

# Power Quality Disturbance Classification Using Convolutional Neural Networks

Ana Camila Ferreira Mamede  
Uberlândia, Brasil  
anacamilamamede@gmail.com

**Abstract**— With the increasing penetration of distributed generation and renewable energy sources in distribution grids, power quality has become a growing concern for both utilities and consumers. Effective mitigation of these issues depends not only on the efficiency of corrective measures but also on the system's ability to accurately detect and classify Power Quality Disturbances (PQDs). This paper presents a Convolutional Neural Network (CNN) model capable of classifying seventeen types of PQDs using only a single cycle of the sinusoidal signal, without the need for manual feature extraction. The proposed model achieves 97% overall accuracy, enabling fast and reliable classification.

**Keywords**— Power quality, power quality disturbances, convolutional neural networks, deep learning, signal classification, distribution grid.

## I. INTRODUCTION

The increasing integration of renewable energy resources, along with factors such as capacitor bank switching, nonlinear loads, atmospheric discharges, and the activation of large motors, can lead to various Power Quality Disturbance (PQD) issues. These disturbances can cause transformer overheating, reduced product quality, decreased equipment lifespan, and false circuit breaker tripping, ultimately affecting the safety and stability of power systems [1]. As a result, the detection, classification, and mitigation of PQDs are of significant importance in modern power networks.

Recently, Convolutional Neural Networks (CNNs) have attracted significant interest in PQD classification. Conventional methods typically depended on manual feature extraction and were highly sensitive to noise, which restricted their performance. In contrast, CNNs can automatically learn hierarchical features from raw data, making them a robust solution to address these limitations.

Zhu et al. [2] explored the use of deep learning for power grid disturbance classification by applying an image embedding technique to convert time-series event data into two-dimensional images. These images were then classified using CNNs, demonstrating superior performance over traditional machine learning approaches such as Support Vector Machines and Decision Trees.

Other researchers have proposed deep learning models that utilize raw PQD data without the need for pre-processing or manual feature selection. These methods improve classification accuracy by leveraging CNNs' ability to detect intricate patterns within the data [3]. Additionally, to account for temporal dependencies in PQD signals, a hybrid model combining Temporal Convolutional Networks (TCN) and CNNs was developed, effectively capturing both spatial and temporal features, leading to enhanced classification accuracy [4].

Another innovative approach integrates variational mode decomposition (VMD) with CNNs, reducing the need for extensive training datasets while maintaining high classification performance [5]. Similarly, deep convolutional networks have been employed to automate PQD classification, eliminating the requirement for manual feature extraction and allowing the model to adapt to diverse disturbance patterns [6].

Further advancements include CNN models trained on synthetic PQD signals containing both single and multiple disturbances. These models have demonstrated robustness in handling complex PQD scenarios, including simultaneous disturbances [7]. In addition to CNNs, other deep learning architectures have been explored. For example, a multi-modal parallel feature extraction and classification model was proposed to capture both temporal and spatial PQD characteristics, further improving classification accuracy [8]. Moreover, the combination of S-transform and CNNs has been investigated, where the S-transform extracts time-frequency features, which are then classified using a CNN. This method has proven effective in accurately distinguishing various types of PQDs [9].

These advancements highlight the transformative impact of CNNs and deep learning in PQD classification. By automating feature extraction and efficiently handling complex disturbance patterns, these models offer a promising avenue for improving power quality monitoring and ensuring electrical system reliability.

This study utilizes an open-access database [10] containing seventeen types of PQDs to develop a CNN model that directly processes raw signals. The proposed model achieves 97% classification accuracy using only a single sine cycle.

This paper is structured as follows: Section II outlines the types of power quality disturbances classified by the CNN model. Section III describes the CNN training process and evaluation metrics. Section IV presents the training performance results, while Section V details the testing performance. Section VI provides an in-depth discussion of the results, and Section VII concludes the study.

## II. POWER QUALITY DISTURBANCES

With the substantial increase in distributed generation across the power system, numerous power quality issues must be addressed to ensure grid stability and performance. Power quality disturbances can occur as single events or as combined disturbances when two events happen simultaneously in the same signal. In this work, seventeen types of power quality disturbances (PQDs) are considered—nine as single events and eight as combined disturbances. These disturbances are

described in Table I, based on the definitions provided in [11, 12], and an example of each is presented in Fig. 1.

