

Power quality disturbance signal segmentation and classification based on modified BI-LSTM with double attention mechanism

Poras Khetarpal¹, Neelu nagpal^{2*}, Pierluigi Siano^{3,4*}, Mohammed AL-Numay⁵

¹ Information Technology Department, Bharati vidyapeeth's college of engineering, New Delhi-110063, India.

² Electrical and Electronics Engineering Department, Maharaja Agrasen Institute of Technology, Delhi-110086, India.

³ Department of Management and Innovation Systems, University of Salerno, 84084 Fisciano, Italy

^{4,5} Electrical Engineering Department, College of Engineering, King Saud University, P.O. Box 800, Riyadh 11421, Saudi Arabia.

* E-mail: psiano@unisa.it ; nagpalneelu1971@ieee.org

Abstract: Abstract—This paper proposes a recurrent neural network (RNN) based model to segment and classify multiple combined multiple power quality disturbances (PQDs) from the PQD voltage signal. A modified bi-directional long short-term memory (BI-LSTM) model with two different types of attention mechanism is developed. Firstly, an attention gate is added to the basic LSTM cell to reduce the training time and focus the memory on important PQD signal part. Secondly, attention layer is added to the BI-LSTM to obtain the more important part of the voltage signal by assigning weightage to the output of the BI-LSTM model. Finally a SoftMax classifier is applied to classify the combined PQD signal in 96 different combinations. The proposed BI-LSTM model with attention gate and attention layer mechanism is compared to different baseline models based on recurrent neural network (RNN) and convolution neural network (CNN). From simulation study, it is inferred that with the proposed method, the PQD signal is easily segmented from the voltage signal which makes the process of PQD classification more accurate with less computation complexity and in less time as compared to alternative approaches.

1 Introduction

With an upsurge in demand of renewable energy, increase in the use of power electronics-based devices and non-linear loads along with the development of smart grid, smart meters and introduction of electric vehicles, the power quality (PQ) issues hold much more importance and needs to be addressed properly. Good PQ hints towards a reliable and stable power system whereas poor PQ may become a serious issue for the end users [1]-[2]. Neglecting this issue may have a significant impact on the economical and reliable operation of smart grid networks, as well as a reduction in the efficiency of industrial production, resulting in economic losses, and can damage or shorten the life of users' sensitive household devices. The reduced or halted production of industries and reduction in the life time of household items and cost of their maintenance leads to economic losses also. The study of [3] has summarized economic research for European countries which encounters financial losses of around 25 billion Euros due to PQ issues, with above 90% of the losses incurred in manufacturing industries. In order to minimize and control the effects of PQD on the power system network, a highly accurate, effective and reliable PQD detection and classification technique is essential to take corresponding action to mitigate them. Generally, due to renewable energy source integration to the grid, transient PQD are produced on the grid side and due the power electronics devices on the load side combined PQD are added to the power system [4]. Voltage PQDs could be generally classified as single and multiple (combined) disturbances. General single PQDs include steady state disturbances like harmonics, notching etc and transient disturbances include voltage swell, voltage sag, voltage impulse etc. [5]. Combined PQDs are combination of two or more single PQDs. Classification of PQD includes mainly segmentation of the PQD from the voltage signal [1], then from the PQD section important and strong features are extracted using feature extraction techniques like Wavelet transform (WT) [6], Hilbert Huang transform (HHT) [7], S transform [8]. Finally, a machine learning

based classifier is used to classify the type of PQDs. Support vector machine (SVM) [9], artificial neural networks (ANN) [9], Fuzzy logic (FL) [10] based classifiers etc. are used for the same. Through deep learning-based approach, the process of feature extraction and classification is optimized and united. Hence, strong, optimal and minimum set of features are obtained automatically rather than using hand crafted features as in feature extraction techniques. Deep learning-based models are generally based on PQD image-based approach using CNN [11] or RNN based methods [12]. In [13], eleven PQDs are classified using S transform and PNN based classifier achieving an accuracy of 93.2%. In [14] nine types of PQDs are classified using a modification of S transform method (Double-Resolution S-transform (DRST)) for essential feature extraction and Directed Acyclic Graph Support Vector Machines (DAG-SVMs) a SVM based method for classification of the disturbance with an average accuracy of 97%. In [15] a combination of S transform and Fuzzy logic is applied for PQD classification of fourteen types of PQDs obtaining an accuracy of 98% with SNR ratio of 0dB. In [7] a classification accuracy of 99% is obtained for 16 types of PQD using HHT as feature extraction method and Weighted bi-directional extreme learning machine (WBELM) as signal classifier. [16] obtained a classification accuracy of 98.95% for classifying 8 types of PQD signals using a combination of WT and SVM. [17] used sparse auto encoder (SAE) for essential feature encoding and independent component analysis (ICA) as a classifier obtaining an accuracy of 98.6 % for PQD classification. [12] and [9] classified 96 types of combined PQDs with maximum six types of single PQDs combined together.

