Indonesian Journal of Electrical Engineering and Computer Science
Vol. 38, No. 1, April 2025, pp. 1~21
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v38.i1.pp1-21

A review of convolutional neural networks for classifying power quality problems using Keras API

Adamu Sa'idu[1,2], Nasiru B. Kadandani[1], Bala Boyi Bukata[1]
[1]Department of Electrical Engineering, Faculty of Engineering, Bayero University, Kano, Nigeria
[2]Department of Electrical/Electronics Technology, School of Secondary Education (Technical), Federal College of Education (Technical) Bichi, Kano State, Nigeria

Article Info

Article history:

Received Jul 4, 2024
Revised Nov 3, 2024
Accepted Nov 23, 2024

Keywords:

Convolutional neural
Deep learning
Distribution corridor
Keras API
Networks
Power quality problems

ABSTRACT

The major causes of electric power quality (PQ) problems are mainly due to the increased utilization of nonlinear loads, capacitor and load switching events, transformer energization, and occurrence of assorted faults at the distribution corridor. The problems often introduce harmonics and other waveform anomalies like voltage sags, voltage swells and interruptions along the power systems. A timely classification of such problems is important in understanding their impact on costly power system economy. The paper explores comprehensive review of PQ issues, operational concept of convolutional neural network (CNN) and its utilization in solving PQ problems. Novel deep learning (DL) approach using variant of DenseNet CNN technique in Keras API platform is deployed to extract the features of, and classify PQ problems. The proposed technique improves classification performance with an accuracy of 99.96%. It shows remarkable improvement over the traditional techniques in the literature which were 73.53% to 99.92% accurate for a period from 2018 to 2023. The most promising part of the method is the improvement shown in the classification performance when compared with that obtained in the literature. The technique can also be applied in real time to cater for real PQ problems.

Corresponding Author:

Adamu Sa'idu
Department of Electrical Engineering, Faculty of Engineering, Bayero University
P. M. B. 3011, Kano, Nigeria
Email: adamugiwa77@gmail.com

1. INTRODUCTION

Power quality (PQ) of an electrical power system network is affected with number of issues [1] that may cause malfunctioning, damage of appliances and even fund losses. According to the IEEE STD 519-1992 document, PQ has been identified on the idea of earthing electronic components in a way that is appropriate to the performance of that equipment and agreeable with the premise wiring system and other connected devices [2]. Other PQ definitions have also been advanced, see for example [3]-[6]. The concept of good or bad PQ depends on the customer's perception. Conversely, if the equipment operation is satisfactorily, the customer labels the power as that of good quality, otherwise it is of bad quality [7]. Between the two extremes, several grades of PQ may exist, again depending on the customer's perception. These grades may be as a result of the escalation of nonlinear devices within the power system, unprecedented faults, as well as power outages. As a result, most electronics equipment and other sensitive devices at the customer's loads tend to either shut down or get completely damaged, thereby causing serious dissatisfaction to the customer. The PQ disturbances affect both the customer as well as the utility by, for example, causing transmission line losses and shutting down of generators. Understanding and tackling PQ events has hence become paramount in improving service quality at customers and utility's distribution corridors. It has been reported in [8] that nonlinear related PQ problems directly affect sensitive distribution system equipment like electric drives, programmable logic controllers and single-loop controllers. Therefore, an envisaged timely identification and classification of such problems is reconsidered in this paper.

In view of this, conventional ways in dealing with PQ issues involve feature extraction method as well as classification process using signal synthesis analysis. Such techniques employed for feature synthesis in literature include fourier transform (FT), short time fourier transform (STFT), wavelet transform (WT), S-transform (ST), wigner ville distribution (WVD), empirical mode decomposition (EMD), independent component analysis (ICA) and variational mode decomposition (VMD) see for instant [9]. Most of these methods are associated with drawbacks for example; FT is simple but unsuitable for chaotic systems with variable disturbances largely because of its risk in time-frequency representation ability. However, STFT can use a sliding window to settle FT's time-frequency localization problem in chaotic systems. But it is restricted with the dimension of the sliding window to be further deployed. WT can improve the time-frequency resolution in PQ disturbance analysis however it is reactive to noise. ST combines STFT with WT and subdues some of the WT drawbacks. Hence, the ST is predominant to FT, STFT and WT approaches specifically in the noise-rich systems. Even though, the generic adaptation of ST restricted its application because of its complexities. Recently, WVD method has gained more attention as one of the analytical devices for nonstationary signals due to its high time-frequency resolution and high performance in the existence of noise again [9]. However, this method requires more time to transfer 1D to 2D image file which causes delays in the training process. The EMD being adaptive in nature is a time-frequency approach for synthesizing nonstationary signals which disintegrate signal into a finite number of intrinsic mode functions (IMFs) [10]. Also, it has some inherent problems like mode mixing and boundary affects that led to improper IMF decomposition ability. EMD-ICA technique was introduced to address these issues. The technique is efficient in removing the mode mixing effect but faces the lack of the amplitude information due to the included ICA. On the other hand, the VMD dismantles a multimodal signal to finite number of band-limited IMFs. When comparing EMD based approaches, the VMD is a more strong and reliable technique to noise as well as sampling errors which generalize a classical Wiener filter into multiples and adaptive bands. The signal processing approaches discussed so far are always employed to extract features from many types of PQ problems.

Furthermore, as soon as feature is extracted, PQ problems classification process starts using specific technique. For instance, in Cai *et al.* [9], PQ disturbances classification techniques include support vector machines (SVM), artificial neural network (ANN), probabilistic neural networks (PNN) and decision tree (DT). Amongst the techniques mention above, SVM is the most widely utilize method with little samples and structural risks involved are reduced. So also, DT is decision making technique in tree-like pattern graph used to highlight the relationship of various features that makes categorization of PQ problems easy [11]. However, SVM and DT techniques gave accumulative errors in the classification process of PQ problems. Tackling this issue, ANN-based classifiers are generally applied via efficient learning process. The approach eliminates the presence of iterations or accumulative errors. The PNN was obtained from Bayesian network and kernel fisher discriminant algorithm. The PNN was accepted to be quicker and more reliable than ANN and its stay alive methods comprising of feature extraction and classification techniques had been demonstrated to be fruitful [12]. Yet, it has three drawbacks firstly; feature extraction methods are not automatic, because various kinds as well as quantities of features have varying effects on the classification performances. Hence, accuracy of the classifier is often undermined as key features could be missed out. Secondly, feature extraction technique and classification stage are two individualistic activities yet variables of classification performances could be enhanced as per PQ signals analysis. Notwithstanding, variables of feature extraction are secured immediately when the operation is accomplished and limits accuracy of the classification results. Hence, the attributes of PQ problems cannot be reconditioned in the due course. Thirdly, conventional methods are shallow (little or slight depth) learning mechanism in nature. The classification performances are lower than deep learning (DL) techniques because the latter have deep layer network and big data support [13].

Later on, various techniques were developed to automatically detect and classifies PQ problems, particularly based on signal processing techniques [14]. Nowadays, few approaches have been put in place to come up with automatic classification of PQ problems utilizing voluminous data involving machine learning (ML) techniques directly. ML is a general term that refers to algorithms which learns from vast amount of data. Recently, ML has received much awareness due to the evolution of more promising algorithms, more training data availability as well as more computational resources globally [15]. It can also be applied for a wide area of applications such as credit-card fraud identification, speech recognition and medical diagnosis. DL point out the common type of ML techniques employed for learning discriminative characteristics from a given data in chronological order using assembled, layer wise structures. Amongst which are CNN, long short-term memory networks (LSTMs), convolutional autoencoders (CAEs) and LSTM autoencoders [16]. DL models demonstrated high anticipated capabilities in image and speech recognition, natural language processing (NLP), and intelligent gamification [17]. Although, there is existence of several excellent review papers in the field, the focus is not on timely feature extraction precision and accurate classification of PQ problems.

According to the aforementioned issues relating to DL capability, this paper is aimed at reviewing convolutional neural network (CNN) based DenseNet architecture for timely feature extracting and classifying PQ problems as one of the contributions in attempts to solve PQ problems at the distribution corner. In this paper, the application of DL technique particularly CNN based Keras API to automatically classify PQ problems will be comprehensively demonstrated. Even though, the proposed DL model has the capability of accurately classifying five different PQ problems is also expected to outperform other models brought forward in the literature. Model validation is also performed so as to authenticate the performance of developed approach. The sound contributions of this paper are: (1) The comprehensive review of CNN for PQ problems classification processes (2) In disparity to the current PQ signals analysis, CNN-Keras model is employed for the PQ problems classification (3) The proposed model of CNN based Keras has superior performances in image categorization (4) The time consuming stages in feature extraction, feature selection and data size minimization in traditional ML based algorithms are hugely reduced and (5) An adaptive moment estimation (Adam) optimization algorithm was employed to get the best hyperparameters for tuning CNN-Keras model.

Remaining parts of the paper were organized as follows. Section 2 reviews PQ problems and section 3 describes a comprehensive overview of the CNN. In section 4 applications of CNN in PQ problems classification were reviewed. Section 5 presents methodology of the proposed technique. Result and discussion are given in section 6, and section 7 highlights the current challenges of the CNN. Finally, some conclusions were drawn in section 8.

2. POWER QUALITY DEFINED

Theoretically, PQ is considered to be a multifaceted electromagnetic phenomenon that disturbs voltage and current signal from ideal waveform which is referred as the PQ problem. The term PQ encompasses any facet related to peak, angle and frequency of voltage and current waveshapes living in a power network [18]. Hence, bad or poor PQ may exist due to transient conditions in the power network or within the connection of nonlinear loads [19]. Utilization of more sensitive loads such as computers, industrial drives, telecommunications and medical equipment in power system network may also leads to PQ problems [20]. PQ problems include voltage sag or dip, voltage swell, power interruptions, voltage flicker, voltage surges, voltage spikes, switching transients, frequency variations, electrical line noise, brownouts, blackouts, notch as contained in Table 1.

For that reason, Anand and Srivastava [21] defined PQ problems as any difficulty displayed in relation to voltage, current or leading to frequency deviations which yields to failure or malfunction of customer appliances. Consequently, PQ problems have resulted in lost time, lost production, production of scraps, lost sales and conveyance delays as well as damaged production equipment. The sources of PQ problems can vary, ranging from natural phenomena such as lightning, floods and earthquakes to manmade induced like energization of capacitor banks and transformers, switching or start-up of large induction motor loads, operation of unsymmetrical non-linear loads, failure of distribution system equipment and wrong connections in distribution substations and consumer’s premises.

Table 1. Description of some PQ events classification, duration and voltage magnitude [22]

S/N	Category	Duration	Voltage magnitude
1	Voltage sag	0.5cycle – 1mins	0.1 – 0.9 pu
2	Voltage swell	0.5cycle – 1mins	0.1 – 1.8 pu
3	Interruption	0.5cycle – 1mins	<0.1 pu
4	Transients		
	a. Impulsive	50nsec – 1msec	
	b. Oscillatory	5µsec – 50msec	<0.8 pu
5	Overvoltage	>1min	1.1 – 1.2 pu
6	Undervoltage	>1min	0.8 – 0.9 pu
7	Voltage imbalance	Steady state	0.5 – 2%
8	Harmonic	Steady state	
9	Notch	Steady state	
10	Noise	Steady state	

From the foregoing discussion, the PQ problems have become much more complicated with proliferation of solid-state controllers [23], usage of which could not be overlooked because of their cost advantages, reduction in size, energy preservation, easy control, low wear and tear and other maintenance advantages they offer to the modern electric system [24]. Table 2 presents some commonly used mathematical expressions for parametric variations describing some PQ problems. However, the solid-state control devices, the customers load, as well as the generation system have all been identified as sources of PQ problems [25]. For this reason, PQ had become an important field of research in electrical engineering. Exceptionally, in radial distribution network (RDN) characterized by elevated power losses leading to high R/X ratio resulting in approximately 10 to 13% losses of the produced power seriously affect the system network [26]. This menace has posed serious challenges to both utilities and equipment manufacturers in meeting the customer's equipment PQ requirements as stipulated by the IEEE STD 519 of 1992. Conversely, several approaches have evolved for the reduction of PQ problems see [27] as an example.

