard, the combinations that include P, Q, and θ achieve higher validation accuracy; the proposed multi-input tensor with V/I/P/Q/ θ is the one with the best accuracy, being 99.5%. On the other hand, although the combinations with four variables, and even with three, achieve accuracy rates over 90%, the non-considered

variables affect the final classification producing misclassification for certain events. For example, the combination of V/I/P/Q has a classification rate of 97%, but the majority of its misclassification is related to the generation tripping cases, which makes it non-viable for the recording of faulted generation sources. Moreover, the training time grows according to the number of combined system variables, being approximately 17 s for single variables, 19 s for two-variable combinations, 21 s for three-variable combinations, 23 s for four-variable combinations, and 25 s for the combination of five variables. However, this time increase is negligible considering that the training is an offline process. Hence, the remarkable effectiveness of the selected variables in the multi-input tensor is clear.

4. Conclusions

This paper addresses PQ event classification. The paper's main contributions to the state of the art can be summarized into two points: the proposed multi-input tensor to identify different PQ events produced by generation/demand changes and faulted lines; and the fast-training classification method based on the LSTM network for PQ event classification.

Traditional approaches for PQ event classification consider voltage and current values as features for their classification. However, this consideration limits the types of issues that can be analyzed, hindering the classification of events with similar characteristics, such as islanding, load variation, or generator tripping. In this regard, the proposed multi-input tensor proved to be effective in identifying the underlying causes of the considered PQ events.

On the other hand, besides the outstanding results achieved by the compared methods based on RNN and CNN, the proposed method, which employs an LSTM network, trained over 10 times faster than the CNN-based method, which is advantageous for retraining the model, allowing the classification model to be constantly updated. In this regard, the evaluation of model degradation is intended to be studied in future work. Finally, the long short-term memory method avoids the vanishing gradient problem present in the RNN, improving the classification accuracy of the proposed method.

Author Contributions: Conceptualization, O.C.-R. and E.B. (Emilio Barocio); methodology, O.C.-R.; software, O.C.-R.; validation, O.C.-R., E.B. (Emilio Barocio) and E.B. (Ernesto Beltran); formal analysis, O.C.-R., E.B. (Emilio Barocio), E.B. (Ernesto Beltran) and R.D.R.-S.; investigation, O.C.-R. and E.B. (Emilio Barocio); resources, O.C.-R. and E.B. (Emilio Barocio); data curation, O.C.-R.; writing—original draft preparation, O.C.-R., E.B. (Emilio Barocio) and E.B. (Ernesto Beltran); writing—review and editing, O.C.-R., E.B. (Emilio Barocio), E.B. (Ernesto Beltran) and R.D.R.-S.; visualization, O.C.-R., E.B. (Emilio Barocio) and E.B. (Ernesto Beltran); supervision, E.B. (Ernesto Beltran) and E.B. (Emilio Barocio); project administration, E.B. (Emilio Barocio). All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Data are contained within the article.

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

Abbreviations

The following abbreviations are used in this manuscript:

LSTM	Long short-term memory
PQ	Power quality
DGS	Distributed generating systems
RNN	Recurrent neuronal networks
CNN	Convolutional neuronal networks

