

Comparative study of CNN and fused 2D CNN-LSTM with CWT and STFT for power quality disturbance classification

Bouchra Ferial Khaldi, Fatma Zohra Dekhandji, Abdelmadjid Recioui

Laboratory of Signals and Systems, Department of Power and Control, Institute of Electrical and Electronic Engineering, University M'hamed Bougara, Boumerdes, Algeria

Article Info

Article history:

Received May 20, 2025

Revised Mar 21, 2026

Accepted Apr 19, 2026

Keywords:

Continuous wavelet transform

Convolutional neural networks

Long short-term memory

Microgrid faults

Power quality disturbances

Short-time Fourier transform

ABSTRACT

The integration of solar and wind energy has increased electricity generation but also introduced power quality disturbances (PQDs) that threaten grid stability. This study examines the detection and classification of five PQD types—voltage sag, swell, interruption, harmonics, and normal conditions—across noisy environments (0, 10, 20, and 30 dB) signal-to-noise ratio (SNR). Traditional methods—support vector machine (SVM), random forest (RF), artificial neural networks (ANN), and 1D convolutional neural networks (1D CNN)—are evaluated on raw signal data, while advanced models—2D CNN and fused 2D CNN-LSTM—utilize time-frequency representations (continuous wavelet transform (CWT) and short-time Fourier transform (STFT)). Results show that deep learning (DL) models achieve high accuracy even in noisy environments, with the fused 2D CNN-LSTM using CWT outperforming all other methods. Noise adversely affects feature extraction, with CWT consistently outperforming STFT under low SNR conditions. These findings demonstrate that combining DL models with robust time-frequency analysis and temporal modeling enhances PQD classification and supports dependable monitoring in smart grid environments.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Bouchra Ferial Khaldi

Laboratory of Signals and Systems, Department of Power and Control

Institute of Electrical and Electronic Engineering, University M'hamed Bougara

Boumerdes, Algeria

Email: b.khaldi@univ-boumerdes.dz

1. INTRODUCTION

The integration of renewable energy sources (RESs) is essential for meeting the growing demand for electricity. However, this integration often causes power quality (PQ) disturbances, such as harmonics, voltage sags, swells, and interruptions, which can reduce grid stability and damage electrical equipment. PQ refers to the ability of the electrical system to provide consistent and reliable electricity, usually with stable sinusoidal waveforms and steady voltage and frequency [1]. Any changes in these ideal conditions, including disturbances in voltage or current, can lead to inefficiencies and system failures. Therefore, continuous monitoring and quick detection of these disturbances are necessary to maintain the stability and performance of electrical systems [2]. Feature extraction is an important part of monitoring, as it involves identifying key characteristics of electrical signals that show disturbances [3]. Traditional methods, such as Fourier transform (FT) and discrete Fourier transform (DFT), work well for analyzing signals with steady characteristics. However, these methods have limitations when applied to nonstationary signals, as they struggle to capture sudden disturbances. To overcome these limitations, more advanced methods, such as short-time Fourier transform (STFT) and s-transform (ST), have been introduced. While these techniques provide better

frequency information over time, they still face difficulties in analyzing fast or changing disturbances. Moreover, feature extraction is significantly impacted by the presence of noise, which can distort key signal characteristics and lead to misclassification. As reported in [4], noisy environments can reduce the accuracy of classification models, highlighting the importance of robust feature extraction techniques for reliable PQD detection. Artificial intelligence (AI), especially machine learning (ML) and deep learning (DL), has become a powerful tool for automating the detection and classification of PQ disturbances [5]. Combining AI models with advanced signal processing methods has further improved disturbance classification. The researchers [6]-[8], wavelet transform (WT) is integrated with classifiers like support vector machine (SVM) and artificial neural networks (ANN), demonstrating strong performance even in noisy conditions. The researchers [9]-[11], discrete wavelet transform (DWT) is combined with Kalman filters and SVM for effective feature extraction and classification, improving computational efficiency and accuracy. ST is used in combination with classifiers like random forest (RF), decision trees (DT), and optimization methods like whale optimization algorithm (WOA), demonstrating enhanced detection of power quality disturbance (PQD) in real-time [12]-[14]. Hilbert-huang transform (HHT) is integrated with AI models, such as neural networks, neuro-fuzzy systems and fuzzy-decision trees, significantly improving adaptability to complex disturbance patterns [15]-[17]. The researchers [18], [19], the gabor transform (GT) is combined with probabilistic neural networks (PNN) and SVM for real-time recognition of PQDs, offering high accuracy and robustness in noisy environments. Convolutional neural networks (CNNs) and long short-term memory (LSTMs) are widely used techniques as well, that have proven successful in classifying different types of PQ events. Channa *et al.* [20] proposed a CNN-LSTM model for classifying PQDs in PV-integrated microgrids, the approach was validated using PSCAD-simulated microgrid signals. Cen *et al.* [21] a fused CNN-LSTM model using FFT for real-time PQD classification is presented and in [22] a comparison of CNN, LSTM, and CNN-LSTM was made for PQD detection using simulated signals. They employed STFT for feature extraction, though it was not tested with noisy signals.

In this paper, we aim to address the challenge of accurately classifying PQDs in noisy environments by developing a DL-based framework focused on enhancing the feature extraction and classification process. PQ events are generated with diverse amplitudes, durations, and timings, while varying noise levels (0, 10, 20, and 30 dB) are introduced to simulate realistic grid conditions. The extreme 0 dB case serves to evaluate the models' robustness under severe noise. Three main model configurations are evaluated: i) a 1D CNN using raw time-domain signals, ii) a 2D CNN using time-frequency representations from continuous wavelet transform (CWT) and STFT, and iii) a fused 2D CNN-LSTM model leveraging both spatial and temporal features. To improve generalization, data augmentation, normalization, and noise injection are applied. Hyperparameters are optimized using a Bayesian-guided manual tuning strategy. The models are tested on synthetic signals and validated using Simulink-generated microgrid disturbances.

The key achievement of this work is the innovation of a robust and noise-resilient DL framework for PQD classification, supported by a comprehensive comparative analysis of traditional and DL models under various noise conditions. Unlike previous studies that focused primarily on ideal or low-noise scenarios, this research systematically evaluates the effect of noise on model performance and generalization. By integrating advanced time-frequency feature extraction techniques CWT and STFT with CNN and hybrid CNN-LSTM architectures, and benchmarking them against the conventional classifiers such as ANN, SVM, and RF, the study establishes clear performance boundaries and highlights the superior adaptability of DL approaches. Additional contributions include: i) the creation of a noise-rich and diverse PQD dataset, ii) the design of a fused CNN-LSTM model for enhanced temporal modeling, and iii) the application of Bayesian-guided tuning for optimized performance. Overall, this work offers new insights into the robustness of PQD classifiers under realistic operating conditions and provides a reliable foundation for PQ monitoring in modern smart grids.

The structure of the paper is organized as follows: section 2 introduces the key methodologies, including CNN, LSTM, CWT, and STFT, and their relevance to PQD classification. Section 3 describes the synthetic data generation and preprocessing steps, focusing on creating a balanced, noise-rich dataset. Section 4 details the architectures of ANN, 1D CNN, 2D CNN, and fused 2D CNN-LSTM, highlighting their integration with CWT and STFT for feature extraction. Section 5 outlines the experimental setup, including hyperparameter tuning and training configurations. Section 6 presents the results and discussion, emphasizing the ablation study that evaluates the independent contributions of feature extraction methods and model components. Section 7 validates the findings through microgrid fault detection simulations, demonstrating the re-al-world applicability and robustness of the proposed methodologies. The last section, section 8, is the conclusion of this research. This paper contributes by systematically identifying the optimal feature extraction techniques and model architectures for PQD classification, validated through rigorous experimentation and real-world simulations.

2. METHOD

This section introduces the main techniques used to develop the proposed framework for PQD classification. It includes a brief overview of CNN, LSTM networks, and two time–frequency feature extraction methods—STFT and CWT. These methods form the foundation of the DL architectures employed in this study to ensure robust feature extraction and accurate classification under noisy conditions.

2.1. Convolutional neural network

CNNs are a powerful DL approach used extensively in pattern recognition and image classification. Their architecture is inspired by the way neurons in the human brain connect [23]. CNNs have two main components: the feature extractor, which transforms input images into feature maps, and the classifier, which typically uses a neural network to process these maps. Their architecture is structured hierarchically, with multiple layers performing specific functions such as convolution, pooling, non-linearity, and full connection. Each layer extracts progressively more abstract features from the input images, leading to highly efficient pattern recognition. The key layers are further illustrated and explained in Figure 1 [24].

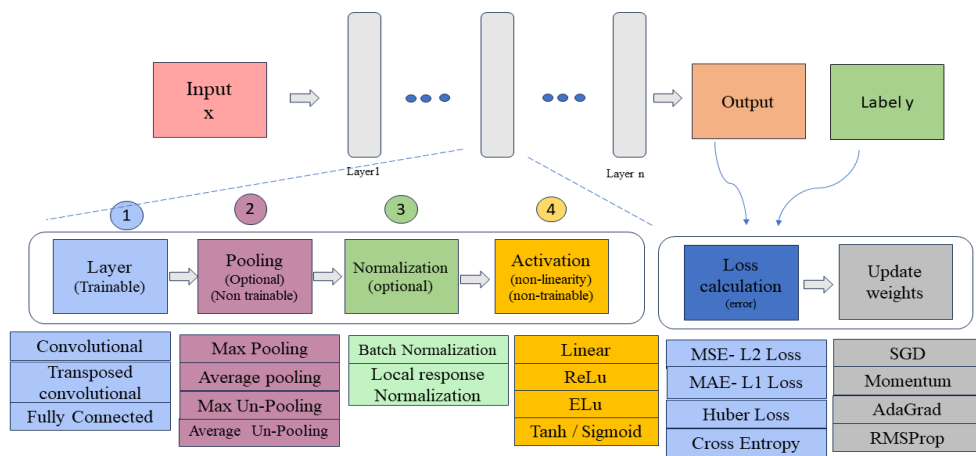


Figure 1. CNN architecture

The central component of CNNs is the convolutional layer, which uses kernel filters that slide over the input image to extract features and execute element-wise multiplication and summing to produce feature maps. Batch normalization improves training stability and convergence, while strides and padding regulate filter movement and output size [23], [24].