Table 2. Expressions and parameter variations of various PQ problems [28]

PQ problem	Mathematical expression	Parameter variation
Pure sinewave	$y(t) = Asin(wt)$	$A = 1; w = 2\pi f$
Voltage sag	$y(t) = A\left(1 - \alpha\left(u(t - t_1) - u(t - t_2)\right)\right) \sin(wt)$	$0.1 \le \alpha \le 0.9;$ $T \le t_2 - t_1 \le 9T$
Voltage swell	$y(t) = A\left(1 + \alpha\left(u(t - t_1) - u(t - t_2)\right)\right) \sin(wt)$	$0.1 \le \alpha \le 0.9;$ $T \le t_2 - t_1 \le 9T$
Interruption	$y(t) = A\left(1 - \alpha\left(u(t - t_1) - u(t - t_2)\right)\right) \sin(wt)$	$0.9 \le \alpha \le 1;$ $T \le t_2 - t_1 \le 9T$
Transient	$y(t) = A\left(\sin(wt) + be^{-\gamma(t-t_1)} \sin\left(w_{tr}(t - t_1)\right)\right)$	$-2 \le b \le 2$ $50 \le \gamma \le 100$ $500Hz \le f_{tr} \le 15ooHz$
Harmonics	$y(t) = A(\alpha_1 \sin(wt) + \alpha_2 \sin(2wt) + \alpha_3 \sin(3wt) + \alpha_5 \sin(5wt) + \alpha_7 \sin(7wt))$	$\alpha_1 = 1.0$ $\alpha_2 - \alpha_7 = 0.0 - 0.3$
Flicker	$y(t) = A\left(1 + \alpha_f \sin\left(\beta_f wt\right)\right)\sin(wt)$	$0.01 \le \alpha_f \le 0.25$ $2Hz \le f_f \le 8Hz$
Noise	$y(t) = \sin(wt) + A(randn(\sin(wt))$	$A = 0.1$

Recently, many grids' stakeholders perform spontaneous PQ monitoring to get relevant data of the power being supplied and equipment performance. The time taken for long term PQ measurements resulted in huge data to be handling. Actual performance of PQ monitoring device relies on its capability to analyze and presents voluminous raw data obtained from monitoring exercise. But notwithstanding, it consumes more time and may not yields perfect result. Orthodox scientific tools are still required to speed up big data interpretation analysis with excellent result [29]. The PQ big data is nothing but huge amount of data sequel to continuous PQ monitoring [30]. Hence, big data is voluminous or large amount of data with specific complexities described by 4V's (i.e. Volume, Velocity, Variety and Veracity) depicted in Table 3.

Table 3. Description of 4Vs for PQ big data [29]

S/N	Parameter	Description
1	Volume	This is amount, size and scale of data that could not be managed without dedicated analytic tools.
2	Velocity	The speed at which data is generated and how fast the data should be processed is termed velocity.
3	Variety	This is heterogeneity of data being utilized. Big data always comes from various sources, which can be different in types, format, semantic and volume.
4	Veracity	Quality of collected data is refers to as veracity. It is concerned with biases, noise and abnormality in the data. Accuracy of any analytic process applied to the data depends greatly on the veracity of the source data.

3. CONVOLUTIONAL NEURAL NETWORK

Within the sphere of DL research, CNN technique is the most utilized straightforward algorithm [31]. Chief significant characteristic of CNN compared to its antecedent is that it automatically recognizes pertinent attributes in the absence of human supervision [32]. The CNNs had been amply employed in wide area of different fields, including computer vision, speech processing, face recognition, and image classification. Table 4 described evolution of CNN architectures and their characteristics over time.

Table 4. Evolution of CNN architectures over time [33]

S/N	CNN Type	Year	Characteristics	Dataset
1	LeNet	1998	It was the first popular CNN architecture, it has seven layers, it uses ReLU activation function	MNIST
2	AlexNet	2012	It is deeper and wider compared to LeNet, It uses RELU, dropout and overlap pooling GPUs NVIDIA GTX 580	ImageNet
3	VGG (Visual geometry group)	2014	It uses small-sized filters and has homogeneous topology	ImageNet
4	GoogLeNet	2015	It is the first CNN architecture to introduce block concept and different filter size. It uses split transform to merge idea	ImageNet
5	InceptionV-3	2015	It is able to take care of bottleneck issue and used small filters instead of large filters. It has better feature representation.	ImageNet
6	Inception V-4	2016	It uses asymmetric filters with split transform and integration concept	ImageNet
7	ResNet	2016	It is robust against overfitting due to symmetry mapping-based skip connections with residual learning	ImageNet, CIFAR-10
8	Xception	2017	It has depth-based convolution followed by point-based convolution	ImageNet
9	Residual Attention Neural Network	2017	It is the first CNN architecture to introduce attention mechanism e.g. Transformer-based CNNS and non-local neural network	CIFAR-10, CIFAR-100, ImageNet
10	ResNext	2017	It instituted pricipality, homogeneous topology and grouped convolution	CIFAR-10, CIFAR-100, ImageNet
11	Squeeze and Excitation Network (SENet)	2017	It modeled interdependencies between feature maps. Incorporates channel-wise attention to adaptively recalibrate feature maps. It enhances feature representation	ImageNet
12	DenseNet	2017	It has blocks of layers, layers connected to each other and Cross layer information flow	CIFAR-10, CIFAR-100, ImageNet
13	MobileNetV1	2018	It has inverted residual structure. Designed for mobile and embedded devices, depth wise separable convolution to reduce computation	ImageNet
14	HRNetV2	2020	High-resolution representation	ImageNet
15	DiceNET	2021	Initiated size-based CNN,incorporating height, width, and depth	

Therefore, CNNs had evolved resulting to various architectures tailored for different tasks in computer vision. Some of the notable CNN variants and their differences are identified in this work refer to Tables 4 and 5 for more information. These variants differ in terms of architecture, depth, skip connections and specialized features. Presently, researchers continue to explore new CNN architectures to improve performance and address specific challenges in computer vision tasks [34]. The architectural variants of CNN such as AlexNet, VGG, GoogLeNet and ResNet for example, are mostly employed in the classification of PQ problems signals. Nonetheless, the structures were facing hurdle in convergence, overfitting issue and vanishing gradient problems. Also, their accuracies and adequate performances criteria still requires improvement and some little modifications in the model. The very good variant to address all these problems is DenseNet. It is recommended for this work because of its simplicity and ease in application. DenseNet is common architecture of CNN for image recognition that has acquired current structure with low parameters having different networks such as DenseNet – 121, DenseNet – 160 and DenseNet – 201. Among the three networks, DenseNet – 121 is specifically recommended for this research because it can reduce the issue of vanishing gradient, advocates feature reuse and lessens parameter involvement, which are major factors for training DL models. Moreso, DenseNet – 121 had proven to be efficient in pinpointing PQ problems on the basis of signal analysis. It has a total of 121 layers for convolutional, transition, classification and dense block layers and is very popular DL architecture in image classification. The DenseNet CNN architecture uses ReLU activation function and dropout to demonstrate effectiveness in classification performance. The variants of CNNs were conspired with neurons in human and animal brains, very much alike to ANN [35]. In human brain, complex series of cells forms the visual cortex and this sequence is simulated by the CNN. DL, in its remarkable success, presently is one of the well-known research areas in the field of ML [36]. Outstanding differences between ML and DL approaches were depicted in Figures 1 and 2 respectively.

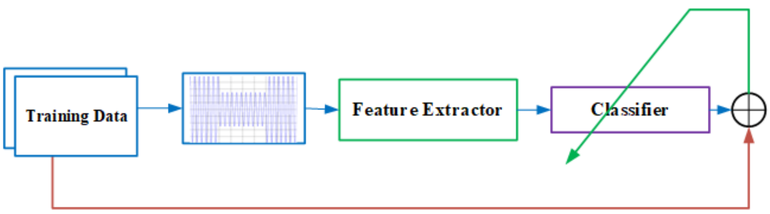

Figure 1. ML technique [37]

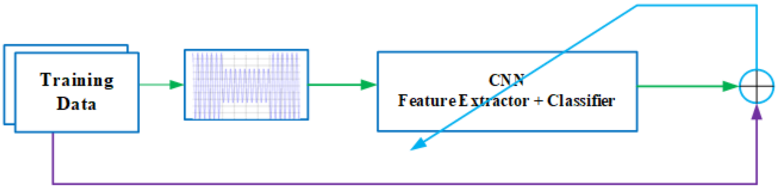

Figure 2. DL technique [37]

Some of the major benefits of CNN are equal descriptions, scattered interconnections and parameters sharing. Fully connected (FC) layer, shared weight and local interconnection of CNN structure was applied to make complete usage of 2D input data structures such as image signals. Nevertheless, CNN can manage PQ big data very fast and produce good results. This operation extremely utilizes small numbers of parameter which speeds up the training processes and accelerates model performance [38]. A common CNN type which is close to the multi-layer perceptron (MLP) consists of convolution layers, activation function, sub-sampling or pooling layers and FC layers as illustrated in Figure 3.

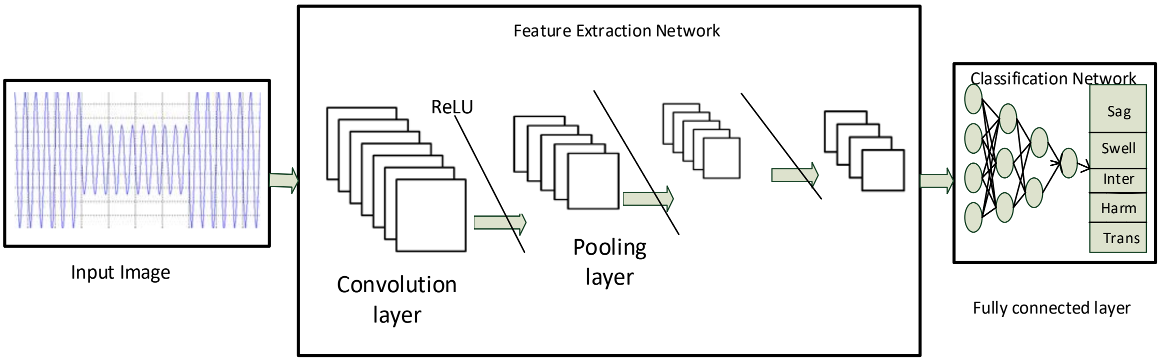

Figure 3. CNN architecture

3.1. Operation of CNN

The basic operation of CNN is presented based on the working principle of each layer in its composition as follows:

Convolutional layer: The most significant layer in CNN training process is the convolutional layer which contains collection of filters called kernels [39]. Input signal to the CNN is expressed as N-dimensional matrix, convolved with the filters to produce the output feature map. Usually, a grid of discrete numbers or values represents the kernel known as the kernel weight. Assigned arbitrarily values act as the weights of the kernel at the initial stage of the CNN training exercise. Also, there were many ways employed to initialize the weights [40]. Consequently, the kernel weights are adjusted at every training time to extract significant features. This filter is also known as feature detector. In (1) describes convolution operation in simplified form. Figure 4 describes an example of convolution operation of 5 by 5 gray – scale image with 3 by 3 random initialized weight kernel that slides with the input image horizontally and vertically to produce the feature map.

$$Y_j^k = \sum_i (W_{ij}^k * x_j^k + b_j^k) \tag{1}$$

where Y_j^k is output feature map, x_j^k input feature map, W_{ij}^k is set of 2D filters and b_j^k is trainable bias parameter.