TABLE I. POWER QUALITY DISTURBANCES DESCRIPTION

| Disturbance                      | Description                                                                                                                                                                                             |
|----------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Pure Sinusoidal                  | An ideal voltage waveform with a smooth, continuous sine wave shape, free from distortions such as harmonics, transients, or notches. It ensures efficient power transfer and minimal electrical noise. |
| Transient                        | A sudden, short-duration voltage or current disturbance caused by lightning, switching, or faults, leading to equipment damage or misoperation.                                                         |
| Harmonics                        | Distortion in the voltage or current waveform due to nonlinear loads, causing overheating and inefficiencies in electrical equipment.                                                                   |
| Interruption                     | Complete loss of voltage or current for a period, typically caused by faults, switchgear operations, or supply failures.                                                                                |
| Sag (Voltage Dip)                | A temporary reduction in voltage magnitude (typically 10-90% of nominal voltage) lasting from milliseconds to a few seconds.                                                                            |
| Swell (Voltage Swell)            | A temporary voltage increase (typically 110-180% of nominal voltage) lasting from milliseconds to a few seconds, often due to load shedding or capacitor switching.                                     |
| Flicker                          | Rapid, repetitive voltage fluctuations that cause visible light flickering, usually caused by fluctuating loads (e.g., welding machines, arc furnaces).                                                 |
| Notch                            | A short-duration distortion in the voltage waveform caused by commutation in power electronic devices.                                                                                                  |
| Oscillatory Transient            | A high-frequency voltage or current fluctuation lasting a short duration, typically caused by switching operations or capacitor bank energization.                                                      |
| Harmonics with Sag               | Harmonics present during a voltage sag, exacerbating waveform distortion and power system instability.                                                                                                  |
| Harmonics with Swell             | Harmonics present during a voltage swell, increasing the risk of equipment malfunction and insulation breakdown.                                                                                        |
| Sag with Harmonics               | A voltage sag that occurs along with harmonic distortion, increasing the risk of equipment misoperation.                                                                                                |
| Sag with Oscillatory Transient   | A voltage sag followed by a transient oscillation, often caused by switching actions or fault clearing.                                                                                                 |
| Swell with Harmonics             | A voltage swell occurring with harmonic distortion, which may cause additional stress on sensitive equipment.                                                                                           |
| Swell with Oscillatory Transient | A voltage swell followed by an oscillatory transient, often resulting from capacitor switching or sudden load disconnection.                                                                            |
| Flicker with Sag                 | Flicker combined with a voltage drop (sag), leading to noticeable light flickering and a temporary reduction in voltage magnitude.                                                                      |
| Flicker with Swell               | Flicker combined with a voltage rise (swell), causing variations in brightness and potential equipment stress.                                                                                          |

While some PQDs—such as flicker—are typically characterized by low-frequency variations best observed over multiple cycles, this study explores the feasibility of identifying them from waveform distortions within a single cycle. The proposed CNN model is trained to capture subtle features and patterns in the voltage waveform. The waveforms in Fig. 1 were randomly extracted from the database [10], used for training and validating the model.