The pooling layer reduces feature map dimensions while preserving key information, minimizing computation and overfitting risks, crucial in PQ classification [24], [25]. Among pooling types, max pooling is preferred for emphasizing prominent features. After pooling, the 2D maps are flattened into a 1D vector for the next layer. Figure 2 illustrates the pooling techniques: Figure 2(a) shows the original matrix, Figure 2(b) shows max pooling, and Figure 2(c) shows average pooling.

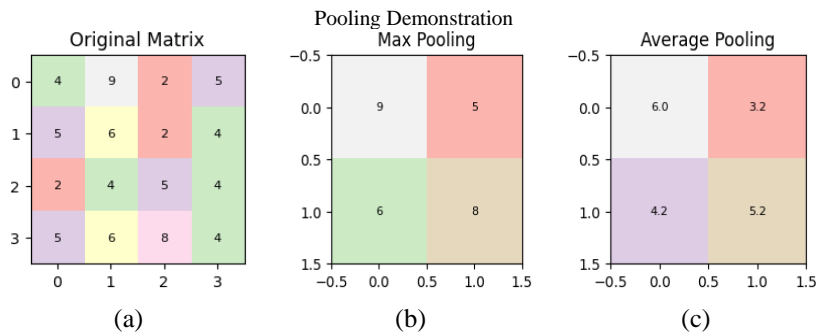


Figure 2. Pooling techniques: (a) original matrix, (b) max pooling technique, and (c) average pooling technique

The fully connected (FC) layer, located at the end of the CNN, connects all neurons from the previous layer to deliver final predictions by analyzing high-level feature combinations [7], [23], [24]. The non-linearity layer applies activation functions to introduce complexity into learning. The rectified linear unit (ReLU) is the most used due to its efficiency, defined as (1):

$$f(x)\text{ReLU} = \max(0, x) \quad (1)$$

It converts all negative inputs to zero, accelerating computation [25].

The loss function evaluates prediction errors, guiding model optimization. Cross-entropy is used for multi-class classification, while root mean square error (RMSE) suits regression tasks [26]. Overall, CNNs achieve high accuracy in PQD image classification through automatic hierarchical feature extraction from time–frequency representations.

2.2. Long short-term memory

A recurrent neural network (RNN) is a type of ANN designed to process sequential data by introducing a feedback loop in its architecture. A drawback of RNNs lies in their inability to retain long-term dependencies effectively due to the issue of the vanishing gradient [27]. One solution to this problem in RNNs is to use a variant of the basic RNN architecture, such as LSTM. LSTMs are designed to handle long-term dependencies in sequential data, which makes them well-suited for analyzing PQ data. LSTMs address the problem of the vanishing gradient by introducing a memory cell, which is a container that can hold information for an extended period of time. The memory cell, depicted in Figure 3, is controlled by three gates: the input gate, the forget gate, and the output gate. These gates decide what information to add to, remove from, and output from the memory cell [28].

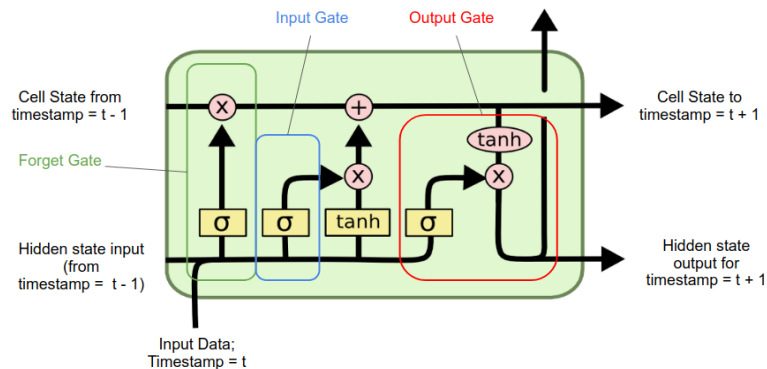


Figure 3. LSTM cell architecture [29]

The forget gate controls which information from the previous cell state should be retained or discarded using a sigmoid function that outputs values between 0 (forget) and 1 (keep). The input gate selects relevant information from the current input, combining a sigmoid function for regulation and a tanh function to generate values between -1 and 1. The output gate defines the next hidden state using both sigmoid and tanh functions to determine which information is passed forward. In this study, LSTMs are used to model the temporal evolution of PQDs. When combined with CNNs, the LSTM layer receives feature sequences extracted by the CNN and learns temporal correlations between them, enhancing the model's ability to recognize disturbances that evolve over time.

2.3. Short time Fourier transform

In response to the limitations of FFT, STFT has been employed to address issues such as extracting both frequency and phase information from signals [5]. Disturbance signals are analyzed in both the frequency and time domains, providing details such as start time, stop time, rise time, and duration through the STFT. These signals are typically non-stationary, with properties like amplitude, frequency, and phase varying over time. The discrete STFT of a signal $x[n]$ is defined as (2) [2]:

$$X[n, \lambda] = \sum_{m=-\infty}^{\infty} x[n+m]w[m]e^{-j\lambda m} \quad (2)$$

Where a window function of size L is represented as $w[m]$. The signal $x[n]$ usually consists of N discrete samples, and the window $w[m]$ is defined by (3):

$$w[l] \neq 0, \text{ if } 0 \leq l \leq L - 1$$

$$w[l] = 0, \text{ if } l < 0 \text{ or } l > L \quad (3)$$

where the size of the window is L , which can be the same or less than the number of samples N of signal $x[n]$. with:

$$X[n, \lambda] = \sum_{m=0}^{L-1} x[n + l]w[l]e^{-j\lambda l/N} \quad (4)$$

The effectiveness of the STFT depends on the choice of the window function, which determines how the signal is localized in time and frequency. A rectangular window, the simplest option, leaves the signal unchanged within the section and sets values outside to zero. However, it causes boundary discontinuities, introducing distortions that spread across the frequency spectrum. To address this, smoother windows like the Hann window are commonly used. The Hann window, a raised cosine function tapering to zero at the boundaries, reduces such distortions but slightly blurs frequencies, trading distortion reduction for spectral accuracy [30]. Figure 4 illustrates the effect of windowing on a sinusoidal signal: Figure 4(a) shows the rectangular window, and Figure 4(b) shows the Hann window. Although STFT effectively captures both time and frequency information, its fixed window size limits resolution adaptability, which motivates the use of CWT for more detailed feature extraction.

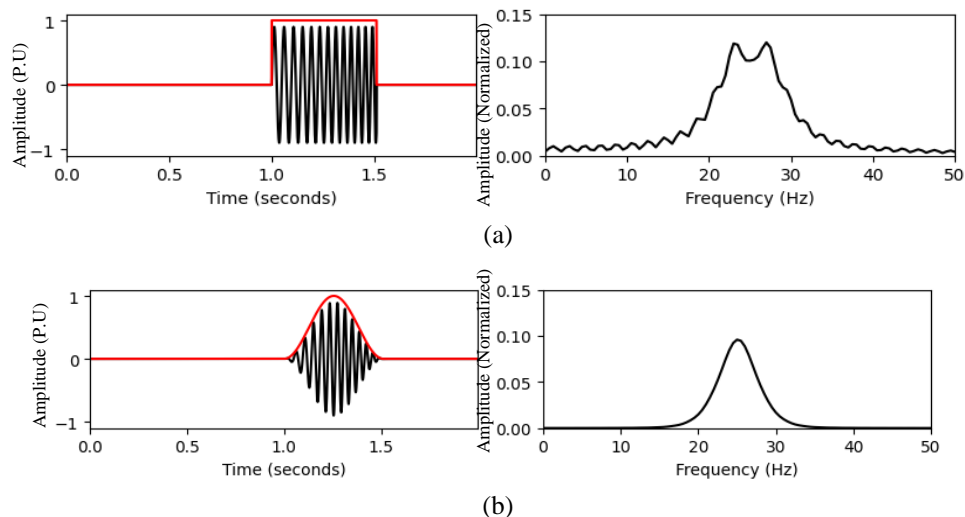


Figure 4. Comparison between rectangular and Hann Windows on a sinusoidal signal; (a) rectangular window and (b) Hann window [9]

2.4. Continuous wavelet transform

In order to address the limitations of the STFT, the WT was introduced. Three primary methods are available for implementing WT: wavelet series, continuous WT, and discrete WT. Every wavelet in continuous WT is produced by translating and scaling operations in a mother wavelet, this latter is an oscillating function with zero average and finite energy. For a continuous time signal $x(t)$, the continuous WT is defined as (5) [31]:

$$CWT_{\psi}x(a, b) = \int_{-\infty}^{\infty} x(t) \Psi_{a,b}^*(t) dt \quad a, b \in R, a \neq 0 \quad (5)$$

where:

$$\Psi_{a,b}^*(t) = \frac{1}{\sqrt{a}} \Psi^*\left(\frac{t-b}{a}\right) \quad (6)$$

The mother wavelet is $\Psi_{a,b}^*(t)$. Scaling and translating parameters are denoted by a and b , respectively. The WT addresses the limitations of the STFT by offering superior time and frequency localization, as wavelets scale effectively across both domains. In this work, CWT is used to generate time–

frequency images that emphasize transient and oscillatory characteristics of PQDs. These images enable CNN-based models to capture fine details and sharp transitions caused by disturbances such as harmonics or interruptions.

