[Phr2] JPCSA4PaperLastCheck

by Turnitin Indonesia

Regulated 2D Grayscale Image for Finding Power Quality Abnormalities in Actual Data

Yeong-Chin Chen[1,*], M Syamsudin[1,2], S S Berutu[2]

1 Department of Computer Science and Information Engineering, Asia University, Taichung 413, Taiwan; 107221004@gm.asia.edu.tw

2 Department of Electrical Engineering, Politeknik Negeri Pontianak, Kalimantan Barat, Indonesia;

3 Department of Information and Technology, Universitas Kristen Imanuel, Yogyakarta 55571, Indonesia; sandinoberutu@gmail.com

* Correspondence: ycchenster@gmail.com

Abstract. It is possible to preserve power quality by classifying and identifying abnormalities. Prior studies focused on enhancing the PQD classification performance in one-dimensional (1D) CNNs. Recently, various image conversion methods have been established to facilitate CNN for PQD classification. PQD is a 1D signal that needs to be converted to a 2D image through data pre-processing since 2D images may include more PQD information than 1D signals. However, the PQD data used for the power quality classifier is synthetic PQD produced using mathematical models with parameter modifications in accordance with IEEE Std. 1159, which places limitations on prior research. This study uses data from the Amrita Honeywell Hackathon 2021 to examine how the response-based 2D deep CNN power quality classifier responds to actual field power quality disruptions. The results of the study show that a 2D deep CNN with regulated 2D grayscale pictures based on a process-regulated 2D image matrix can classify real data power quality disturbances with accuracy, precision, recall, and F1-score of 98.80%, 98.99%, and 98.60%, respectively. Additionally, 2D images can potentially contain more PQD data than 1D signals, enhancing identification performance on actual data.

1. Introduction

Power quality is becoming a more important consideration, especially with the global adoption of the smart grid idea to define future electrical firms. Electric power utilities and customers are anticipated to obtain optimum voltage and current waveforms at rated power frequency. However, distributed energy generating is one of the main causes of power quality disturbances (PQDs). Therefore, such disruptions must be identified before the proper mitigation mechanisms can be established to enhance power quality in a practical distribution network.

As of now, categorization and disturbance detection have been found to be crucial steps in preserving power quality. In the context of the smart grid, it is feasible to develop a power quality system based on the Internet of things [1] and deploy it along with the distribution network with the aim of informing utilities about consumption and disruptions via a two-way communication infrastructure. A general architecture can be seen in Figure 1.

Deep learning-based intelligent algorithms have recently been employed to categorize PQD. Convolutional neural networks (CNN) are among the most efficient techniques and are frequently employed in PQD classification research [2]. In earlier studies, the focus was on enhancing PQD classification performance in one-dimensional (1D) CNNs. Different image conversion techniques have been developed in recent years to facilitate the usage of CNN for PQD classification, which is a 1D signal that needs data pre-processing to convert to a 2D image as 2D images can include more PQD information than 1D signals. While [3] uses the signal-waveform image to train the PQD classifier immediately, and [4] transforms the sag signal into the PQD image using the space phasor diagram. However, the PQD data used for the power quality classifier is synthetic PQD produced using mathematical models with parameter alterations in accordance with IEEE Std. 1159, therefore, previous research has limitations. In order to depict a system based on historical data, models are abstractions of reality. Existent systems are intricate and made up of numerous interconnected parts. As a result, the quality of the model declines if the existent system experiences considerable changes.

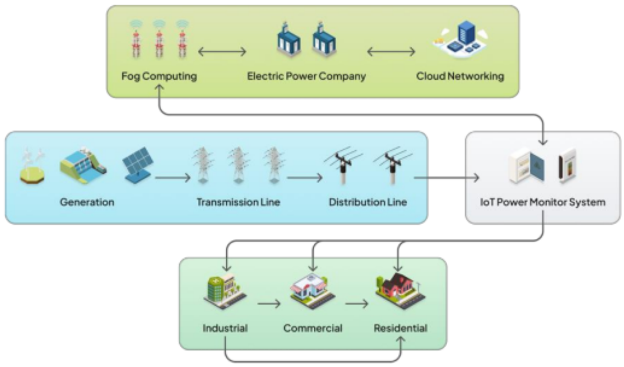


Figure 1. IoT-PMS in smart grid system

This pilot research examines how the response-based 2D deep CNN power quality classifier responds to actual field power quality disruptions using data from the Amrita Honeywell Hackathon 2021. The findings can be utilized to prove the hypothesis that a system for detecting power quality issues based on 2D deep CNN can improve the performance of identification accuracy on actual data.