However, in order not miss or loses some vital information at the extreme edge of the input image, a padding technique is applied. This is highly important in determining border size data in relation to input signal which when employed, dimension of the input image will increase and consequently, the dimension of the output feature map will also increase as described in Figure 5.

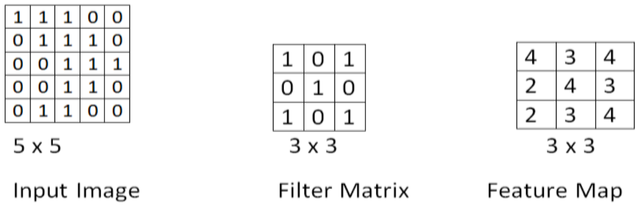

Figure 4. Convolution operation of input image with filter

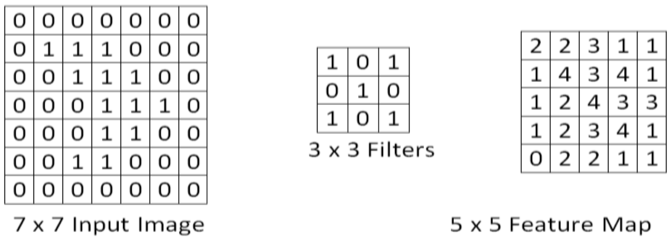

Figure 5. Convolution operation using padding technique

Activation function: Mapping input data to output is the principal operation of all kinds of activation function in neural network. That's to say activation function makes decision as to either or not to fire a neuron in reference to specific input signal by producing the correct output. The Activation function introduces nonlinearity to the output matrix [41]. It also has the ability to distinguish distinctly dominant feature which allows error back-propagation to be applied to train the network. Sigmoid, Tanh and ReLU together with its variants like leaky ReLU, noisy ReLU and parametric linear unit are commonly used activation functions in CNN and other deep neural networks. The most commonly applied activation function is ReLU because it changes all the values of inputs into positive numbers and has lower computational loads when compared with other activation functions.

$$f(x) = \max(0, x) \tag{2}$$

Pooling layer: This layer indicates movement of two – dimensional filter with each channel of feature map thereby summarizing the features lying within the range operated by the filter [42]. Major operation of this layer is the subsampling of feature maps [43]. It also downsizes or shrinks large size feature map to create smaller maps while maintaining major part of superior information (or features) in each and every steps of the pooling operation [44]. Many pooling techniques are available for utilization in the pooling layers. The most commonly used methods are tree pooling, gated pooling, average pooling, min pooling, max pooling, global average pooling (GAP), and global max pooling. The familiarized and repeatedly applied pooling techniques are max and average pooling again in [44]. For the PQ problems waveforms average pooling is applied because it is more sensitive to noise signal [45]. Max pooling approach has better performance capability than average pooling method. Hence, it is utilized in this work. Figure 6 illustrates these two pooling operations.

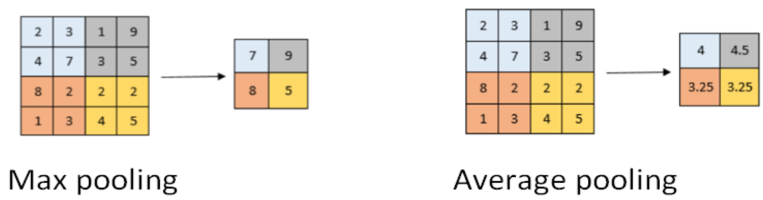

Figure 6. Max and average pooling operation

Many at times, the overall performance of CNN is reduced due to pooling operation. This indicates the major shortfall of this layer as it helps to estimate availability of certain feature in the input data and also pays more attention entirely on finding correct position of that feature. Hence, CNN model misses out very important information which leads to the application of padding technique to retain the information. The whole series of convolution operation, nonlinearity function and pooling exercise is repeated number of times to obtain the flatten vector before moving to the final layer which is FC layer for classification.

Fully connected layer: Usually FC layer is positioned at the end of CNN architecture. Within the layer, each neuron has a linkage to all neurons of the subsequent layer (hence name FC layer). It is described as the CNN classifier which mimic the basic method of multiple-layer perceptron neural network in feed-forward ANN. Input to FC layer comes from the last pooling technique or convolutional layer. The input to the FC layer is in vector form which is generated from the feature map after flattening. The output of the FC layer gives the final CNN output as illustrated in Figure 3. Loss functions are employed in the output layer to estimate the predicted error produced across training samples in the CNN model. In (3) clearly describes this layer.

$$Z = f(W^T q + b) \tag{3}$$

where q and Z are input and output respectively, W describes matrix with connections weights and b denotes bias term vector.

3.2. Loss function

So far, the previous sections have described distinguished layer types in CNN architecture. Moreso, classification is successful which represents the final layer of the CNN architecture. The loss or cost functions are applied in the output layer to approximate the predicted error produced during training in the CNN model. The produced error reveals variation between actual outputs and predicted one. Many loss functions such as cross – entropy or SoftMax function, Euclidean loss function and hinge loss function were employed in different CNN applications as explained in [45]. SoftMax activation function is the commonly utilized function for measuring CNN performance. It is also known as log loss function and has output probability of p $\in \{0, 1\}$. This output layer employs the SoftMax activation to generate the output within the probability distribution.

3.3. CNN regularization technique

In the CNN model analysis, over-fitting signifies the key issue involved in developing good generalization. A model is considered be over-fitted, under-fitted or just-fitted as illustrated in Figure 7. It is just - fitted if it operates good on training and testing data. Different intuitive aspects such as drop-out, drop-weights, data augmentation and batch normalization were occasionally applied to help regularization in order to avoid over-fitting.

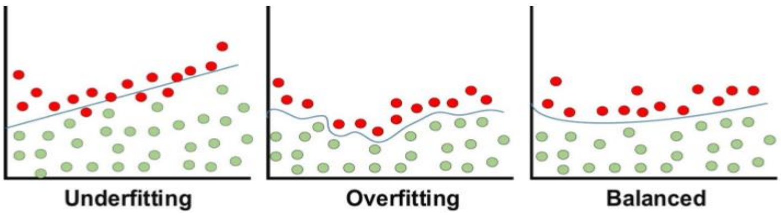

Figure 7. Description of over – fitting, under – fitting and just - fitted issues

Application of batch normalization method guarantees good performance of the output activation function due to the unit Gaussian distribution behaviours. Also, subtracting mean from a given input and dividing by standard deviation normalizes the output at each step which mostly prevents the problem of vanishing gradient from rising.

3.4. CNN optimization techniques

Optimization techniques are CNN learning processes. Two major things that involved in the learning process are the learning algorithm selection (optimizer) and the application of many enhancement techniques (such as AdaDelta, Adagrad, and momentum). The latter along with the learning algorithm improves the output of the training result. The network parameters shall always be updated through every epochs so as to lessens the error [46], [47]. In order to reduce the impact of the learning error, the algorithm repeatedly updates the network parameters at each and every iteration [48]. Moreover, in order to upgrade the parameters rightly, there is need to estimate the gradient function (slope) by utilizing first-order derivative with respect to the network characteristics. Again, parameters are upgraded in the other direction of the gradient to lessen the learning error. Various gradient based learning algorithms such as batch gradient descent (BGD), stochastic gradient descent (SGD), mini – batch gradient descent (MBGD), momentum and adaptive moment estimation (Adam) are available and commonly employed as highlighted in [49]. Adam is a learning approach designed particularly for training deep neural networks.

3.5. CNN model architecture

In order to obtain good result, selection of CNN architecture is another very critical issue for enhancing performance of various CNN designs. Many adjustments have been put in place for CNN architectures recently. Specifically, the new modifications in CNN architectures were achieved based on the utilization of network depth. Table 5 presents brief overview of most popular CNN architectures, starting from AlexNet model in 2012 and ending at the high resolution (HR) model in 2020. Considering the architectural features (such as input size, depth, and robustness) is the main aspect in assisting engineers or researchers to choose the most appropriate architecture for their propose application.

Table 5. Overview of most popular CNN architectures [17]

Model	Depth	Dataset	Error rate	Input size	Year
AlexNet	8	ImageNet	16.4	227 x 227 x 3	2012
VGG	16, 19	ImageNet	7.3	224 x 224 x 3	2014
GoogLeNet	22	ImageNet	6.7	224 x 224 x 3	2015
ResNet	152	ImageNet	3.57	224 x 224 x 3	2016
DenseNet	121, 160, 201	CIFAR10, CIFAR100, ImageNet	3.46, 17.18, 5.54	224 X 224 X 3	2017
MobileNet-v2	53	ImageNet	-	224 x 224 x 3	2018
HRNetV2	-	ImageNet	5.4	224 x 224 x 3	2020

4. APPLICATION OF CNN IN CLASSIFYING PQ PROBLEMS

This section portrays a literature survey on the classification of PQ problems using DL approaches. In a study of DL approach for detection and categorization of PQDs with windowed signals described using voltage signals only is conducted in [1]. The approach uses overlapped windowed signals with different SNR which performs satisfactorily with a value above 97% accuracy, even though the operation of the classifier reduces as the noise interference in the input improves. The use of Wigner – Ville distribution via deep convolutional neural networks (WVD – CNN) for identification of PQDs is proposed by [9] whereby the WVD was employed to transport 1D voltage disturbance signals into 2D voltage image files and CNN based model was developed and processed using image data to obtain optimized parameters for PQDs classification. Nine types of synthetic signals and three real world measurements were employed to test the model which gives best classification accuracy of 99.67%. The result obtained ensures the effectiveness of the proposed model. Research on novel technique for multiples PQ disturbances classification employing multi – task CNN (MT - CNN) approach was developed to actualize multi – label classification of multiple PQDs [11]. The method extracts more significant features and yields better recognition rate of 94.63% and it has very strong capability to resist overfitting issue. The technique could largely improve accuracy rate for representing multiple PQDs under different signal to noise ratio conditions.

Another research on PQDs monitoring and classification employing improved principal component analysis (PCA) and CNN for wind grid distribution systems was presented whereby the statistical features were extracted via improved PCA (IPCA) whereas features like mean, standard deviation, energy were extracted using 1D – CNN. The method classifies PQDs with maximum classification accuracy of 99.92% which was tested with noise and noiseless environment [12]. Also, a study of PQ disturbance classification incorporating compressed sensing and deep convolutional neural network (CS – DCNN) was suggested whereby data is compressed to reduce the requirement for acquisition device memory which increase the transmission rate [15]. The deep CNN was used for feature extraction and classification without any delay and data pre – processing operation. The model indicated good classification performance in noise data with accuracy of 99.50% and a loss of 0.02.

