

Article

Interpretable DWT-1DCNN-LSTM Network for Power Quality Disturbance Classification

Shuangquan Yang, Tao Shan and Xiaomei Yang * 

College of Electrical Engineering, Sichuan University, Chengdu 610065, China;
yangshuangquan@stu.scu.edu.cn (S.Y.); shan_tao@stu.scu.edu.cn (T.S.)

* Correspondence: yangxiaomei@scu.edu.cn

Abstract: The proportion of new energy sources, such as wind, photovoltaic and hydropower, in the power grid is increasing year by year. In addition, a large number of nonlinear loads are connected to the grid, resulting in frequent power quality disturbances (PQDs), which pose challenges to the stability and reliability of the power system. Accurate identification of these disturbances is crucial for effective grid management and protection. Although deep learning methods have high accuracy, their lack of interpretability can limit their acceptance in engineering applications. Traditional signal analysis has a good physical foundation, but it is not integrated with deep learning to a sufficient degree. To address these issues, we propose the DWT-1DCNN-LSTM network as an interpretable model for PQD classification. This method effectively decomposes the time-domain signals into sub-signals in different frequency bands by employing the Discrete Wavelet Transform (DWT), which enhances the anti-interference capability of the classification model. This approach enhances the interference resilience of the classification model through the incorporation of the Discrete Wavelet Transform (DWT), which effectively decomposes time-domain signals into sub-signals across different frequency bands. The one-dimensional Convolutional Neural Network (1DCNN) then extracts local features, while the Long Short-Term Memory network (LSTM) analyzes temporal dependencies of the transformed sub-signals. Experimental validation with simulated datasets demonstrates that the DWT-1DCNN-LSTM model achieves an accuracy of 99.27%, outperforming the DWT-1DCNN, 1DCNN-LSTM, LSTM, and CNN models by 1.59%, 1.13%, 1.44%, and 6.48%, respectively. The robustness provided by the DWT module makes the model well suited for PQDs in environments with large disturbances, helping to detect and mitigate PQDs in a timely manner and ultimately contributing to improved power quality and system reliability.

Keywords: power quality disturbances; interpretable; discrete wavelet transform; one-dimensional convolutional neural network; long short-term memory network



Academic Editor: Frede Blaabjerg

Received: 2 December 2024

Revised: 5 January 2025

Accepted: 6 January 2025

Published: 7 January 2025

Citation: Yang, S.; Shan, T.; Yang, X. Interpretable DWT-1DCNN-LSTM Network for Power Quality Disturbance Classification. *Energies* **2025**, *18*, 231. <https://doi.org/10.3390/en18020231>

Copyright: © 2025 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

Within the framework of the “dual carbon” goals, constructing a new power system is essential for promoting the green and low-carbon transformation of energy in China and for establishing a sustainable energy system. It also represents a necessary choice for implementing a new energy security strategy and ensuring national energy security. As the installed capacity of renewable energy sources, including wind, solar, and hydropower, steadily increases, the integration of nonlinear loads into the power grid is also on the rise. This trend has resulted in frequent occurrences of PQDs within the power system [1,2]. On the one hand, PQDs can harm precision instruments, thereby decreasing product quality

and delaying industrial production schedules. On the other hand, frequent PQDs can cause overload or overheating of grid equipment, resulting in a higher rate of equipment failure and increased maintenance costs. Therefore, the identification and classification of PQDs are of significant importance for the safe and stable operation of the power grid and the advancement of social production [3].

The classification task of PQDs is typically divided into two stages: feature extraction and classification. Numerous methods exist for feature extraction from PQDs, with common techniques including Fourier Transform (FT) [4], Short-time Fourier Transform (STFT) [5,6], Stockwell Transform (SF) [7–9], and Empirical Mode Decomposition (EMD) [10]. Although FT provides spectral information for the entire time domain, it is primarily suited for analyzing stationary signals and fails to capture local information. STFT segments the signal using a window function, enabling the acquisition of frequency domain information over specific time intervals. However, the fixed-size window limits the granularity of spectral analysis. While the S-transform alleviates the limitation of fixed window widths, its frequency resolution varies across different frequency bands, making it difficult to effectively balance time-frequency resolution throughout the entire disturbance frequency range and complicating the accurate analysis of disturbances across multiple frequency bands. Conversely, EMD is sensitive to data length and noise, suffers from modal overlap and distortion, and the computational process of EMD is complicated.

DWT [11–13] utilizes a multi-resolution algorithm to decompose the signal into sub-bands of different frequency bands, which has more advantages in the processing of non-smooth signals. DWT is able to capture the local transient features of the signal in depth with adaptively selected wavelet bases, overcoming the resolution limitation of STFT due to fixed window width by adaptively selecting wavelet bases to deeply capture local transient features of the signal. Compared with STFT, DWT maintains consistent time-frequency resolution over the entire frequency band, ensuring more accurate analysis of various perturbations. Compared with EMD, DWT is less sensitive to data length and noise, and its computational process is simpler and more efficient than EMD.

Traditional classification methods for PQDs include Support Vector Machines (SVMs) [14,15] and decision trees (DTs) [16–18]. SVMs require high-quality feature extraction and often struggle with high-dimensional data, whereas DTs tend to have poor generalization ability and are prone to overfitting. Furthermore, traditional machine learning methods demonstrate limited information capture capabilities, making precise classification of PQDs under composite disturbances challenging.

In the study of PQD classification methods, the traditional classification methods are oriented towards obtaining key features by improving the feature extraction algorithm and then using the extracted features to classify the PQDs. Mohd S. N. B. Marzuki et al. [19] transformed the signal using ST and then extracted six standard deviation features from the ST matrix in different ways using a modified algorithm and then fed the features into SVM for classification. The innovation of this method lies in the algorithm of feature extraction, but by extracting the features artificially, it is prone to feature loss which affects the results of classification. Xiaomei et al. [3] use the improved DWT transform tunable Q-factor wavelet transform (TQWT) to decompose the PQDs signal into different sub-bands, and then the features are constructed for the four optimal decompositions, which ultimately improves the signal's immunity to interference and recognition accuracy. The feature extraction part of the traditional classification methods in the research relies heavily on the experience of experts, while the key information is easily lost in the feature processing stage due to improper processing, which leads to less accurate classification results.

Fueled by the rapid advancement of computer processing technologies, deep learning (DL) has evolved from a nascent concept to a mature technology, establishing itself as

a mainstream approach in various fields. In this context, scholars in the field of power quality research have begun to use DL methods for fault detection and classification of time-series signals [20,21]. Zamrooth et al. [22] used the Red Deer Optimization optimization algorithm to select the optimal number of hidden layers of the deep recurrent neural network (DRNN) to improve the DRNN network's classification accuracy of PQDs and also reduce the computation time of the network. But for the extracted features, we cannot understand. Zhanbei et al. [23] used LSTM and a self-designed lightweight residual network to extract one-dimensional features of PQDs and spatial features of 2D image modalities, respectively, and finally used feature fusion before classification, which not only improved the classification accuracy of the network but also reduced the parameters of the model. Feature fusion is a good approach, but the image analysis intervention will increase the computation of the network; at the same time, we can not understand the specific workings of the network. Although the deep neural network used by the DL method has strong data mining ability and analysis ability by relying on a large number of network parameters and nonlinear features, the internal operation mechanism of the DL method is more complex and difficult to understand, and the extracted features can not be understood by human beings, so it is called a black box. This hinders the further development of DL methods [24]. Improving the interpretability of the network is the key to the continued development of DL methods, so a large number of scholars have conducted research on the interpretability of DL methods. Sun Yanling et al. [25] replaced the first layer of convolutional kernel in the network with Morelet wavelet and Laplace wavelet kernel, which made the network theoretically interpretable. Qian Chen et al. [26] embedded physically meaningful time-frequency transformation methods as trainable preprocessing layers into the traditional convolutional layer and proposed interpretable DL methods that can be used for fault diagnosis. Yanling Sun and Qian Chen were able to interpret the first layer of features extracted by the network in response, but their research as a pair mainly focused on the interpretability study of shallow features.