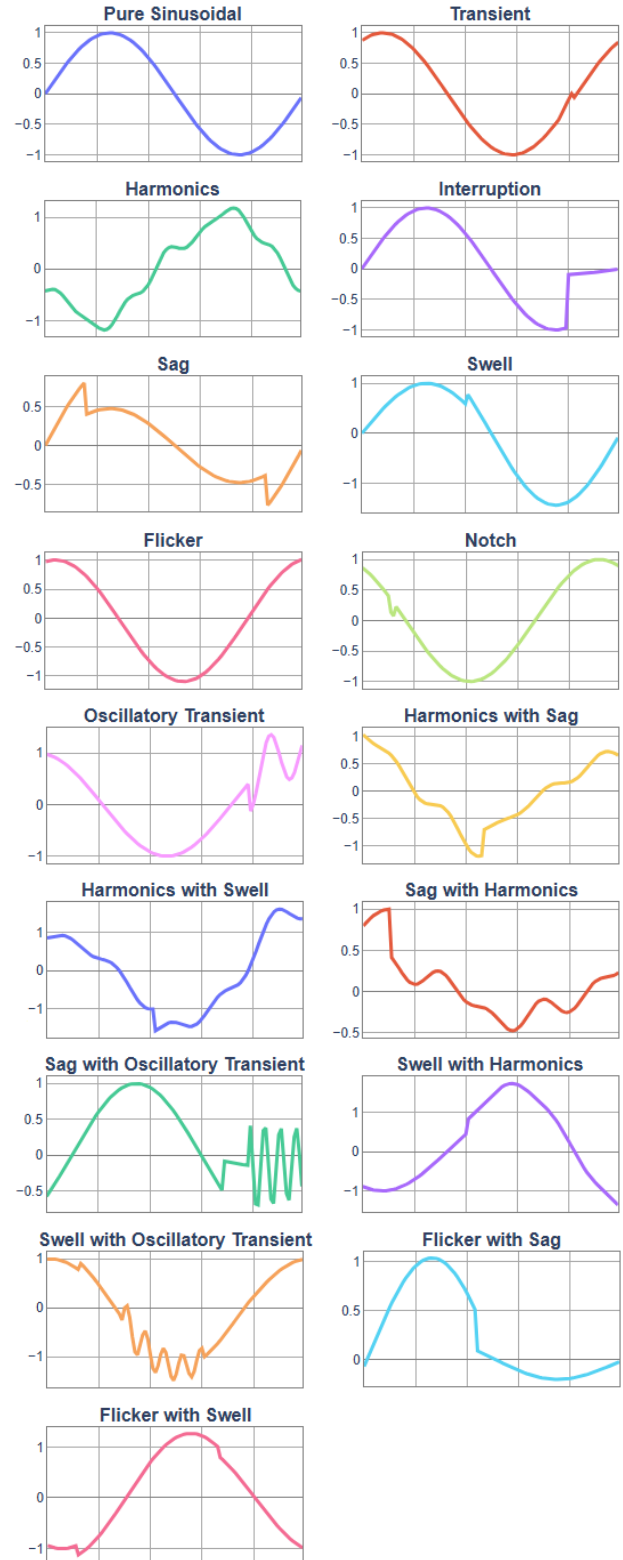


Fig. 1. Signals considered for the evaluation of the proposed methodology.

### III. CNN MODEL TRAINING PROCESS AND EVALUATION METRICS

The process of training the convolutional neural network (CNN) model follows the standard workflow used in machine learning, consisting of six key steps: data collection, data preprocessing and cleaning, model selection, model training, model evaluation, and model tuning and optimization.

The first step, data collection, involves finding a suitable database containing power quality disturbance signals. In this study, a publicly available dataset [10] was used. This dataset consists of 17,000 signals, with 1,000 signals per class, and each signal contains 100 samples. The amplitude of each signal is already scaled between -1 and 1, eliminating the need for additional data cleaning or normalization.

A key objective of this work is to train the CNN directly on raw signals, without manually extracting features or performing additional transformations. Instead, the selected CNN model takes the raw 100-sample signals as input, allowing the network to learn relevant patterns automatically.

To ensure robust training and evaluation, the dataset is split into three subsets:

- Training set: Used to train the model.
- Validation set: Used to assess the model's performance during training, helping to fine-tune hyperparameters and prevent overfitting.
- Testing set: Used to evaluate the final model's performance on unseen data, ensuring it generalizes well to new inputs.

By following this approach, the CNN model is validated using data it has never encountered during training, which provides a more reliable estimate of real-world performance.

The architecture of the CNN model is illustrated in Fig. 2. The entire implementation was carried out using Python, with the TensorFlow library utilized to construct and train the convolutional neural network.

```
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Conv1D, MaxPooling1D, Flatten, Dense,
Dropout, BatchNormalization

# Define CNN model
model = Sequential([
    Conv1D(filters=32, kernel_size=3, activation='relu', input_shape=(100, 1)),
    BatchNormalization(),
    MaxPooling1D(pool_size=2),

    Conv1D(filters=64, kernel_size=3, activation='relu'),
    BatchNormalization(),
    MaxPooling1D(pool_size=2),

    Conv1D(filters=96, kernel_size=3, activation='relu'),
    BatchNormalization(),
    MaxPooling1D(pool_size=2),

    Conv1D(filters=128, kernel_size=3, activation='relu'),
    BatchNormalization(),
    MaxPooling1D(pool_size=2),
    Dropout(0.3),

    Flatten(),
    Dense(128, activation='relu', kernel_regularizer=l2(0.005)),
    Dropout(0.5),
    Dense(num_classes, activation='softmax') # Multi-class classification
])

# Compile model
model.compile(optimizer='adam', loss='categorical_crossentropy', metrics=['accuracy'])
```

Fig. 2. CNN architecture.

