electronics

MDPI

Article

Pretrained Configuration of Power-Quality Grayscale-Image Dataset for Sensor Improvement in Smart-Grid Transmission

Yeong-Chin Chen [1,*], Mariana Syamsudin [1,2] and Sunneng S. Berutu [3]

[1] Department of Computer Science and Information Engineering, Asia University, Taichung 413, Taiwan
[2] Department of Electrical Engineering, Politeknik Negeri Pontianak, Pontianak 78124, Indonesia
[3] Department of Information and Technology, Universitas Kristen Imanuel, Yogyakarta 55571, Indonesia
* Correspondence: ycchenster@gmail.com

Abstract: The primary source of the various power-quality-disruption (PQD) concerns in smart grids is the large number of sensors, intelligent electronic devices (IEDs), remote terminal units, smart meters, measurement units, and computers that are linked by a large network. Because real-time data exchange via a network of various sensors demands a small file size without an adverse effect on the information quality, one measure of the power-quality monitoring in a smart grid is restricted by the vast volume of the data collection. In order to provide dependable and bandwidth-friendly data transfer, the data-processing techniques' effectiveness was evaluated for precise power-quality monitoring in wireless sensor networks (WSNs) using grayscale PQD image data and employing pretrained PQD data with deep-learning techniques, such as ResNet50, MobileNet, and EfficientNetB0. The suggested layers, added between the pretrained base model and the classifier, modify the pretrained approaches. The result shows that advanced MobileNet is a fairly good-fitting model. This model outperforms the other pretraining methods, with 99.32% accuracy, the smallest file size, and the fastest computation time. The preprocessed data's output is anticipated to allow for reliable and bandwidth-friendly data-packet transmission in WSNs.

Keywords: grayscale PQD image dataset; pretrained methods; sensor network

Citation: Chen, Y.-C.; Syamsudin, M.; Berutu, S.S. Pretrained Configuration of Power-Quality Grayscale-Image Dataset for Sensor Improvement in Smart-Grid Transmission. *Electronics* **2022**, *11*, 3060. https://doi.org/10.3390/electronics11193060

Academic Editor: Jahangir Hossain

Received: 17 August 2022
Accepted: 23 September 2022
Published: 26 September 2022

1. Introduction

The process of developing and delivering power to end users has remained relatively stable over the years. In a traditional grid (Figure 1), the power systems are built on a few controlled and massive power sources, and primarily on hydroelectric or fossil-fuel-based energy-production systems, with a vast transmission network supplying power to customers through a distribution system. The electricity supplier creates a consumption plan based on the historical data from their customers, and orders electricity from the power plant based on that plan. This was possible because the fluctuations in energy use were low in the past, and the transmission system was generally reliable. This is significantly different from today, when large fluctuations in electricity usage make the transmission system more unreliable. As a result, a technological upgrade from the conventional grid is required to change the existing grid into a high-performance grid with huge potential.

Due to this transition, the smart grid has attracted much research interest over the past ten years. The emergence of smart grids becomes a solution when traditional networks are no longer adequate for implementation in the power system. Traditional electricity systems are single-directional-power and communication-transfer systems, and they include the integration and contribution of every distributed energy resource in the smart-grid-environment bidirectional electricity and data flow [1]. Smart networks use information other than historical data, which constantly monitors what is going on in the network and handles the flow of electricity directly. The software that collects, analyzes, and independently decides how energy will be distributed is at the heart of the smart grid. The

information gathered by energy suppliers from many sources is thereby processed in a single location, making the power grid far more predictable, adaptable, and trustworthy. The smart grid collects information from smart meters and other intelligent sensors, including IoT devices (IoT—Internet of things).

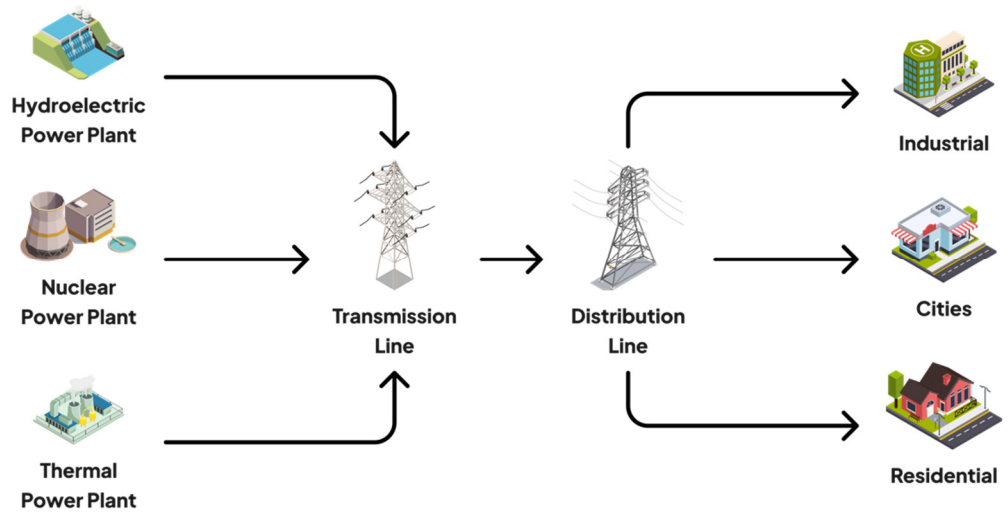


Figure 1. Traditional grid systems.

Because the traditional power system is being transformed into a more efficient and reliable smart grid, this shift places increased strain on a couple of centuries of power-grid infrastructure, necessitating further expenditure to guarantee safe and consistent electricity delivery to consumers. The smart grid is made up of a vast number of sensors, intelligent electronic devices (IEDs), remote terminal units, smart meters, measurement units, and computers that are linked by a pervasive network, which is the ultimate source of the various power-quality-disturbance (PQD) concerns. International standards for categorizing the electrical disturbances that affect the grid or user have been developed because the PQ is a crucial prerequisite for smart grids [2]. The term power quality refers to nonstationary disturbances that cause major malfunctions in electrical equipment [3]. The methods and threshold values that define electrical disturbances, such as overvoltage, undervoltage (sag or dip), fluctuation, harmonic distortion, etc., are laid forth in the PQ definition provided by the standards IEEE-1159 [4] and EN-50160 [5].

On the one hand, among the available methods to monitor the above disturbances, data acquisition with traditional wired systems may have considerable repercussions for the operator's safety [6]. On the other hand, using a wireless sensor network (WSN) eliminates these shortcomings and makes the data acquisition much safer [7]. One receiving station can acquire multiple data sensors remotely. On the contrary, most wired sensor networks use lengthy cables to deliver the acquired data to the central computer. These cables are subjected to wear and tear, leading to channel losses. Thus, the use of WSNs in data acquisition does not only contribute to its safety, but also to its economy.

A WSN is a system that comprises several computational and sensor units dispersed throughout a monitored environment. WSNs have been used to automate the usage of computers, sensors, and wireless communication equipment for both academic and commercial applications throughout the past few decades. ZebraNet, for instance, was created to track wildlife [8]. The purpose of CitySense is to provide weather and air-quality reports [9]. The SensorMap portal was created to provide services for genetic monitoring [10]. The design of specialized systems, such as the ones mentioned above, has received more research focus to meet the application-dependent service needs [11].

WSNs are employed for a wide range of purposes, which frequently need real-time data transfer. A well-known obstacle to WSN implementation is bandwidth restriction, which results in sample-rate and sensor-number limitations. This can be resolved by

decreasing the extra communication load using compression techniques and an occurrence-communication method [12].