In practical power system, presence of multiple or combined PQD is very much common. Although, multiple PQDs could be present together which requires better and sensitive detection and classification technique of PQD [18]. Most important boost for PQD classification algorithm is proper segmentation of the disturbance part. Proper segmentation and feature extraction of the PQD are far more important for good PQD classification accuracy rather

than the use of complex classifiers alone [19]. In this paper, RNN based approach for PQD segmentation and classification is used. The model uses BI-LSTM with two different types of attention mechanism. one is the attention gate embedded in the basic LSTM cell and other is the overall weighing mechanism to implement the most weighted decision mechanism. This attention gate applied in the LSTM cells improves effective and decisive key information extraction from future and past states of the PQD signal and overall weighing mechanism improves overall decision capability of the model [20]. In [19], the forget gate is removed from the LSTM cell totally to save computation time. Contrarily, in the proposed model, the attention mechanism is applied to the forget gate also so that essential abstract information is captured and LSTM cell is updated. The main contribution of the suggested work is as follows:

- An attention mechanism-based learning framework is applied to automatically learn the strong, optimal and non-redundant features of the PQD present in the voltage signal rather than working on complete signal, the model focuses on the PQD event part of the voltage signal, automatically removing unimportant features and saving computation time and cost.
- The present work introduces the improved BI-LSTM model with attention mechanism and attention gate mechanism for efficient and accurate PQDs. Further, the presence of attention mechanism in gate and attention layers reduce the computation time of the model and the PQD is segmented easily from the voltage signal.
- Extensive case study for performance analysis of the present PQD classifier by simulating the various scenarios to determine the accuracy with and without phase shift at different signal to noise ratio (SNR).
- Comparative analysis with the reported work to justify the superiority of the proposed PQ assessment framework to produce efficient and accurate results.

The paper is structured as follows: After introducing the basic concept and related literature review in section 1, the proposed model is discussed with the help of illustrating diagrams in section 2. Section 3 presents the model parameter and data set generation used followed by the result outcome in section 4. Based on the result analysis, the conclusion of the work is drawn in section 5.

2 THE PROPOSED MODEL

The proposed model is a modified BI-LSTM based network where the basic LSTM cell is modified with the application of attention gate as shown in Fig. 1. Also, the output of the BI-LSTM layers is fed to attention layer which helps the network to concentrate on the specific PQD part from the voltage signal. This model is named as BI-LSTM with attention gate and attention layer mechanism (BI-LSTMAGAL). The architecture of LSTM and BI-LSTM based models is presented in the following subsections.