The most common evaluation metrics for classification problems are precision, recall, F1-score, and accuracy. These metrics are derived from the confusion matrix, which consists of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN).

Precision measures the proportion of correctly predicted positive instances among all instances classified as positive. A higher precision indicates fewer false positives, meaning the model is more reliable when it classifies a sample as belonging to a particular class. It is calculated by (1).

$$\text{Precision} = \frac{TP}{TP + FP} \quad (1)$$

Recall (also known as sensitivity or true positive rate) measures the proportion of actual positive instances that were correctly classified, it is given by (2). A high recall means the model captures most of the relevant instances, minimizing false negatives.

$$\text{Recall} = \frac{TP}{TP + FN} \quad (2)$$

F1-score is the harmonic mean of precision and recall, balancing the trade-off between the two. It is particularly useful when dealing with imbalanced datasets, where one class has significantly more samples than another. A high F1-score indicates both strong precision and recall, making it a robust metric for evaluating model performance. It is computed as:

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (3)$$

Accuracy represents the overall proportion of correctly classified instances across all classes, it is calculated by (4).

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (4)$$

These metrics provide a comprehensive assessment of model performance, ensuring that both correct classifications and misclassifications are properly considered. In this work, the CNN model is evaluated using these metrics to measure its effectiveness in classifying power quality disturbances.

### IV. TRAINING PERFORMANCE RESULTS

The dataset is divided into three subsets: 70% for training, 20% for validation, and 10% for testing. This means that for each class, out of 1,000 signals, 700 are used to train the CNN model, 200 are used for validation during training, and 100 are used for final evaluation to assess the model's performance on unseen data.

The training and validation accuracy over epochs is shown in Fig. 3, illustrating how the model's performance improves as it learns from the data. Fig. 4 presents the training and validation loss curves, which indicate how well the model is minimizing the error. A decreasing training loss indicates successful learning, while a stable validation loss suggests good generalization and helps identify potential overfitting. Ideally, the validation loss should closely follow the training loss without a significant gap; a sharp divergence could signal overfitting. The plots show that after epoch 60, both accuracy and loss stabilized, indicating training convergence.

The confusion matrix for the validation set, shown in Fig. 5, provides insights into the model's classification performance across different power quality disturbances, highlighting the most frequently misclassified classes. In a confusion matrix, diagonal elements represent correct classifications, while off-diagonal elements indicate misclassifications. Table II presents the evaluation metrics,

including precision, recall and F1-score, offering a comprehensive assessment of the model's effectiveness.

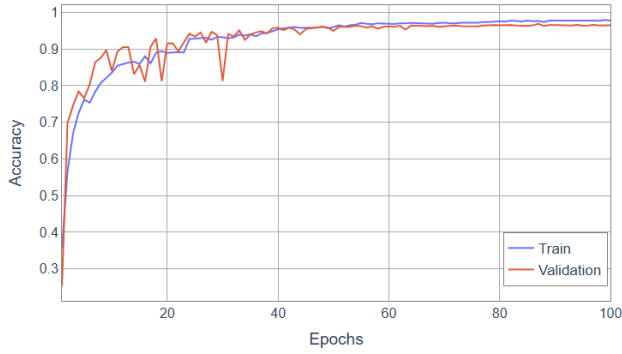


Fig. 3. Training and validation accuracy over epochs.

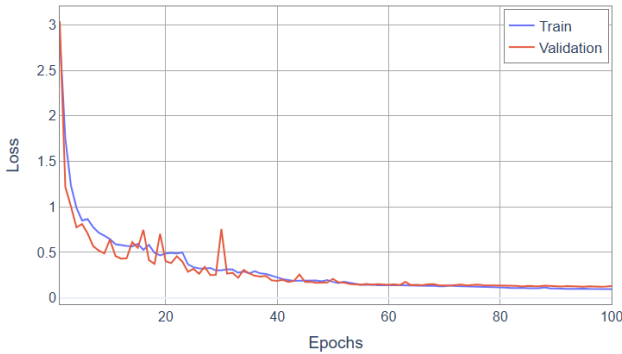


Fig. 4. Training and validation loss over epochs.