2. Research Methodology

2.1. Data and Appliances

The data utilized in this study is published for the Amrita Honeywell Hackathon 2021's power quality analysis and comparison purposes. The dataset employed in this study is made up of signals classified into the following five power quality categories: normal, third harmonic wave, fifth harmonic wave, voltage dip, and transient. Each signal is described by 128 sampling data points (N_s), two sampling cycles (*Nc*), and a nominal value of fundamental frequency (*f*) in the range of 59.5 and 60.5 Hz. The following describes the power quality condition concerning the output class value with a total of 12,000 raw data: Normal (1,998), 3rd harmonic wave (2,000), 5th harmonic wave (3,000), voltage dip (2,000) and transient (3,000).

The hardware utilized is a MacBook Air with the following specifications: 8 GB of RAM, an M1 chip with a 16-core Neural Engine, a 7-core GPU, and an 8-core CPU with four performance cores and four efficiency cores. The learning model executed on a Google Colab accelerated GPU and internet connection using Wi-Fi 802.11ax Wi-Fi 6.

2.2. Data Preparation

Before being used in model training, preprocessing data is a crucial stage in its preparation and transformation, when the range of the data samples varies, normalization is a popular data processing technique where numerical column values are altered to have a uniform scale [5]. It is essential to to scale the data into a value range of -1 to 1 in normalization process before using it to reconstruct the dataset into two dimensions since the power quality distribution one-dimensional dataset value has a broad range, specifically between -7,185 and 11,997. Before being trained in the deep learning model in this research, raw power quality distribution data were processed at the data pre-processing phase, which comprise of two phases: signal synchronization (SS) and image regulation (IR), as shown in Figure 2.

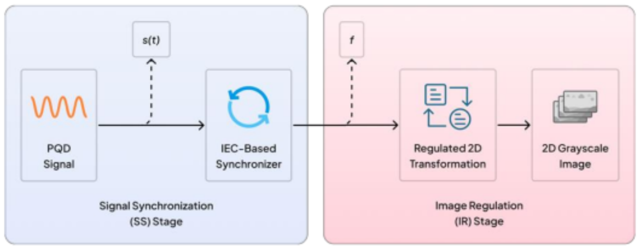
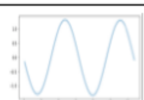
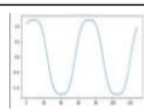
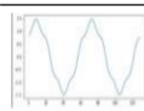
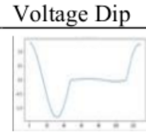
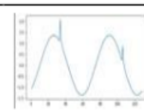


Figure 2. The stage of data preparation

In the SS stage, the regulated cycle duration is determined using the fundamental frequency received from the IEC (Std. 61000-4-7) based synchronizer in line with the fundamental frequency variation. The 2D grayscale picture matrix would then be controlled in the IR stage once the PQD signal had been correctly divided using the acquired fundamental frequency. The following are the essential phases in data preprocessing: Identify the submatrix dimension first. The square submatrix (number of the rows (N_{row}) is exactly as many columns (N_{col}) selected. Second, split the PQD signal into a number of cycles. The N_c cycles of the PQD signal are determined by the f value. Thirdly, create submatrices from the divided cycles. To create a controlled matrix, combine the submatrices in step four. Finally, create a 2D grayscale image from the controlled matrix. The grayscale image is produced by converting the matrix's components to the grayscale color range (0–255). The resulted image resolution is N_{row} x N_{col} pixels [6]. Table 1 shows the results of a regulated 2D grayscale image created using the previously discussed method.

Table 1. Power quality disturbances signal form and 2D grayscale image

	Normal	3rd Harmonic	5th Harmonic	Voltage Dip	Transient
PQD 1-D Signal					
PQD 2-D Image					

The disturbance classification stage would next process the regulated feature image to carry out the PQD identification.

2.3. *Method*

The research method in this study separated into 5 stages as shown in Figure 3. First, compile a CSV file with the power quality signal dataset from 5 classes. Second, the signal dataset is initially normalized during the data preprocessing stage. Third, create an image format with a metric size of 64 x 64 using the normalized dataset. In the fourth step, a ratio of 80% of the dataset utilized for model learning was used for model training data, while 20% was used for model validation. In many disciplines, using machine learning or deep learning models to solve issues typically involves dividing information into ratios [5]. Furthermore, each category received 200 testing data, for 1,000 testing data spread throughout 5 classes.