Some computations must be performed by the smart meter online to identify the PQ, while others need an off-line strategy, such as disturbance propagation. As a result, the smart-sensor network must perform some computations while relying on a big-data postprocessor for others. A general smart sensor in a smart-grid system is shown in Figure 2.

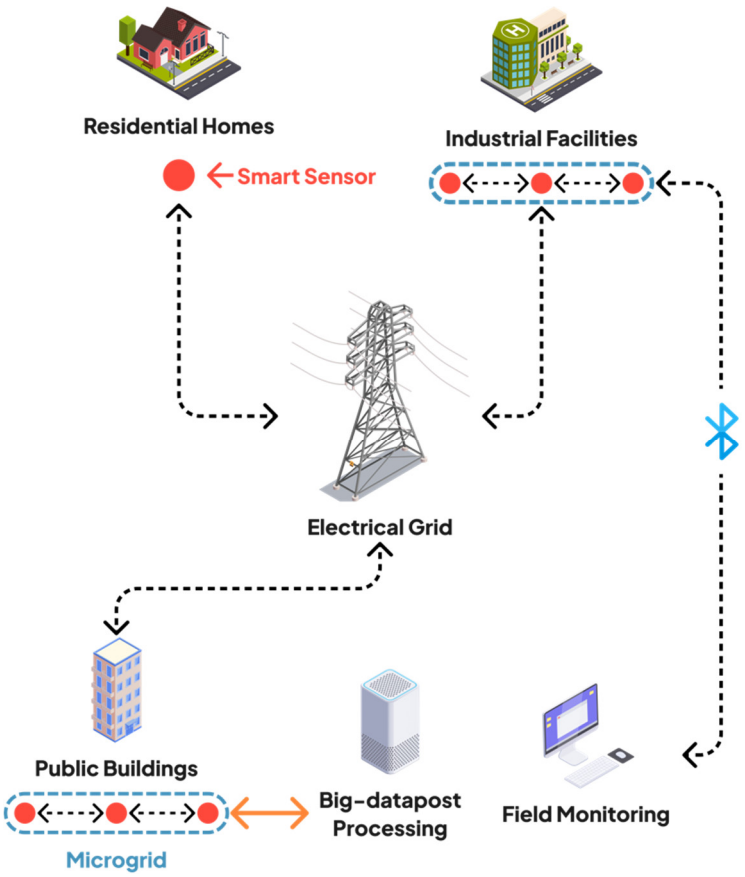


Figure 2. Reprinted from Ref. [13]: Smart sensor in smart-grid system.

Smart sensors can be discreetly installed inside several structures, including private residences, commercial buildings, and public buildings. These data are then communicated to other sensors, controller devices, centralized management platforms, distributed computing platforms, and others using transmission standards, such as Wi-Fi, Bluetooth, near-field communication (NFC), radio-frequency identification (RFID), LTE/5G, and others, for data aggregation and analysis. However, in this model, the smart sensor incorporates a wireless Bluetooth communication module, large storage device, and data-gathering module. A number of factors led to the selection of Bluetooth as a transmission method, including its low power consumption (when compared with other high-data-rate wireless communication systems, such as Wi-Fi), simplicity, widespread use worldwide, and ability to function in any environment (for example, the absence of a Wi-Fi connection) [14]. A microprocessor, internal data bus, real-time clock, universal serial bus (USB), and various soft cores for signal processing are also included. A mobile device, such as a smartphone or tablet, is used to integrate and remotely monitor the network of smart sensors in the system. A large amount of data is produced by the system, which is subsequently processed further in a big-data center. As the smart sensors are nonobtrusive, the system does not need to be powered off during the data-flow communication. This makes the system very simple to

use. It is also incredibly adaptable because it supports a wide range of current and voltage levels, and a variety of programmable soft-core-based processing capabilities.

While a different approach was used in [15], the implementation of fifth-generation (5G) networks in smart grids would lead to the development of novel "edge" and "fog" technology models at the utility level, together with smart control and automation. This strategy, however, is unfriendly to nations that still struggle to offer 5G services uniformly across their nations. By 2025, it is expected that 5G services will only make up 5% of all Internet connections in developing nations, such as Indonesia, despite its increased coverage and speed. This is the lowest among the Asia-Pacific countries, as compared with South Korea, which leads at 67%, as indicated in the 2020 GSMA Mobile Economy Asia Pacific report [16].

Because we are using a transmission standard that is massively used today, we need a model that is applicable to insufficient-quality transmission networks (e.g., remote areas). One measure of power-quality monitoring is constrained by the enormous volume of the data collection because real-time data sharing over a network of numerous sensors requires a small file size without compromising the information quality. Delivering power-quality-monitoring services is difficult as a result. This problem is addressed by evaluating the effectiveness of the data-processing techniques for precise power-quality monitoring in WSNs using 2D regulated grayscale PQD image data from recent research findings, which utilize the basic convolutional neural network (CNN). Furthermore, providing pretrained PQD data using deep-learning techniques, such as ResNet50, MobileNet, and EfficientNetB0, provide dependable and bandwidth-friendly data transfer.

The following are the primary contributions of this work:

- An experimental evaluation of dataset-pretraining methodologies was conducted for online PQD classification on WSN nodes with constrained computing capabilities, constrained internal storage, and low energy consumption. To the best of our knowledge, the earlier PQD research only provided ResNet data for pretraining. While this was occurring, it was difficult or challenging to locate references to the implementation of MobileNet and EfficientNetB0 pretraining on PQDs;
- The study investigates how responsive response-based 2D-depth CNN power-quality classifiers lead to substantive improvements in the field power quality. Because the PQD data utilized for the power-quality classifiers were synthetic, and the PQDs were developed using a mathematical model with parameter changes in line with IEEE Std. 1159, the earlier research had difficulty identifying the real disturbances because the model was an abstraction from reality.

The following are the contents of this study: Section 1 includes the introduction; Section 2 describes the prior works, which provide the background for the current topic; Section 3 proposes data and solutions for addressing the presented situation; Section 4 describes the data-driven validation and discusses the results; Section 5 concludes the paper and presents future work.

2. Related Work

It has been determined that categorization and abnormality detection are crucial methods for preserving power quality. The computational speed of the algorithm for classifying and detecting disturbances is the most critical aspect to take into account in the context of the smart grid, aiming to send information on the consumption and disruptions to utilities via a two-way communication infrastructure. In other words, the computational speed must be compatible with the bandwidth and data-transfer speed.

One-dimensional (1D) and two-dimensional (2D) datasets are two novel dataset-based methodologies for finding and classifying PQDs. In past research, the main goal was to improve the PQD-classification performance in 1D convolutional neural networks (CNNs). The most effective method, CNNs, are frequently used in PQD-classification research [17]. The current technique uses a 1D CNN algorithm and principal component analysis (PCA) to categorize data using 1D PQDs. The wind-grid distribution system, a wind-energy-based

renewable energy system conceived and developed to distribute electricity to the grid, uses this technology [18]. However, past research on the problem of the training time was limited because of the vast data volume. Large data files are generated in PQ monitoring as a result of the high sample rate and amount of measurement points [19].

A data-compression approach is necessary to shorten the amount of time that the calculations take during the training stage [20]. Signal-compression algorithms have been proposed to reduce the amount of data that needs to be saved. Recently, there has been some scientific interest in CNN compression. In order to save storage costs and enable a fast Fourier transform to speed up the computing, this work [21] suggests replacing the traditional linear projection on the completely linked layer with circular projection. A different study [22] aimed to reduce the network's total number of parameters and operations. The pruning approach can significantly reduce the computing workload and parameter size. However, significant PQD data would be lost due to the compression process.