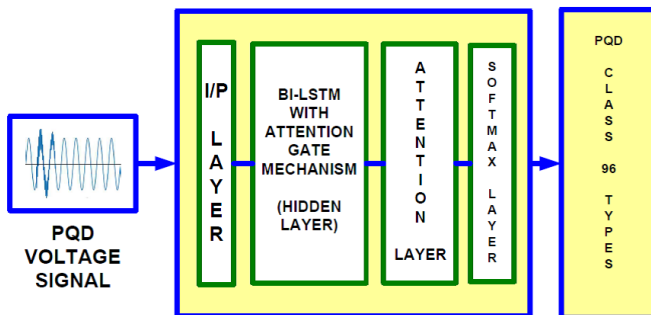


Fig. 1: Proposed BI-LSTM based model

2.1 LSTM AND BI-LSTM

Neural network-based modelling approaches have capacity to model nonlinear and complex relations and are better suited to process data which is sequential [21]. A type of neural network, recurrent neural networks (RNN), have internal memory that can be used to exploit past data. As a result, they are frequently employed in sequential data processing. Although RNN can improve sequential data processing, long-term data retention is impossible due to the vanishing gradient problem, so LSTM was introduced. The LSTM framework was proposed in 1995, has quickly grabbed attention of researchers in the field of signal and sequence analysis. The concept of weighted data plays a major role in LSTM. A LSTM unit, generally comprises of a memory cell and of 3 gates (forget gate, input gate, and output gate). These gates control the selective flow of non-linear information [22]. The state of an LSTM cell is preserved over time and is updated or modified whenever the cell receives an input. By default, the update

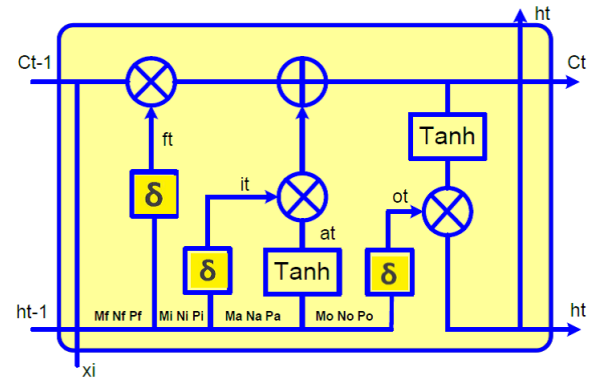


Fig. 2: A basic LSTM cell with peephole connection

of LSTM cell output state is determined by the current input and previous state of hidden output. With the addition of peephole connection, the previous state of the cell is also taken as a parameter. Fig. 2 shows a basic LSTM cell with flow of data between inputs and gates. At any current time, t , $C_{(t-1)}$ and $h_{(t-1)}$ indicate output of the cell at $(t-1)$ time (previous output of the cell), and hidden state at time $(t-1)$ of the cell respectively (previous hidden state of the cell). Input gate (i_t), output gate (O_t), and forget gate (f_t) represent the three gates of LSTM cell. The gates in a cell decide which information is necessary, i.e. i_t is responsible for the control of current incoming information. It controls the integration of current data with the memory unit. O_t plays the role of determining what part of present output (C_t) is useful to h_t and what is not required. C_t might have some information which is useless to h_t , and thus, not all data in C_t is related to hidden output. As a result, O_t generates a hidden state h_t . f_t determines what data from previous state $C_{(t-1)}$ is to be kept and discards the data which is not required. For example, in a PQD signal only the important data is remembered and rest unnecessary data is rejected by this gate. This rejection of unimportant data is done so that memory retention is optimized. The memory cell is updated at each time instant, t with the help of (a_t) as shown in Eqn. (1d). H_t and c_t are updated with each time instant t as per equations (1e) and (1f).

Bi-directional LSTM (BI-LSTM) is a step ahead of traditional unidirectional LSTM. Two independent LSTMs are combined to form BI-LSTM, these LSTM work independent of each other and information from the PQD is finally merged together as shown in Fig. 3. This framework is also equipped with feedback connections and hence the data is correlated easily [23]. As the name suggests BI-LSTM consists of two data layers, a forward LSTM layer and a backward LSTM layer. Taking reference from a voltage signal which may have PQD, As shown in Fig. 2, referring the index of time from 1 to $t-1$, with time t as the current time index, h_{tf} is calculated which

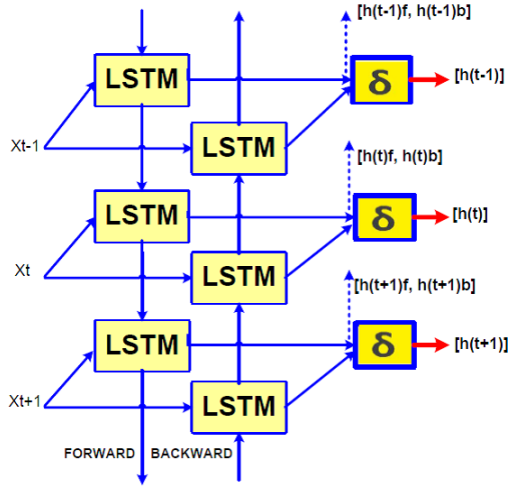


Fig. 3: A basic LSTM cell with peephole connection

is the forward layer LSTM hidden sequence of the output. Similarly, h_{tb} which is the backward layer hidden sequence output is calculated from the index of time $t+1$ to signal end. These hidden states of forward and backward layer are updated with standard LSTM learning procedure. Equation (1g) gives the combined output of the Bi-LSTM network.