Another work in the same year by [17] proposed identification of PQ disturbances (PQDs) using phase space recognition with convolutional neural networks (PSR – CNN). The PSR transform 1D voltage signal into 2D voltage image file and the CNN does the classification automatically with high accuracy of 99.80%. The performance showed that the model is able to produce more improves results without human involvement when compared to existing methods. Ekici *et al.* [20] introduced optimized Bayesian CNN for

PQ problems classification. The result proved that the proposed approach superseded some ML algorithms such as decision trees and random forests regarding accuracy. The whole accuracy attained by the method was reported to be 99.80%. The approach was evaluated using publicly available PQ dataset. Rodriguez *et al.* [38] presented another study on PQ disturbance classification via deep convolutional auto – encoders and stacked LSTM RNNs. In this research, a strong algorithm that contains deep CAE and stacked LSTM RNNs was used. High classification accuracy of 98.7% was reached and training time was reduced from 5 seconds to 3 seconds per training iteration. The usage of DL with 2D wavelet scalograms for PQD classification was described by [45]. Continuous wavelet transform (CWT) was used to produce scalogram of 2D that express signal pattern of PQ event through time – frequency representation. CNN is employed to categorize the data in accordance to the image disturbance with high accuracy of 97.67% in noiseless signals. Xue *et al.* [47] developed a deep CNN with spectrogram using microgrid for PQ disturbance classification. Spectrogram technique was used to restructure the PQDs signals and the CNN used for classification. The method divides the PQDs into different semantic features for detecting single and multiples signals over contrary time scales in a sample. The method proved to be robust to noise with high classification accuracy of 99.60% but merging of CNN and RNN will improve the classification performance.

Research conducted by [48] described how signal processing and DL techniques were used for PQ problems managing and classification. The study introduces innovative method employing compressive sensing (CS), singular spectrum analysis (SSA), WT and DNN for monitoring classification of single and combined PQ disturbances (PQDs). The SSA – CS – DNN algorithm proves to be bet way of PQDs classification with high accuracy of 99.85%. Ahajjam *et al.* [50] presented research on electrical PQDs classification using temporal – spectral images with deep CNNs. The study described new approach of PQD detection and classification technique involving fusing temporal with spectral images and deep CNNs (FTSI – CNN). The technique reduces feature dimension while retaining time – frequency information to attain good performance accuracy of 99.67%. Good FTSI design will further improve accuracy while maintaining its low complexity. Mohammadi *et al.* [51] conducted a research of PQ disturbances classification via full – convolutional Siamese network and k – nearest neighbor. In the study, combined algorithm of k – NN and fully – convolutional Siamese were put in place to classify PQDs by learning small samples with higher than 80% accuracy. Multitude convolutional layers and connection layers are there to develop Siamese network and output reaction judge's category of the signal. To ensure high enough accuracy, the sampling frequency should be more than 1275 Hz. Liu *et al.* [52] presented a complex PQ disturbance classification using curvelet transform and DL technique. In this novel approach, SSA, curvelet transform (CT) and deep CNNs were applied to sense and categorize PQDs. Good classification result was obtained and compared to SVM and other classifiers in which the current technique superseded in terms of accuracy with 99.52%.

Another new approach for PQDs classification through sparse autoencoders (SAE) based on DNN was presented to extract features and classify PQDs [53]. A good feature extraction performance was realized via SAE and high classification accuracy of 93% was attained using DNN. Manan *et al.* [54] developed another DL approach in the field of PQ disturbance classification where CWT was used to generate the coefficient matrix and later the coefficients were changed to image file using feature matrix and given to CNN as input for classification. Higher classification accuracy of 99.60% and noise immunity were recorded in this technique. Mishra *et al.* [55] described the power of their approach in correctly detecting and classifying multiple PQDs using temporal DL. Encode – decode temporal CNN (EDT - CNN) technique that merges feature selection with classification in a single block was employed. A good classification performance of 99.52% was realized in noisy data. Ramalingappa and Manjunatha [56] developed hybrid approach with complex wavelet phasor model and customized CNN to represent PQ issues. The dataset involved in the study were collected from the power grid in India. The method attained good accuracy of 99.33% in classifying PQ problems. Mohan *et al.* [57] proposed another architecture called Deep Power for PQ disturbances classification. The proposed method was evaluated on the publicly available UCI Electric PQ dataset and the result showed remarkable achievement with high accuracy of 99.71%. Sahani and Dash [58] recommended the application of FPGA based deep CNN for PQ event identification. The proposed approach was evaluated using synthesized and experimental data collected from process adaptive VMD data. Overall accuracy obtained from this method is reported as 96.75%. Qiu *et al.* [59] developed a different method using multifusion CNN based automatic classification framework of complex PQ disturbances. The method fused information based on the PQ event's time domain and frequency domain features. The method is evaluated on a dataset collected from PQ monitoring system and the result showed remarkable achievement with overall accuracy of 98.46%. Yigit *et al.* [60] presented an automatic detection model of PQD using CNN structure with gated recurrent unit. A matrix of PQD signals was obtained using STFT which contains reconstruction of signal in time and frequency domain which are good input for CNN. Characteristics are automatically distilled using CNN without any preprocessing technique and the GRU is employed to classify features. The function of the method is tested on a dataset which contains a total of seven single and combine defects with high classification accuracy of 98.44%. Wang *et al.* [61] proposed a novel approach via symmetrized dot pattern (SDP) algorithm plus CNN (SDP – CNN) for PQ fault identification. The voltage variation was extracted and introduced into SDP for fault recognition using CNN. The PQ problem was accurately recognized by the CNN with high accuracy rate of 94.20%. Aziz *et al.* [62] proposed novel CNN approach for fault classification in photovoltaic arrays. This approach extracts 2D CNN features from 2D scalogram created by PV system data so that it can efficiently detect and classify system faults. The method achieved high detection accuracy of 73.53% and outperforms existing techniques.

Another novel technique was introduced to classify PQD using segmented and modified S – transform (SMST), DCNN and multiclass SVM (MSVM) [63]. SMST is employed to analyze the PQDs and obtained 2D contour maps. The DCNN automatically extracts features using 2D contour maps and lastly MSVM identifier does the PQD classification. After performing extensive simulations, the method showed better performance with 98.86% extracting accuracy which displayed the model effectiveness and robustness on 21 PQ waveforms.

In brief, results reported in this review showed that CNN-based models could attain high classification accuracy of PQ problems as indicated in Table 6. The accuracies achieved by these models ranged from 73.53% to 99.92% within a period of six years (2018 to 2023) depend on the dataset and technique used. The method proposed in this work could still be used to PQ problems analysis to automate and improve timely feature extraction precision with high efficient classification performance. This determines the mitigation technique to be applied to control the classified PQ problems.

Various DL libraries that supply in - built classes of neural networks that have rapid numerical calculation and automatic estimation of gradients for both CPU and GPU are detailed in [64]. Specifically, the DL networks libraries and their sources are given in Table 7. They are the platforms used for DL implementation.

Table 6. Comparison of various techniques used for solving PQ problems

S/N	Method	FE Tech	No. of PQD	Accuracy	Year	Reference
1	WS – DL	Manual	15	97%	2021	[1]
2	WVD – CNN	Automatic	12	99.67%	2019	[9]
3	MT – CNN	Automatic	7	94.63%	2019	[11]
4	IPCA – CNN	Manual	12	99.92%	2019	[12]
5	CS – DCNN	Manual		99.50%	2019	[15]
6	PSR – CNN	Manual	10	99.80%	2019	[17]
7	OB – CNN	Automatic		99.80%	2020	[20]
8	CAE – SLSTM – RNN	Automatic	9	98.70%	2020	[38]
9	CWT – Scalogram – CNN	Manual		97.67%	2023	[45]
10	SPECTOGRAM – CNN	Manual	13	99.60%	2020	[47]
11	CS – SSA – WT – DNN	Manual	15	99.85%	2019	[48]
12	FTSI – CNN	Manual	9	99.67%	2020	[50]
13	FCSN – KNN	Manual	7	80%	2019	[51]
14	SSA –FDCT – DCNN	Manual	31	99.52%	2018	[52]
15	SAE – DCNN	Automatic	7	93%	2021	[53]
16	CWT – CNN	Manual	16	99.60%	2019	[54]
17	EDT – CNN	Manual	29	99.52%	2019	[55]
18	CWP – customized CNN	Manual		99.33%	2022	[56]
19	DL – Based Approach	Automatic		99.71%	2018	[57]
20	FPGA - CNN	Manual		96.75%	2020	[58]
21	MF – CNN	Automatic		98.46%	2020	[59]
22	STFT – CNN	Manual	9	98.44%	2021	[60]
23	SDP – CNN	Manual		94.20%	2023	[61]
24	Scalogram – CNN	Manual		73.53%	2020	[62]
25	SMST – DCNN – MSVM	Manual	21	98.86%	2023	[63]
26	CNN – FLC – DSTATCOM	Automatic	5			Proposed Method

Table 7. DL networks libraries and their sources [65]

S/N	DL Library	Source	Flexibility	Efficiency
1	PyTorch	https://pytorch.org/	High	Medium
2	TensorFlow	https://www.tensorflow.org/	High	High
3	MatConvNet	http://www.vlfeat.org/matconvnet/	Low	High
4	Keras	https://keras.io/	Low	High
5	Theano	http://deeplearning.net/software/theano/	Medium	Medium
6	Caffe	https://caffe.berkeleyvision.org/	Low	High
7	Julia	https://julialang.org/	High	High

5. METHODOLOGY OF THE PROPOSED CNN–KERAS API MODEL

The need to have optimal PQ problems classification at the distribution level based DenseNet CNN architecture is of paramount importance for proper and stable power supply and utilization. This DenseNet architecture targets to recognize PQ problem using densely connecting layers. High performance computing, big data management and high accuracy motivates the application of this CNN technique as it learns features and tasks directly from a given data. It provides a simple communication model for enhancing information flow between layers. Thence, CNN based on DenseNet provides option for correct PQ problems image categorization as well as capability of mitigating vanishing gradient problem. This novel method of DenseNet CNN – Keras architecture automatically extract features, classify and control the PQ problems via distribution static compensator (DSTATCOM) is proposed for the task and depicted in Figure 8. The method composed of two parts in which part one is the classification whereby the CNN based Keras API is utilized for that purpose and the second part is the control of the classified PQ problems. The technique could be employed for the generation of synthetic data so as to magnify the volume of the training dataset [66]. Fuzzy logic control based DSTATCOM is employed to do the mitigation. Therefore, this work is targeted to categorize PQ problems based on the proposed CNN classifier.

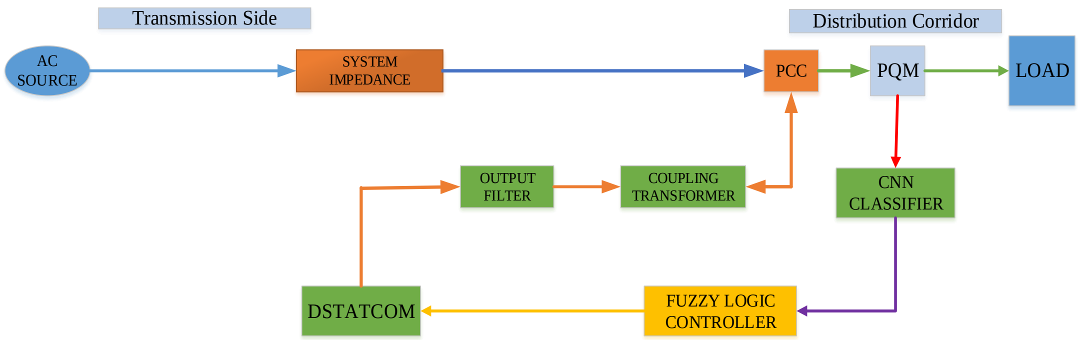

Figure 8. Complete block diagram of CNN – FLC - DSTATCOM approach

5.1. Data generation

For PQ problems analysis, normally disturbance signals are obtained either by synthesis, simulation, laboratory/experimental set up, real grids or public dataset for further analysis. Real time PQ problems signals are not easy to capture. For this reason, therefore in this work, five types of PQ problems are blended in MATLAB using mathematical models according to the IEEE-1159 [67] as indicated in Table 2. PQ problems are simply produced and appear very close to actual situation. The major significances of utilizing parametric equation are that it is easy to modify signal parameters in a vast realm and in a desired manner and it is also simpler to get the samples in a large quantity. The PQ signals produced by mathematical models could be simply utilized in the classification of PQ problems to extract their discriminative features.