TABLE II. EVALUATION METRICS FOR VALIDATION SET

| Class                            | Precision | Recall | F1-score |
|----------------------------------|-----------|--------|----------|
| Pure Sinusoidal                  | 0.98      | 1.00   | 0.99     |
| Transient                        | 0.92      | 1.00   | 0.96     |
| Harmonics                        | 1.00      | 1.00   | 1.00     |
| Interruption                     | 0.99      | 0.99   | 0.99     |
| Sag                              | 0.99      | 0.98   | 0.99     |
| Swell                            | 1.00      | 0.99   | 0.99     |
| Flicker                          | 1.00      | 1.00   | 1.00     |
| Notch                            | 1.00      | 0.99   | 0.99     |
| Oscillatory Transient            | 0.99      | 0.92   | 0.95     |
| Harmonics with Sag               | 1.00      | 1.00   | 1.00     |
| Harmonics with Swell             | 1.00      | 1.00   | 1.00     |
| Sag with Harmonics               | 0.99      | 0.98   | 0.99     |
| Sag with Oscillatory Transient   | 0.99      | 0.99   | 0.99     |
| Swell with Harmonics             | 1.00      | 1.00   | 1.00     |
| Swell with Oscillatory Transient | 1.00      | 0.98   | 0.99     |
| Flicker with Sag                 | 1.00      | 0.99   | 1.00     |
| Flicker with Swell               | 0.99      | 0.99   | 0.99     |

The model achieved an overall accuracy of 99% on the validation set, demonstrating highly reliable performance across all categories.

Despite this near-perfect accuracy, some misclassifications were observed. Specifically, 8% of the "Oscillatory Transient" instances were misclassified as

"Transient." This suggests that the model has some difficulty distinguishing between these two disturbance types, likely because their waveforms can exhibit similar characteristics, especially when the oscillatory transient has a small amplitude. The overlap in their signal patterns may cause the model to confuse them, highlighting a potential area for further improvement in feature extraction or model fine-tuning.

## V. TESTING PERFORMANCE RESULTS

The model's performance on the test set is critical for evaluating its generalization capability. The test set consists of 100 signals per class, totaling 1,700 signals. The confusion matrix, illustrating the relationship between predicted and actual PQD classifications, is shown in Fig. 6. Table III presents the evaluation metrics, including precision, recall, and F1-score. The model achieved an overall accuracy of 96.89%.

TABLE III. EVALUATION METRICS FOR TEST SET

| Class                            | Precision | Recall | F1-score |
|----------------------------------|-----------|--------|----------|
| Pure Sinusoidal                  | 0.98      | 1.00   | 0.99     |
| Transient                        | 0.88      | 1.00   | 0.93     |
| Harmonics                        | 0.98      | 1.00   | 0.99     |
| Interruption                     | 0.99      | 1.00   | 1.00     |
| Sag                              | 0.99      | 0.98   | 0.98     |
| Swell                            | 0.99      | 1.00   | 1.00     |
| Flicker                          | 0.98      | 1.00   | 0.99     |
| Notch                            | 1.00      | 1.00   | 1.00     |
| Oscillatory Transient            | 0.91      | 0.86   | 0.89     |
| Harmonics with Sag               | 0.96      | 0.97   | 0.97     |
| Harmonics with Swell             | 0.99      | 0.99   | 0.99     |
| Sag with Harmonics               | 0.92      | 0.96   | 0.94     |
| Sag with Oscillatory Transient   | 0.97      | 0.91   | 0.94     |
| Swell with Harmonics             | 0.93      | 0.98   | 0.96     |
| Swell with Oscillatory Transient | 1.00      | 0.87   | 0.93     |
| Flicker with Sag                 | 1.00      | 0.96   | 0.98     |
| Flicker with Swell               | 1.00      | 0.99   | 0.99     |

## VI. DISCUSSION OF RESULTS

Figures 5 and 6, along with Tables II and III, demonstrate the model's strong classification performance, with high precision, recall, and F1-scores across most classes. The high classification accuracy and low false negative rates further confirm its reliability. However, certain classes, particularly Oscillatory Transient, Harmonics with Sag, and Sag with Harmonics, exhibit a decline in performance on the test set compared to the validation set. The next subsections provides a detailed analysis of these key findings, based on Figures 5 and 6 and Tables II and III.