Because 2D images can include more PQD information than 1D signals, an image-conversion approach has been developed in recent years to make it easier to use CNNs for PQD classification. PQD signals are 1D signals that require data preprocessing in order to transform into 2D images. In one study [23], the PQD classifier was immediately trained using the signal-waveform picture, and in [24], the sag signal was converted into the PQD image using the space-phasor diagram.

While prior studies used a three-channel format comprising data for red, green, and blue (RGB)-hued data, [25] displays an image-transformation matrix in which the PQD signal's sample points are rearranged in the matrix before being turned into a grayscale image. Nevertheless, as [26] transitions, certain crucial elements are utterly lost. According to Karasu's method in [25,27], rearranging the picture-transformation matrix leads to classification errors when the fundamental frequency deviates from its nominal value. Because the fundamental frequency varies, the time locations of the PQDs decrease. The approach has a training accuracy of 98.69%, while Zheng's method [28] has a training accuracy of 97.98%.

The fundamental-frequency variation was detected, and the image matrix was controlled by the IEC-based synchronizer to enhance the classification performance. The controlled 2D grayscale image can maintain the signal's information and waveform characteristics. The results of the testing and field measurements showed that the suggested strategy was more effective than the previously used approaches, and it could boost the PQD-classification effectiveness, with an accuracy of greater than 99.79 percent [29].

The optimum PQD-classification method is still being researched to enhance the system reliability in power systems. Many researchers use enhanced CNN architecture, and specifically the residual neural network (ResNet), to perform multiple PQD analyses. According to the research [30], ResNet18 outperforms other CNN designs in terms of accuracy (95.77 percent) when compared with other classifiers, such as the basic CNN, deep CNN (DCNN), and GoogLeNet. In comparison, the MobileNetV2 classifier is built and tested to classify the surface-water quality. The testing findings reveal that the classifier performs admirably and can be easily implemented on edge devices [31]. Squeeze-and-excite blocks from MobileNetV2 and inverted bottleneck residual blocks serve as the basis for the EfficientNetB0 network. EfficientNet transfers well and achieves state-of-the-art accuracy on Flowers (98.8%), CIFAR-100 (91.7%), and three other transfer-learning datasets with fewer parameters [32].

3. Research Methodology

3.1. Data-Transmission Method

An IoT-based power-monitoring system [33] is presented in this section. As depicted in Figure 3, the Arduino ATMega chip is utilized as the controller for data acquisition and data transmission. For the data acquisition, ZMPT101B and ACS712 sensors are employed to obtain the voltage and current values, whereas the NodeMCU device is used for the data transmission from the Arduino to the server. The Arduino first re-

ceives these values through the analog–digital-conversion (ADC) process. Then, from the Arduino, data are transmitted to the NodeMCU device using the universal-asynchronous-receiver/transmitter (UART) transmission system. This device utilizes the TCP connection for transferring data to the server. On the server side, the Go server is employed to store the voltage and current data. Meanwhile, the user interface is developed by using PHP programming. Finally, the information on the power-usage performance can be displayed using a multiplatform device.

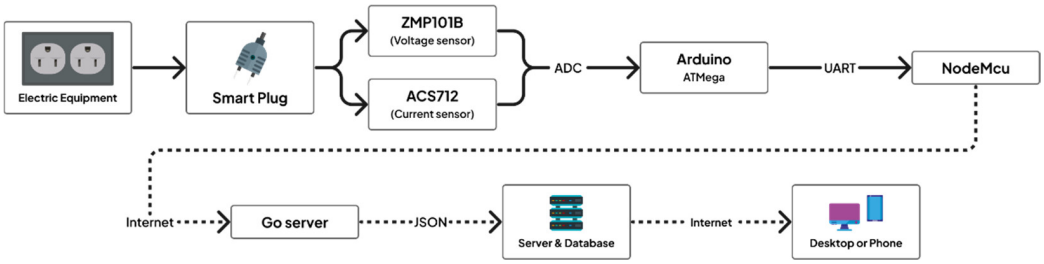


Figure 3. The power-monitoring system.

To reduce the Wi-Fi transmission burden, the data acquired from the data-acquisition layer are classified using a pretrained dataset. In this approach, the data size is smaller and there is a faster algorithm computation time than the basic CNN method. It assumes that the data load reduces and the speed of transmission increases.

3.2. Transfer Learning

In order to stack more layers and achieve higher overall accuracy, the DCNN is proficient at recognizing low-, medium-, and high-level features in images. Because the deep-neural-network architecture is comprehensive and the design is complex, a valuable technique known as transfer learning can be employed for a specific type of task. Transfer learning (TL) is a strategy for solving other similar problems by employing a pretrained model on a dataset as a starting point and adjusting and updating its parameters to fit the new dataset. When performing the model-training process, the TL model will assist in reducing the amount of data utilized, the calculation procedure, and the calculation time [34].

3.3. Pretrained Deep-Learning Network

With fixed weights for the specific application, a pretrained network has already learned to retrieve powerful and valuable features from natural photos. When the dataset is small and the application domain is related, pretrained networks can be deployed. Moreover, it takes time and computing power to train a CNN from the beginning. According to one study [35], employing weights from a distant task may improve the performance compared with randomly initialized weights.

There is currently a large number of pretrained CNNs, such as ResNet, MobileNet, EfficientNet, etc. In some cases, several pretrained networks deliver an exceptional performance. The current study investigates the ideal CNN network configuration for PQD classification considering its excellent performance. The pretraining network was chosen based on its ease of use and most excellent performance in prior iterations of the ILSVRC (ImageNet Large Scale Visual Recognition Challenge) competition. Other factors considered include the network's time and space complexity, the error rate displayed in the ILSVRC competition and more [36].

Table 1 [37] highlights the performance of ImageNet, an image database that is arranged according to the WordNet hierarchy, with hundreds of millions of images representing each node of the hierarchy. The data demonstrate that the EfficientNetB0 significantly outperforms other pretrained models. EfficientNetB0, in particular, uses 5.3 M parameters, has a running time of 4.9 m for each inference step (GPU), and achieves 93.3 percent in the

top-5 accuracy, in comparison with MobileNet, which computes 4.3 M parameters in 3.4 m per inference step (GPU), but only achieves 89.5 percent top-5 accuracy. The widely used ResNet50 has a top-5 accuracy of 92.1 percent using 25.6 M and a 4.6 m per inference step (GPU). A previous study by Marques et al. [38] revealed that one of the existing DCNNs, such as the EfficientNet model, might perform brilliantly in a number of medical-picture classifications, such as chest X-rays. With a 99.62% accuracy rate, the improved EfficientNet outperformed a number of the well-known DCNNs that had been previously released, including ResNet [39] and MobileNet [40].

Table 1. Summary of ImageNet performance.

Model	Size (MB)	Top-1 Accuracy	Top-5 Accuracy	Parameters	Depth	Time Per Inference Step (CPU)	Time Per Inference Step (GPU)
ResNet50	98	74.9%	92.1%	25.6 M	107	58.2	4.6
MobileNet	16	70.4%	89.5%	4.3 M	55	22.6	3.4
EfficientNetB0	29	77.1%	93.3%	5.3 M	132	46.0	4.9

4. Experiment Setup

The experiments were carried out to evaluate the performances of the three different pretrained CNNs, ResNet50, MobileNet, and EfficientNetB0, when PQD classification was applied to the images from the signal-power-quality dataset of the Amrita Honeywell Hackathon 2021 [41].