$$i_t = (N_i x_t + P_i h_{(t-1)} + M_i c_{(t-1)} + b_i) \quad (1a)$$

$$O_t = \delta(N_o x_t + P_o h_{(t-1)} + M_o c_{(t-1)} + b_o) \quad (1b)$$

$$f_t = \delta(N_f x_t + P_f h_{(t-1)} + M_f c_{(t-1)} + b_f) \quad (1c)$$

$$a_t = \delta(N_a x_t + P_a h_{(t-1)} + M_a c_{(t-1)} + b_a) \quad (1d)$$

$$c_t = f_t c_{(t-1)} + i_t a_t \quad (1e)$$

$$h_t = o_t * \tanh(C_t) \quad (1f)$$

$$h_t = o_t * \tanh(C_t) \quad (1g)$$

$$h_t = \delta(W_h [\vec{h}_t, \vec{h}_t] + b_h) \quad (1h)$$

W, P and M represents the weights to be learned by the BI-LSTM network and b_f represents the bias values to be learned by BI-LSTM. Sigmoid (σ) indicates activation of sigmoid function. Equation (1e) is modified with i_t being replaced with $(1 - f_t)$, giving more weightage to forget gate. Hence, the state update equation (1e) is modified to equation (2):

$$c_t = f_t c_{(t-1)} + (1 - f_t) a_t \quad (2)$$

2.2 Self-Attention Gate

A self-attention mechanism (self-attention gate) is added in the LSTM cell to focus and capture the essential data or information and then update the cell state as shown in Fig. 4. Basically, a gate is added into the structure of LSTM cell which introduces the self-attention mechanism in the cell. Attention gate parameters are W_f and V_f respectively.

$$f_t = \delta(V_f \tanh(W_f c_{(t-1)})) \quad (3)$$

With this self-attention mechanism in the gate of LSTM, the training parameters of the cell are reduced and hence, the speed of Bi-LSTM network is improved significantly. The training process is upgraded and refined with this gate.

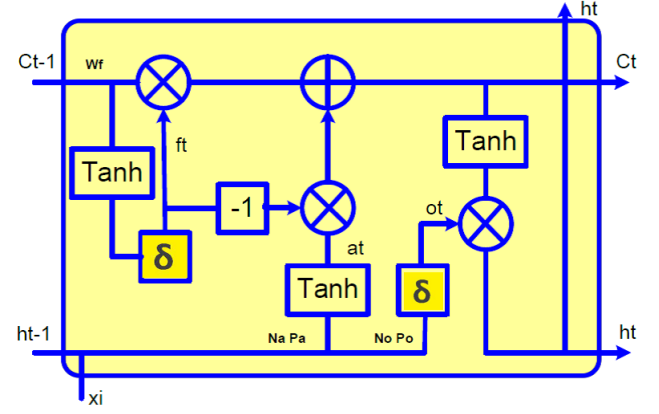


Fig. 4: LSTM cell with attention gate

2.3 Self-Attention Layer

Attention layer is introduced in the network to help the network to concentrate automatically on the important aspects of PQD signal that have significant effect on signal classification. Weight coefficients are added to decide and derive important information from PQD sequence. Higher weights indicate higher importance of the point of sequence in PQD signal. For a PQD signal, there might be a presence of unwanted irregularities and noise in the signal. All the hidden states from the hidden vector of BI-LSTM are not required for classification. Hence, a soft attention mechanism is applied as shown in Fig. 3 [24] which automatically learns and assigns weights to the elements of hidden vector. These weights assign importance to the hidden states and hence, PQD signal classification is obtained in less time and learning time is reduced.