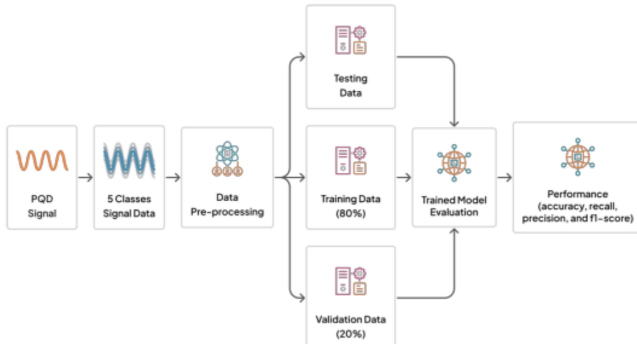


Figure 3. Stages of research method

In the fifth stage, the training model is evaluated to determine the performance of the 2D deep CNN model in recognizing PQDs. The model structure of 2D deep CNN applied is composed of 4 convolutional layers, 2 maxpooling layers and 1 dropout layer before 2 fully connected layers are shown in Figure 4. Therefore, the accuracy, recall, precision, and f1-score are used to evaluate the model's performance results. [7].

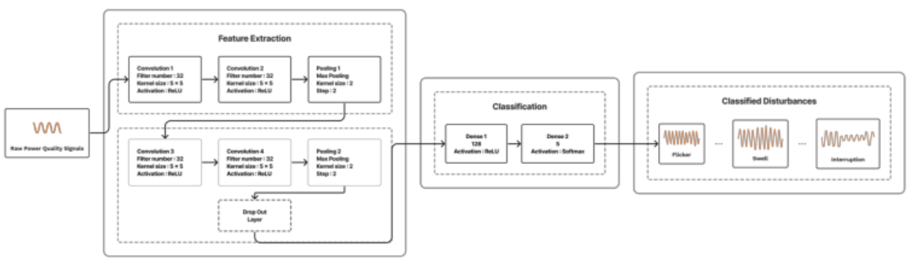


Figure 4. CNN model architecture

3. Result

Figure 5 shows the graphs among both the model of training and validation. According to the graph, training and validation accuracy are, respectively, 97.34% and 96.75%. To perform the fitting accuracy, the dropout layer value was adjusted to 0.35. The model demonstrates that when it is applied, there is neither (very slightly) over-fitting nor under-fitting of the model.

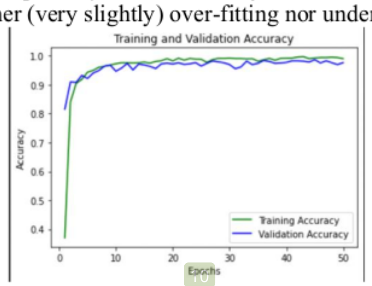


Figure 5. (a) Fitting graph of training and validation accuracy

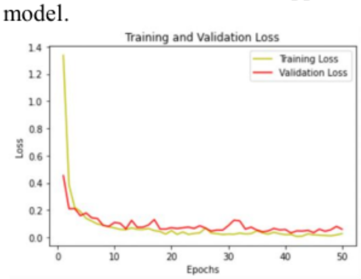


Figure 5. (b) Fitting graph of training and validation loss

Table 3 presents the performance evaluation for comparing CNN classification performance using real data and generated data.

Table 3. Performance of deep CNN classification method for synthetic and real data of PQD

Index	1D Signal (with SNR 20dB)	Regulated 2D Image Matrix (synyhetic data) [6]	Regulated 2D Image Matrix (real data)
Accuracy(%)	98.25	99.97	98.99
Precision(%)	98.28	99.81	98.60
Recall(%)	98.21	99.80	98.80
F1-score(%)	98.28	99.80	98.80

Table 3 demonstrates how a deep CNN classifier may identify five real data power quality disruption classifications. The outcome demonstrates that the 2D deep CNN method is superior in handling synthetic data, while the accuracy is marginally lower when processing real data at 98.99%. The small amount of data is one of the elements affecting the data's resilience. However, the outcome outperforms the 1D CNN classification method's performance for synthetic data tainted with 20 dB SNR noise.

The accuracy level provides details about a model's accuracy, or it can be argued that the performance of the model improves with increasing accuracy. Parameters of the training process, such as learning rate, batch size, and a number of epochs, were changed. By choosing the right parameters, the training procedure can be improved.

4. Conclusions

The study's findings showed that a 2D deep CNN with regulated 2D grayscale images based on a process-regulated 2D image matrix has the ability to classify real data power quality disturbances with the accuracy of 98.99%, and precision of 98.60%, recall and F1-score each with a value of 98.80%. Moreover, 2D images can contain more PQD information than 1D signals, enhancing the accuracy of identification performance on actual data.