3. DATA GENERATION AND PREPROCESSING

3.1. Synthetic data generation

To comprehensively evaluate the robustness and generalization ability of PQD classification models, a balanced synthetic dataset was constructed representing five PQD types: voltage sag, swell, interruption, harmonics, and normal conditions. Synthetic data allow precise control over signal parameters and noise levels, enabling systematic evaluation of model robustness under conditions that may be rare or difficult to capture in real measurements. Each disturbance signal was generated using parameterized mathematical models, with key parameters (frequency, amplitude, phase, disturbance onset, and duration) randomized within realistic ranges to simulate the natural variability observed in practical settings. The mathematical formulas and the specific parameter ranges used for each PQD type are summarized in Table 1. This approach ensures diversity and prevents overfitting to fixed signal patterns. To emulate real-world measurement conditions, Gaussian white noise was added to each signal at four SNR levels: 0 dB, 10 dB, 20 dB, and 30 dB. The inclusion of 0 dB SNR provides a challenging stress test of model robustness, while the upper SNRs reflect more common operating conditions. For each disturbance class and SNR level, an equal number of samples were generated to ensure class balance.

- $V(t)$: instantaneous voltage signal
- t : time vector (0 to 1 s, sampled at 1 kHz)
- f : frequency (Hz)
- φ : phase (radians)
- A : amplitude (per unit)
- Sag depth/swell depth: scaling factor during disturbance
- onset/duration: random window within the signal
- h_2, h_3 : harmonic coefficients

Table 1. Mathematical formulations and parameter ranges used for synthetic PQD data generation [4]

PQD	Formula	Key parameters	Description
Normal	$v(t)=A \cdot \sin(2\pi ft + \varphi)$	f : 48–52 Hz, φ : $0-2\pi$, A : 0.8–1.2 p.u.	Pure sinusoidal voltage
Sag	$v(t)=A \cdot \sin(2\pi ft + \varphi)$; sag window= $v(t) \times$ sag depth	f : 48–52 Hz (frequency), φ : $0-2\pi$ (phase), A : 0.8–1.2 p.u. (amplitude), Sag depth: 0.2–0.5, onset: random (s), duration: random (s)	Voltage dip for a short time
Swell	$v(t)=A \cdot \sin(2\pi ft + \varphi)$; swell window= $v(t) \times$ swell depth	f : 48–52 Hz, φ : $0-2\pi$, A : 0.8–1.2 p.u., Swell depth: 1.2–1.5, onset: random (s), duration: random (s)	Voltage increase for a short time
Interrupted	$v(t)=A \cdot \sin(2\pi ft + \varphi)$; interruption window= $v(t)=0$	f : 48–52 Hz, φ : $0-2\pi$, A : 0.8–1.2 p.u., onset: random (s), duration: random (s)	Complete voltage loss briefly
Harmonic	$v(t)=A$ $[\sin(2\pi ft + \varphi) + h_2 \cdot \sin(4\pi ft + \varphi) + h_3 \cdot \sin(6\pi ft + \varphi)]$	f : 48–52 Hz, φ : $0-2\pi$, A : 0.8–1.2 p.u., h_2 : 0.4–0.6, h_3 : 0.2–0.4	Signal with 2nd and 3rd harmonics

3.2. Data preprocessing

Prior to model training, all synthetic data underwent systematic preprocessing to ensure consistency and suitability for DL models. For the raw signal dataset, each voltage signal was normalized to the [0, 1] range, facilitating stable and efficient training. Class labels were one-hot encoded to support multi-class classification tasks. The dataset was then split into training and validation subsets using stratified sampling, ensuring balanced class distributions in each set. For the image-based datasets (CWT and STFT), each signal

was transformed into a 2D time-frequency representation. The resulting images were resized to a uniform dimension and normalized to the [0, 1] range. Images were organized according to their class and SNR level for efficient loading during training. All preprocessing steps were designed to maintain class balance and prevent information leakage between training and validation data.

4. MODELS ARCHITECTURES

This study systematically compares ML and DL approaches for PQD classification using both raw voltage signals and feature extraction techniques such as CWT and STFT, with an emphasis on robustness under varying noise conditions. We implemented a classical ANN for baseline comparison, a 1D convolutional neural network (1D CNN) for extracting temporal features from raw signals, and both 2D CNN and 2D CNN-LSTM architectures for capturing spatial and temporal features from time-frequency images, evaluating performance across SNRs of 0, 10, 20, and 30 dB to simulate both practical and extreme electrical noise environments. The workflow used in this study is illustrated in Figure 5, which summarizes the models and techniques employed.

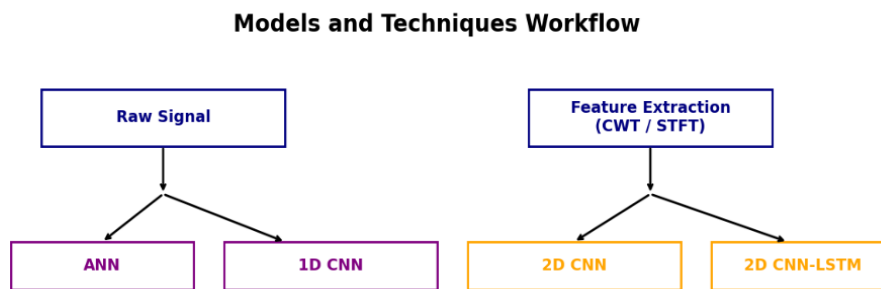


Figure 5. Block diagram of the workflow

4.1. Artificial neural network

The ANN model serves as a classical baseline for PQD classification using raw voltage signals. It processes normalized signals of length 1,000 and applies FC layers with batch normalization and LeakyReLU activation for non-linear mapping. As shown in Figure 6, the final SoftMax layer outputs probabilities for five disturbance classes.

4.2. 1D convolutional neural network

The 1D CNN model, selected for its ability to efficiently extract temporal features from raw voltage signals, also uses the 1,000-point voltage signal as input. It consists of convolutional layers with max pooling for feature extraction, followed by a FC layer and a SoftMax output layer for classification into five categories, as summarized in Figure 7. The model size is about 7.95 MB. The network achieves fast inference, requiring around 0.1 milliseconds per sample.

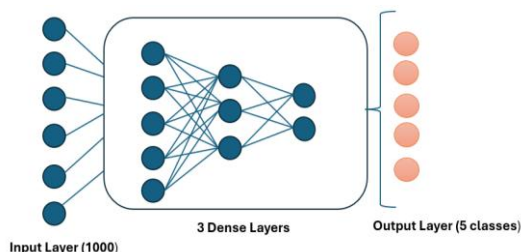


Figure 6. ANN architecture

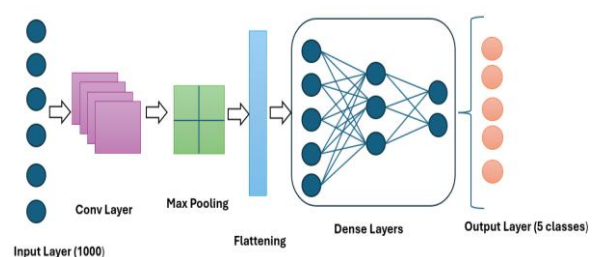


Figure 7. 1D CNN architecture

4.3. 2D convolutional neural network

The 2D CNN model is utilized to capture spatial features from time-frequency images for improved classification accuracy. It processes 128×128 RGB images generated from CWT or STFT representations of

voltage signals. The architecture consists of convolutional layers, each followed by max pooling for hierarchical feature extraction, then a FC and a SoftMax output for five-class classification. The network has a model size of 1.22 MB. The average inference time is about 3.05 milliseconds per image. The architecture is presented in Figure 8.

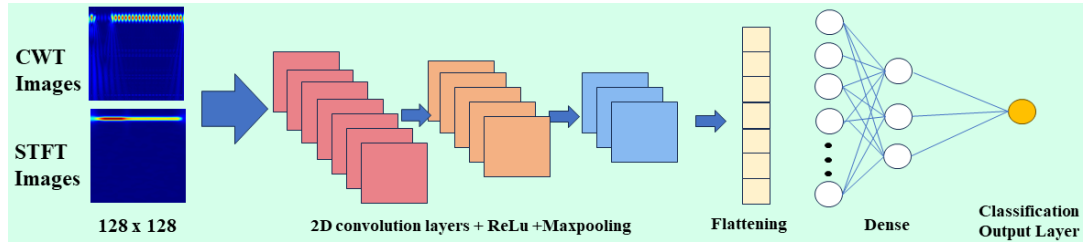


Figure 8. 2D CNN architecture

4.4. Fused 2D CNN-LSTM

Incorporating LSTM after CNN layers enables the model to capture temporal dependencies across the extracted spatial features, improving classification of disturbances with dynamic temporal characteristics. The fused 2D CNN-LSTM model also accepts 128x128 RGB images as input. This model, as presented in Figure 9, is composed of convolutional layers with max pooling for spatial feature extraction, followed by flattening, reshaping, and a subsequent LSTM layer. The output is processed by a FC layer and a SoftMax output for five-class classification. The model size is about 64.67 MB, with an average inference time of about 3.19 milliseconds per image. Table 2 summarizes the models architectures details.

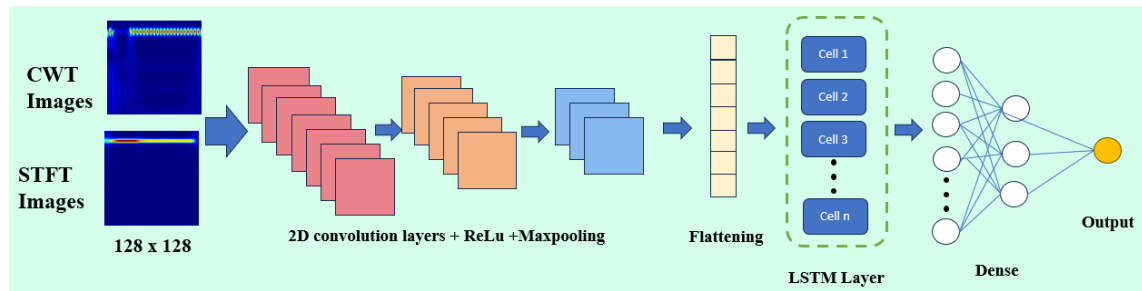


Figure 9. Fused 2D-CNN-LSTM architecture

Table 2. Summary of model architectures and key components

Model	Input type	Feature extraction	Main layers	Dropout	Activation	Output	Trainable params	Model size	Batch size (inference)
ANN	1D (1000,)	None	Dense (128, 64, 5), batch normalization, LeakyReLU	Yes	LeakyReLU	5	137,477	0.54 MB	128
1D CNN	1D (1000, 1)	None	3×Conv1D, MaxPool, Flatten, Dense (128, 5)	Yes	ReLU	5	2,083,937	7.95 MB	128
2D CNN	2D (128×128×3)	CWT/STFT	4×Conv2D, MaxPool, Flatten, Dense (128, 5)	No	ReLU	5	320,213	1.22 MB	64
2D CNN-LSTM	2D (128×128×3)	CWT/STFT	3×Conv2D, MaxPool, Flatten, LSTM, Dense (128, 5)	Yes	ReLU	5	16,953,669	64.67 MB	32

5. EXPERIMENTAL SETUP

All experiments were conducted on the synthetic dataset described in section 3, comprising five PQD classes (voltage sag, swell, interruption, harmonics, and normal) and four SNR levels (0 dB, 10 dB, 20 dB, and 30 dB). Each class and SNR combination contained 400 samples, resulting in a total of 8,000 signals. To ensure balanced representation, each signal was normalized and split into training (80%) and validation (20%) sets using stratified sampling, maintaining consistent class and noise-level distribution.