4.1. Data and Hardware

The dataset used in this study consists of signals divided into the following five power-quality categories: normal, third harmonic wave, fifth harmonic wave, voltage dip, and transient. A nominal fundamental frequency (f) in the range from 59.5 to 60.5 Hz, two sampling cycles (Nc), and 128 sample data points (N_s) are used to describe each signal. With a total of 11,998 pieces of raw data, the following power-quality conditions are described in relation to the output-class values: normal (1998), third harmonic wave (2000), fifth harmonic wave (3000), voltage dip (2000), and transient (3000).

The used hardware was a MacBook Air with the following features: 8 GB of RAM, a 16-core Neural Engine on the M1 chip, a 7-core GPU, and an 8-core CPU with four performance cores and four efficiency cores. The learning model was implemented utilizing a Google Colab-accelerated GPU and a Wi-Fi 802.11ax Wi-Fi 6 connection to the Internet.

4.2. Data Preprocessing

Preprocessing data are essential in preparation and modification, before being utilized in model training. When the range of the data samples fluctuates, normalization is a frequently used data-processing technique in which the numerical column values are changed to have a uniform scale. Before using the data to reconstruct the dataset into two dimensions, the data must be scaled into a value range from −1 to 1 as part of the normalization process, as the power-quality distribution of a one-dimensional-dataset value has a wide range, and especially between −7185 and 11,997. The raw power-quality-distribution data in this study were handled in the data-preprocessing phase, which includes the two steps of signal synchronization (SS) and image regulation (IR), as shown in Figure 4.

In the SS stage, the fundamental frequency obtained from the IEC (Standard 61000-4-7)-based synchronizer is used to calculate the regulated cycle duration in accordance with the fundamental-frequency variation. After the PQD signal has been properly separated using the acquired fundamental frequency, the 2D grayscale-picture matrix is controlled in the IR stage. The crucial step in data preparation is determining the submatrix dimension. The square submatrix has precisely as many columns (N_{col}) chosen as rows (N_{row}). The PQD signal should then be divided into several cycles. The f value determines the N_c cycles of the PQD

signal. Third, take the divided cycles and generate submatrices. Step four should merge the submatrices to produce a controlled matrix. Finally, take the controlled matrix and turn it into a 2D grayscale image. The matrix's components are converted to the grayscale color space to create the grayscale image (0–255). $N_{row} \times N_{col}$ pixels are the size of the final image [30]. The output of a controlled 2D grayscale image produced using the previously mentioned technique is shown in Table 2.

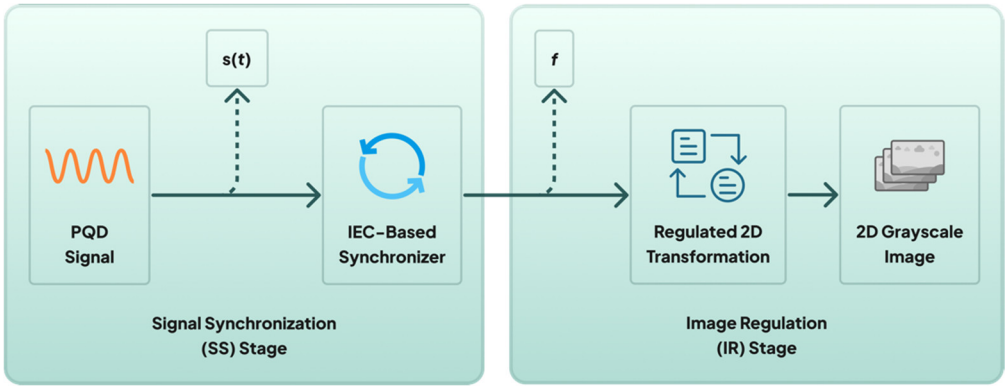
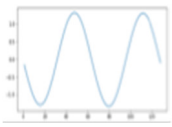
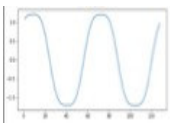
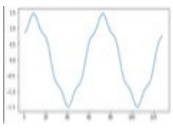
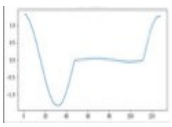
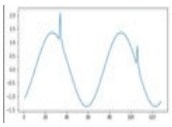


Figure 4. The stages of data preparation.

Table 2. Power-quality-disturbance signal form and 2D grayscale image.

	Normal	Third Harmonic	Fifth Harmonic	Voltage Dip	Transient
PQD 1D Signal					
PQD 2D Image					

The disturbance-classification stage then processes the regulated feature image to complete the PQD identification.

4.3. Stages of Research

According to Figure 5, the research process for this study was divided into six parts. First, a CSV file was created containing the power-quality signal dataset for the first five classes. In the data-preprocessing stage, the signal dataset was initially normalized. Third, the normalized dataset produced an image format with a metric resolution of 64 × 64. In the fourth step, the model-training data accounted for 80% of the dataset used for the model learning, while the model-validation data made up 20% of the dataset. In many areas, splitting information into ratios is a common practice when utilizing machine-learning or deep-learning models to solve problems. Additionally, 200 test results were given to each category, totaling 1000 test results divided among five classes.

The model-training stage is the fifth stage. ResNet50, MobileNet, NASNetMobile, and EfficientNetB0 transfer-learning models (trained models) for PQD classification were employed in this work because they are among the best-performing transfer-learning models (trained models) that are commonly utilized by researchers for image classification. The performance of the 2D-deep CNN model in recognizing PQDs was evaluated based on the training model in the sixth stage. Figure 5 illustrates the model structure of the applied 2D-deep CNN, which consists of four convolutional layers, two max-pooling layers, and

one dropout layer before two fully connected layers. Therefore, the model's performance outcomes were assessed using the accuracy, recall, precision, and F1 score [42].

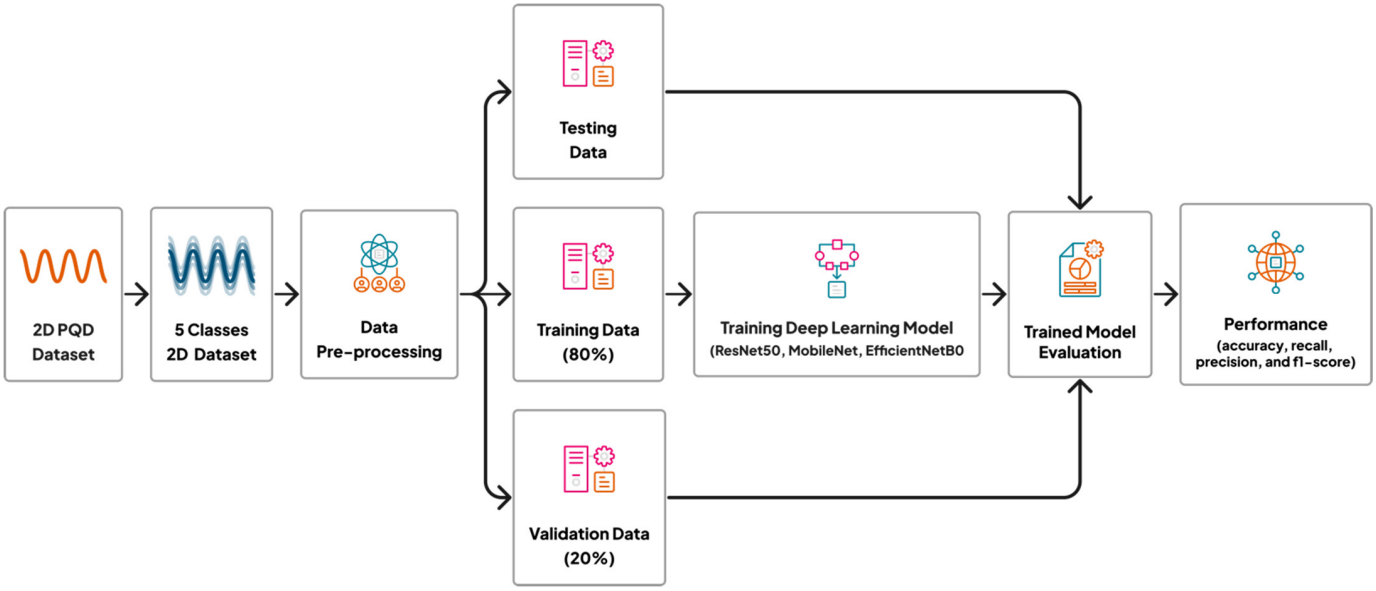


Figure 5. Stages of research method.