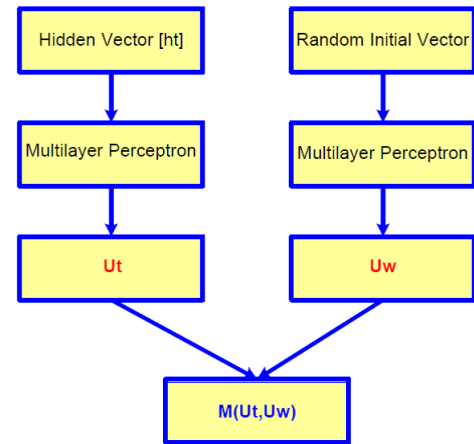


Fig. 5: Self-Attention-Based BI-LSTM architecture

2.4 Basic Self-Attention Mechanism

The hidden vector h_t is first fed to a Multi-Layer Perceptron (MLP) and a new representation u_t is obtained as in (4a). A weight value to symbolize and assign the importance to states is calculated for h_t with the help of u_t and context vector (u_w). u_w is a high dimensional representation and helps in judging the importance of h_t states. It is initialized randomly and similarity of u_t and u_w is used to measure importance of each state of h_t which is learnt jointly

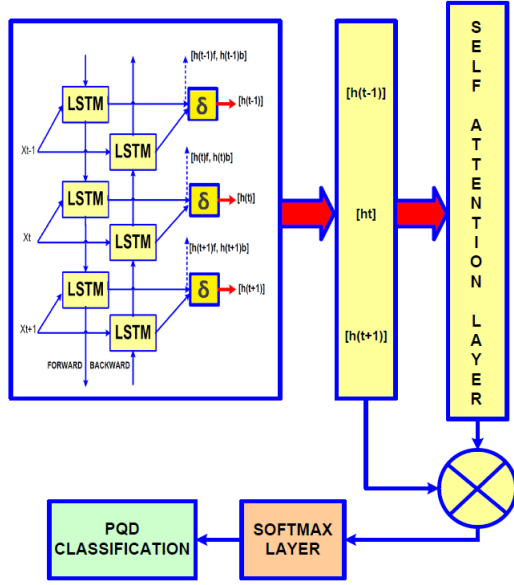


Fig. 6: Detailed proposed BI-LSTM model

Table 1 Model parameters

Parameter	Value
BI-LSTM	1
Hidden layer size	80
Optimizer	Adam Gradient Descent
Batch size	16
Learning rate	0.001
Learning rate decay period	10
Learning rate decay factor	0.5

in training process. Fig. 5 shows the self-attention learning mechanism. Similarity between u_w and u_t is represented by function M and weighted mean for h_t is obtained by softmax function.

$$u_t = \tanh(W_w h_t + b_w) \quad (4a)$$

$$\partial_t = \frac{\exp(u_t^T u_w)}{\sum_t \exp(u_t^T u_w)} \quad (4b)$$

$$s = \sum_t \partial_t h_t \quad (4c)$$

2.5 SoftMax Classifier

SoftMax classifier is used as the classifier for the proposed model. High level representation is produced by the model for PQD. The hidden vector of each state is multiplied by the corresponding weight and hence, vector s is obtained and it is regarded as the optimal feature set which is when fed to the SoftMax classifier gives the classification result for PQD type.

$$\tilde{y} = \text{soft max}(W_s s + b_s) \quad (5)$$

The detailed proposed BI-LSTM based model is shown in Fig. 6.

3 MODEL PARAMETER SETTING, DATASET GENERATION

3.1 Model Parameters

The Self-Attention improved BI-LSTM model proposed in this paper is composed of one enhanced BI-LSTM layer with attention

gate, a weighing layer based on attention mechanism and a SoftMax classifier. Input PQD signal is fed to the input BI-LSTM layer at each timestamp of t and dropout rate of 0.1 is set in order to avoid overfitting problems. Model was trained for 20 epochs in the experiments. Batch size of 16 was considered. initially the learning rate was set to 0.001. For training the model, Adam gradient descent method is used. This method optimizes the process of gradient descent method and helps reaching the global minimum efficiently. To evaluate the model performance in a better way, the model was trained with ten-fold cross validation method. In order to better measure the performance of our model, the 10-fold cross-validation method was used to train the model, and finally take the average score of ten times cross-validation to obtain evaluation result. Table 1 gives the values of the parameters set for the model.