Three data representations were evaluated:

- Raw voltage signals (1,000 points per signal).
- Time-frequency images obtained via CWT using the amor wavelet.
- Time-frequency images obtained via STFT using a Hann window.

For the image-based data, resulting scalograms and spectrograms were resized to 128×128 pixels. Data preprocessing included normalization for all inputs and random spatial transformations for images to improve generalization.

Four model architectures were compared:

- A classical ANN.
- A 1D CNN for temporal feature extraction from raw signals.
- A 2D CNN for spatial feature extraction from time-frequency images.
- A fused 2D CNN-LSTM combining spatial and temporal features.

All models were implemented in Tensor Flow 2.12.0 and trained on Google Colab using an NVIDIA Tesla T4 GPU. Training was conducted for up to 20 epochs with early stopping (patience 5), and batch sizes were set according to model complexity and GPU memory, as summarized in Table 3.

Table 3. Summary of model architectures and inference performance

Model	Optimizer	Learning rate	Epochs	Loss function	Regularization	Batch size (inference)	Inference time (ms/sample)
ANN	Adam	Auto-adjusted by Adam	35	Categorical cross-entropy	Dropout (0.3, 0.2)	128	0.04
1D CNN	Adam	Auto-adjusted by Adam	20	Categorical cross-entropy	Dropout (0.5)	64	0.10
2D CNN	SGD	0.01	20	Categorical cross-entropy	None	64	3.05
2D CNN-LSTM	SGD	0.01	20	Categorical cross-entropy	Dropout (0.2)	32	3.19

Hyperparameter selection—including learning rate, number of filters, kernel sizes, and dense units—was performed using a combination of Bayesian optimization (via scikit-optimize) and manual fine-tuning. Bayesian optimization allowed the efficient exploration of promising parameter ranges by updating prior beliefs with new experimental evidence, while manual refinement further improved classification accuracy and reduced validation loss. Particle swarm optimization (PSO) was initially considered but proved difficult to integrate, leading to the exclusive use of Bayesian optimization.

To ensure unbiased evaluation and prevent overfitting, all hyperparameter tuning was performed using the validation set only, with the test set kept strictly separate until the final assessment. This systematic approach guarantees that the selected models are both accurate and computationally efficient, capable of robust PQD classification across diverse noise levels and data representations.

6. RESULTS AND DISCUSSION

In this section, the previously described models are systematically trained and evaluated for the task of PQD classification under varying noise conditions. To ensure experimental rigor and fair comparison, all models were trained using the Adam optimizer with categorical cross-entropy loss; however, the number of training epochs and batch sizes were tailored to each architecture's convergence properties and computational requirements. This configuration reflects best practices in DL for balancing model complexity, training stability, and computational efficiency. The evolution of training and validation accuracy and loss for each model is illustrated in Figure 10, Figure 10(a) for the ANN, Figure 10(b) for the 1D CNN, Figure 10(c) for the 2D CNN, and Figure 10(d) for the 2D CNN-LSTM. Similarly, the corresponding training and validation loss curves are presented in Figure 11, Figure 11(a) for the ANN, Figure 11(b) for the 1D CNN, Figure 11(c) for the 2D CNN, and Figure 11(d) for the 2D CNN-LSTM.

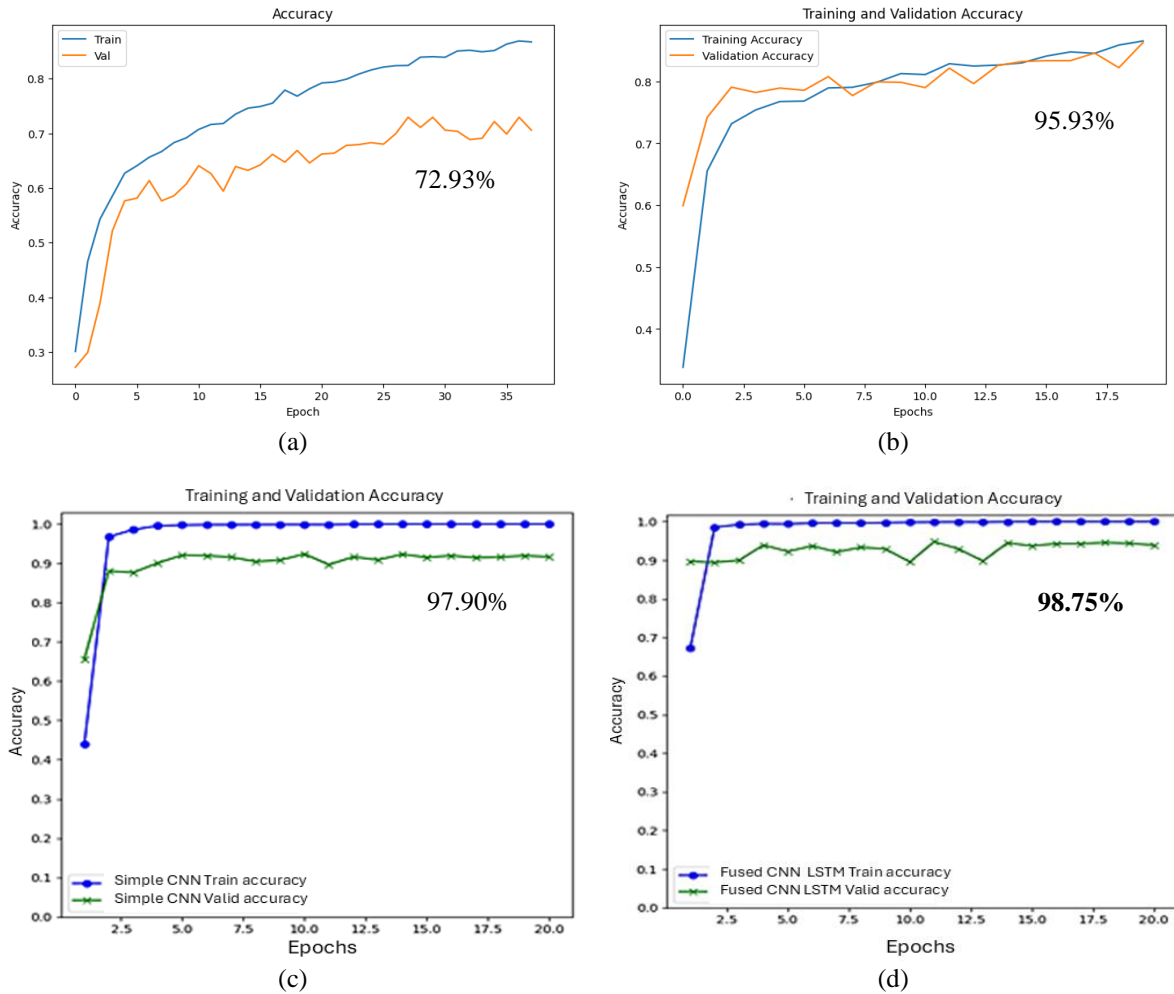


Figure 10. Training and validation curves: (a) ANN, (b) 1D CNN, (c) 2D CNN, and (d)-2D CNN-LSTM

Examining these curves together provides a comprehensive view of the learning dynamics and generalization capabilities of each architecture under standardized training conditions. The ANN model (Figures 10 (a) and 11(a)) shows a gradual increase in validation accuracy and a corresponding decrease in loss; however, both curves exhibit noticeable fluctuations, indicating limited generalization capability, likely due to its simpler structure and sensitivity to noise. In contrast, the CNN-based models (Figures 10(b)–(d) and 11(b)–(d)), including the 1D CNN, 2D CNN, and the fused 2D CNN-LSTM, display rapid improvements in accuracy and sharp reductions in loss during the initial training epochs, with both metrics stabilizing at higher accuracy and lower loss values. The relatively small gaps between training and validation curves for these deep models suggest effective regularization and robust generalization to unseen data. Notably, the fused 2D CNN-LSTM achieves the highest validation accuracy and consistently low loss, highlighting the benefit of combining spatial and temporal feature extraction.

Together, these results underscore the importance of model architecture and feature extraction strategies for achieving robust PQD classification performance, particularly under challenging and noisy conditions, as further analyzed in subsequent sections. The comparison of the confusion matrices for each model is presented in Figure 12(a) (ANN), Figure 12(b) (1D CNN), Figure 12(c) (2D CNN), and Figure 12(d) (fused 2D CNN-LSTM).

The ANN Figure 12(a) demonstrates notable confusion between the 'normal', 'sag', and 'swell' classes, as well as some misclassification of 'Interrupted' events, reflecting its limited feature extraction capacity. The 1D CNN Figure 12(b) shows a clear improvement, with most classes being accurately identified and only minor confusion remaining, mainly between 'normal' and 'swell', and between 'Sag' and 'Interrupted'. In contrast, Figure 12(c) the 2D CNN, and Figure 12(d) 2D CNN-LSTM deliver near-perfect classification, with almost all samples correctly assigned to their true categories and only a handful of misclassifications visible. These results indicate that increasing model complexity reduces class confusion

and enhances the ability to distinguish between categories, highlighting the effectiveness of deep spatial and spatiotemporal feature extraction for precise PQD classification in challenging scenarios.