4.4. Proposed Layer

This study modified pretrained methods to dig for more information from the PQD dataset. The proposed layers are placed between the pretrained base model and the classifier. Figure 6 illustrates the proposed layers for EfficientNetB0, MobileNet, and ResNet50, composed of a global average pooling 2D (GAP2D) layer, dropout layer, and batch normalization. To avoid cases of extreme overfitting caused by the advanced feature management, a pooling layer was introduced. The GAP2D layer might drastically decrease the number of parameters by scaling the input tensor's height, width, and depth from the base model. By switching to the dense layer at this stage, which can overwhelm the classifier, the massive influx of characteristics is controlled. The feature maps were not entirely diminished by the GAP2D layer. Instead, it calculated the average of all the spatial data and retained the most complex patterns required to determinate the image [43].

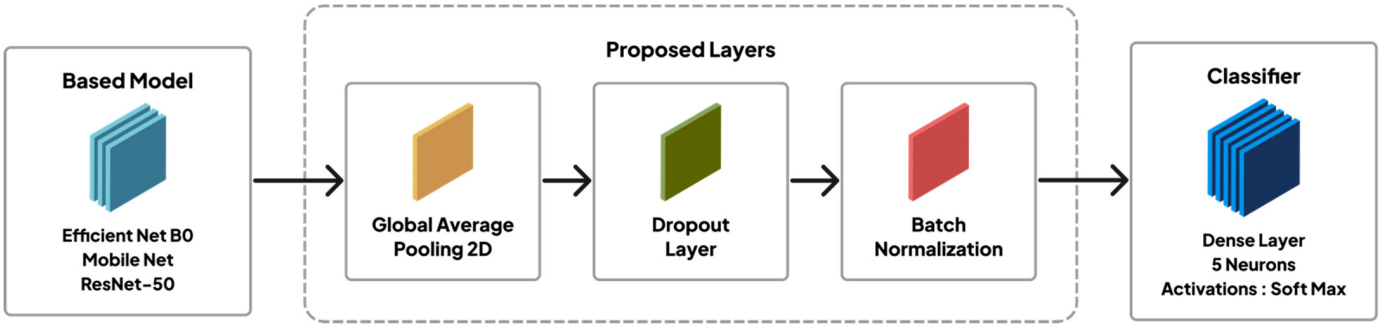


Figure 6. Proposed layers for advanced pretrained model.

The feature sets from the GAP2D layer are directed to a dropout layer and batch-normalization layer. Dropout is a technique that approximates the concurrent training of many neural networks with diverse topologies. Several layer outputs are discarded or randomly "dropped out" during training. To determine the probability at which the gradient outputs are dropped out, or the inverse likelihood for which the gradient outputs are preserved, a new hyperparameter is offered. A typical value is 0.5 for maintaining

a node's output in a hidden layer, and a value close to 1.0, such as 0.8, for maintaining a node's input in a visible layer. In comparison, batch normalization is an algorithmic method that makes the training of deep neural networks (DNNs) faster and more stable. Furthermore, the model connected the classifier with SoftMax activation and five neurons representing the five given labels.

4.5. Hyperparameter Value

The hyperparameter settings and loss function used for the task to yield effective results are described in this section.

A DL model's performance is measured not solely in terms of accuracy, but also in terms of loss. [44]. The model seeks to achieve its lowest rate of mistakes because a model with a smaller computed loss is more effective [45]. The cross-entropy-loss function is used in this work to obtain the average measure of the difference between the expected and forecast values. The loss measurement for binary classification is shown in Equation (1), where y represents the binary values of 0 or 1, and p is the probability [46].

$$CE = -(y \log(p) + (1 - y) \log(1 - p)) \quad (1)$$

The Adam optimizer was used in this work to provide optimal loss reduction during training. This optimization approach works as an adaptive gradient-descent function, allowing for faster weight loss towards the local minima [47]. When compared with alternative optimizers, such as stochastic gradient descent (SGD) [48] or RMSProp [49], because of its simplicity of implementation, effective memory usage, and quicker learning stage, the Adam optimizer was selected.

The settings for the hyperparameters are shown in Table 3. With the other given hyperparameters, a low learning rate (LR) is effective. The 32-batch size provided enough load to transport data across the network without using all the computing memory. Furthermore, we chose durations within 50 epochs to train each model incrementally to see how it would perform.

Table 3. Hyperparameters specified for training.

Hyperparameter	Value
Learning Rate	0.0004
Batch Size	32
Optimizer	Adam
Dropout	0.5
Epoch	50

5. Results

This section reviews the results gained from the prepared dataset throughout the validation and training stages. The outcomes of evaluating the deep-learning networks EfficientNetB0, MobileNet, and ResNet50 in the PQD-classification task on the actual PQD image dataset are presented in Table 4.

Table 4. Performance of pretrained deep-learning network for 2D grayscale-image PQD dataset.

Network	Accuracy (%)	
	Training	Validation
EfficientNetB0	99.55	98.58
MobileNet	98.90	97.46
ResNet50	99.03	96.85
Basic CNN	97.34	96.75

The training dataset is employed to calculate the training learning. Because validation learning is determined from a hold-out validation dataset and from how effectively the model generalizes, it offers information about how well the model is learning. Compared with MobileNet and ResNet50, the data from the table show that EfficientNetB0 produced very accurate deep-learning solutions for identifying 2D grayscale images from the PQD dataset, with a training accuracy of 99.55% and validation accuracy of 98.58%.

The training and validation losses that decrease to the point of stability, with a small gap between the two final loss values in Figure 7, show that the EfficientNetB0 training process is a good-fitting model.

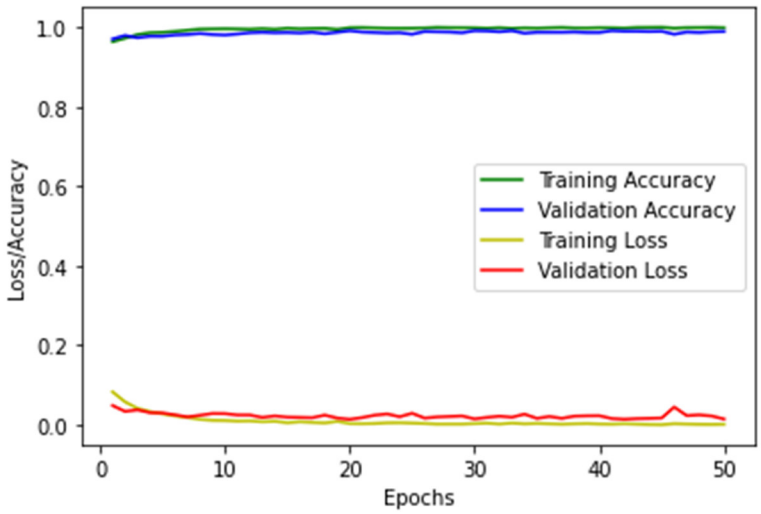


Figure 7. Training and validation progress of EfficientNetB0.