### A. Pure Sinusoidal, Transient and Harmonics

- **Pure Sinusoidal:** The model correctly classified all true positives, with only one false positive (a Sag misclassified as Pure Sinusoidal). This performance is slightly better than on the validation set, where one instance was misclassified.

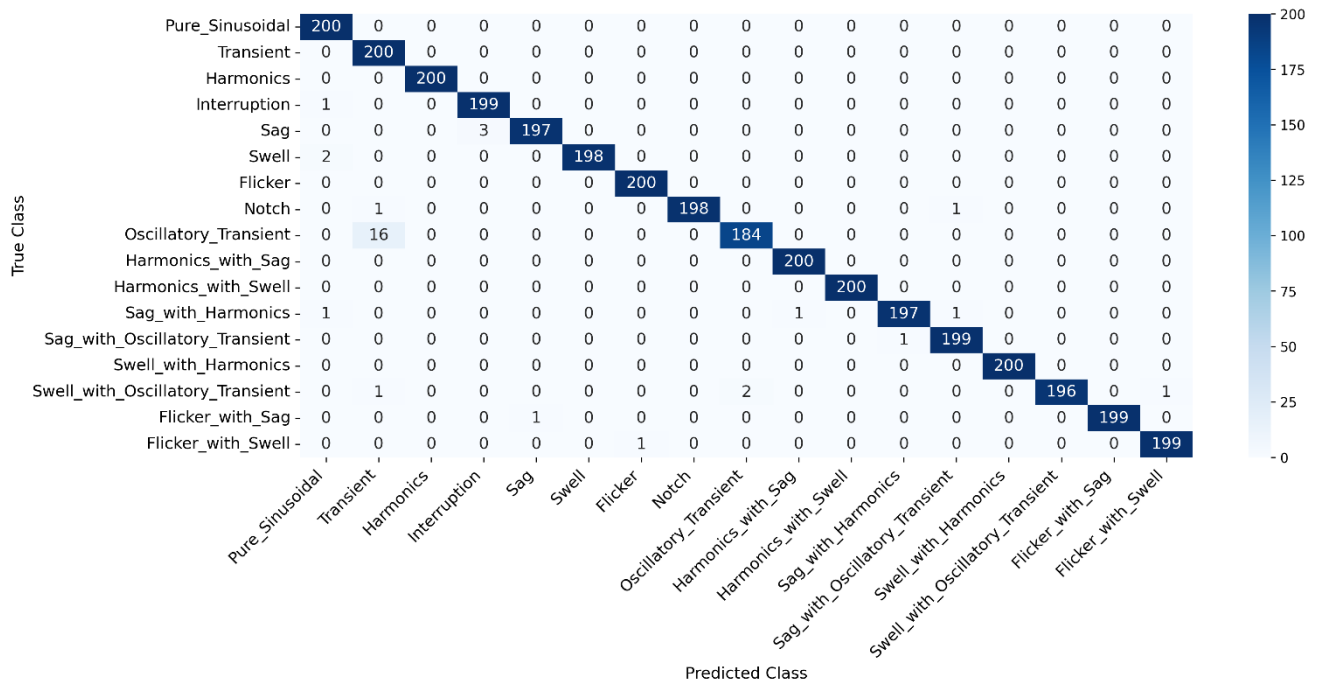


Fig. 5. Confusion matrix of predicted and actual power quality disturbances – Validation set.