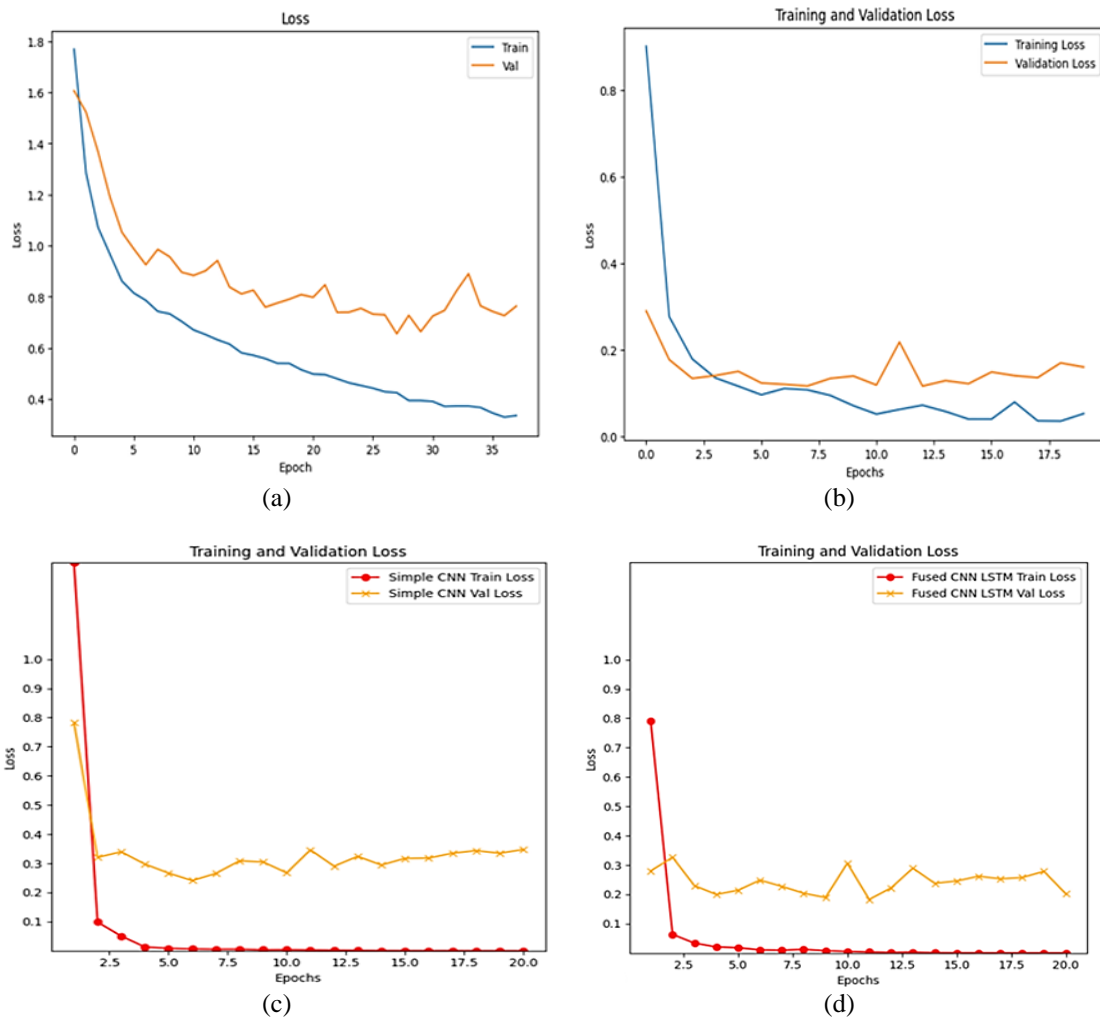


Figure 11. Training and validation loss curves: (a) ANN, (b) 1D CNN, (c) 2D CNN, and (d)-2D CNN-LSTM

The graphs in Figure 13, provide a clear comparison of the noise impact on the performance of various models across different PQD types (sag, swell, interrupted, normal, and harmonic) and SNR levels (0 dB, 10 dB, 20 dB, and 30 dB). Each subplot represents a specific model, offering an insightful look into the trends and findings. The ANN model, as shown in Figure 13(a) exhibits significant sensitivity to noise, with low accuracy for most PQD types at 0 dB. As the SNR increases, the accuracy improves, peaking at 84.29% for normal PQD at 30 dB, but overall performance remains inconsistent. The 1D CNN model in Figure 13(b) performs better, showing steady improvement as noise decreases, especially for normal and harmonic PQD types, which achieve 100% accuracy at higher SNR levels. However, raw signal processing without feature extraction limits its effectiveness under noisy conditions. The 2D CNN models (CWT and STFT) demonstrate remarkable stability and robustness to noise Figures 13(c) and (d). For most PQD types, accuracy is near-perfect at low SNR levels and reaches 100% across the board as SNR increases, highlighting the strength of feature extraction techniques like CWT and STFT. Among the models, 2D CNN-LSTM with CWT, presented in Figure 13(e) stands out as the best performer, maintaining 100% accuracy across all PQD types starting from 10dB and showing exceptional resilience to noise. While for Figure 13(f), the 2D CNN-LSTM model with STFT also achieves 100% accuracy for most PQD types, but it shows slight variability under lower SNR conditions, particularly for sag and swell.

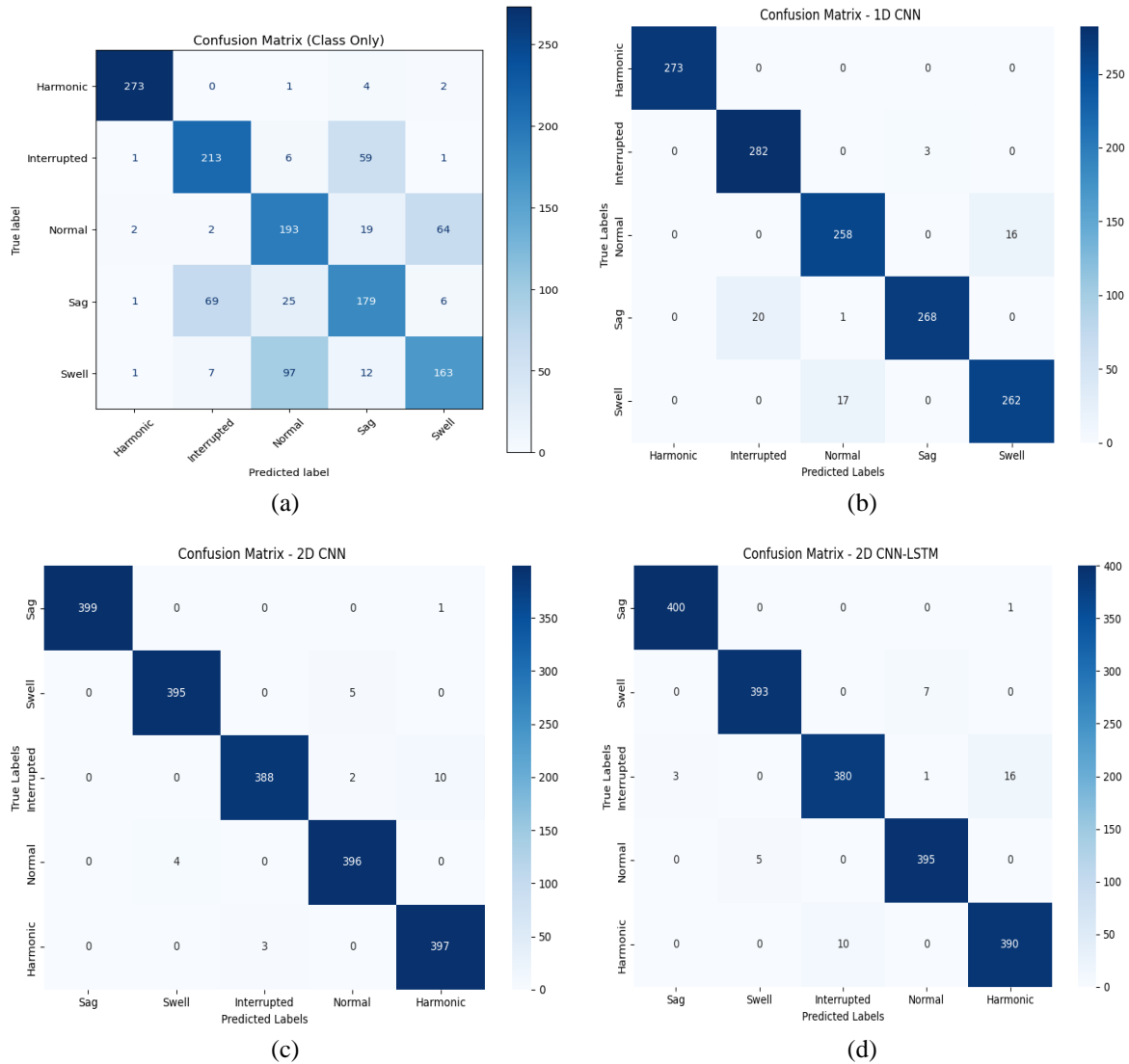


Figure 12. Comparison between confusion matrices of the models: (a) ANN, (b) 1D CNN, (c) 2D CNN, and (d) 2D CNN-LSTM

Based on the previous graphs, comparing accuracy across SNR levels, we observed that integrating feature extraction techniques with DL models significantly improved PQD detection results. To further investigate which technique performs better, we evaluated recall and precision metrics for models using CWT and STFT across PQD types, the results are presented in Figure 14. Figure 14(a) presents the results for the 2D CNN model, while Figure 14(b) shows the corresponding performance for the fused 2D CNN-LSTM model.