Figure 8 depicts an example of MobileNet overfitting. This may arise if the model is trained for an inordinately extended period. The 30th epoch may be the inflection point in the validation loss, as experience after that point illustrates the dynamics of overfitting after that point.

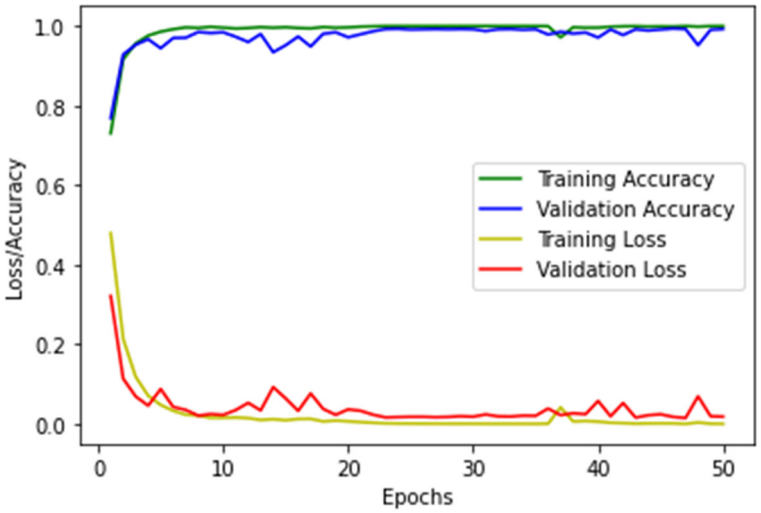


Figure 8. Training and validation progress of MobileNet.

Figure 9 shows that the training and validation learning curves of ResNet50 demonstrate a training dataset that may be too small compared with the validation dataset. Both learning curves can identify this situation for the training loss and validation loss, showing improvement, but after the 35th epoch, a slight gap remains between both curves.

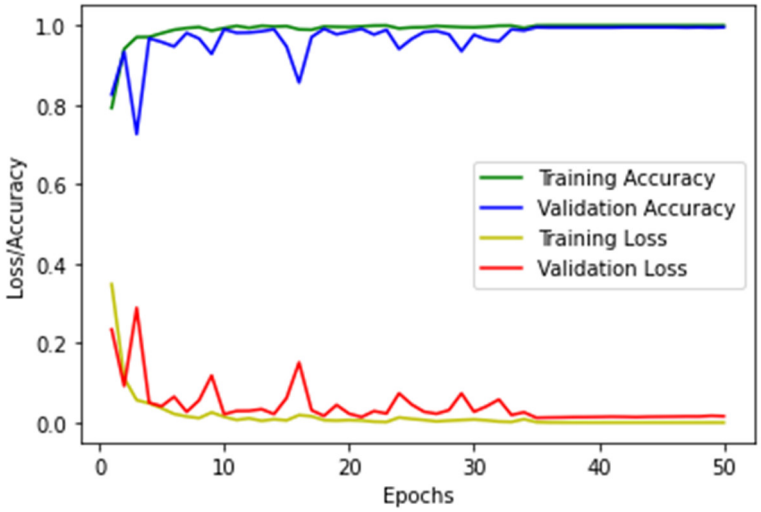


Figure 9. Training and validation progress of ResNet50.

The graphs for both the training and validation models are displayed in Figure 10. The graphs show that the accuracies during training and validation were 97.34% and 96.75%, respectively. The value of the dropout layer was changed to 0.35 to perform the fitting accuracy. The model shows that it neither overfit (slightly) nor underfit (very significantly) when implemented.

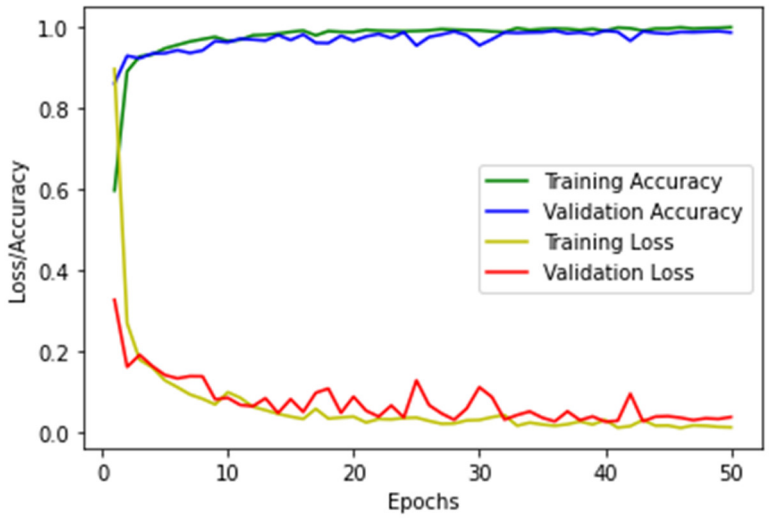


Figure 10. Training and validation progress of basic CNN.

A specialized metric called the confusion matrix (CM) shows how well a trained model can forecast from a given validation dataset. A third harmonic, fifth harmonic, normal, transient, and voltage dip are the true-class and ground-truth labels shown by the CM's corresponding rows and columns. For each validation sample, the anticipated results show the percentages of accurate and inaccurate predictions or classifications. Following the values, the accuracy, precision, recall, and F1 score of each model are computed. The recall value indicates how many times the model was able to detect a specific category. Precision is the frequency with which a model correctly predicts an actual class. The total number of accurate predictions made from all available samples is the accuracy. The F1 score is the weighted average of the recall and precision values [50].

Table 5 computes the total performance of the classification using a pretrained model based on its accuracy, precision, recall, and F1 score using the CM matrix. From the examined finding, ResNet50 achieved the highest accuracy of 99.55%, followed by MobileNet

with 99.32%, and EfficientNetB0 with 99.10%. As shown in the architectures of and methods of implementing all the models, ResNet50 has more parameters (25.6 MB) to be used, and so it will show a better performance compared with EfficientNetB0 and MobileNet. ResNet50, however, requires a more extensive data capacity (98 MB).

Table 5. Comparison of classification performances.

Network	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
EfficientNetB0	99.10	98.60	99.00	98.80
MobileNet	99.32	99.00	99.20	99.20
ResNet50	99.55	99.20	99.40	99.40
Basic CNN	98.99	98.60	98.80	98.80

The most severe resource limitation on WSNs during implementation for smart-grid objectives is restricted battery energy. Transmission power control and data-packet-size optimization are effective strategies for increasing the network lifetime and lowering the energy consumption [51]. As a result, this study suggests that the MobileNet and EfficientNetB0 pretraining models be used for 2D grayscale-image PQD-data classification. The MobileNet basic model requires only 16 MB of data size. The EfficientNetB0 base model requires 29 MB of data size. Furthermore, the compute times for both models are 22.6 m and 46.00 m, respectively, compared with the ResNet50 compute time of 58.2 m per inference step (CPU), and the base model requiring 98 MB of data size. At the same time, the three models' accuracies differ slightly.