Since the data's robustness significantly impacts the accuracy value of the applied model, research with real data is constrained by the small number of data samples obtained for the training process. Future work will require a transfer learning mechanism and a way to integrate several superior models to outperform the existing model.

References

[1] E. Elbouchikhi, M. F. Zia, M. Benbouzid, and S. El Hani, "Overview of signal processing and machine learning for smart grid condition monitoring," *Electronics,* vol. 10, no. 21, p. 2725, 2021.

[2] H. Sindi, M. Nour, M. Rawa, Ş. Öztürk, and K. Polat, "A novel hybrid deep learning approach including combination of 1D power signals and 2D signal images for power quality disturbance classification," *Expert Systems with Applications,* vol. 174, p. 114785, 2021.

[3] R. Zhu, X. Gong, S. Hu, and Y. Wang, "Power quality disturbances classification via fully-convolutional Siamese network and k-nearest neighbor," *Energies,* vol. 12, no. 24, p. 4732, 2019.

[4] A. Bagheri, M. H. Bollen, and I. Y. Gu, "Improved characterization of multi-stage voltage dips based on the space phasor model," *Electric Power Systems Research,* vol. 154, pp. 319-328, 2018.

[5] H.-C. Chen *et al.*, "Using Deep Learning Methods to Predict Walking Intensity from Plantar Pressure Images," in *International Conference on Applied Human Factors and Ergonomics*, 2021: Springer, pp. 270-277.

[6] C.-I. Chen, S. S. Berutu, Y.-C. Chen, H.-C. Yang, and C.-H. Chen, "Regulated Two-Dimensional Deep Convolutional Neural Network-Based Power Quality Classifier for Microgrid," *Energies,* vol. 15, no. 7, p. 2532, 2022.

[7] A. Tharwat, "Classification assessment methods," *Applied Computing and Informatics,* 2020.

[Phr2] JPCSA4PaperLastCheck

ORIGINALITY REPORT

15%	10%	13%	7%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Cheng-I Chen, Sunneng Sandino Berutu, Yeong-Chin Chen, Hao-Cheng Yang, Chung-Hsien Chen. "Regulated Two-Dimensional Deep Convolutional Neural Network-Based Power Quality Classifier for Microgrid", Energies, 2022 Publication	2%
2	ndltd.ncl.edu.tw Internet Source	2%
3	zenodo.org Internet Source	2%
4	Submitted to Asia University Student Paper	1%
5	Dabeeruddin Syed, Ameema Zainab, Ali Ghrayeb, Shady S. Refaat, Haitham Abu-Rub, Othmane Bouhali. "Smart Grid Big Data Analytics: Survey of Technologies, Techniques, and Applications", IEEE Access, 2021 Publication	1%

6	Submitted to Universitas Muhammadiyah Yogyakarta Student Paper	1 %
7	"Advances in Physical, Social & Occupational Ergonomics", Springer Science and Business Media LLC, 2021 Publication	1 %
8	Hsing-Chung Chen, Sunardi, Yih-Kuen Jan, Ben-Yi Liau, Chih-Yang Lin, Jen-Yung Tsai, Cheng-Tsung Li, Chi-Wen Lung. "Chapter 35 Using Deep Learning Methods to Predict Walking Intensity from Plantar Pressure Images", Springer Science and Business Media LLC, 2021 Publication	1 %
9	www.mdpi.com Internet Source	1 %
10	Submitted to University of East London Student Paper	1 %
11	arxiv.org Internet Source	<1 %
12	Francesco Salamini, Hakan Özkan, Andrea Brandolini, Ralf Schäfer-Pregl, William Martin. "Genetics and geography of wild cereal domestication in the near east", Nature Reviews Genetics, 2002 Publication	<1 %

13	ideas.repec.org Internet Source	<1 %
14	Lal Khan, Ammar Amjad, Kanwar Muhammad Afaq, Hsien-Tsung Chang. "Deep Sentiment Analysis Using CNN-LSTM Architecture of English and Roman Urdu Text Shared in Social Media", Applied Sciences, 2022 Publication	<1 %

Exclude quotes	Off	Exclude matches	Off
Exclude bibliography	Off		

[Phr2] JPCSA4PaperLastCheck

GRADEMARK REPORT

FINAL GRADE

/1

GENERAL COMMENTS

Instructor

PAGE 1

PAGE 2

PAGE 3

PAGE 4

PAGE 5

PAGE 6