3.2 Dataset Generation

A rich data set with the presence of required PQDs is required for proper segmentation and classification of PQD. PQD occurrence is not location or time specific, and thus cannot be guaranteed. As a result, the neural network is trained using synthetic PQ event signals generated in MATLAB using numerical equations depicting PQD signals. PQDs generated from these numerical models closely resemble real PQDs, so these generated signals are used for training and testing the proposed model.

Mathematical modelling of nine PQD voltage signals is done using MATLAB as referred from [25]. A database of 1000 voltage signals for each type of PQD is generated. Nine basic PQD classes including pure signal (D1) are generated which are voltage interruption (D2), voltage sag (D3), voltage swell (D4), voltage harmonics (D5), oscillatory transients (D6), voltage flicker (D7), voltage notch (D8) and voltage spikes (D9). The signals are generated with sampling frequency of 3.2 KHz with a fundamental frequency of 50Hz. Different combinations of these nine basic PQDs with each other is shown in table 2 giving total 96 different combinations of PQDs [9] with percentage of classification accuracy obtained with noiseless as well as noisy condition. Fig. 7 shows some of the PQD waveform images generated through the simulation.

4 RESULTS

The simulation study is carried out using Pytorch 0.4.0, a deep learning framework for implementing the experiments with NVIDIA GeForce GTX 1080 as the processor. Based on the results, Table 2 is prepared that represents the combination of multiple PQDs. In this table, there are six columns with each column adds to a combination of PQDs. Here, six columns of PQD means six different types of PQD combination are present. First column in the table represents a presence of single type of PQD. Second column represents combination of two types of PQDs. It is indicated in this table, ninety-six different types of non-mutually exclusive combinations of PQDs are presented. 1000 signals for each type of disturbance is produced with 750 signals used for training and 250 signals for testing. Cell of the column gives the presence of particular PQD. A signal to noise (SNR) ratio of 20 dB and 40 dB is considered with noiseless disturbance and percentage of classification accuracy is presented in Table 2. Here, voltage flicker is represented by "VF", voltage harmonic and inter-harmonics is represented by "VHI", voltage oscillatory transient is represented by "VOT", voltage notch is represented by "VN", impulsive transient is represented by "VIT", voltage interruption is represented by "VI". Further more details of the table is explained as follows: 1) Column 1 represents the presence of eight single PQDs including voltage swell, voltage sag, voltage flicker, voltage harmonic and inter-harmonic, voltage notch, voltage interruption, impulsive transient of voltage and oscillatory transient of voltage.

2) Column two represents twenty sets of PQDs combinations. A combination of two PQDs is represented by each cell of column two. One PQD from adjacent cell of column one from cell of column two itself. Column three represents combination of thirty multiple PQDs with each cell of column three along with adjacent cells of