Based on the preceding analysis, enhancing the interpretability of DL methods is crucial for their further development. This paper introduces an explainable DWT-1DCNN-LSTM neural network that combines the physical interpretability of the DWT [27] with the data mining and analytical capabilities inherent in DL methods. Wavelet analysis is a powerful tool for signal denoising. Integrating DWT into the neural network not only addresses the network's shortcomings in signal analysis and reduces the number of model parameters but also enhances its ability to filter out noise and irrelevant information, thereby improving robustness. The short-term memory characteristic of the Long Short-Term Memory network [28] effectively extracts temporal characteristics from signals. The network deeply integrates DWT and neural network and extends the learning space from the time domain to the more useful wavelet domain so as to break through the bottleneck of the interpretability of the DL method in the classification task. The DWT-CNN-LSTM takes the discrete temporal signals as inputs and firstly alternates between the signal decomposition based on the DWT and the local feature learning based on the CNN, then uses the single-layer LSTM neural network to learn the time domain features of decomposed feature signals, and then uses the final feature information for average pooling before full connectivity to obtain the fault categories' feature information; then, a single-layer LSTM neural network is used to learn the time-domain features of the decomposed feature signals, and the final feature information is average pooled and then fully connected to obtain the fault category.

Compared with traditional methods, our proposed method integrates feature extraction and classification into the network using DL, which can accurately classify PQDs while retaining the maximum information; compared with existing deep learning methods, our

method can make a reasonable interpretation of the extracted features, which can help people to understand the features, and at the same time, it also provides a new direction for the study of DL interpretability.

The main contributions of this paper are as follows.

1. We deeply integrate the basic theory of signal processing and deep learning methods.
2. The addition of LSTM and DWT to CNN enables the network to capture the features of the signal from the time and wavelet domains, respectively, thus increasing the diversity of extracted features.
3. When researching the interpretability of DL methods, we go beyond the local concerns and focus on the whole network.
4. By adding DWT, the DWT-CNN-LSTM is able to perform classification tasks even in the presence of large noise.

This paper is organized as follows: Section 2 provides a detailed description of the DWT-1DCNN-LSTM network architecture, including the functionalities and interactions of each module. Section 3 outlines the dataset selection. Section 4 presents the experimental results, analyzing the model's performance under various conditions and comparing it to existing methods. Finally, Section 5 summarizes the main findings of this research and suggests potential future directions and improvements.

2. Theory and Methodology

2.1. DWT-1DCNN-LSTM Network Architecture

To enhance the interpretability of DL methods, improve the network's robustness against interference, and simultaneously increase the accuracy of classification tasks for PQDs, this paper proposes an explainable DWT-1DCNN-LSTM neural network. The architecture of the DWT-1DCNN-LSTM network is illustrated in Figure 1. This network comprises three main components: the DWT module, a 1DCNN feature extraction module, and the LSTM module.

Initially, the network decomposes the signal into low-frequency and high-frequency components using the DWT. Subsequently, convolutional kernels are utilized to extract local information from both low-frequency and high-frequency signals, yielding a richer set of feature information. By repeatedly applying this operation, the network can extract even more comprehensive feature information. Finally, the LSTM captures the temporal characteristics of the feature signals, and a global average pooling operation is applied to reduce the dimensionality of the outputs from the LSTM. The classification of PQDs is then conducted through the fully connected layer and the Softmax layer.

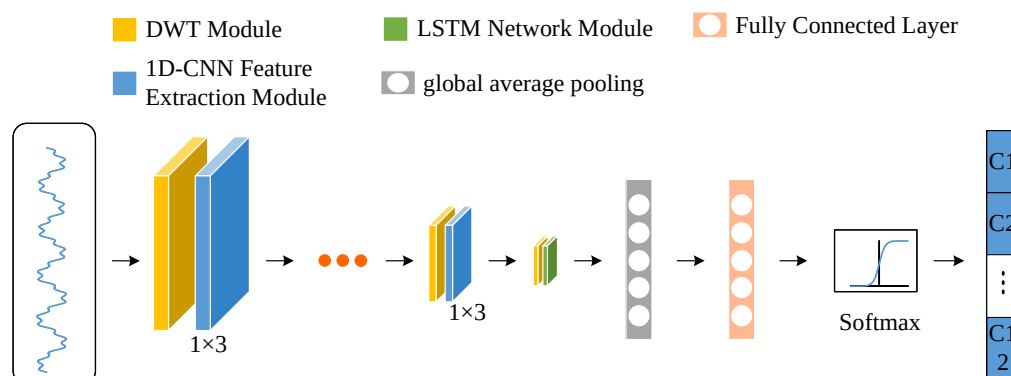


Figure 1. The network architecture of DWT-1DCNN-LSTM.

2.2. DWT Module

Wavelet Transform (WT) is commonly used for denoising and is well suited for non-stationary signals, as it helps to extract various features closely related to oscillation type recognition. The continuous WT of the signal $x(t)$ is given by:

$$W(v, \tau) = \frac{1}{\sqrt{v}} \int_{-\infty}^{+\infty} x(t) \hat{\psi}\left(\frac{t - \tau}{v}\right) dt \quad (1)$$

where $\hat{\psi}$ is the wavelet function, and τ and v are the translation factor and the scaling factor, respectively.

To apply WT to the discrete signal $x(n)$, it is necessary to discretize the WT to obtain the DWT. The DWT transformation is performed using the multiresolution algorithm proposed by Mallat [29], with the specific decomposition process illustrated in Figure 2. Essentially, this involves decomposing the signal using a low-pass filter $h(n)$ and a high-pass filter $g(n)$. After obtaining the outputs from each filter, downsampling by a factor of 2 is applied to remove redundant information. The computational steps are described in Equations (2) and (3):

$$x_{\text{dwt},h}[n] = \sum_{k=0}^{K-1} x[2n - k]h[k] \quad (2)$$

$$x_{\text{dwt},g}[n] = \sum_{k=0}^{K-1} x[2n - k]g[k] \quad (3)$$

where K is the length of the filter, $h[n]$ and $g[n]$ are the parameters of the low-pass and high-frequency filters, respectively, and $x_{\text{dwt},h}[n]$ and $x_{\text{dwt},g}[n]$ are the outputs obtained by down-sampling the outputs of the low-frequency and high-frequency filters with a downsampling factor of 2.

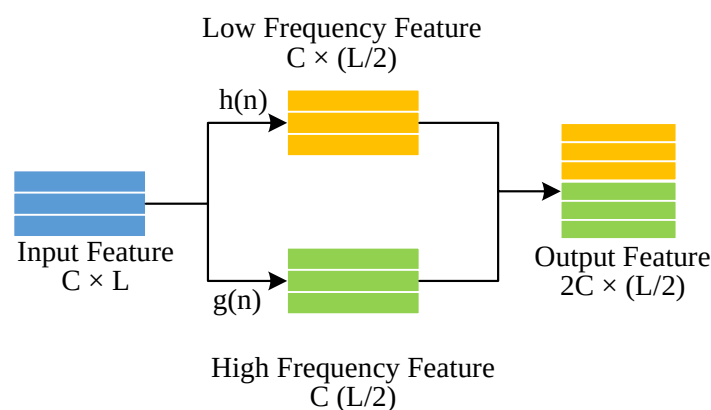


Figure 2. DWT module.

2.3. The 1DCNN Feature Extraction Module

The advantage of 1D-CNN, a class of feed-forward neural networks with convolutional computation and a deep structure, lies in the local connection and weight sharing, which not only reduces the number of weights to make the network easy to optimize but also reduces the complexity of the network. The 1D-CNN feature extraction module constructed in this paper is shown in Figure 3.

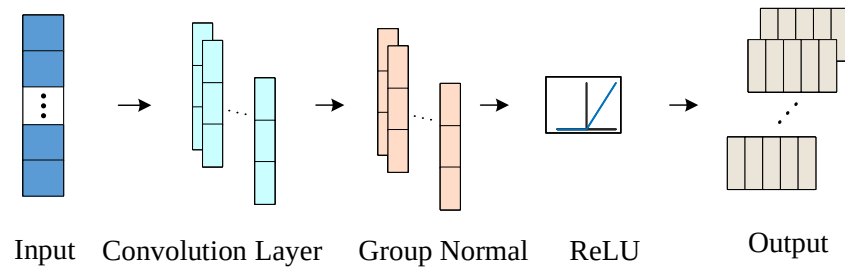


Figure 3. The 1DCNN feature extraction module.