5.2. Data preprocessing

The amplitude of the synthesized PQ signals is standardized to 1 p.u. The basic fundamental frequency is tuned at 50 Hz and sampling frequency of synthesized signal is equally set at 6.4 kHz. Ten (10) cycles were sampled every synthesized signal and each cycle contains 128 sampling points totaling 1280 sample point signals. The amplitude of the synthesized PQ signals is then presented in form of matrix. Every image has a dimension of 200 x 200 input images to the CNN model.

5.3. Development of CNN training model

The developed DenseNet CNN variant has one input layer, three convolutional layers, three max-pooling layers, FC layer and output layer as illustrated in Figure 9. Numbers of convolution filters in the three different convolutional layers are 32, 64 and 64 respectively are involved in the development process. Features of PQ problems signals are extracted utilizing convolution operator which maintains the spatial features of the input matrix. In a nutshell, kernel moves over an input image using (4) to generate feature map using the stride size of 1.

$$C_{ij}^{k} = \sum_{m=0}^{2} \sum_{n=0}^{2} W_{m,n}\, X_{i+m,j+n} + b^{k} \tag{4}$$

where W is weight of kernel, X is input image, b is a bias term, C is result of convolution operation, k is number of kernels, i,j and m,n are location labels of original image and convolution kernel matrices respectively. This proposed CNN model is executed in Keras platform to obtained results.

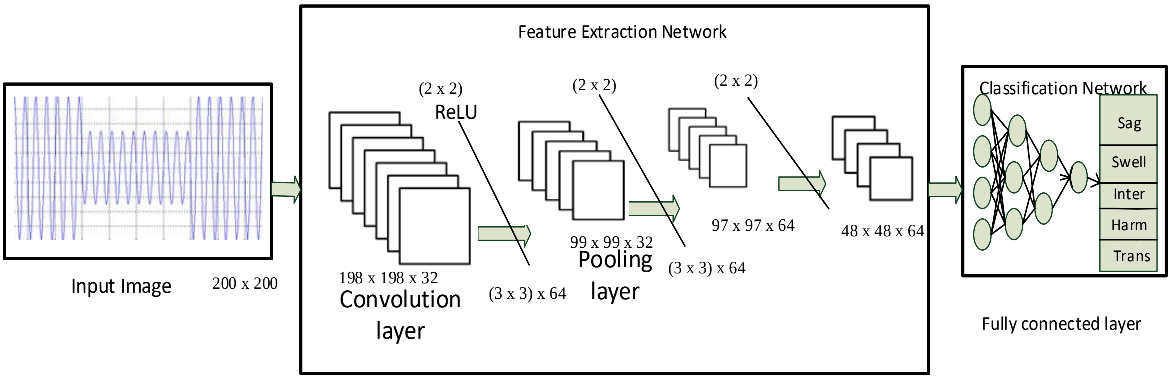

Figure 9. Structure of the proposed model

5.4. Operational architecture of the CNN model

To ensure the proper procedure of feature analysis, zero padding approach is employed to retain information of input signal. Progressively, convolution exercise secures local dependencies in the image and various feature maps are produced by various kernels. Then, the activation function operation is conducted at the end of each convolutional layer. This introduces nonlinearity in CNN model for learning nonlinear property of PQ problem data. However, rectified linear unit (ReLU) is adapted as activation function because of its quick training speed and lack of gradient vanishing problem. Next pooling layer is adduced to the ReLU function in order to reduce dimension and number of parameters of the network. This decreases the time for training and eradicates overfitting effectively. Output of pooling layer is then flatten and presented to FC layer in vector form. Lastly, this layer utilizes softmax activation function to approximate the classification result which is then fed to final output layer. Table 8 illustrates complete structure of proposed CNN model. Sequential operation of 2D Conv + ReLU + Maxpooling is repeated number of times to obtain the discriminative features before taking to FC layer for classification. This function consists of three major operations that is batch normalization (BN), activation function (ReLU), pooling and convolution (CONV). The output or feature map of each step is obtained using (5).

$$FM = N - f + 1 \tag{5}$$

Where FM is feature map, N is dimension of the input matrix and f is filter or kernel.

Table 8. Architecture of the proposed CNN model

Layers	Filter/Kernel size	Input size	Output size
2D Conv1 + ReLU	(3, 3) x 32	200 x 200	198 x 198 x 32
2D Conv2 + ReLU + Maxpooling	(3, 3) x 64 (2 x 2)	198 x 198 x 32	99 x 99 x 32
2D Conv3 + ReLU + Maxpooling	(3, 3) x 64 (2 x 2)	99 x 99 x32	97 x 97 x 64
FC1 + SoftMax	64	48 x 48 x 64	46 x 46 x 64
FC2 + SoftMax	64	46 x 46 x 64	23 x 23 x 64

5.5. Implementation of the CNN – Keras model

The developed CNN model is taken to keras API platform for simulation in which dataset is loaded and divides into training, validation and testing sets. Keras API Anaconda Navigator through python and Jupyter notebook software is used to train the model. It is high – level DL application programming interface (API) introduced by Google for executing neural networks. It is documented in Python programming language which makes the execution of neural networks easy and supports multiple backend neural networks computation. Therefore, the proposed classification model is trained in few steps after which the Keras API libraries was installed through anaconda. Firstly, anaconda navigator is opened and Jupyter notebook is lunched then Python 3 is selected in the pop – up menu appeared at the righthand corner of the notebook from which the pretrained model written python code was loaded and simulation starts. To train the CNN – Keras model the following steps were adopted.

a) Identify CNN model architecture (number of hidden layers, number of neurons, various activating functions and so on)
b) Compile() method to set model's optimizer and metrics
c) Provide Fit() method to train the model
d) Execute Evaluate() method to evaluate the trained model
e) Predict () method to make predictions.

These steps are implemented by uploading required libraries and modules, reading data as well as conducting basics data checks, creating arrays for features and response variable as well as creating the training and test datasets.

6. RESULTS AND DISCUSSION OF THE PROPOSED MODEL

The PQ problems were generated and synthesized in MATLAB using mathematical models. The PQ problems generated appeared to be very much similar to the actual situations as shown in Figure 10 which makes it easy to modify signal parameters in vast range and in desired manner and also obtain large quantity of data easily.

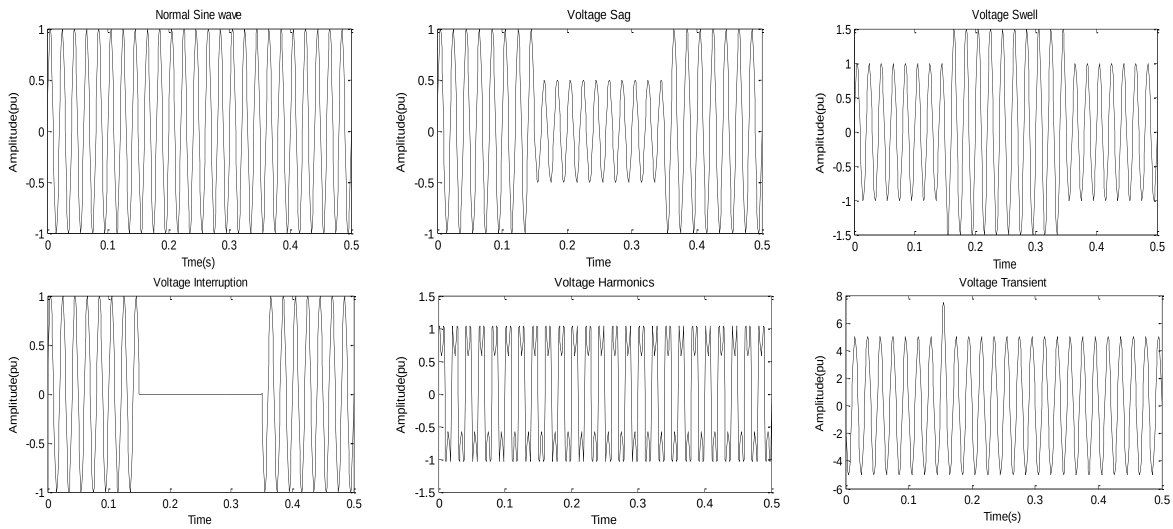

Figure 10. Samples of PQ problems generated in MATLAB

6.1. Training samples of the CNN - Keras approach

The training data for CNN model is divided into three percentages so as to obtain good result. Detailed description of the sample signals based on percentage of training, validation and testing is shown in Table 9. Within five PQ problems classified, there are 2560 sample signals for each problem where 60% of the signals were utilized for training and remaining 40% was used for both validation and testing sets.

A classification model is established for training in Keras API using python codes which results in obtaining good classification performance. Figure 11 described classification result of proposed DenseNet CNN based Keras model.

Table 9. Description of PQ sample signals used in the model

Class category	PQ problem	Training (60%)	Validation (20%)	Testing (20%)
Type_1	Voltage Sag	1536	512	512
Type_2	Voltage Swell	1536	512	512
Type_3	Interruption	1536	512	512
Type_4	Transients	1536	512	512
Type_5	Harmonics	1536	512	512

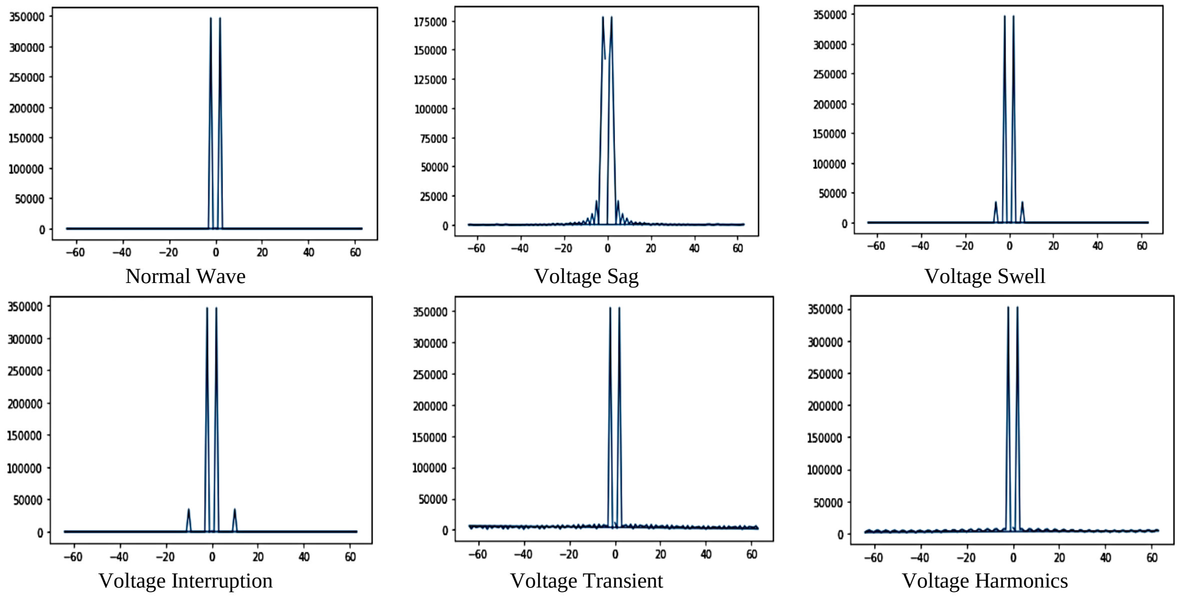

Figure 11. Samples of PQ signals classified by the CNN – Keras model

6.2. CNN–Keras based classification results

Totally, 20 epochs were exactly used for the model's convergence during training process to obtain 1280 parameters of the model. The hardware used for simulation is Intel (R) core i3 6006U 2.00GHz4GB RAM 64bit OS x 64 base processor. The result of CNN – Keras model is described in Table 10. This indicates how model behaves based on the given dataset during training and validation.