6. Conclusions and Future Work

Previously, an experimental evaluation of ResNet-dataset-pretraining approaches for online PQD classification on WSN nodes with limited processing capabilities and internal storage and low energy consumption was carried out. However, references to the implementation of the MobileNet and EfficientNetB0 pretraining on PQDs are still challenging to find. As a result, this work took the initiative and examined contemporary CNNs for classifying and detecting PQDs utilizing the MobileNet- and EfficientNetB0-data-pretraining approaches. The pretrained techniques are modified by the proposed layers, which are inserted between the pretrained base model and the classifier. This research also looks into how responsive response-based 2D-depth CNN power-quality classifiers can lead to significant improvements in the field power quality.

Because ResNet50 has a greater number of parameters (25.6 MB), upon evaluation with 11,998 raw data images, the CNN classifier utilizing ResNet50 achieved the highest accuracy of 99.55%, followed by MobileNet with 99.32%, and EfficientNetB0 with 99.10%. However, in the training performance, the accuracy of EfficientNetB0 was better than the other pretrained methods, with a training accuracy of 99.55% and validation accuracy of 98.58%, which is a good fit, as demonstrated by the training and validation losses that decreased to the point of stability, with a minimal gap between the two final loss values (0.97%). In contrast, the gaps of MobileNet and ResNet50, respectively, were 1.44% and 2.18%.

As a result, compared with the basic deep CNN classification technique, the transfer-learning-based EfficientNetB0, MobileNet, and ResNet50 could efficiently improve the classification of the 2D-deep CNN using regulated 2D grayscale images. MobileNet surpasses the other pretraining methods evaluated, in general, because MobileNet has the smallest file size and fastest computation time (Table 1) to obtain a fairly good validation accuracy (Table 4). Setting the number of iterations to less than 30 epochs shows that the model works well for making new predictions based on data that have not been seen before. It also has a high test accuracy, which is slightly lower than ResNet50 (Table 5). It can be said that the model can be accurately trained.

The result of the preprocessed data is anticipated to enable dependable and bandwidth-friendly data-packet transmission in wireless sensor networks (WSNs), where the primary

challenges are the restricted node energy and storage space. In addition, the smart grid may access reliable Internet data that become essential for preventing power disruptions and outages through online power-system-condition monitoring, diagnostics, and protection. The modest number of data samples obtained during the training procedure limits research with real data. As a result, we advocate conducting more studies that include collecting additional data from other smart grids using various sensor devices.

Author Contributions: Conceptualization, Y.-C.C. and M.S.; data curation, S.S.B.; methodology, Y.-C.C. and M.S.; software, S.S.B. and M.S.; formal analysis, Y.-C.C. and M.S.; resources, M.S. and S.S.B.; writing—original draft preparation, M.S; writing—review and editing, Y.-C.C. and S.S.B.; supervision, Y.-C.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Simulated data can be provided upon request.

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

References

1. Ibrahim, M.S.; Dong, W.; Yang, Q. Machine learning driven smart electric power systems: Current trends and new perspectives. *Appl. Energy* **2020**, *272*, 115237. [CrossRef]
2. Lu, R.; Liang, X.; Li, X.; Lin, X.; Shen, X. EPPA: An efficient and privacy-preserving aggregation scheme for secure smart grid communications. *IEEE Trans. Parallel Distrib. Syst.* **2012**, *23*, 1621–1631.
3. Jerin, A.R.A.; Prabaharan, N.; Kumar, N.M.; Palanisamy, K.; Umashankar, S.; Siano, P. Smart grid and power quality issues. In *Hybrid-Renewable Energy Systems in Microgrids*; Elsevier: Amsterdam, The Netherlands, 2018; pp. 195–202.
4. *IEEE Std 1159TM-2019 (Revision of IEEE Std 1159-2009)*; P. E. W. Group. IEEE Recommended Practice for Monitoring Electric Power Quality. The Institute of Electrical and Electronics Engineers, Inc.: Piscataway, NJ, USA, 2019.
5. Masetti, C. Revision of European Standard EN 50160 on power quality: Reasons and solutions. In Proceedings of the 14th International Conference on Harmonics and Quality of Power-ICHQP 2010, Bergamo, Italy, 26–29 September 2010; pp. 1–7.
6. Sridhar, S.; Rao, K.U.; Jade, S. Detection and classification of PQ disturbances in the supply to induction motor using wavelet transform and feedforward neural network. In Proceedings of the 2015 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT), Coimbatore, India, 5–7 March 2015; pp. 1–5.
7. Akyildiz, I.; Su, W.; Sankarasubramaniam, Y.; Cayirci, E. Wireless sensor networks: A survey. *Comput. Netw.* **2002**, *28*, 393–422. [CrossRef]
8. Juang, P.; Oki, H.; Wang, Y.; Martonosi, M.; Peh, L.S.; Rubenstein, D. Energy-efficient computing for wildlife tracking: Design tradeoffs and early experiences with ZebraNet. In Proceedings of the 10th International Conference on Architectural Support for Programming Languages and Operating Systems, San Jose, CA, USA, 5–9 October 2002; pp. 96–107.
9. Murty, R.N.; Mainland, G.; Rose, I.; Chowdhury, A.R.; Gosain, A.; Bers, J.; Welsh, M. Citysense: An urban-scale wireless sensor network and testbed. In Proceedings of the 2008 IEEE Conference on Technologies for Homeland Security, Waltham, MA, USA, 12–13 May 2008; pp. 583–588.
10. Kolli, R.; Doshi, P. Optima: Tool for ontology alignment with application to semantic reconciliation of sensor metadata for publication in sensormap. In Proceedings of the 2008 IEEE International Conference on Semantic Computing, Washington, DC, USA, 4–7 August 2008; pp. 484–485.
11. Lim, Y.; Kim, H.-M.; Kang, S. A design of wireless sensor networks for a power quality monitoring system. *Sensors* **2010**, *10*, 9712–9725. [CrossRef] [PubMed]
12. Beuchert, J.; Solowjow, F.; Trimpe, S.; Seel, T. Overcoming bandwidth limitations in wireless sensor networks by exploitation of cyclic signal patterns: An event-triggered learning approach. *Sensors* **2020**, *20*, 260. [CrossRef]
13. Morales-Velazquez, L.; de Jesus Romero-Troncoso, R.; Herrera-Ruiz, G.; Morinigo-Sotelo, D.; Osornio-Rios, R.A. Smart sensor network for power quality monitoring in electrical installations. *Measurement* **2017**, *103*, 133–142. [CrossRef]
14. Aram, S.; Troiano, A.; Pasero, E. Environment sensing using smartphone. In Proceedings of the 2012 IEEE Sensors Applications Symposium Proceedings, Brescia, Italy, 7–9 February 2012; pp. 1–4.
15. Almasarani, A.; Majid, M. 5G-Wireless sensor networks for smart grid-accelerating technology's progress and innovation in the kingdom of Saudi arabia. *Procedia Comput. Sci.* **2021**, *182*, 46–55.
16. Telkomsel. The Impact of 5G in Indonesia's 2030. Available online: https://telkomseliot.com/en/news-insight/the-impact-of-5g-in-indonesias-2030 (accessed on 7 August 2022).
17. Sindi, H.; Nour, M.; Rawa, M.; Öztürk, Ş.; Polat, K. A novel hybrid deep learning approach including combination of 1D power signals and 2D signal images for power quality disturbance classification. *Expert Syst. Appl.* **2021**, *174*, 114785. [CrossRef]