Each module contains a convolutional layer, a group normalization (GN) layer [30], and a ReLU function activation layer. The convolutional layer is used to capture the local features of the signal, and the GN divides the channel into groups and normalizes the data by calculating the mean and variance within the group; the GN is computed independently of the batch size, and its accuracy is stable over a larger range of batches. Nonlinear transformations are introduced through the ReLU function. The specific computational steps for the convolutional layer are as follows.

The low-frequency mapping and high-frequency mapping groups x_{dwt} of the DWT module are used as inputs to the convolutional layer, and the specific computational procedure is as follows:

Firstly, the convolution operation is performed on the input signal

$$z_i^l = w_i^l * x_{\text{dwt}}^l + b_i^l \quad (4)$$

where x_{dwt} represents the input of the l th convolutional layer, and w_i^l and b_i^l represent the weight and offset of the i th convolutional kernel of the l th layer, respectively. $*$ the convolution operation.

Next, the GN layer normalizes the output of the convolution kernel to improve the training speed as well as to mitigate the gradient explosion, and the output value of the i th channel of the l -th layer is as follows.

$$\bar{z}_i^l = \frac{z_i^l - \mu_b^l}{\sqrt{(\delta_b^l)^2 + \varepsilon}} \quad (5)$$

where μ_b^l and δ_b^l denote the group normalized mean and variance of the channel corresponding to input z_i^l , respectively, and ε denotes a small constant preventing the denominator from being zero.

Finally, a nonlinear mapping relation is introduced through the activation layer:

$$y_i^l = \text{ReLU}(\bar{z}_i^l) \quad (6)$$

2.4. LSTM Network Module

After the DWT and the convolution operation of the 1DCNN, the number of features of the input signal increases while the effective length of the signal decreases. The feature information encapsulates both the frequency domain characteristics and the local properties of the input signal. Given that the sampled PQD signals possess temporal characteristics, the time-domain features also represent significant factors influencing classification outcomes. Therefore, this network employs LSTM networks to learn the temporal features between the input and output. LSTM is a specialized type of recurrent neural network, with its structure depicted in Figure 4.

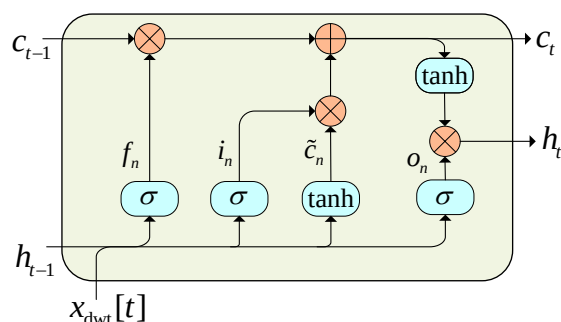


Figure 4. LSTM network module.

To address the limitations of traditional recurrent neural networks (RNNs) in terms of long-term memory retention, the LSTM architecture incorporates three critical gating units within the hidden layers: the forget gate, the input gate, and the output gate. These gates, respectively, determine which information to forget from the memory at time step $t - 1$, manage the input of the current signal and the output from the previous time step at time t , and regulate the output of the memory at time t . The utilization of these three gates effectively mitigates issues related to gradient instability, long-term dependencies, and memory constraints that are prevalent in standard RNNs. The computational process of the LSTM network is as follows:

Step 1: Utilize the low-frequency mapping and high-frequency mapping combination from the DWT module, along with the output state at time $t - 1$, to compute the forget gate f_t , input gate i_t , output gate o_t , and candidate memory $\tilde{c}_t$. The calculation steps are illustrated in Equations (7)–(10).

$$f_t = \sigma(W_f \bullet [h_{t-1}, x_{\text{dwt}}[t]] + b_f) \quad (7)$$

$$i_t = \sigma(W_i \bullet [h_{t-1}, x_{\text{dwt}}[t]] + b_i) \quad (8)$$

$$o_t = \sigma(W_o \bullet [h_{t-1}, x_{\text{dwt}}[t]] + b_o) \quad (9)$$

$$\tilde{c}_t = \tanh(W_c \bullet [h_{t-1}, x_{\text{dwt}}[t]] + b_c) \quad (10)$$

where $\sigma(\cdot)$ is the sigmoid function, $\tanh(\cdot)$ is the hyperbolic tangent function, and W and b are the parameters of the LSTM cell; $\bullet$ represents the dot product of vectors

Step 2: The forget gate utilizes selective forgetting on the memory at time $t - 1$, while the input gate selectively inputs the output of the candidate memory at time t . The combined result yields the memory at time t . The calculation steps are shown in Equation (11).

$$c_t = f_t \times c_{t-1} + i_t \times \tilde{c}_t \quad (11)$$

Step 3: The output gate is utilized to selectively output the state memory at time t , resulting in h_t . The calculation process is illustrated in Equation (12).

$$h_t = o_t \times \tanh(c_t) \quad (12)$$

To reduce the spatial dimensions of the data, a global average pooling operation is applied to the LSTM layer, followed by a fully connected layer and a softmax output. The label with the highest probability is then selected as the classification result.

2.5. Detailed Parameter Settings of DWT-1DCNN-LSTM Neural Network

For the parameters of the convolutional layer, we refer to reference [31]; firstly, the number of layers of the convolutional kernel and the number of layers of the CNN in the initial model are set to 3 and 6, and then according to the principle of a single variable in the experiment, combined with the correctness of the classification results, the number of parameters, and the training time, the parameters of the convolutional layer are determined, and the output of the LSTM is set to the same output to ensure that the number of features is unchanged. The overall architectural parameters of the model are finally shown in Table 1.

Table 1. Parameters of the DWM-1DCNN-LSTM.

Layer	Type	Kernel	Channel	Output
1	Input	—	—	5000×1
2	DWA	—	—	2500×2
3	1DCNN	1×3	24	2498×24
4	DWA	—	—	1249×48
5	1DCNN	1×3	48	1247×48
6	DWA	—	—	624×96
7	1DCNN	1×3	96	622×96
8	DWA	—	—	311×192
9	1DCNN	1×3	192	309×192
10	DWA	—	—	155×384
11	1DCNN	1×3	384	153×384
12	DWA	—	—	77×768
13	LSTM	—	—	77×768
	Global			
14	Average Pooling	—	—	768
15	FC + Softmax	—	12	12

“—” indicates “parameter does not exist”.

2.6. Presentation of Other Hyperparameters

In the experimental process, the setting of hyperparameters has a great impact on the experimental results, so we set reasonable hyperparameters. The hyperparameters set in this experiment include learning rate, training batch size, group normalization size, Dropout and training period also include optimizer settings and loss function settings. Setting the learning rate too small will lead to slow convergence, and setting the learning rate too large will lead to dispersion of the training process; setting the training batch size too large will lead to overfitting of the model and also increase the memory consumption of the device, resulting in the experiment not being able to be carried out correctly, and setting the training batch size too large will lead to prolongation of the training time. If the group normalization is set to 1, it will be equivalent to the batch normalization process, and if the setting is set too small, it will be equivalent to layer normalization, and both too large and too small will lose the group normalization. Dropout is generally set between 0 and 1. Setting it too large will lead to underfitting of the model while setting it too small will lead to overfitting of the model; if the training period is too large, it will be a waste of time, and if it is too small, the training process will not be able to converge.

In the process of the experiment, we take the test accuracy of the model as a reference standard, adjust the parameters, and finally determine the parameters used in this experiment. The learning rate is set to 0.0001; the training Batch is set to 200; the group normalization size is set to 6; a total of five layers of dropout is set; the first three layers of the dropout are set to 0.1; and the next two layers are set to 0.2 The training period is set to 100; the Adam optimization algorithm updates the parameters of the network model,

which usually handles non-smooth objectives well, converges faster, and is commonly used as the default optimizer. Therefore, we used the Adam optimization algorithm, and currently for classification tasks, the cross-entropy loss function is generally used, so the cross-entropy loss function is also used in this experiment.