Table 10. Training and validation results

Epoch	Loss	Accuracy	Validation loss	Validation accuracy	Loss improvement	Accuracy improvement
1	1.0996	0.5451	1.1394	0.5254	-0.0398	0.0197
2	0.4547	0.8044	0.3843	0.8600	0.0704	-0.0556
3	0.1576	0.9572	0.4421	0.8429	-0.2845	0.1143
4	0.0933	0.9681	0.3788	0.7804	-0.2855	0.1877
5	0.0087	0.9996	0.0053	0.9881	0.0034	0.0115
6	0.0033	0.9994	0.0016	0.9892	0.0017	0.0102
7	0.0016	0.9992	0.0087	0.9975	-0.0071	0.0017
8	0.0018	0.9980	0.0005	0.9952	0.0013	0.0028
9	0.0008	0.9983	0.0005	0.9938	0.0003	0.0045
10	0.0004	0.9987	0.0002	0.9961	0.0002	0.0026
11	0.0006	0.9990	0.0006	0.9968	0.0000	0.0022
12	0.0004	0.9995	0.0007	0.9796	-0.0003	0.0199
13	0.0703	0.9994	0.0003	0.9689	0.0700	0.0305
14	0.0393	0.9860	0.0030	0.9883	0.0363	-0.0023
15	0.0155	0.9961	0.0004	0.9786	0.0151	0.0175
16	0.0003	0.9968	0.0002	0.9889	0.0001	0.0079
17	0.0002	0.9996	0.0001	0.9975	0.0001	0.0021
18	0.0001	0.9996	0.0008	0.9975	-0.0007	0.0021
19	0.0002	0.9996	0.0006	0.9975	-0.0004	0.0021
20	0.0001	0.9996	0.0017	0.9975	-0.0016	0.0021

The performance metrics used to analyze the classification performance of this work are accuracy and loss (error). Accuracy measures the percentage of classifications made by the model which is depicted in the classification accuracy curve shown in Figure 12.

From Figure 10, the highest classification accuracy and validation accuracy reached are 99.96% and 99.75% respectively. This is due to the fact that the CNN based Keras model learns very fast and can effectively work based on the given dataset. Hence, the model performed well as it attained the highest classification accuracy and validation accuracy as indicated in Figure 12. The loss or cost function is another metrics for determining ML model performance. The selection of loss function relies on the problem to be solved and the nature of data being used. It is used during model's training to update parameters in such a way that minimizes difference between predicted and actual output. The loss and validation loss curves of the trained model is indicated in Figure 13. The lower the loss, the better the model is performing. The lowest loss (error) of 0.0001 is attained during model's training. This indicates that the model performs better throughout the training process.

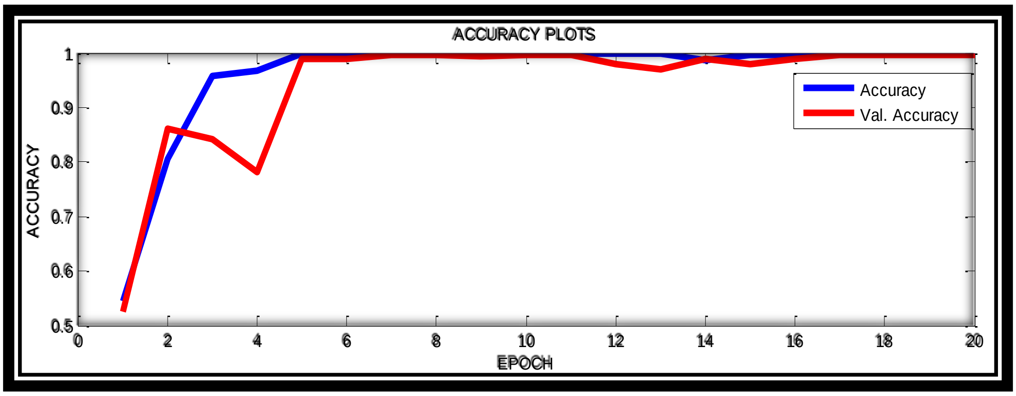

Figure 12. CNN – keras classification accuracy curves

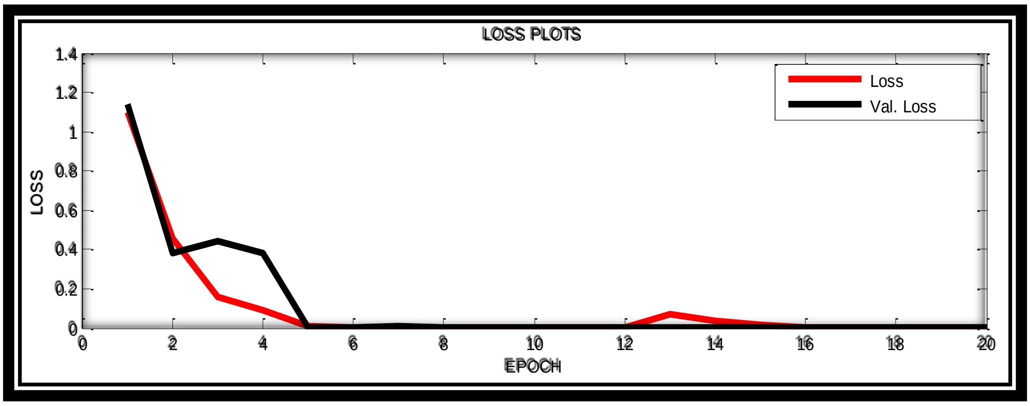

Figure 13. Training performance loss curves

6.3. Discussion of the results

This study is basically centered on DL technique to categorize PQ problems that specifically employed the utilization of DenseNet CNN – Keras approach to timely and automatically extract feature maps and classify PQ problems. The simulation activities proved that the DenseNet CNN – Keras method had achieved high classification accuracy with lowest possible error. The approach showed a remarkable improvement in accuracy up to 0.21% over traditional techniques and some other CNN approaches as shown in Table 10. The approach excelled in the aspects of:

a) Number of optimal features of DenseNet CNN - Keras method is lessen and improved by envisaging feature map of convolutional layers in the model. Notably, number of parameters is reduced to nearly 50% from 2,560 to 1,280.
b) The method used in this work explores the timely as well as effectiveness of PQ problems features extraction and classification. Within only 20 epochs the model converged with excellent classification performance.
c) The proposed technique has improved classification and validation accuracies of 99.96% and 99.75% at the same lowest error of 0.0001as shown in Table 10.
d) For real time analysis, classification accuracy could be improved by including small amount of real time measured data in the training procedure to sieve parameters of the model. This will lead to excellent performance of learning and adaptation which is recommended for further study.

e) The approach employed is actualized under offline conditions. Even tough, online execution is also possible and considered in future research.

Also, comparison of classification result with other automated methods is described in Table 11. The main contribution of this novel technique is the significant improvement realized in the classification performance when compared with other techniques. This proved that the approach can distilled features automatically with high classification accuracy. Although, model misbehaved from epoch 2 to 5 due to the nature of the dataset as indicated in Figures 12 and 13, it can equally be applied in PQ problems analysis especially when there are voluminous datasets.

Based on the foregoing analysis, the PQ problems could easily be categorized with high accuracy by DenseNet CNN – Keras technique. However, the feature maps shall be extracted using disturbance signals automatically in the absence of human intervention. Although, number of optimal parameters were decreased and enhanced by envisaging the feature maps of convolutional layers in the model in order to get good result. It proves that the proposed method has excellent performance in learning and adaptation. The technique employed in this paper is actualized using offline situations. Moreover, online implementation is also possible and considered in future research.

Table 11. Performance comparison of the proposed approach with other automatic methods

S/N	Method	FE technique	Number of PQ problems	Accuracy (%)	Reference
1	WVD – CNN	Automatic	9	99.67	[9]
2	MT – CNN	Automatic	6	94.63	[11]
3	CS – DCNN	Automatic	15	99.50	[15]
4	PSR – CNN	Automatic	10	99.80	[17]
5	Bayesian CNN	Automatic	5	99.80	[20]
6	TSI – CNN	Automatic	9	97.84	[50]
7	CNN – Keras	Automatic	5	99.96	Proposed method

7. CURRENT CHALLENGES AND FUTURE DIRECTIONS OF CNN

CNNs models had attained good classification performance on PQ data. However, there some challenges, where CNN models need to be improved upon. Some of the drawbacks experienced in the process of training CNN models are highlighted below.

7.1. Data augmentation

One of the major shortcomings of CNN is its inability to show good performance generally when the data is not voluminous. In 2012 AlexNet addressed this issue to some extent by innovating the idea of data augmentation [65]. Data augmentation could assist CNN in learning vast internal representations which could ultimately leads to greater performance. This technique could be involved to obtain synthesized data to magnify the size of the training dataset. The above fact can help to mitigate the data availability limitation.

7.2. Efficient training

Effective learning of CNN requires strong hardware such as graphical processing units (GPUs). Moreover, it required to employs CNN in embedded and smart devices to operate efficiently. Some applications of DL in embedded systems are wound intensity correction, law enforcement in smart cities [68].

7.3. Hyper–parameter selection

Selection of rightful hyperparameter largely influences performance of CNN. Small changes of any hyperparameter values could affect the whole performance of the model. Hence, tactful selection of hyperparameters is the main design issue that needs to be tackled via some appropriate optimization strategies.

7.4. Supervised learning mechanism

Supervised learning mechanism is mostly applied in many CNN applications. Therefore, availability of huge and synthesized data is demanded for its proper learning [69]. While in the same vein, human being can learn and generalizes from few examples.

7.5. Feature visualization

Every layer of the CNN works towards extracting better as well as specific features related to assigned task. Meanwhile in some tasks, it is crucial to identify nature of features extracted by the CNN before classification. The idea of feature visualization in CNN could assist in this direction. Hinton noticed that lower layers should handover their knowledge only to relevant neurons of subsequent layer. In this regard, Hinton introduced a very amazing capsule network approach [70].

7.6. Training CNN on noisy image data

It was observed that during learning process of CNN in noise image data could cause an increment of misclassification error. Any increase in small quantity of random noise in the input data may fool the process in such a way that the model can only categorize the original signals and slantly disturbed the design procedure.

7.7. Interpretability and explainability

CNNs models are generally like a "black box" that is to say it is hard to explain the results and actualize how the model reached conclusion. Therefore, it is difficult to itemize the main cause of PQ problems and takes proper corrective action. Approaches like feature visualization and attention mechanisms could improve interpretability of CNN models. This attribute helps to identify the most significant features for classification and provides awareness to the existing causes of PQ problems.

Future work for this paper includes evaluating the proposed approach (CNN–Keras) on other datasets and investigating the robustness of the approach under different operating conditions. Also, the approach can be executed online and employed for multiple PQ problems classification which is challenging practically. HRNetV2 CNN architecture can equally be used to implement the approach for improved performance in the future. Lastly, PQ mitigation actions can be identified in accordance with the information obtained by the method using fuzzy logic controller via DSTATCOM (CNN–FLC–DSTATCOM).