18. Mengi, O.O.; Altas, I.H. A new energy management technique for PV/wind/grid renewable energy system. *Int. J. Photoenergy* **2015**, *2015*, 1–19. [CrossRef]
19. He, S.; Tian, W.; Zhang, J.; Li, K.; Zhang, M.; Zhu, R. A high efficient approach for power disturbance waveform compression in the view of heisenberg uncertainty. *IEEE Trans. Ind. Inform.* **2018**, *15*, 2580–2591. [CrossRef]
20. Shen, Y.; Abubakar, M.; Liu, H.; Hussain, F. Power quality disturbance monitoring and classification based on improved PCA and convolution neural network for wind-grid distribution systems. *Energies* **2019**, *12*, 1280. [CrossRef]
21. Lu, Y.; Kumar, A.; Zhai, S.; Cheng, Y.; Javidi, T.; Feris, R. Fully-adaptive feature sharing in multi-task networks with applications in person attribute classification. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 5334–5343.
22. Yang, T.-J.; Chen, Y.-H.; Sze, V. Designing energy-efficient convolutional neural networks using energy-aware pruning. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 5687–5695.
23. Zhu, R.; Gong, X.; Hu, S.; Wang, Y. Power quality disturbances classification via fully-convolutional Siamese network and k-nearest neighbor. *Energies* **2019**, *12*, 4732. [CrossRef]
24. Bagheri, A.; Bollen, M.H.; Gu, I.Y. Improved characterization of multi-stage voltage dips based on the space phasor model. *Electr. Power Syst. Res.* **2018**, *154*, 319–328. [CrossRef]
25. Karasu, S.; Saraç, Z. Investigation of power quality disturbances by using 2D discrete orthonormal S-transform, machine learning and multi-objective evolutionary algorithms. *Swarm Evol. Comput.* **2019**, *44*, 1060–1072. [CrossRef]
26. Bagheri, A.; Gu, I.Y.; Bollen, M.H.; Balouji, E. A robust transform-domain deep convolutional network for voltage dip classification. *IEEE Trans. Power Deliv.* **2018**, *33*, 2794–2802. [CrossRef]
27. Karasu, S.; Sarac, Z. Classification of power quality disturbances by 2D-Riesz Transform, multi-objective grey wolf optimizer and machine learning methods. *Digit. Signal Process.* **2020**, *101*, 102711. [CrossRef]
28. Zheng, Z.; Qi, L.; Wang, H.; Pan, A.; Zhou, J. Recognition method of voltage sag causes based on two—Dimensional transform and deep learning hybrid model. *IET Power Electron.* **2020**, *13*, 168–177. [CrossRef]
29. Chen, C.-I.; Berutu, S.S.; Chen, Y.-C.; Yang, H.-C.; Chen, C.-H. Regulated Two-Dimensional Deep Convolutional Neural Network-Based Power Quality Classifier for Microgrid. *Energies* **2022**, *15*, 2532. [CrossRef]
30. Abd Jamlus, N.U.I.; Shahbudin, S.; Kassim, M. Power Quality Disturbances Classification Analysis Using Residual Neural Network. In Proceedings of the 2022 IEEE 18th International Colloquium on Signal Processing & Applications (CSPA), Virtual. 7–13 May 2022; pp. 442–447.
31. Hao, J.; Tao, Y. Surface water quality classification based on MobileNetV2. *J. Phys. Conf. Ser.* **2020**, *1646*, 012049. [CrossRef]
32. Tan, M.; Le, Q. Efficientnet: Rethinking model scaling for convolutional neural networks. In Proceedings of the International Conference on Machine Learning, Long Beach, CA, USA, 10–15 June 2019; pp. 6105–6114.
33. Chen, Y.-C.; Berutu, S.S.; Wang, Y.-H. Smart meter development for cloud-based home electricity monitor system. *J. Electron. Sci. Technol.* **2020**, *18*, 378–390.
34. Wu, Y.; Wu, Q.; Dey, N.; Sherratt, S. Learning models for semantic classification of insufficient plantar pressure images. *Int. J. Interact. Multimed. Artif. Intell.* **2020**, *6*, 51–61. [CrossRef]
35. Yosinski, J.; Clune, J.; Bengio, Y.; Lipson, H. How transferable are features in deep neural networks? *Adv. Neural Inf. Process. Syst.* **2014**, *27*, 3320–3328.
36. Islam, M.M.; Tasnim, N.; Baek, J.-H. Human gender classification using transfer learning via Pareto frontier CNN networks. *Inventions* **2020**, *5*, 16. [CrossRef]
37. Keras. Available Models. Available online: https://keras.io/api/applications/ (accessed on 4 July 2022).
38. Marques, G.; Agarwal, D.; de la Torre Díez, I. Automated medical diagnosis of COVID-19 through EfficientNet convolutional neural network. *Appl. Soft Comput.* **2020**, *96*, 106691. [CrossRef] [PubMed]
39. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 26 June–1 July 2016; pp. 770–778.
40. Zhang, C.; Patras, P.; Haddadi, H. Deep learning in mobile and wireless networking: A survey. *IEEE Commun. Surv. Tutor.* **2019**, *21*, 2224–2287. [CrossRef]
41. Kotla, J.R.; Narayan, A. Power Quality Classification Dataset—1. Available online: https://www.kaggle.com/datasets/jaideepreddykotla/powerqualitydistributiondataset1 (accessed on 18 May 2022).
42. Tharwat, A. Classification assessment methods. *Appl. Comput. Inform.* **2020**, *17*, 168–192. [CrossRef]
43. Al-Sabaawi, A.; Ibrahim, H.M.; Arkah, Z.M.; Al-Amidie, M.; Alzubaidi, L. Amended convolutional neural network with global average pooling for image classification. In Proceedings of the International Conference on Intelligent Systems Design and Applications, Virtual. 12–15 December 2020; pp. 171–180.
44. Dogo, E.; Afolabi, O.; Nwulu, N.; Twala, B.; Aigbavboa, C. A comparative analysis of gradient descent-based optimization algorithms on convolutional neural networks. In Proceedings of the 2018 International Conference on Computational Techniques, Electronics and Mechanical Systems (CTEMS), Belgaum, India, 21–22 December 2018; pp. 92–99.
45. Sun, S.; Cao, Z.; Zhu, H.; Zhao, J. A survey of optimization methods from a machine learning perspective. *IEEE Trans. Cybern.* **2019**, *50*, 3668–3681. [CrossRef]

46. Montalbo, F.J.P.; Alon, A.S. Empirical analysis of a fine-tuned deep convolutional model in classifying and detecting malaria parasites from blood smears. *KSII Trans. Internet Inf. Syst. (TIIS)* **2021**, *15*, 147–165.
47. Yi, D.; Ahn, J.; Ji, S. An effective optimization method for machine learning based on ADAM. *Appl. Sci.* **2020**, *10*, 1073. [CrossRef]
48. Sutskever, I.; Martens, J.; Dahl, G.; Hinton, G. On the importance of initialization and momentum in deep learning. In Proceedings of the International Conference on Machine Learning, Atlanta, GA, USA, 16–21 June 2013; pp. 1139–1147.
49. Hinton, G.; Srivastava, N.; Swersky, K. Neural networks for machine learning lecture 6a overview of mini-batch gradient descent. *Cited On* **2012**, *14*, 2.
50. Hossin, M.; Sulaiman, M.N. A review on evaluation metrics for data classification evaluations. *Int. J. Data Min. Knowl. Manag. Process* **2015**, *5*, 1–11.
51. Kurt, S.; Yildiz, H.U.; Yigit, M.; Tavli, B.; Gungor, V.C. Packet size optimization in wireless sensor networks for smart grid applications. *IEEE Trans. Ind. Electron.* **2016**, *64*, 2392–2401. [CrossRef]