3. Data Simulation

PQDs are modeled in accordance with the IEEE 1159 standard [32], and simulations are conducted using MATLAB software (R2020b). The mathematical models of the six types of single PQDs are presented in Table 2. Composite PQDs are formed from the single PQDs. Ultimately, a total of 12 classes of signals are generated: 6 classes of single PQDs and 6 classes of composite PQDs.

Table 2. Mathematical network model for single-signal simulation.

PQDs	Mathematical Model	Parameter	Label
normal	$x(t) = A \sin(\omega t + \theta) * (u(t - t_1) - u(t - t_2))$	$0.9 \leq A \leq 1.1; 0 \leq \theta \leq 2\pi; 0 \leq t_2 - t_1 \leq 0.2;$	C1
flicker	$x(t) = A \sin(\omega t + \theta) \left(1 + \partial_f * \sin(\beta \omega t)\right)$	$0.1 \leq \partial_f \leq 0.2; 0.1 \leq \beta \leq 0.4$	C2
harmonic	$x(t) = A[\partial_1 \sin(\omega t) + \partial_3 \sin(3\omega t) + \partial_5 \sin(5\omega t) + \partial_7 \sin(7\omega t)]$	$0.05 \leq \partial_3, \partial_5, \partial_7 \leq 0.15; \sum \partial_i^2 = 1$	C3
interrupt	$x(t) = A \sin(\omega t + \theta)(1 - \partial_{INT}(u(t - t_1) - u(t - t_2)))$	$0.9 \leq \partial_{INT} \leq 1; T \leq t_2 - t_1 \leq 9T;$	C4
sag	$x(t) = A \sin(\omega t)(1 - \beta(\mu(t - t_1) - \mu(t - t_2)))$	$0.1 \leq \beta \leq 0.9; T \leq t_2 - t_1 \leq 9T;$	C5
swell	$x(t) = A \sin(\omega t)(1 + \beta(\mu(t - t_1) - \mu(t - t_2)))$	$0.1 \leq \beta \leq 0.8; T \leq t_2 - t_1 \leq 9T$	C6

Composite PQDs are created by the superposition of single PQDs. Specifically, the combinations are as follows: Flicker with Harmonic Disturbance consists of C2 + C3; Flicker with Sag Composite Disturbance consists of C2 + C5; Flicker with Harmonic Composite Disturbance consists of C2 + C6; Interruption with Harmonic Composite Disturbance consists of C3 + C4; Sag with Harmonic Composite Disturbance consists of C3 + C5; and Swell with Harmonic Composite Disturbance consists of C3 + C6.

The data simulation and acquisition are conducted with a sampling frequency of 25 kHz and a number of sampling points of 5000. The result of sampling is shown in Figure 5. The ratio of the training set and test set is 3:1. The specific quantities are shown in Table 3.

Table 3. Number of datasets.

Label	Training Set	Test Set	Label	Training Set	Test Set
C1	10,000	3333	C7(C2 + C3)	10,000	3333
C2	10,000	3333	C8(C2 + C5)	10,000	3333
C3	10,000	3333	C9(C2 + C6)	10,000	3333
C4	10,000	3333	C10 (C3 + C4)	10,000	3333
C5	10,000	3333	C11(C3 + C5)	10,000	3333
C6	10,000	3333	C12(C3 + C6)	10,000	3333

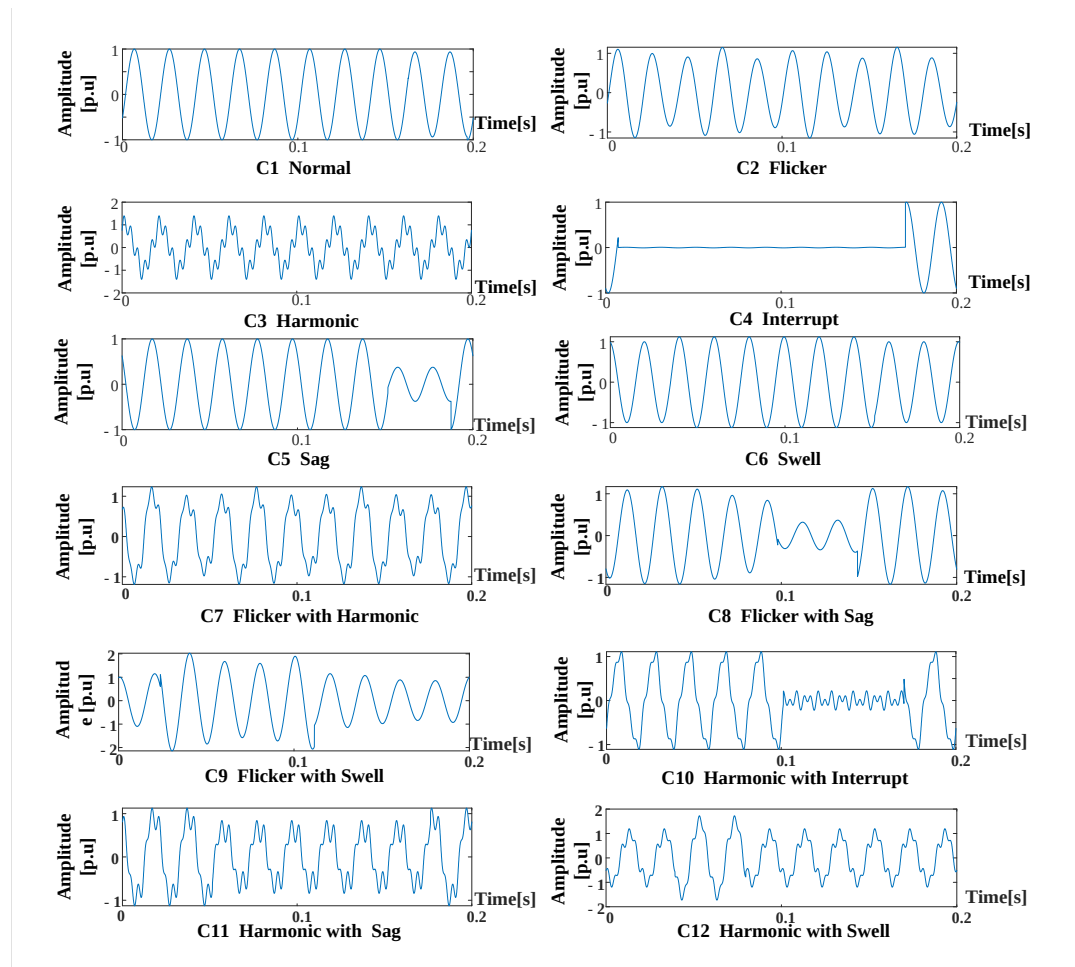


Figure 5. Schematic diagram of disturbance signal simulation.

4. Experiment and Discussion

4.1. Evaluation Metrics

In our experiment, precision (*Pre*), recall (*Rec*), F1-score (*F1*), and accuracy (*Acc*) are used as evaluation metrics for the network, defined as follows:

$$Pre = \frac{TP}{TP + FP} \quad (13)$$

$$Rec = \frac{TP}{TP + FN} \quad (14)$$

$$F1 = \frac{2 * Pre * Rec}{Pre + Rec} \quad (15)$$

$$Acc = \frac{TP + TN}{TP + TN + FN + FP} \quad (16)$$

where *TP*, *FP*, *FN*, and *TN* true positives, false positives, false negatives, and true negatives, respectively.

$$SSR = \sum_{i=1}^k \sum_{j=1}^{n_i} (\bar{a}_i - \bar{a})^2 \quad (17)$$

$$SSE = \sum_{i=1}^k \sum_{j=1}^{n_i} (a_{ij} - \bar{a})^2 \quad (18)$$

$$F = \frac{SSR/(k-1)}{SSE/(n-k)} \quad (19)$$

where SSR represents the sum of squared deviations between groups of the dependent variable, k is the number of levels of the independent variable; $\bar{a}_i$ is the sample mean of the dependent variable at the i -th level of the independent variable; n_i is the sample size at the i -th level of the independent variable; $\bar{a}$ is the mean value of the independent variable; SSE denotes the sum of squares of the deviations of each sample of data from the mean of the group in which it is embedded a_{ij} is the j -th sample value at the i -th level of the independent variable; $k-1$ is the degree of freedom of SSR ; $n-k$ is the degree of freedom of SSE .