8. CONCLUSION

This comprehensive review of CNN based Keras application to PQ problems classification reflects promising outcomes of 99.96% classification performance. For the efficiency, accuracy and computational speed CNN based Keras approach showed excellent performance leading to minimal error (loss) of 0.0001 and high accuracy of 99.96% for PQ problems classification which address some limitations in the existing methods. Also, CNN techniques had not been exhausted squarely in such applications as far as their abilities may allow. They could be expanded further to categorize PQ problems yielding enhanced CNN performance, robustness and interpretability.

The experimental results show that, DenseNet CNN – Keras approach of PQ analysis could be utilized in real time situation because of their built – in performance characteristics such as swift estimation, greater accuracy as well as increased efficiency. It also proved that applying DenseNet CNN – Keras technique can significantly improves the PQ problems classification performance and eventually increasing the reliability as well as stability of the power systems creating better protective and secured system.

ACKNOWLEDGMENTS

Authors wish to use this opportunity to extend their warmth appreciation to the Tertiary Education Trust Fund (TETFUND) Nigeria for its contribution to this effort.

REFERENCES

[1] W. L. Rodrigues, F. A. S. Borges, A. O. de Carvalho Filho, and R. de A. L. Rabelo, "A deep learning approach for the detection and classification of power quality disturbances with windowed signals," *SN Computer Science*, vol. 2, no. 2, Jan. 2021, doi: 10.1007/s42979-020-00435-1.
[2] E. Hossain, M. R. Tur, S. Padmanaban, S. Ay, and I. Khan, "Analysis and mitigation of power quality issues in distributed generation systems using custom power devices," *IEEE Access*, vol. 6, pp. 16816–16833, 2018, doi: 10.1109/ACCESS.2018.2814981.
[3] M. O. Okeloa, "Detection and classification of power quality events using discrete wavelet transform and support vector machine," Feb. 2015, doi: 10.1109/PEDSTC.2015.7093333.
[4] M. Nivetha and D. Karunakaran, "A review of power quality analysis, techniques, methods and controlling," *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering*, vol. 5, no. 3, pp. 2049–2054, 2016.
[5] F. Ucar, O. F. Alcin, B. Dandil, and F. Ata, "Power quality event detection using a fast extreme learning machine," *Energies*, vol. 11, no. 1, p. 145, Jan. 2018, doi: 10.3390/en11010145.
[6] Y. Han, X. Ning, P. Yang, and L. Xu, "Review of power sharing, voltage restoration and stabilization techniques in hierarchical controlled DC microgrids," *IEEE Access*, vol. 7, pp. 149202–149223, 2019, doi: 10.1109/ACCESS.2019.2946706.
[7] K. Thirumala, A. C. Umarikar, and T. Jain, "A new classification model based on SVM for single and combined power quality disturbances," in *2016 National Power Systems Conference, NPSC 2016*, Dec. 2017, pp. 1–6, doi: 10.1109/NPSC.2016.7858889.
[8] C. O'Donovan, C. Giannetti, and G. Todeschini, "A novel deep learning power quality disturbance classification method using autoencoders," in *ICAART 2021 - Proceedings of the 13th International Conference on Agents and Artificial Intelligence*, 2021, vol. 2, pp. 373–380, doi: 10.5220/0010347103730380.

[9] K. Cai, W. Cao, L. Aarniovuori, H. Pang, Y. Lin, and G. Li, "Classification of power quality disturbances using wigner-ville distribution and deep convolutional neural networks," *IEEE Access*, vol. 7, no. Vmd, pp. 119099–119109, 2019, doi: 10.1109/ACCESS.2019.2937193.

[10] R. Shilpa, S. S. Prabhu, and P. S. Puttaswamy, "Analysis of power quality disturbance (PQD) using empirical mode decomposition (EMD) and SVM classifier," in *International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE)*, Dec. 2016, pp. 1468–1472, doi: 10.1109/ERECT.2015.7498978.

[11] Y. Dong, H. Cao, G. Xu, C. Yue, and X. Ding, "A novel method for multiple power quality disturbances classification using a multi-task convolution neural network," *2019 4th International Conference on Power and Renewable Energy, ICPRE 2019*, pp. 274–278, 2019, doi: 10.1109/ICPRE48497.2019.9034702.

[12] Y. Shen, M. Abubakar, H. Liu, and F. Hussain, "Power quality disturbance monitoring and classification based on improved PCA and convolution neural network for wind-grid distribution systems," *Energies*, vol. 12, no. 7, 2019, doi: 10.3390/en12071280.

[13] E. Balouji and O. Salor, "Classification of power quality events using deep learning on event images," in *3rd International Conference on Pattern Analysis and Image Analysis, IPRIA 2017*, Apr. 2017, pp. 216–221, doi: 10.1109/PRIA.2017.7983049.

[14] W. L. Rodrígues Jr, F. A. S. Borges, A. O. de Carvalho Filho and R. A.. L. Rabelo, "A Deep Learning Approach for the Detection and Classification of Power Quality Disturbances with Windowed Signals," *SN Computer Science*, vol. 2, no. 64, Pp. 1 – 14, 2021, doi: 10.1007/s42979-020-00435-1

[15] S. Wang and H. Chen, "A novel deep learning method for the classification of power quality disturbances using deep convolutional neural network," *Applied Energy*, vol. 235, pp. 1126–1140, Feb. 2019, doi: 10.1016/j.apenergy.2018.09.160.

[16] L. Alzubaidi *et al.*, *Review of deep learning: concepts, CNN architectures, challenges, applications, future directions*, vol. 8, no. 1. Springer International Publishing, 2021.

[17] K. Cai, T. Hu, W. Cao, and G. Li, "Classifying power quality disturbances based on phase space reconstruction and a convolutional neural network," *Applied Sciences (Switzerland)*, vol. 9, no. 18, 2019, doi: 10.3390/app9183681.

[18] P. R and R. H. R, "Enhancement of power quality using fuzzy logic controlled dstatcom," *International Journal of Advances in Signal and Image Sciences*, vol. 6, no. 1, p. 21, 2020, doi: 10.29284/ijasis.6.1.2020.21-28.

[19] S. B. Pandu *et al.*, "Power quality enhancement in sensitive local distribution grid using interval type-ii fuzzy logic controlled DSTATCOM," *IEEE Access*, vol. 9, pp. 59888–59899, 2021, doi: 10.1109/ACCESS.2021.3072865.

[20] S. Ekici, F. Ucar, B. Dandil, and R. Arghandeh, "Power quality event classification using optimized Bayesian convolutional neural networks," *Electrical Engineering*, vol. 103, no. 1, pp. 67–77, Jul. 2021, doi: 10.1007/s00202-020-01066-8.

[21] V. Anand and S. K. Srivastava, "Causes, effects and solutions of poor quality problems in the power systems," *Journal of Engineering Research and Applications www.ijera.com*, vol. 4, no. 5, pp. 67–74, 2014, [Online]. Available: www.ijera.com.

[22] Y. Han, Y. Feng, P. Yang, L. Xu, Y. Xu, and F. Blaabjerg, "Cause, classification of voltage sag, and voltage sag emulators and applications: a comprehensive overview," *IEEE Access*, vol. 8, pp. 1922–1934, 2020, doi: 10.1109/ACCESS.2019.2958965.

[23] J. C. Bravo-Rodríguez, F. J. Torres, and M. D. Borrás, "Hybrid machine learning models for classifying power quality disturbances: A comparative study," *Energies*, vol. 13, no. 11, 2020, doi: 10.3390/en13112761.

[24] G. M. Babu, V. Kalyani, and J. Narender, "A voltage controlled mode fact devices for power quality improvement and protection," *International Journal of Engineering Development and Research (IJEDR)*, vol. 4, no. 4, 2016.

[25] P. Kumar, K. Narmitha, and D. Aruma, "Analysis of dynamic voltage restorer with PI and Fuzzy logic based controller for voltage sag mitigation in distribution system," *Advance Research Journal of Multidisciplinary Discoveries*, pp. 40–46, 2017.

[26] A. M. Abd El-Hameid, A. A. Elbaset, M. Ebeed, and M. Abdelsattar, "Literature review and power quality issues," in *Enhancement of Grid-Connected Photovoltaic Systems Using Artificial Intelligence*, Springer Nature Switzerland, 2023, pp. 5–37.

[27] B. B. Bukata and Y. Li, "A novel model - free prediction of power quality problems via DSTATCOM," *in 18th international conference on automation and computing (ICAC), loughborough, UK*, pp. 1–6, 2012.

[28] C. S. Reddy, R. Divya, and M. G. Nair, "Power quality event classification using machine learning techniques," *International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)*, vol. 8, no. 2, pp. 30–36, 2021.

[29] R. A. de Oliveira and M. H. J. Bollen, "Deep learning for power quality," *Electric Power Systems Research*, vol. 214, no. PA, p. 108887, 2023, doi: 10.1016/j.epsr.2022.108887.

[30] I. S. Samanta *et al.*, "A comprehensive review of deep-learning applications to power quality analysis," *Energies*, vol. 16, no. 11, p. 4406, May 2023, doi: 10.3390/en16114406.

[31] M. Khodayar, G. Liu, J. Wang, and M. E. Khodayar, "Deep learning in power systems research: A review," *CSEE Journal of Power and Energy Systems*, vol. 7, no. 2, pp. 209–220, 2021, doi: 10.17775/CSEEJPES.2020.02700.

[32] S. Hożyń, "Convolutional neural networks for classifying electronic components in industrial applications," *Energies*, vol. 16, no. 2, p. 887, Jan. 2023, doi: 10.3390/en16020887.

[33] M. Krichen, "Convolutional neural networks: a survey," *Computers*, vol. 12, no. 8, pp. 1–41, 2023, doi: 10.3390/computers12080151.

[34] A. P. Koli and A. A. Mahajan, "A review of convolutional neural network architectures and its applications," *International Journal of Science Technology and Management (IJSTM)*, vol. 11, no. 8, pp. 50–59, 2022.

[35] L. Deng and D. Yu, "Deep learning: methods and applications," *Foundations and trends® in signal processing*, vol. 7, pp. 197–387, 2017.

[36] F. Xiao, T. Lu, M. Wu, and Q. Ai, "Maximal overlap discrete wavelet transform and deep learning for robust denoising and detection of power quality disturbance," *IET Generation, Transmission and Distribution*, vol. 14, no. 1, pp. 140–147, Dec. 2020, doi: 10.1049/iet-gtd.2019.1121.

[37] P. Kim, *MATLAB deep learning: with machine learning, neural networks and artificial intelligence*. Berkeley, CA: Apress, 2017.

[38] M. A. Rodriguez, J. Felipe Sotomonte, J. Cifuentes, and M. Bueno-Lopez, "Power quality disturbance classification via deep convolutional auto-encoders and stacked LSTM recurrent neural networks," *SEST 2020 - 3rd International Conference on Smart Energy Systems and Technologies*, 2020, doi: 10.1109/SEST48500.2020.9203082.

[39] S. Sony, K. Dunphy, A. Sadhu, and M. Capretz, "A systematic review of convolutional neural network-based structural condition assessment techniques," *Engineering Structures*, vol. 226, no. August 2020, p. 111347, 2021, doi: 10.1016/j.engstruct.2020.111347.