References

1. Parag, Y.; Ainspan, M. Sustainable microgrids: Economic, environmental and social costs and benefits of microgrid deployment. *Energy Sustain. Develop.* **2019**, *52*, 72–81. [[CrossRef](#)]
2. Alkahtani, A.; Alfalahi, S.; Athamneh, S.; Al-Shetwi, A.; Mansor, M.; Hannan, M.; Agelidis, V. Power quality in microgrids including supraharmonics: issues, standards, and mitigations. *IEEE Access* **2020**, *8*, 127104–127122. [[CrossRef](#)]
3. Baptista, J.; Faria, P.; Canizes, B.; Pinto, T. Power Quality of Renewable Energy Source Systems: A New Paradigm of Electrical Grids. *Energies* **2022**, *15*, 3195. [[CrossRef](#)]
4. Chen, C.; Chen, Y.; Chin, Y.; Chen, H. Integrated power-quality monitoring mechanism for microgrid. *IEEE Trans. Smart Grid* **2018**, *9*, 6877–6885. [[CrossRef](#)]
5. The Electric Power Research Institute (EPRI). *The Cost of Power Disturbances to Industrial and Digital Economy Companies Product 1006274*; EPRI: Washington, DC, USA, 2001.
6. Targosz, R.; Manson, J. Pan-European power quality survey. In Proceedings of the 2007 9th International Conference on Electrical Power Quality and Utilisation, Barcelona, Spain, 9–11 October 2007; pp. 1–6.
7. Chawada, G.; Shaik, A.; Shaik, M.; Padmanaban, S.; Prakash, O.; Kaliannan, P. Comprehensive review on detection and classification of power quality disturbances in utility grid with renewable energy penetration. *IEEE Power Energy Soc.* **2020**, *8*, 146807–146830. [[CrossRef](#)]
8. Mahela, O.P.; Shaik, A.; Gupta, N. A critical review of detection and classification of power quality events. *Renew. Sustain. Energy Rev.* **2015**, *41*, 495–505. [[CrossRef](#)]
9. Garcia, C.I.; Grasso, F.; Luchetta, A.; Piccirilli, M.C.; Paolucci, L.; Talluri, G. A Comparison of Power Quality Disturbance Detection and Classification Methods Using CNN, LSTM and CNN-LSTM. *Appl. Sci.* **2020**, *10*, 6755. [[CrossRef](#)]
10. Elvira-Ortiz, D.A.; Saucedo-Dorantes, J.J.; Osornio-Rios, R.A.; Morinigo-Sotelo, D.; Antonino-Daviu, J.A. Power Quality Monitoring Strategy Based on an Optimized Multi-Domain Feature Selection for the Detection and Classification of Disturbances in Wind Generators. *Electronics* **2022**, *11*, 287. [[CrossRef](#)]
11. 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](#)]
12. Guerrero-Sánchez, A.E.; Rivas-Araiza, E.A.; Garduño-Aparicio, M.; Tovar-Arriaga, S.; Rodríguez-Resendiz, J.; Toledano-Ayala, M. Novel Methodology for Classifying Electrical Disturbances Using Deep Neural Networks. *Technologies* **2023**, *11*, 82. [[CrossRef](#)]
13. Samanta, I.S.; Panda, S.; Rout, P.K.; Bajaj, M.; Piecha, M.; Blazek, V.; Prokop, L. A Comprehensive Review of Deep-Learning Applications to Power Quality Analysis. *Energies* **2023**, *16*, 4406. [[CrossRef](#)]
14. 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](#)]
15. Achlerkar, D.; Samantaray, S.R.; Manikandan, M.S. Variational mode decomposition and decision tree based detection and classification of power quality disturbances in grid-connected generation system. *IEEE Trans. Smart Grid* **2016**, *3*, 3122–3132. [[CrossRef](#)]
16. Abdelsalam, A.; Abdelaziz, Y.; Kamh, M. A generalized approach for power quality disturbances recognition based on Kalman filter. *IEEE Power Energy Soc.* **2021**, *9*, 93614–93628. [[CrossRef](#)]
17. Cortes-Robles, O.; Barocio, E.; Obushevs, A.; Segundo, R. Fast-training feedforward neural network for multi-scale power quality monitoring in power systems with distributed generation sources. *Measurement* **2021**, *170*, 108690. [[CrossRef](#)]
18. Rai, P.; Londhe, N.D.; Raj, R. Fault classification in power system distribution network integrated with distributed generators using CNN. *Electr. Power Syst. Res.* **2020**, *192*, 106914. [[CrossRef](#)]
19. Abasi, M.; Saffarian, A.; Joorabian, M.; Seifossadat, S. Fault classification and fault area detection in GUPFC-compensated double-circuit transmission lines based on the analysis of active and reactive powers measured by PMUs. *Measurement* **2021**, *169*, 108499. [[CrossRef](#)]
20. Yu, J.; Hou, Y.; Lam, A.; Li, V. Intelligent fault detection scheme for microgrids with wavelet-based deep neural networks. *IEEE Trans. Smart Grid* **2019**, *10*, 1694–1703. [[CrossRef](#)]
21. Wang, G.; Gao, F.; Liu, J.; Li, Q.; Zhao, Y. Design consideration and performance analysis of a hybrid islanding detection method combining voltage unbalance/total harmonic distortion and bilateral reactive power variation. *Cpss Trans. Power Electron. Appl.* **2020**, *5*, 86–100. [[CrossRef](#)]
22. Pouryekt, A.; Ramachandaramurthy, V.; Mithulananthan, N.; Arulampalam, A. Islanding detection and enhancement of microgrid performance. *IEEE Syst. J.* **2018**, *12*, 3131–3141. [[CrossRef](#)]
23. Radhakrishnan, P.; Ramaiyan, K.; Vinayagam, A.; Veerasamy, V. A stacking ensemble classification model for detection and classification of power quality disturbances in PV integrated power network. *Measurement* **2021**, *175*, 109025. [[CrossRef](#)]
24. Chamchuen, S.; Siritaratiwat, A.; Fuangfoo, P.; Suthisopapan, P.; Khunkitti, P. Adaptive Salp Swarm Algorithm as Optimal Feature Selection for Power Quality Disturbance Classification. *Appl. Sci.* **2021**, *11*, 5670. [[CrossRef](#)]
25. Cuculić, A.; Draščić, L.; Panić, I.; Čelić, J. Classification of Electrical Power Disturbances on Hybrid-Electric Ferries Using Wavelet Transform and Neural Network. *J. Mar. Sci. Eng.* **2022**, *10*, 1190. [[CrossRef](#)]
26. Beniwal, R.K.; Saini, M.K.; Nayyar, A.; Qureshi, B.; Aggarwal, A. A Critical Analysis of Methodologies for Detection and Classification of Power Quality Events in Smart Grid. *IEEE Access* **2021**, *9*, 83507–83534. [[CrossRef](#)]