Higher F indicates greater between-group differences relative to within-group differences, implying that the model is more sensitive to disturbances and less resistant to them.

4.2. Experimental Results

The network was trained using the simulated dataset, yielding training and testing accuracy results, as illustrated in Figure 6.

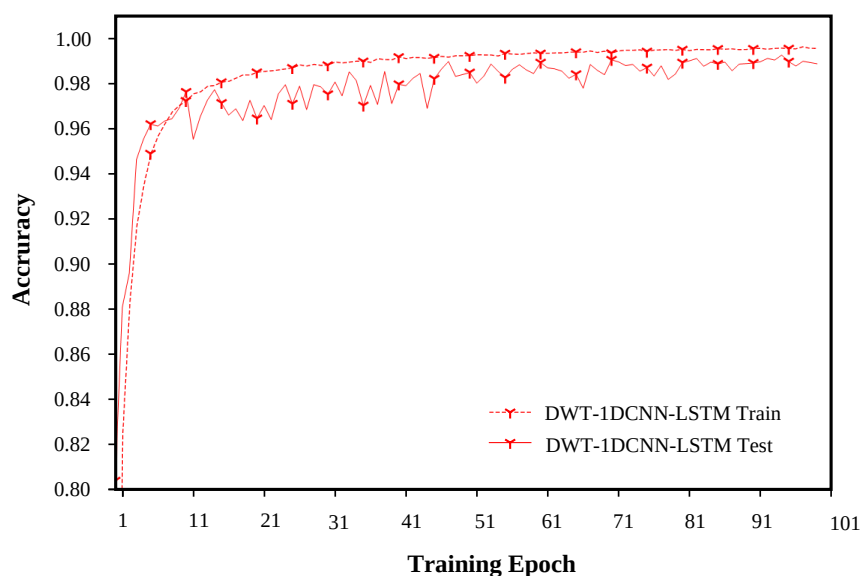


Figure 6. Training process of DWT-1DCNN-LSTM in a 40 dB environment.

Figure 6 demonstrates that the testing and training accuracies of the DWT-1DCNN-LSTM network quickly surpassed 90%, with a minimal difference between them, indicating a strong generalization capability of the network. The network achieved an accuracy of 99.27% at the 100th training iteration. Furthermore, the changes in testing accuracy throughout the training process were relatively smooth, ultimately converging toward stability.

By visualizing the data, the confusion matrix was obtained (see Figure 7). When the signal-to-noise ratio is 40 dB, the DWT-CNN-LSTM network has high recognition ability for all 12 types of PQDs, and the recognition rate even reaches 100% for C2, C3, C4 and C7. In addition, from Figure 7, it can be seen that the recognition accuracy of the DWT-CNN-LSTM network is higher than 98.89% and 97.89% for single and composite PQDs, respectively, which indicates that the model has a slightly weaker rate of recognition for composite PQDs. The reasons for recognition error are mainly as follows.

For single PQDs, the amplitude of normal signal C1 is from 0.9 to 1.1, while the amplitude of Sag C5 is from 0.1 to 0.9, the amplitude of Swell is from 1.1 to 1.8, and the amplitude of interrupt C4 is from 0 to 0.1. Therefore, after adding the noise, the amplitude of the signal can easily break through the corresponding amplitude thresholds, which will make the C1 be recognized as C5 or C6, and at the same time the amplitude upper limit of C5 is close to the amplitude lower limit of C1, and the amplitude lower limit of C5 is close to the upper limit of the amplitude of C4; then C6 will be easily recognized as C1 and C4. At the same time, the normalization process in the network will further highlight the relative change in the amplitude of C5 and C6 rather than the specific value; then, the features of C5 and C6 are easily confused in the recognition process.

For composite PQDs, the single PQDs that make up C8 and C9 are predominantly amplitude-varying signals, and after mixing, they are easily recognized as other signals with amplitude characteristics. However, when the composite PQDs contain harmonics, the frequency domain characteristics of the signal will affect the classification of the signal, while the amplitude characteristics are confused; therefore, when there are harmonics in the signals of C10, C11, and C12, the DWT will increase the attention to the harmonics while the attention to the aspect of amplitude change is appropriately reduced, so that the C10, C11, and C12 can be easily confused between them.

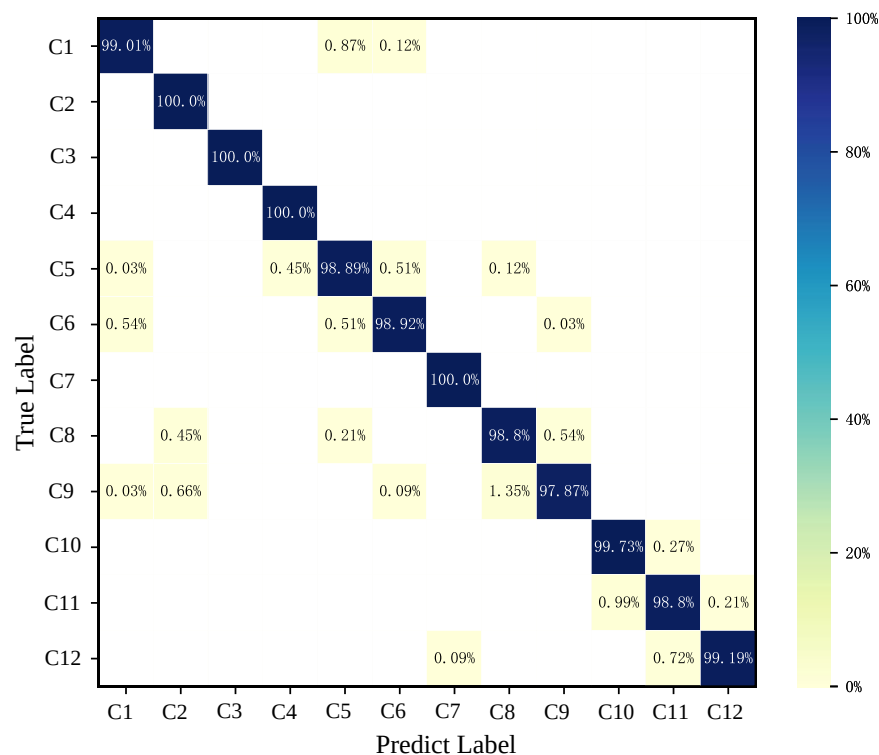


Figure 7. Confusion matrix of classification results in a 40 dB environment.

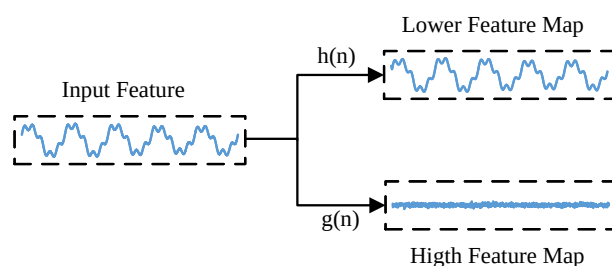
Utilizing the classification results, metrics such as *Pre*, *Rec*, and *F1* were calculated (see Table 4). The table shows that four out of the twelve categories achieved a precision of 100%. Furthermore, the Recall for the twelve categories remained around 99%, with *F1*-scores exceeding 98.37%. These results indicate that the network is capable of effectively addressing the challenges associated with the classification tasks of PQDs.

Table 4. *Pre*, *Rec*, and *F1* of the DWT-1DCNN-LSTM network in a 40 dB environment.