[40] D. Mnyanghwalo, H. Kundaeli, E. Kalinga, and N. Hamisi, "Deep learning approaches for fault detection and classifications in the electrical secondary distribution network: Methods comparison and recurrent neural network accuracy comparison," *Cogent Engineering*, vol. 7, no. 1, p. 1857500, Jan. 2020, doi: 10.1080/23311916.2020.1857500.

[41] W. L. Rodrigues Junior, F. A. Silva Borges, R. D. 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," *Proceedings of the International Joint Conference on Neural Networks*, vol. 2019-July, no. July, pp. 1–6, 2019, doi: 10.1109/IJCNN.2019.8852287.

[42] A. K. Ozcanli, F. Yaprakdal, and M. Baysal, “Deep learning methods and applications for electrical power systems: A comprehensive review,” *International Journal of Energy Research*, vol. 44, no. 9, pp. 7136–7157, Mar. 2020, doi: 10.1002/er.5331.

[43] J. Wang, Z. Xu, and Y. Che, “Power quality disturbance classification based on compressed sensing and deep convolution neural networks,” *IEEE Access*, vol. 7, pp. 78336–78346, 2019, doi: 10.1109/ACCESS.2019.2922367.

[44] G. Habib and S. Qureshi, “Optimization and acceleration of convolutional neural networks: A survey,” *Journal of King Saud University - Computer and Information Sciences*, vol. 34, no. 7, pp. 4244–4268, Jul. 2022, doi: 10.1016/j.jksuci.2020.10.004.

[45] R. S. Salles and P. F. Ribeiro, “The use of deep learning and 2-D wavelet scalograms for power quality disturbances classification,” *Electric Power Systems Research*, vol. 214, no. PA, p. 108834, 2023, doi: 10.1016/j.epsr.2022.108834.

[46] E. Galván and P. Mooney, “Neuroevolution in Deep Neural Networks: Current Trends and Future Challenges,” *IEEE Transactions on Artificial Intelligence*, vol. 2, no. 6, pp. 476–493, 2021, doi: 10.1109/TAI.2021.3067574.

[47] H. Xue, A. Chen, D. Zhang, and C. Zhang, “A novel deep convolution neural network and spectrogram based microgrid power quality disturbances classification method,” *Conference Proceedings - IEEE Applied Power Electronics Conference and Exposition - APEC*, vol. 2020-March, pp. 2303–2307, 2020, doi: 10.1109/APEC39645.2020.9124252.

[48] H. Liu *et al.*, “Signal Processing and Deep Learning Techniques for Power Quality Events Monitoring and Classification,” *Electric Power Components and Systems*, vol. 47, no. 14–15, pp. 1332–1348, 2019, doi: 10.1080/15325008.2019.1666178.

[49] W. Rawat and Z. Wang, “Deep convolutional neural networks for image classification: A comprehensive review,” *Neural Computation*, vol. 29, no. 9, pp. 2352–2449, Sep. 2017, doi: 10.1162/NECO_a_00990.

[50] 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,” *2020 International Wireless Communications and Mobile Computing, IWCMC 2020*, pp. 1701–1706, 2020, doi: 10.1109/IWCMC48107.2020.9148438.

[51] M. Mohammadi, M. Afrasiabi, S. Afrasiabi, and B. Parang, “Detection and classification of multiple power quality disturbances based on temporal deep learning,” *Proceedings - 2019 IEEE International Conference on Environment and Electrical Engineering and 2019 IEEE Industrial and Commercial Power Systems Europe, EEEIC/I and CPS Europe 2019*, pp. 0–4, 2019, doi: 10.1109/EEEIC.2019.8783378.

[52] 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, 2019, doi: 10.3390/en12244732.

[53] H. Liu, F. Hussain, Y. Shen, S. Arif, A. Nazir, and M. Abubakar, “Complex power quality disturbances classification via curvelet transform and deep learning,” *Electric Power Systems Research*, vol. 163, no. April, pp. 1–9, 2018, doi: 10.1016/j.epsr.2018.05.018.

[54] N. A. Manan, S. Shahbudin, M. Kassim, R. Mohamad, and F. Y. Abdul Rahman, “Power quality disturbances classification using sparse autoencoder (SAE) based on deep neural network,” *ISCAIE 2021 - IEEE 11th Symposium on Computer Applications and Industrial Electronics*, pp. 19–22, 2021, doi: 10.1109/ISCAIE51753.2021.9431822.

[55] P. K. Mishra, U. Subudhi, and S. Jain, “Power quality disturbances classification with deep learning approach,” *Proceedings - 2019 International Conference on Information Technology, ICIT 2019*, pp. 273–278, 2019, doi: 10.1109/ICIT48102.2019.00055.

[56] L. Ramalingappa and A. Manjunatha, “Power quality event classification using complex wavelets phasor models and customized convolution neural network,” *International Journal of Electrical and Computer Engineering*, vol. 12, no. 1, pp. 22–31, 2022, doi: 10.11591/ijece.v12i1.pp22-31.

[57] N. Mohan, K. P. Soman, and R. Vinayakumar, “Deep power: deep learning architectures for power quality disturbances classification,” in *Proceedings of 2017 IEEE International Conference on Technological Advancements in Power and Energy: Exploring Energy Solutions for an Intelligent Power Grid, TAP Energy 2017*, Dec. 2018, pp. 1–6, doi: 10.1109/TAPENERGY.2017.8397249.

[58] M. Sahani and P. K. Dash, “FPGA-based deep convolutional neural network of process adaptive VMD data with online sequential RVFLN for power quality events recognition,” *IEEE Transactions on Power Electronics*, vol. 36, no. 4, pp. 4006–4015, 2021, doi: 10.1109/TPEL.2020.3023770.

[59] W. Qiu, Q. Tang, J. Liu, and W. Yao, “An automatic identification framework for complex power quality disturbances based on multifusion convolutional neural network,” *IEEE Transactions on Industrial Informatics*, vol. 16, no. 5, pp. 3233–3241, 2020, doi: 10.1109/TII.2019.2920689.

[60] E. Yiğit, U. Özkaya, Ş. Öztürk, D. Singh, and H. Gritli, “Automatic detection of power quality disturbance using convolutional neural network structure with gated recurrent unit,” *Mobile Information Systems*, vol. 2021, pp. 1–11, Jul. 2021, doi: 10.1155/2021/7917500.

[61] M. H. Wang, S. H. Chi, and S. Der Lu, “Power quality fault identification method based on SDP and convolutional neural network,” *Applied Sciences (Switzerland)*, vol. 13, no. 4, 2023, doi: 10.3390/app13042265.

[62] F. Aziz, A. Ul Haq, S. Ahmad, Y. Mahmoud, M. Jalal, and U. Ali, “A novel convolutional neural network-based approach for fault classification in photovoltaic arrays,” *IEEE Access*, vol. 8, pp. 41889–41904, 2020, doi: 10.1109/ACCESS.2020.2977116.

[63] M. Liu, Y. Chen, Z. Zhang, and S. Deng, “Classification of power quality disturbance using segmented and modified S-transform and DCNN-MSVM hybrid model,” *IEEE Access*, vol. 11, no. January, pp. 890–899, 2023, doi: 10.1109/ACCESS.2022.3233767.

[64] K. R. Sree Jyothi, P. Venkatesh Kumar, and J. JayaKumar, “A review of different configurations and control techniques for DSTATCOM in the distribution system,” *E3S Web of Conferences*, vol. 309, p. 1119, 2021, doi: 10.1051/e3sconf/202130901119.

[65] A. Khan, A. Sohail, U. Zahoora, and A. S. Qureshi, “A survey of the recent architectures of deep convolutional neural networks,” *Artificial Intelligence Review*, vol. 53, no. 8, pp. 5455–5516, 2020, doi: 10.1007/s10462-020-09825-6.

[66] H. Lu *et al.*, “Wound intensity correction and segmentation with convolutional neural networks,” *Concurrency and Computation: Practice and Experience*, vol. 29, no. 6, Aug. 2017, doi: 10.1002/cpe.3927.

[67] Institute of Electrical and Electronics Engineers, *IEEE Std 1159 - IEEE Recommended Practice for Monitoring Electric Power Quality.*, vol. 2009, no. June. 2009.

[68] C. Szegedy *et al.*, “Going deeper with convolutions,” in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Jun. 2015, vol. 07-12-June-2015, pp. 1–9, doi: 10.1109/CVPR.2015.7298594.

[69] G. Hinton, S. Sabour, and N. Frosst, “Matrix capsules with EM routing,” *6th International Conference on Learning Representations, ICLR 2018 - Conference Track Proceedings*, pp. 1–15, 2018.

[70] C. Szegedy, J. Bruna, D. Erhan, and I. Goodfellow, “Intriguing properties of neural networks,” pp. 1–10, 2014.

BIOGRAPHIES OF AUTHORS

Adamu Sa'idu received his B. Eng. and M. Eng. degrees in Electrical Engineering from Bayero University Kano Nigeria in 2008 and 2015 respectively. Adamu is currently working with Federal College of Education (Technical) Bichi Kano Nigeria in the department of Electrical/Electronic Technology Education. He is a registered Member of the Council for the Regulation of Engineering in Nigeria (COREN). Adamu is a qualified and certified teacher having full Membership with Teachers Registration Council of Nigeria (TRCN). His research area of interest includes electrical power and machine, electrical load forecasting, power quality control, artificial intelligence, machine learning, and deep learning. He can be contacted through email: adamugiwa77@gmail.com.

Nasiru B. Kadandani received his Ph.D. degree in Electrical Engineering from Newcastle University, Newcastle upon Tyne, England, United Kingdom, in 2021, specializing in Power Electronics. He received his Master and Bachelor degrees in Electrical Engineering from Bayero University, Kano, Nigeria, in 2010 and 2004 respectively. Dr. Kadandani is presently working as an Associate Professor in the Department of Electrical Engineering, Bayero University, Kano Nigeria. He has authored/coauthored more than 50 journal articles and conference proceeding papers. His current research interests include multilevel converters, solid state transformers, grid integration of renewable energy, power quality control and electric vehicle charging systems. Dr Kadandani is a registered Member of the Council for the Regulation of Engineering in Nigeria (COREN), the Nigerian Society of Engineers (NSE) and the Institute of Electrical and Electronics Engineers (IEEE). He was the Nigerian representative during the 9th International Training on Wind Turbine Technology and Applications held at the Centre for Wind Energy Technology (C-WET), Chennai, India, in 2012. He can be contacted via email: nbkadandani.ele@buk.edu.ng.

Bala Boyi Bukata received his Ph.D. degree in Electronics and Electrical Engineering from the University of Glasgow (UoG) UK, in 2012. He received his Masters and a Diploma in Electrical Engineering Degree from Bayero University, Kano in Nigeria, and Isle of Wight College, Newport England, respectively. Dr. B. B. Bukata is currently an Associate Professor in the Department of Electrical Engineering Bayero University, Kano. His current research interests include Smart Distribution Grid design and implementation, Virtual Power Plants (VPP) design, application of AI to power system control and simulation, Distributed Energy Resources (RES), Integration of Renewable Energy Systems (RES), and Power Quality Control and Assessment. He devised a ground breaking research for DSTATCOM automation using the famous CAutoD mechanism pioneered by the Glasgow Intelligent Systems Group while he was at the University of Glasgow. Bala is an Associate of Advanced Higher Education Academy (AHEA) in the UK. He is also a Registered COREN (Reg. COREN) Engineer, Member Nigerian Society of Engineers (MNSE), Member of the Institute of Engineering and Technology (MIET), as well as a member of the Institute of Electronics and Electrical Engineering (MIEEE), UK Chapter. He can be contacted via email: bbbukata.ele@buk.edu.ng.