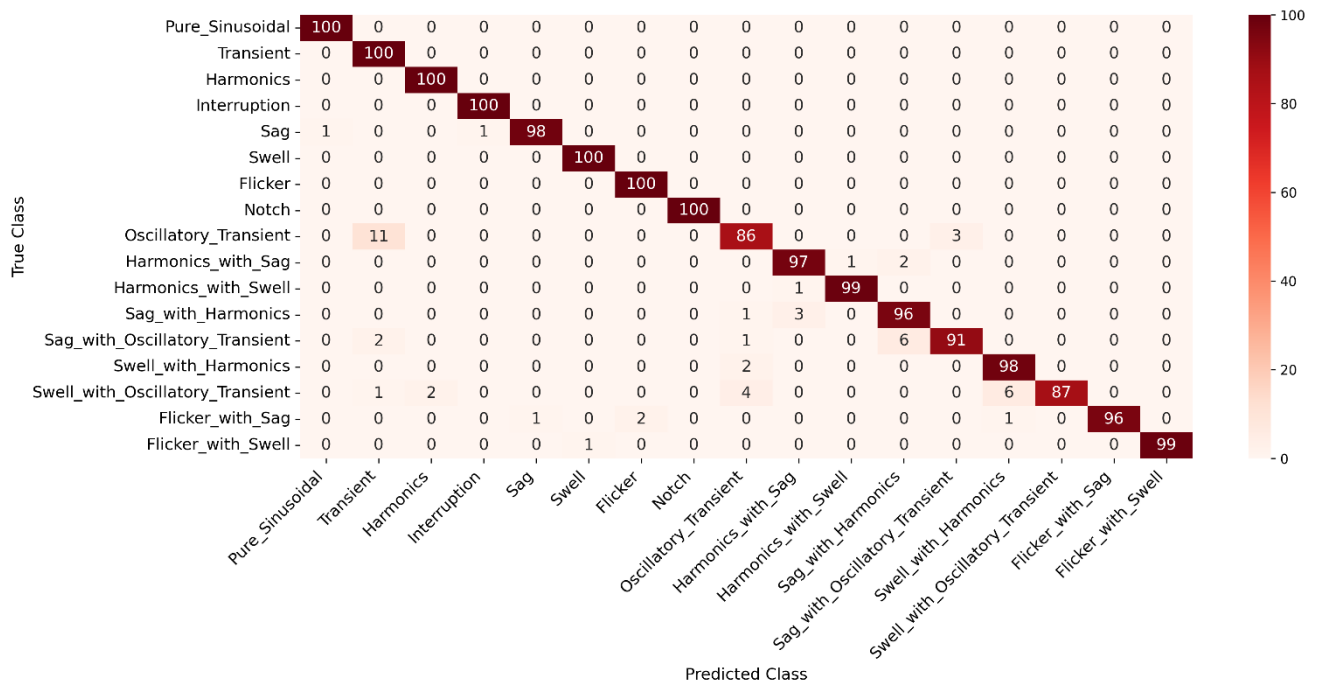


Fig. 6. Confusion matrix of predicted and actual power quality disturbances – Test set

- **Transient:** The recall remains at 1.00, indicating that all transient disturbances were correctly identified. However, precision decreased from 0.92 to 0.88, meaning the model produced more false positives, mistakenly classifying other disturbances as Transients.
- **Harmonics:** Precision slightly dropped to 0.98 due to two Swell with Oscillatory Transient signals being misclassified as Harmonics. Despite this, recall remained at 1.00, ensuring that all actual Harmonics disturbances were detected.

#### B. Classes with Improved Performance on the Test Set

- **Interruption:** The test set results were even better than on validation, with the model correctly classifying all 100 interruptions and producing only one false positive.
- **Notch:** Performance improved on the test set, correctly classifying all 100 Notch signals without any false positives or negatives.

#### C. Classes with Notable Drops in Performance

- **Oscillatory Transient:**

- Precision dropped from 0.99 to 0.91, and recall from 0.92 to 0.86. This indicates an increase in both false positives and false negatives.
- The confusion matrix shows that 11 instances were misclassified as Transient and 3 as Sag with Oscillatory Transient, highlighting difficulties in distinguishing between these disturbances.
- Harmonics with Sag:
  - Performance dropped due to misclassifications: One signal was incorrectly classified as Harmonics with Swell and two signals were misclassified as Sag with Harmonics.
  - This suggests that the model struggles to distinguish between these three similar waveform patterns.
- Sag with Harmonics: Precision dropped due to three misclassifications as Harmonics with Sag, showing a similar confusion as above.

#### D. Consistently Well-Classified Disturbances

- Sag, Swell, Flicker and Harmonics with Swell: These classes maintained nearly the same performance between validation and test sets, demonstrating that the model generalizes well for these disturbances.