Label	Pre	Rec	F1
C1	0.9901	0.9940	0.9920
C2	1.0000	0.9890	0.9945
C3	1.0000	1.0000	1.0000
C4	1.0000	0.9955	0.9978
C5	0.9889	0.9842	0.9865
C6	0.9892	0.9928	0.9910
C7(C2 + C3)	1.0000	0.9991	0.9996
C8(C2 + C5)	0.9880	0.9853	0.9867
C9(C2 + C6)	0.9787	0.9942	0.9864
C10(C3 + C4)	0.9973	0.9902	0.9937
C11(C3 + C5)	0.9880	0.9901	0.9890
C12(C3 + C6)	0.9919	0.9979	0.9949

4.3. Interpretability Analysis

Although DL methods can achieve excellent classification results, many classification networks used in DL are often regarded as black boxes, which do not meet the reliability requirements in critical fields. To enhance the interpretability of the network, this study integrates the DWT deeply into the neural network, enabling the theory of frequency domain analysis to be seamlessly embedded within deep learning. DWT is currently widely employed in the time-frequency analysis of one-dimensional signals, allowing for the decomposition of signals into high-frequency and low-frequency components. As illustrated in Figure 8, the input signal is decomposed into two sub-signals after a single DWT transformation, effectively doubling the number of output features relative to the input signal. Subsequently, one-dimensional convolutional kernels from CNNs are applied to the signal through a sliding window operation. At each step, only the portion of the signal covered by the current convolutional kernel is considered, allowing the model to iteratively capture local information from the feature signals. LSTM networks, which utilize units such as forget gates, input gates, output gates, and memory cells, can retain temporal information from n time steps prior, thereby extracting richer time-domain features.

**Figure 8.** DWT signal exploded view.

The training processes for the 1DCNN-LSTM model (with the DWT module removed) and the DWT-1DCNN-LSTM model are illustrated in Figure 9. Obviously, after removing the DWT module, on the one hand, the classification accuracy of the model was reduced, and on the other hand, the model showed an obvious overfitting phenomenon, and these fully illustrate that the addition of the DWT module significantly improves the model's prediction accuracy and anti-jamming ability.

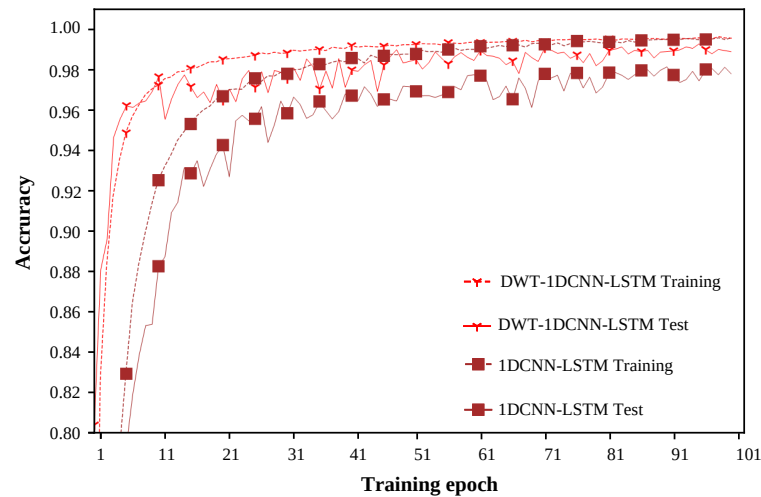


Figure 9. Training process of 1DCNN-LSTM model and DWT-1DCNN-LSTM model.

For a deeper understanding of the interpretability, we visualized the features captured in the signals (see Figure 10), and the fusion features were obtained by summing and normalizing the output signals of the last CNN layer of the CNN-LSTM and DWT-CNN-LSTM networks. As can be seen in Figure 10, both networks capture the change in signal amplitude, and the amplitude of the fused features is also relatively large where the PQD signal plummets. However, the DWT-CNN-LSTM also captures the frequency domain features of the signal compared to the CNN-LSTM network. This indicates that with the addition of DWT, the network focuses on both local and frequency domain features. This can help the network to extract more classification features.

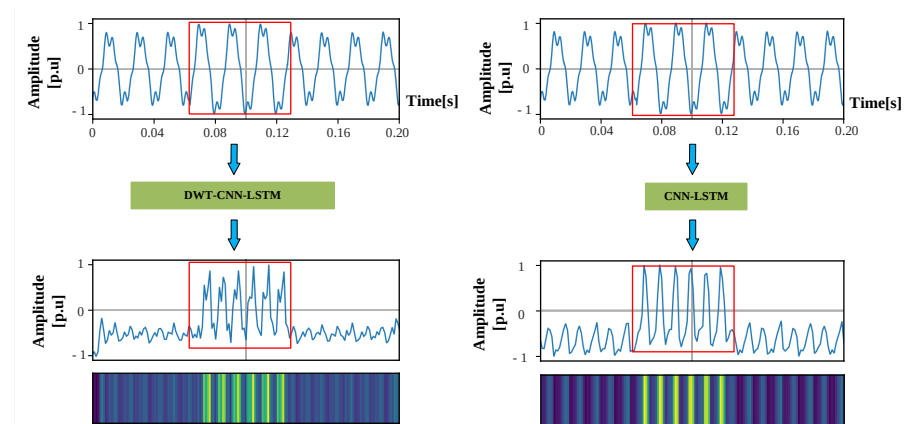


Figure 10. Feature visualization of DWT-CNN-LSTM and CNN-LSTM models.

4.4. Comparative Analysis

To validate the effectiveness of the DWT-1DCNN-LSTM network combination, a comparative analysis was conducted between the DWT-1DCNN-LSTM network, the DWT-1DCNN network (with the LSTM module removed), the LSTM network (with the DWT and 1DCNN modules removed), and the CNN network presented in reference [31]. The DWT-1DCNN-LSTM and DWT-1DCNN networks were subject to an ablation experiment, while the LSTM and 1DCNN networks underwent a comparative evaluation. Simulation data were used to train all four networks, resulting in the training processes illustrated in Figure 11 and various performance metrics, including training accuracy (see Table 5).

As illustrated in Figure 11, there is a significant gap between the training accuracy and testing accuracy of the DWT-1DCNN, 1DCNN-LSTM, and CNN models. This gap

indicates that overfitting occurs during the training of these models, suggesting that they excel at memorizing the training data but lack the ability to generalize to unseen data. In contrast, the DWT-1DCNN-LSTM model exhibits a smaller difference between training accuracy and testing accuracy, indicating that it achieves a better balance between learning and generalization during the training process. This observation suggests that the DWT-1DCNN-LSTM model effectively extracts features from the training data while maintaining good adaptability to the test data.

As shown in Table 5, the test Acc of the DWT-1DCNN-LSTM network is 99.27%, which is 6 percentage points higher than that of the CNN, approximately 1.5 percentage points higher than the DWT-1DCNN and LSTM, and 1.1 percentage points higher than the 1DCNN-LSTM. The Pre of the CNN network's prediction for composite disturbances is overall lower than that for single disturbances, indicating that the CNN network performs poorly in predicting composite disturbances.

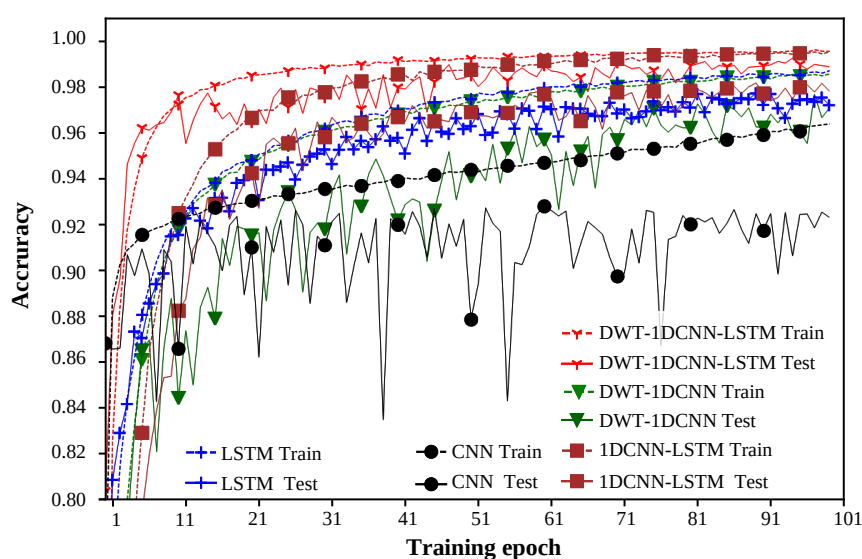


Figure 11. Training process of DWT-1DCNN-LSTM model, DWT-1DCNN model, 1DCNN-LSTM model, LSTM model and CNN model.