The graphs illustrate the comparative performance of CWT and STFT feature extraction techniques across 2D CNN and fused 2D CNN-LSTM. In the 2D CNN (left graph), CWT consistently achieves higher precision and recall metrics across all PQD classes, particularly excelling in sag and normal, with near-perfect values. Meanwhile, STFT shows lower precision and recall for Swell and Interruption, indicating its limited ability to accurately classify these disturbances. In the Fused 2D CNN-LSTM architecture (right graph), the fusion significantly enhances the overall metrics for both techniques, with CWT achieving perfect precision and recall for sag and harmonic. However, STFT continues to lag behind CWT, especially in recall for swell and harmonic, where the metrics drop below 0.95. These findings validate the robustness of CWT in handling noisy conditions and extracting meaningful features for PQD classification, making it the superior feature extraction technique across all architectures and metrics. To further evaluate the performance of the proposed models, a comparative analysis was conducted against conventional ML methods reported in the literature [31]. The comparison results are summarized in Table 4.

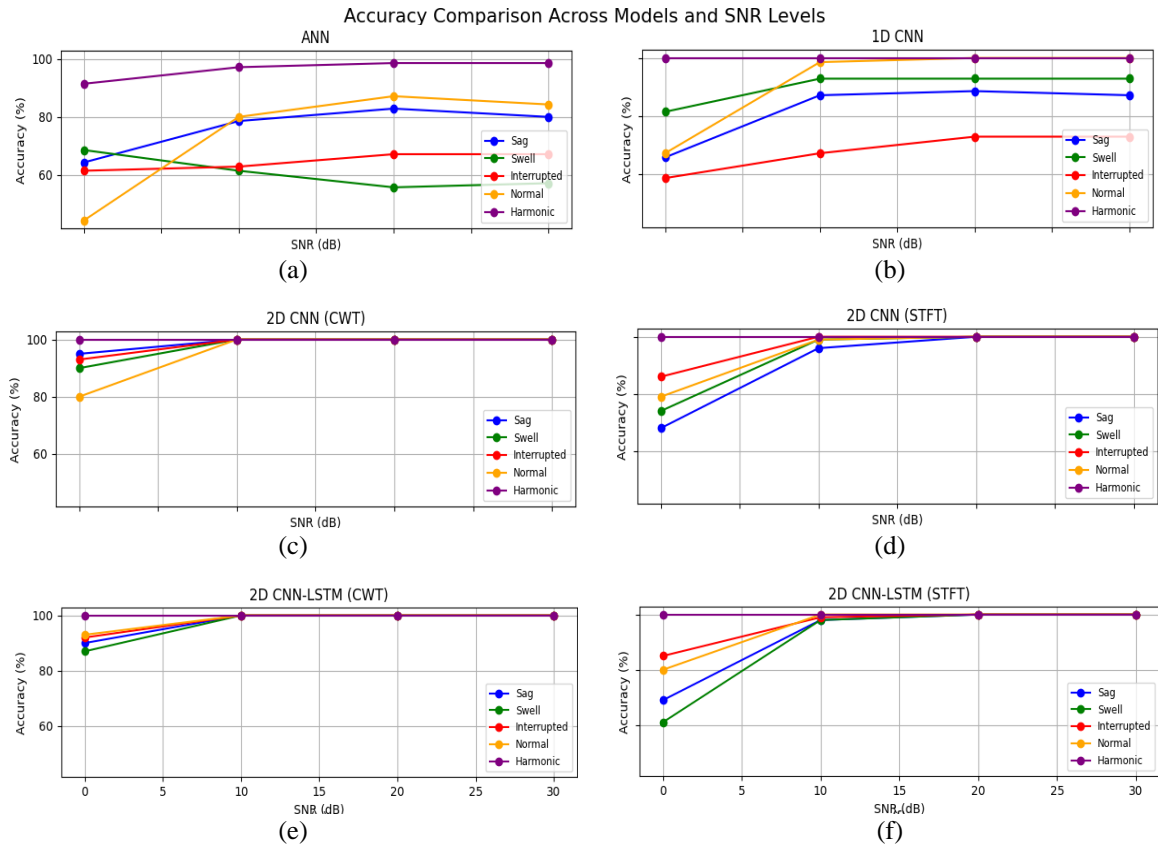


Figure 13. Accuracy comparison across SNR levels for different PQD types: (a) ANN, (b) 1D CNN, (c) 2D CNN with CWT, (d) 2D CNN with STFT, (e) 2D CNN-LSTM with CWT, and (f) 2D CNN-LSTM with STFT

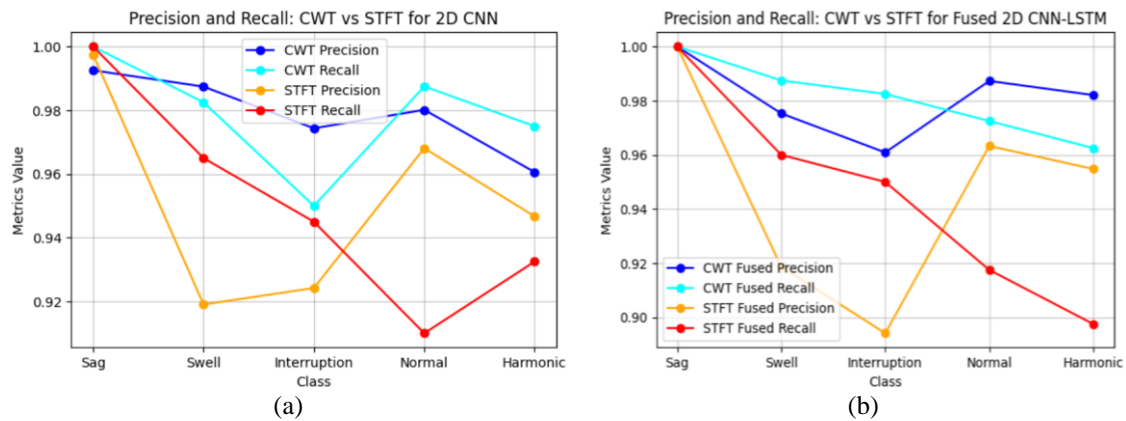


Figure 14. Precision and recall comparison for: (a) 2D CNN and (b) fused 2D CNN-LSTM models using CWT and STFT

The table clearly illustrates that DL architectures significantly outperform traditional ML methods like SVM and RF [31]. While SVM and RF achieve respectable accuracies under noiseless conditions (92.5% and 88.7%, respectively), their performance is limited and highly sensitive to variations in the signal. In contrast, the proposed 1D CNN, 2D CNN, and fused 2D CNN-LSTM models not only achieve higher accuracy overall but also maintain robustness under noisy conditions (0–30 dB). The progression from 1D CNN (95.93%) to 2D CNN (97.90%), and finally to fused 2D CNN-LSTM (98.75%), demonstrates the added benefit of feature extraction through time–frequency representations and temporal modeling. These results confirm that DL models are more effective, resilient, and reliable for PQD classification across different environments.

Table 4. DL vs ML accuracy for PQD classification

Method	Feature type	Disturbances considered	Accuracy (%)	Environment	Remarks
SVM [31]	Statistical features	Sag, swell, interruption, harmonic, and transient.	92.5	Noiseless	Limited robustness.
RF [31]	Statistical features	Sag, swell, interruption, harmonic, and transient.	88.7	Noiseless	lacks temporal learning.
1D CNN (proposed)	Raw signal	Sag, swell, interruption, harmonic, and normal.	95.93	Noisy (0–30 dB)	Learns directly from raw data; higher accuracy under noise.
2D CNN (proposed)	CWT/STFT images	Sag, swell, interruption, harmonic, and normal.	97.90	Noisy (0–30 dB)	Robust time–frequency feature extraction; CWT superior to STFT.
Fused 2D CNN-LSTM (proposed)	CWT/STFT images+temporal modeling	Sag, swell, interruption, harmonic, and normal.	98.75	Noisy (0–30 dB)	Best overall; combines spatial–temporal learning with high noise resilience.

In summary, the findings confirm that DL models outperform traditional ML methods like SVM and RF in handling noisy conditions. Incorporating feature extraction techniques significantly enhances model accuracy, and feeding raw signals without extraction can lead to lower reliability. Precision and recall comparisons further reinforce the superiority of CWT over STFT. CWT consistently achieves higher precision, minimizing false positives, particularly for PQD types like Sag and harmonic, while recall values showcase its ability to accurately identify true positives across all classes, even under noisy conditions. The fused 2D CNN-LSTM architecture amplifies these advantages, with CWT achieving near-perfect precision and recall for most PQD types, including sag, swell, and harmonic. Based on the results, 2D CNN-LSTM with CWT is the most effective model, showcasing unparalleled accuracy, robustness, and reliability across all noise levels and PQD types.

7. MICROGRID FAULT DETECTION SIMULATION

To validate the previous findings, a microgrid was simulated in Simulink, where various faults were introduced to replicate power system disturbances. These disturbances were characterized by different magnitudes and durations, defined by different start and end times. The simulated circuit and corresponding signals are depicted in the Figures 15 and 16. Based on the findings presented in the previous section, the CWT demonstrated superior performance over the STFT in feature extraction. Therefore, CWT was employed in this section to preprocess microgrid signals effectively.