27. Kaushik, R.; Mahela, O.P.; Bhatt, P.K.; Khan, B.; Garg, A.R.; Alhelou, H.H.; Siano, P. Recognition of Islanding and Operational Events in Power System With Renewable Energy Penetration Using a Stockwell Transform-Based Method. *IEEE Syst. J.* **2022**, *16*, 166–175. [[CrossRef](#)]
28. Mahela, O.P.; Khan, B.; Alhelou, H.H.; Siano, P. Power Quality Assessment and Event Detection in Distribution Network With Wind Energy Penetration Using Stockwell Transform and Fuzzy Clustering. *IEEE Trans. Ind. Inform.* **2020**, *16*, 6922–6932. [[CrossRef](#)]
29. Khoa, N.M.; 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](#)]
30. Prakash, O.; Sharma, Y.; Ali, S.; Khan, B.; Padmanaban, S. Estimation of islanding events in utility distribution grid with renewable energy using current variations and stockwell transform. *IEEE Access* **2021**, *9*, 69798–69813.
31. Eristi, H.; Demir, Y. A new algorithm for automatic classification of power quality events based on wavelet transform and SVM. *Expert Syst. Appl.* **2010**, *37*, 4094–4102 [[CrossRef](#)]
32. Hassini, K.; Fakhfakh, A.; Derbel, F. Optimal Placement of μ PMUs in Distribution Networks with Adaptive Topology Changes. *Energies* **2023**, *16*, 7047. [[CrossRef](#)]
33. Wang, S.; Chen, H. A novel deep learning method for the classification of power quality disturbances using deep convolutional neural networks. *Appl. Energy* **2019**, *235*, 1126–1140. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.