Table 5. Pre of different networks.

Label	40 dB				
	DWT-1DCNN-LSTM	DWT-1DCNN	1DCNN-LSTM	LSTM	CNN
C1	0.9901	0.8965	0.9688	0.9712	0.9982
C2	1.0000	0.9952	0.9949	1.0000	0.9976
C3	1.0000	0.9967	1.0000	1.0000	1.0000
C4	1.0000	0.9997	0.9988	0.9886	0.9967
C5	0.9889	0.9721	0.9772	0.9424	0.8281
C6	0.9892	0.9976	0.9877	0.9694	0.8668
C7(C2 + C3)	1.0000	0.9997	0.9967	0.9997	0.9964
C8(C2 + C5)	0.9880	0.9502	0.9547	0.9592	0.8020
C9(C2 + C6)	0.9787	0.9619	0.9547	0.9520	0.8650
C10(C3 + C4)	0.9973	0.9820	0.9820	0.9925	0.9847
C11(C3 + C5)	0.9880	0.9796	0.9787	0.9643	0.8809
C12(C3 + C6)	0.9919	0.9910	0.9832	0.9958	0.9184
Acc	0.9927	0.9768	0.9814	0.9773	0.9279

In contrast, the DWT-1DCNN-LSTM network shows little difference in precision between composite and single disturbances, with both exhibiting high precision, suggesting that the DWT-1DCNN-LSTM network has superior classification ability. The results of the ablation experiments indicate that the combination of DWT, 1DCNN, and LSTM is the most effective, demonstrating significant improvement. Furthermore, the results of the comparative analysis demonstrate that DWT-1DCNN-LSTM achieves better classification performance on this simulation dataset.

4.5. Robustness Analysis

In order to verify the robustness of the model and the effect of DWT addition on the robustness of the network, we tested DWT-CNN-LSTM with CNN-LSTM and CNN three times in 50 dB, 40 dB, and 30 dB environments, respectively, and the results are shown in Table 6. From the table, it can be seen that the overall prediction accuracies of DWT-CNN-LSTM are all relatively high. Taking the average test accuracy as an indicator, when the signal-to-noise ratio is increased from 50 dB to 30 dB, the average accuracy of DWT-CNN-LSTM, CNN-LSTM, and CNN decreases by 0.0035, 0.0121, and 0.0071, respectively. The average accuracy of CNN-LSTM and CNN decreased 4 times and 2 times more than that of DWT-CNN-LSTM, respectively, which indicates that the model's prediction results are more affected under different noise.

Table 6. Acc of DWT-1DCNN-LSTM in different environment.

Model	Noise		
	50 dB	40 dB	30 dB
DWT-CNN-LSTM	0.9934	0.9927	0.9901
	0.9930	0.9926	0.9896
	0.9928	0.9920	0.9890
average Acc	0.9931	0.9924	0.9896
CNN-LSTM	0.9844	0.9768	0.9724
	0.9842	0.9763	0.9719
	0.9838	0.9754	0.9716
average Acc	0.9940	0.9927	0.9925
CNN	0.9292	0.9279	0.9217
	0.9287	0.9274	0.9210
	0.9282	0.9268	0.9201
average Acc	0.9287	0.9274	0.9210

Further analysis of the data in Table 6 through variance analysis yielded the F-values presented in Table 7. The DWT-CNN-LSTM model exhibited the lowest F-value, which generally indicates that the model is insensitive to noise and possesses strong noise immunity. In contrast, the F-value for the CNN-LSTM model combined with CNN was higher than that of DWT-CNN-LSTM, suggesting that inter-group variability is greater than intra-group variability, highlighting the model's higher sensitivity to noise and relatively weaker noise immunity.

Table 7. ANOVA for DWT-CNN-LSTM, CNN-LSTM, CNN.

Label	DWT-CNN-LSTM	CNN-LSTM	CNN
F	57.9691	452.2412	129.7187

To further demonstrate that the addition of DWT improves the network's noise immunity, we compare the test results of the DWT-CNN-LSTM network with those of the CNN-LSTM network in 4 dB, 0 dB and −4 dB environments, as shown in Figure 12.

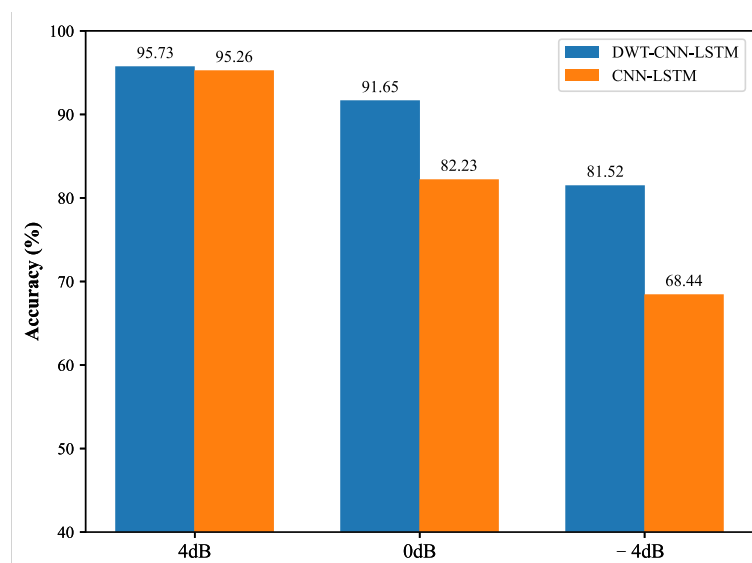


Figure 12. Test Acc of DWT-CNN-LSTM and CNN-LSTM under different noise environments.

From Figure 12, it can be seen that after removing the DWT, the DWT-CNN-LSTM network is 0.47% more accurate than the CNN-LSTM test when the noise is 4 dB, and the DWT-CNN-LSTM network is 9.42% more accurate than the CNN-LSTM test when the noise increases to 0 dB. When the noise continues to increase to −4 dB, the DWT-CNN-LSTM network is 13.08% more accurate than the CNN-LSTM test. This also shows that the addition of DWT enhances the model's noise immunity.

5. Conclusions

In the context of the new energy grid-connected PQDs problem, in order to effectively improve the accuracy of PQDs signal classification, anti-interference ability, and the interpretability of neural networks. This paper proposes an interpretable DWT-CNN-LSTM network for PQDs classification tasks. The following conclusions are drawn after analysis and research:

- (1) It is demonstrated by ablation that the new classification network formed by fusing three modules, DWT, CNN and LSTM, can improve the classification accuracy of PQDs signals.
- (2) Through the visual analysis of the features, it is shown that the introduction of DWT improves the focus on high-frequency signals, and the combination of the physical meaning of the DWT can help people to understand the features extracted by the network, thus increasing the interpretability of the network.
- (3) In 4 dB, 0 dB, and −4 dB environments, CNN-LSTM improves the network's immunity to interference by 0.47%, 9.42%, and 13.08%, respectively, after the addition of DWT, which indicates that the addition of DWT makes the network more robust to the PQDs classification task.

The strong anti-disturbance ability of the model proposed in this paper can realize the accurate classification of disturbance scores in noisy environments, so as to guarantee the quality of the power grid and provide better power grid services for everyone. Since the dataset used in this experiment is obtained by simulation, we can simulate a large number of datasets as needed to train and test the model during the experiment, which

is still somewhat different from the actual situation. In the real situation, the number of complex PQDs is less than the number of single PQDs, which gives rise to the problem of data imbalance. In this case, the accuracy of the model's evaluation of a small number of samples will be reduced, so in the next step, we will continue to carry out research on the phenomenon of PQDs data imbalance.