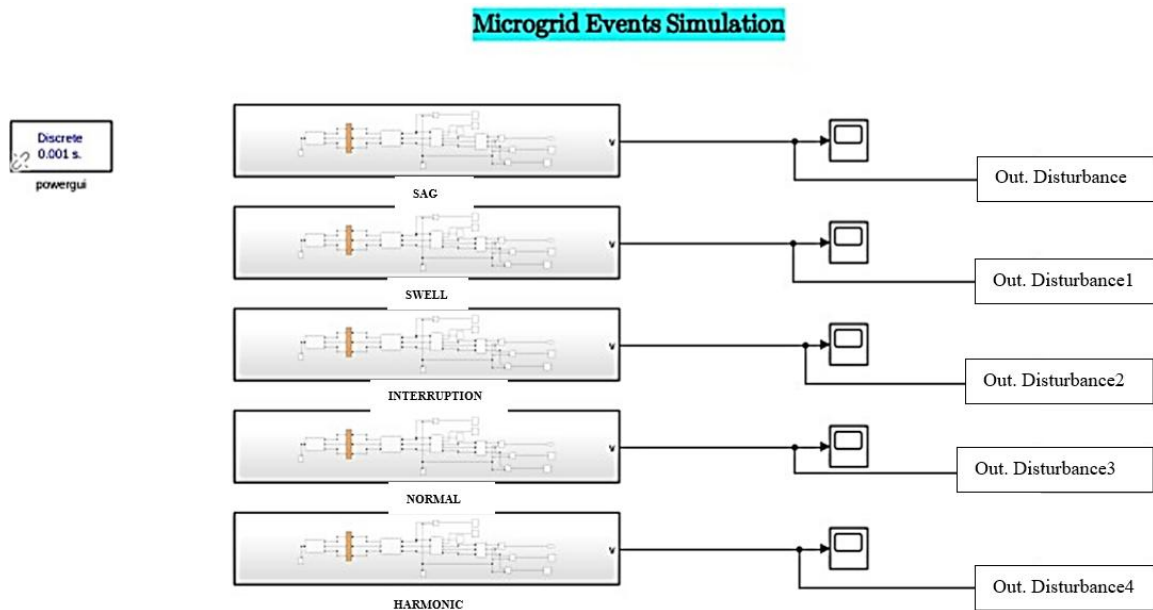


Figure 15. Simulink circuit of MG disturbances

To further analyze the classification performance, Figures 19 and 20 present the predictions for each segmented signal. Figure 19 shows the classification results obtained using the CNN model, where errors in distinguishing certain disturbances are evident. Figure 20, on the other hand, demonstrates the improved performance of the Fused 2D CNN-LSTM model, which correctly classifies all segments with minimal misclassification.

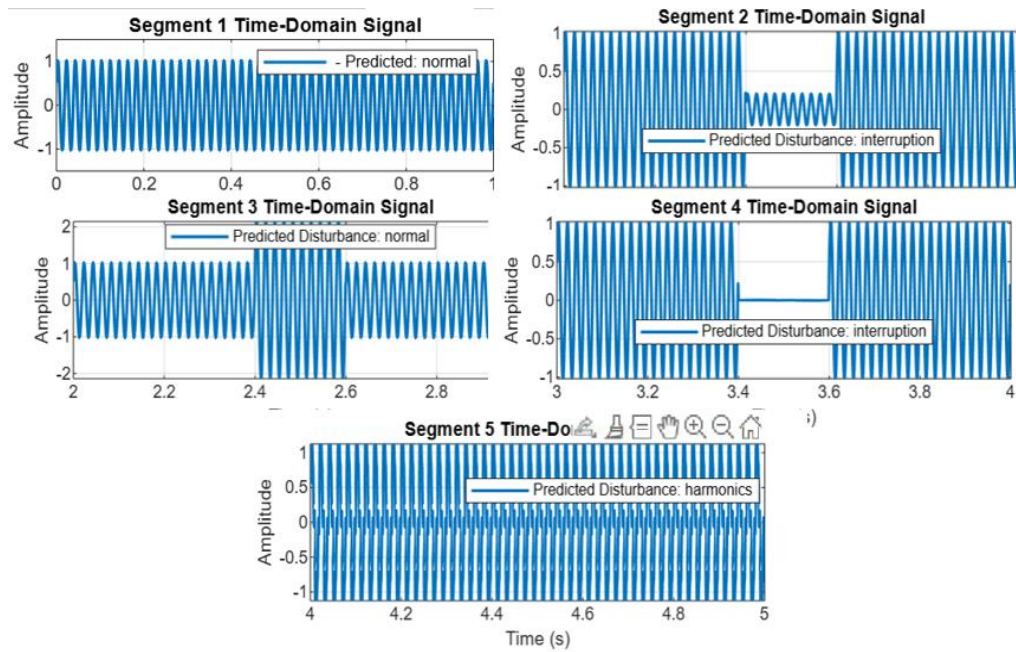


Figure 19. CNN classification of combined event signal

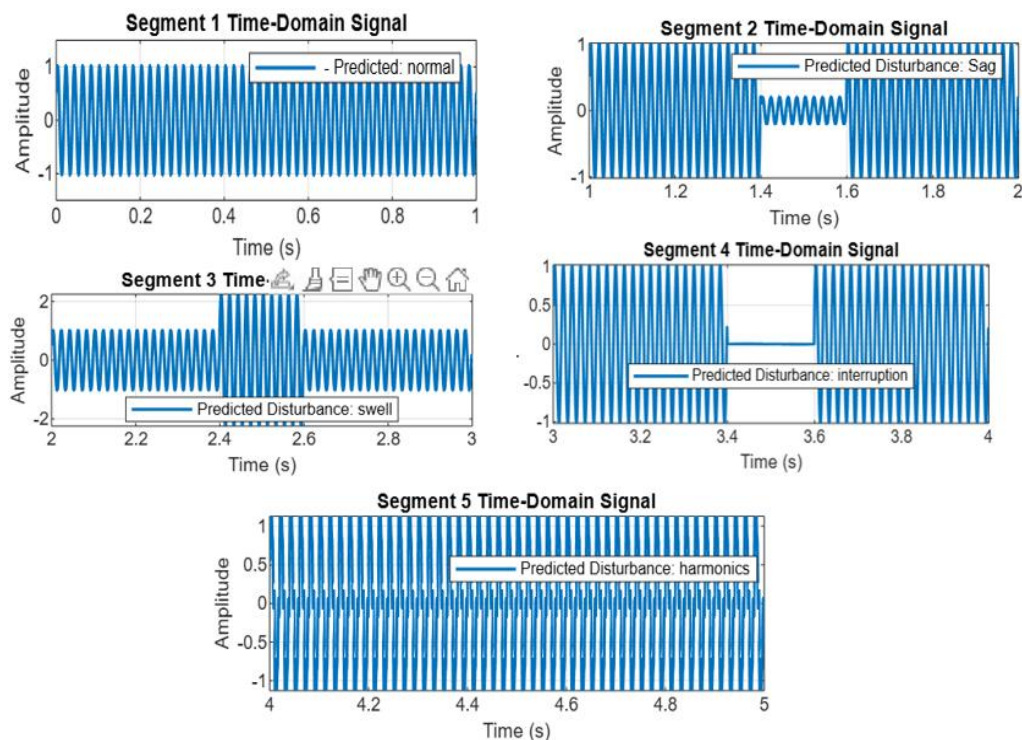


Figure 20. Fused 2D-CNN-LSTM classification of combined event signal

These results confirm that integrating CNN with LSTM enhances classification accuracy, particularly for complex, time-varying disturbances. The ability of the fused model to extract spatial features through CNN and capture temporal dependencies through LSTM makes it a more reliable approach for PQD classification.

8. CONCLUSION

This study investigated the classification of PQD by evaluating various ML and DL models under different noise conditions. Through an ablation study, the work highlighted the relative effectiveness of feature extraction techniques and modeling approaches, showing that the fused 2D CNN-LSTM architecture outperformed other models and that the CWT provided more robust feature representations compared to STFT and raw signals. The findings were further validated through microgrid simulations, confirming the robustness and practical relevance of the proposed methodology under realistic operating conditions. Overall, this work highlights the benefits of combining effective feature extraction with advanced DL architectures for PQD classification, offering guidance for future research and practical implementations in smart grid monitoring and PQ analysis. Future work may focus on deploying these methods on hardware for real-time monitoring, enhancing the efficiency and dependability of disturbance detection in operational systems.

FUNDING INFORMATION

Authors state no funding involved.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Bouchra Feriel Khaldi	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Fatma Zohra Dekhandji	✓	✓		✓		✓		✓		✓	✓	✓	✓	✓
Abdelmadjid Recioui		✓		✓		✓				✓	✓	✓	✓	✓

C : Conceptualization

M : Methodology

So : Software

Va : Validation

Fo : Formal analysis

I : Investigation

R : Resources

D : Data Curation

O : Writing - Original Draft

E : Writing - Review & Editing

Vi : Visualization

Su : Supervision

P : Project administration

Fu : Funding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY




The data that support the findings of this study are available from the corresponding author, [initials: BFK], upon reasonable request.

REFERENCES




- [1] Y. Naderi, S. H. Hosseini, S. G. Zadeh, B. Mohammadi-Ivatloo, J. C. Vasquez, and J. M. Guerrero, "An overview of power quality enhancement techniques applied to distributed generation in electrical distribution networks," *Renewable and Sustainable Energy Reviews*, vol. 93, pp. 201–214, 2018, doi: 10.1016/j.rser.2018.05.013.
- [2] G. S. Chawda *et al.*, "Comprehensive Review on Detection and Classification of Power Quality Disturbances in Utility Grid with Renewable Energy Penetration," *IEEE Access*, vol. 8, pp. 146807–146830, 2020, doi: 10.1109/ACCESS.2020.3014732.
- [3] E. G. Rodríguez, E. R. Archundia, J. A. G. Gnechi, O. I. C. Reyes, J. C. O. Rojas, and A. M. Patiño, "Detection and extraction of optimal features from power quality disturbances based on wavelet coefficient reconstruction and noise-robust methods," *Computers and Electrical Engineering*, vol. 127, part. B, 2025, doi: 10.1016/j.compeleceng.2025.110615.
- [4] M. S. Priyadarshini and M. Sushama, "Modeling of Power Quality Disturbances using Parametric Equations in MATLAB," *International Journal of Engineering Research in Electrical and Electronic Engineering (IJEREEE)*, vol. 1, no. 8, 2015.
- [5] B. F. Khaldi, F. Z. Dekhandji, and A. Recioui, "Power Quality Disturbances: A review of Detection, Classification, Optimization,