### VII. CONCLUSIONS

This work proposed a simplified approach to power quality disturbance (PQD) classification, aiming to develop a model that accurately classifies disturbances without relying on feature extraction. Instead, the model uses only the normalized waveform of a single cycle, contrasting with other methods that incorporate feature extraction, image embedding techniques, and hybrid models.

To achieve this, a convolutional neural network (CNN) was selected due to its strong track record in handling similar classification problems. The results demonstrated the effectiveness of this approach, with the model achieving an overall accuracy of 97%, highlighting its capability to distinguish various PQDs with high precision.

This study confirms that deep learning can effectively classify PQDs without the need for manual feature engineering or extensive coding, simplifying implementation while maintaining high accuracy. Future work could focus on optimizing the model architecture, evaluating its performance on real-world noisy data, improving classification accuracy

for specific disturbance types, and extending this approach into a practical tool for engineers. Such a tool could process real-time data and perform real-time classification, aiding in the identification of power quality disturbances in operational environments.

### REFERENCES

- [1] N. Khosravi *et al.*, "Improvement of power quality parameters using modulated-unified power quality conditioner and switched inductor boost converter by the optimization techniques for a hybrid AC/DC microgrid," *Sci. Rep.*, vol. 12, p. 21675, 2022. doi: 10.1038/s41598-022-26001-8.
- [2] Y. Zhu, C. Liu, and K. Sun, "Image embedding of PMU data for deep learning towards transient disturbance classification," in *Proc. IEEE Int. Conf. Energy Internet (ICEI)*, Beijing, China, 2018, pp. 169–174.
- [3] W. L. Rodrigues Junior, F. A. Silva Borges, R. de A. Lira Rabelo, B. V. A. de Lima, and J. E. Almeida de Alencar, "Classification of power quality disturbances using convolutional network and long short-term memory network," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Budapest, Hungary, 2019, pp. 1–6.
- [4] Z. Yang, W. Liao, K. Liu, X. Chen, and R. Zhu, "Power quality disturbances classification using a TCN-CNN model," in *Proc. Asia Conf. Power Electr. Eng. (ACPEE)*, Shanghai, China, 2022, pp. 2145–2149.
- [5] W. Deng, D. Xu, Y. Xu, and M. Li, "Detection and classification of power quality disturbances using variational mode decomposition and convolutional neural networks," in *Proc. IEEE Annu. Comput. Commun. Workshop Conf. (CCWC)*, Las Vegas, NV, USA, 2021, pp. 1514–1518.
- [6] M. A. Rodriguez, J. F. Sotomonte, J. Cifuentes, and M. Bueno-López, "Power quality disturbance classification via deep convolutional auto-encoders and stacked LSTM recurrent neural networks," in *Proc. Int. Conf. Smart Energy Syst. Technol. (SEST)*, Eindhoven, Netherlands, 2020, pp. 1–6.
- [7] M. A. Ahajjam, D. B. Licea, M. Ghogho, and A. Kobbane, "Electric power quality disturbances classification based on temporal-spectral images and deep convolutional neural networks," in *Proc. Int. Wireless Commun. Mobile Comput. (IWCMC)*, Limassol, Cyprus, 2020, pp. 1701–1706.
- [8] Z. Tong, J. Zhong, J. Li, J. Wu, and Z. Li, "A power quality disturbances classification method based on multi-modal parallel feature extraction," *Sci. Rep.*, vol. 13, no. 1, Oct. 2023. doi: 10.1038/s41598-023-44399-7.
- [9] J. Li, H. Liu, D. Wang, and T. Bi, "Classification of power quality disturbance based on S-transform and convolution neural network," *Front. Energy Res.*, vol. 9, Jun. 2021. doi: 10.3389/fenrg.2021.708131.
- [10] S. Aziz, "SEED power quality disturbance dataset," *Kaggle*, 2023. [Online]. Available: <https://www.kaggle.com/datasets/sumairaziz/seed-power-quality-disturbance-dataset>. [Accessed: Feb. 16, 2025].
- [11] *IEEE Standard 519-2022: IEEE Recommended Practice and Requirements for Harmonic Control in Electric Power Systems*, IEEE Std 519-2022, 2022.
- [12] R. C. Dugan, M. F. McGranaghan, S. Santoso, and H. W. Beaty, *Electrical Power Systems Quality*, 3rd ed. New York, NY, USA: McGraw-Hill, 2012.