Author Contributions: Conceptualization, X.Y.; methodology, S.Y. and X.Y.; software, S.Y. and T.S.; validation, S.Y. and T.S.; formal analysis, S.Y.; investigation, S.Y. and X.Y.; resources, X.Y.; data curation, S.Y. and T.S.; writing—original draft preparation, S.Y.; writing—review and editing, S.Y. and X.Y.; visualization, S.Y. and X.Y.; supervision, S.Y., T.S. and X.Y.; project administration, S.Y., T.S. and X.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data generated by MATLAB-Simulink have been clearly stated in Section 3.

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

References

1. 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. [\[CrossRef\]](#)
2. Liu, Y.; Jin, T.; Mohamed, M.A.; Wang, Q. A novel three-step classification approach based on time-dependent spectral features for complex power quality disturbances. *IEEE Trans. Instrum. Meas.* **2021**, *70*, 3000814. [\[CrossRef\]](#)
3. Yang, X.; Guo, L.; Xiao, X.; Zhang, J. Classification of multiple power quality disturbances based on TQWT and random forest feature selection algorithm. *Power Syst. Technol.* **2020**, *44*, 3014–3020.
4. Xue, L.; Zhuo, T.; Wang, R.; Yang, J. Accelerated reasoning method based on fourier transform. *J. Comput. Appl.* **2023**, *43*, 22.
5. Dhoriyani, S.L.; Kundu, P. Comparative group THD analysis of power quality disturbances using FFT and STFT. In Proceedings of the 2020 IEEE First International Conference on Smart Technologies for Power, Energy and Control (STPEC), Nagpur, India, 25–26 September 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 1–6.
6. Zhu, Q.; Yichen, Y.; Zhan, Y.; Li, j.; Hua, R. IGBT open-circuit fault diagnosis of nodular multilevel converter sub-module based on short-time fourier transform and deep networks. *Trans. China Electrotech. Soc.* **2024**, *39*, 3840–3854.
7. Zhao, F.; Yang, R. Power quality disturbance recognition using S-transform. In Proceedings of the 2006 IEEE Power Engineering Society General Meeting, Montreal, QC, Canada, 18–22 June 2006; IEEE: Piscataway, NJ, USA, 2006; pp. 1–7.
8. Guo, F.; Wang, P.; Wang, Y.; Ren, P.; Zhang, Y. Research on improved s transform for the feature extraction of power quality disturbance signal. *IEEE Access* **2020**, *8*, 137910–137917. [\[CrossRef\]](#)
9. Kumar, R.; Kumar, R.; Marwaha, S.; Singh, B. S-Transform based detection of multiple and multistage power quality disturbances. In Proceedings of the 2020 IEEE 9th Power India International Conference (PIICON), Sonapat, India, 28 February–1 March 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 1–5.
10. Guo, L.; Wei, B.; Han, X.; Li, W. Capacity Optimal Configuration of Hybrid Energy Storage in Hybrid AC/DC Micro-grid Based on Ensemble Empirical Mode Decomposition. *High Volt. Eng.* **2020**, *46*, 527–537.
11. Saleh, S.; Moloney, C.; Rahman, M.A. Implementation of a dynamic voltage restorer system based on discrete wavelet transforms. *IEEE Trans. Power Deliv.* **2008**, *23*, 2366–2375. [\[CrossRef\]](#)
12. Gong, J. Research on the characteristics of wavelet selection in power quality signal de-noising. *Electr. Meas. Instrum.* **2022**, *59*, 129–135.
13. Zhang, K. Research on Time Frequency Analysis Method Based on Wavelet Transform and Module Implementation. Master's Thesis, University of Electronic Science and Technology of China, Chengdu, China, 2022.
14. Janik, P.; Lobos, T. Automated classification of power-quality disturbances using SVM and RBF networks. *IEEE Trans. Power Deliv.* **2006**, *21*, 1663–1669. [\[CrossRef\]](#)
15. Ji, X.; Zhang, H.; Jin, Y.; Yan, X.; Wang, H. Based on SVM power quality disturbance classification algorithm. In Proceedings of the 27th Chinese Control and Decision Conference (2015 CCDC), Qingdao, China, 23–25 May 2015; IEEE: Piscataway, NJ, USA, 2015; pp. 3618–3621.
16. 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. [\[CrossRef\]](#)

17. Achlerkar, P.D.; Samantaray, S.R.; Manikandan, M.S. Variational mode decomposition and decision tree based detection and classification of power quality disturbances in grid-connected distributed generation system. *IEEE Trans. Smart Grid* **2016**, *9*, 3122–3132. [\[CrossRef\]](#)
18. Guo, J.; Li, K.; He, S.; Zhang, m. A real time power quality disturbance classification based on improved incomplete S-transform and decision tree. *Power Syst. Prot. Control* **2013**, *41*, 8.
19. Marzuki, M.S.N.B.I.; Abidin, A.F.; Jusoh, M.Z.; Johari, D. Single and Double PQ Disturbances Classification Based on Stockwell Transform and Support Vector Machine. In Proceedings of the 2021 IEEE International Conference in Power Engineering Application (ICPEA), Putrajaya, Malaysia, 8–9 March 2021; IEEE: Piscataway, NJ, USA, 2021; pp. 210–215.
20. 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. [\[CrossRef\]](#)
21. Xiao, X.; Li, K. Multi-label classification for power quality disturbances by integrated deep learning. *IEEE Access* **2021**, *9*, 152250–152260. [\[CrossRef\]](#)
22. Dawood, Z.; Babulal, C. Red deer optimized recurrent neural network for the classification of power quality disturbance. *Electr. Eng.* **2023**, *105*, 1937–1953. [\[CrossRef\]](#)
23. Tong, Z.; Zhong, J.; Li, J.; Wu, J.; Li, Z. A power quality disturbances classification method based on multi-modal parallel feature extraction. *Sci. Rep.* **2023**, *13*, 17655. [\[CrossRef\]](#)
24. Wang, X.; Dou, J.; Liu, Z.; Liu, C.; Pu, T.; He, J. Review and prospect of explainable artificial intelligence and its application in power systems. *Autom. Electr. Power Syst.* **2024**, *48*, 169–191.
25. Sun, Y.; Sun, X.; Jia, X.; Song, Y.; Yu, C. Research on Fault Diagnosis of Low Speed and Heavy Load Bearing Based on Interpretable Neural Network Model. *Bearing* 1–10. Available online: <https://link.cnki.net/urlid/41.1148.TH.20231127.1538.002> (accessed on 5 January 2025).
26. Chen, Q.; Dong, X.; Tu, G.; Wang, D.; Cheng, C.; Zhao, B.; Peng, Z. TFN: An interpretable neural network with time-frequency transform embedded for intelligent fault diagnosis. *Mech. Syst. Signal Process.* **2024**, *207*, 110952. [\[CrossRef\]](#)
27. Wang, H.; Liu, Z.; Peng, D.; Zuo, M.J. Interpretable convolutional neural network with multilayer wavelet for Noise-Robust Machinery fault diagnosis. *Mech. Syst. Signal Process.* **2023**, *195*, 110314. [\[CrossRef\]](#)
28. He, J.; Liu, Y.; Zhu, Y.; Yang, Y.; Lu, Q.; Hu, Z. Physics-informed TPA-LSTM Network Applied to Multi-step Wind Speed Prediction of Offshore Wind Farms Under Typhoon Conditions. *Ower Syst. Technol.* **2023**, *47*, 4152–4161. [\[CrossRef\]](#)
29. Mallat, S.G. A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Trans. Pattern Anal. Mach. Intell.* **1989**, *11*, 674–693. [\[CrossRef\]](#)
30. Wu, Y.; He, K. Group normalization. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 3–19.
31. Machlev, R.; Chachkes, A.; Belikov, J.; Beck, Y.; Levron, Y. Open source dataset generator for power quality disturbances with deep-learning reference classifiers. *Electr. Power Syst. Res.* **2021**, *195*, 107152. [\[CrossRef\]](#)
32. *IEEE Std 1159-2019 (Revision of IEEE Std 1159-2009)*; IEEE Recommended Practice for Monitoring Electric Power Quality. IEEE: Piscataway, NJ, USA, 2019; pp. 1–98. [\[CrossRef\]](#)

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.