- and Mitigation Techniques,” *Algerian Journal of Signals and Systems*, vol. 9, no. 4, pp. 261–286, 2024, doi: 10.51485/ajss.v9i4.243.
- [6] G. Beichang, W. Xiao, L. Yingying, Z. Zhang, and Y. Fan, “Influence Research of Renewable Energy Application on Power Quality Detection,” in *IOP Conference Series: Earth and Environmental Science*, vol. 168, no. 1, 2018, doi: 10.1088/1755-1315/168/1/012033.
- [7] S. A. Deokar and L. M. Waghmare, “Integrated DWT-FFT approach for detection and classification of power quality disturbances,” *International Journal of Electrical Power and Energy Systems*, vol. 61, pp. 594–605, 2014, doi: 10.1016/j.ijepes.2014.04.015.
- [8] A. Yilmaz and G. Bayrak, “Real-Time Disturbance Detection Using STFT Method in Microgrids,” *Academic Perspective Procedia*, vol. 2, no. 3, pp. 1115–1121, 2019, doi: 10.33793/acperpro.02.03.124.
- [9] N. M. Khoa and L. V. Dai, “Detection and classification of power quality disturbances in power system using modified-combination between the stockwell transform and decision tree methods,” *Energies*, vol. 13, no. 14, 2020, doi: 10.3390/en13143623.
- [10] F. R. Zaro, “Power Quality Disturbances Detection and Classification Rule-Based Decision Tree,” *Wseas Transactions on Signal Processing*, vol. 17, pp. 22–27, 2021, doi: 10.37394/232014.2021.17.3.
- [11] V. Pandya, S. Agarwal, O. P. Mahela, and S. Choudhary, “Recognition of Power Quality Disturbances Using Hybrid Algorithm Based on Combined Features of Stockwell Transform and Hilbert Transform,” in *2020 IEEE International Students’ Conference on Electrical, Electronics and Computer Science, SCEECS 2020*, 2020, doi: 10.1109/SCEECS48394.2020.4.
- [12] T. Ravi and K. S. Kumar, “Detection and Classification of Power Quality Disturbances Using Stock Well Transform and Improved Grey Wolf Optimization-Based Kernel Extreme Learning Machine,” *IEEE Access*, vol. 11, pp. 61710–61727, 2023, doi: 10.1109/ACCESS.2023.3286308.
- [13] H. I. Hussein, A. Alazawi, A. Rodríguez, and F. Muñoz, “Performance Evaluation of ST-Based Methods for Simulating and Analyzing Power Quality Disturbances,” *International Journal on Smart Sensing and Intelligent Systems*, vol. 16, no. 1, 2023, doi: 10.2478/ijssis-2023-0011.
- [14] Z. Cai, F. Ning, W. Li, and T. A. Gulliver, “Power quality signal analysis for the smart grid using the Hilbert-Huang Transform,” in *IEEE Pacific RIM Conference on Communications, Computers, and Signal Processing - Proceedings*, pp. 296–301, 2013, doi: 10.1109/PACRIM.2013.6625492.
- [15] S. Naderian and A. Salemnia, “Detection and classification of power-quality events using discrete Gabor transform and Support Vector Machine,” in *6th Annual International Power Electronics, Drive Systems, and Technologies Conference, PEDSTC 2015*, pp. 544–549, 2015, doi: 10.1109/PEDSTC.2015.7093333.
- [16] M. S. S. M. Basir, M. I. Abdullahi, and A. R. Abdullah, “Window Optimisation of Power Quality Signal Detection using Gabor Transform,” *ASM Science Journal*, vol. 14, pp. 1–10, 2021, doi: 10.32802/asmscj.2020.596.
- [17] J. B. V. Reddy, P. K. Dash, R. Samantaray, and A. K. Moharana, “Fast tracking of power quality disturbance signals using an optimized unscented filter,” *IEEE Transactions on Instrumentation and Measurement*, vol. 58, no. 12, pp. 3943–3952, 2009, doi: 10.1109/TIM.2009.2020835.
- [18] P. K. Sundaram and G. Girishkumar, “Power Quality Disturbance Classification based on Kalman filter and Adaptive Neural Fuzzy Inference System (ANFIS),” *Journal of Physics: Conference Series*, vol. 1831, no. 1, 2021, doi: 10.1088/1742-6596/1831/1/012009.
- [19] K. R. Reddy, P. R. Reddy, and P. Sujatha, “Power Quality Classification of disturbances using Discrete Wavelet Packet Transform (DWPT) with Adaptive Neuro-Fuzzy System,” *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, vol. 12, no. 3, pp. 4892–4902, 2021, doi: 10.17762/turcomat.v12i3.1995.
- [20] I. A. Channa, D. Li, F. H. Dahri, G. E. M. Abro, and F. Zahid, “A Novel Deep Learning Model for Classifying Power Quality Problems in PV-integrated Microgrids using CNN-LSTM,” in *1st International Conference on Innovative Engineering Sciences and Technological Research, ICIESTR 2024 - Proceedings*, 2024, doi: 10.1109/ICIESTR60916.2024.10798309.
- [21] S. Cen, D. O. Kim, and C. G. Lim, “A fused CNN-LSTM model using FFT with application to real-time power quality disturbances recognition,” *Energy Science and Engineering*, vol. 11, no. 7, pp. 2267–2280, 2023, doi: 10.1002/ese3.1450.
- [22] C. I. Garcia, F. Grasso, A. Luchetta, M. C. Piccirilli, L. Paolucci, and G. Talluri, “A comparison of power quality disturbance detection and classification methods using CNN, LSTM and CNN-LSTM,” *Applied Sciences (Switzerland)*, vol. 10, no. 19, pp. 1–22, 2020, doi: 10.3390/app10196755.
- [23] A. Khan, A. Sohail, U. Zahoora, and A. S. Qureshi, “A survey of the recent architectures of deep convolutional neural networks,” *Artificial Intelligence Review*, vol. 53, no. 8, pp. 5455–5516, 2020, doi: 10.1007/s10462-020-09825-6.
- [24] Purwono, A. Ma’arif, W. Rahmانيar, H. I. K. Fathurrahman, A. Z. K. Frisky, and Q. M. U. Haq, “Understanding of Convolutional Neural Network (CNN): A Review,” *International Journal of Robotics and Control Systems*, vol. 2, no. 4, pp. 739–748, 2022, doi: 10.31763/ijrcs.v2i4.888.
- [25] L. Alzubaidi *et al.*, “Review of deep learning: concepts, CNN architectures, challenges, applications, future directions,” *Journal of Big Data*, vol. 8, no. 1, p. 53, Mar. 2021, doi: 10.1186/s40537-021-00444-8.
- [26] C3.ai, “Root Mean Square Error (RMSE),” [Online]. Available: <https://c3.ai/glossary/root-mean-square-error>. (Accessed: Sep. 10, 2024).
- [27] Shipra Saxena, “LSTM | Introduction to LSTM | Long Short-Term Memory Algorithms,” *Analytics Vidhya*, 2023, [Online]. Available: <https://www.analyticsvidhya.com/blog/2021/03/introduction-to-long-short-term-memory-lstm/>. (Accessed: Jun. 8, 2023).
- [28] GeeksforGeeks “Deep Learning | Introduction to Long Short-Term Memory,” [Online]. Available: <https://www.geeksforgeeks.org/deep-learning-introduction-to-long-short-term-memory/>. (Accessed: Jun. 8, 2023).
- [29] S. Helonde, “Mathematical Analysis of Power Quality Disturbances in Electrical Distribution Systems: Detection, Classification, and Mitigation Strategies,” *Panamerican Mathematical Journal*, vol. 34, no. 1, pp. 81–94, 2024, doi: 10.52783/pmj.v34.i1.908
- [30] AudioLabs Erlangen, “STFT Window Function”. [Online]. Available: https://www.audiolabs-erlangen.de/resources/MIR/FMP/C2/C2_STFT-Window.html. (Accessed: Dec. 2024).
- [31] C. V. Venkatesh and S. D. V. S. S. Siva, “Techniques used for Detection of Power Quality Events – a Comparative Study,” in *16th National Power Systems Conference*, no. 3, pp. 307–312, 2010.




BIOGRAPHIES OF AUTHORS

Bouchra Ferial Khaldi    received the Bachelor degree in Electrical Engineering and Electronics from the Institute of Electrical Engineering and Electronics (IGEE), University of Boumerdes, Algeria, and the Master degree in Power Engineering from the same institute in 2021. She is currently pursuing the Ph.D. degree at the same institution, focusing on power quality in smart grids and smart metering systems. Her research interests include artificial intelligence, machine learning, deep learning, and their applications in power quality analysis. She can be contacted at email: b.khaldi@univ-boumerdes.dz.



Fatma Zohra Dekhandji    holds the position of associate professor at the Institute of Electrical Engineering and Electronics, University of Boumerdes, Algeria. In 2016, she obtained a Ph.D. degree in Electrical Engineering with a focus on Electrical and Electronic Engineering from the Institute of Electrical Engineering and Electronics at the University of Boumerdes. She also has a Master (Magister) degree in Electronic System Engineering, which she earned in 2007 at the University of Boumerdes, Institute of Electrical Engineering and Electronics. She completed her engineering studies at the University of Boumerdes' Institute of Electrical Engineering and Electronics in June 2003. Since January 2008, she has served as the leader of a research team at Laboratory Signals and Systems. Some of her areas of interest include power quality, electric motors, smart grids, and artificial intelligence for power systems. She can be contacted at email: fzdekhandji@univ-boumerdes.dz.



Abdelmadjid Recioui    is a full Professor at the Institute of Electrical Engineering and Electronics University of Boumerdes, Algeria. He obtained a Ph.D. degree in electrical and electronic engineering option telecommunications from the Institute of Electrical Engineering and Electronics, University of Boumerdes in 2011. He holds also Master (Magister) degree in Electronic System Engineering which has been achieved at the Institute of Electrical Engineering and Electronics, University of Boumerdes in 2006. In June 2002, he finished his engineering studies at the institute of Electrical Engineering and Electronics, University of Boumerdes. He is a research director at the laboratory signals and systems from January 2008 to present. His research interests include: antennas, wireless communication systems, antenna array synthesis and design, capacity enhancement, system optimization, smart antennas, power system protection, power system optimization, and power system communications. He can be contacted at email: a_recioui@univ-boumerdes.dz.