Table 2 Comination of multiple PQDs

Combination of Disturbances					
1	2	3	4	5	6
VSAG(99.8,2)	VHI(99.8,10))	VOT(99.5, 30)	VIT(99.3, 60)	VN(98.7 , 85)	VF(98.7, 96)
				VF(99.9, 86)	
			VN(98.9, 61)	VF(99.4, 87)	
			VF(99.8, 62)		
		VIT(99.9, 31)	VN(99.8, 63)	VN(99.1, 88)	
			VN(99.5, 64)		
			VN(99.8, 32)	VF(99.8, 65)	
		VOT(99.4, 11)	VF(99.8, 33)		
			VIT(99.4, 34)	VN(99.0, 66)	VF(98.7, 89)
				VF(99.4, 67)	
			VN(99.8, 35)	VF(99.5, 68)	
			VF(99.8, 36)		
			VN(99.4, 37)	VF(98.7, 69)	
		VIT(99.2, 12)	VF(99.3, 38)		
		VN(99.7, 13)	VF(99.4, 39)		
		VF(99.8, 14)			
VSWEL(99.9, 3)	VHI(99.8, 15)	VOT(99.8, 40)	VIT(98.8, 70)	VN(99.4, 90)	VF(99.5, 97)
				VF(98.8, 91)	
			VN(99.8, 71)	VF(98.4, 92)	
			VF(99.5, 72)		
		VIT(99.4, 41)	VN(98.8, 73)	VF(99.2, 93)	
			VF(98.7, 74)		
			VN(99.7, 42)	VF(98.5, 75)	
		VOT(98.2, 16)	VF(99.8, 43)		
			VIT(99.5, 44)	VN(98.4, 76)	VF(98.8, 94)
				VF(98.7, 77)	
			VN(99.8, 45)	VF(98.8, 78)	
			VF(99.5, 46)		
		VIT(99.4, 17)	VN(99.9, 47)	VF(99.3, 79)	
			VF(99.4, 48)		
			VN(99.8, 18)	VF(9.8, 49)	
		VF(98.8, 19)			
VF(99.8, 4)	VHI(99.8, 20)	VOT(98.7, 50)	VIT(98.5, 80)		
			VN(99.2, 81)		
			VIT(98.5, 51)	VN(98.8, 82)	
			VN(99.9, 52)		
		VOT(99.7, 21)	VIT(99.0, 53)	VN(99.8, 83)	
			VN(99.8, 54)		
			VIT(99.3, 22)	VN(99.1, 55)	
		VOT(99.8, 24)	VN(99.8, 23)		
			VIT(98.2, 56)	VN(98.4, 84)	
			VN(98.8, 57)		
			VIT(99.3, 25)	VN(99.4, 58)	
			VN(99.8, 26)		
		VOT(99.5, 6)	VIT(99.4, 27)	VN(98.8, 59)	
			VN(99.8, 28)		
			VIT(99.2, 7)	VN(99.8, 29)	
		VN(99.4, 8)			
		VN(99.7, 9)			

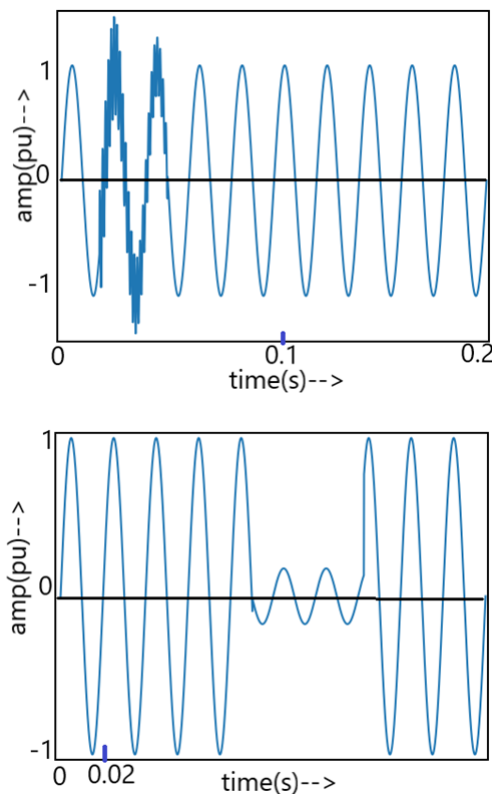


Fig. 7: PQD signals generated through MATLAB (a) Oscillatory transient (b) Voltage sag

Table 3 Classification accuracy of nine PQD signals at different SNR

Type of PQD	Average accuracy (%) 0 dB	Average accuracy (%) 20 dB	Average accuracy (%) 40 dB
Pure (D1)	100	99.99	99.98
Interruption (D2)	100	99.70	99.50
Sag (D3)	100	99.80	99.60
Swell (D4)	100	99.90	99.70
Harmonics (D5)	99.90	99.80	99.70
Oscillatory transient (D6)	99.70	99.50	99.30
Flicker (D7)	99.90	99.80	99.70
Notch (D8)	99.60	99.40	99.20
Spikes (D9)	99.80	99.70	99.60
Overall accuracy (%)	99.87	99.73	99.50

column one and column two representing a combination of disturbance. Similarly, column four represents twenty-five multiple PQDs (combination of four types of PQDs), column five represents eleven multiple PQDs (combination of five PQDs), column six represents two multiple PQDs (combination of six types of PQDs).

3) The highlighted cell in table could be interpreted as a combination of multiple PQ disturbances with classification accuracy of 99.5 percent with SNR of 20dB. Also, (99.5, 97) could be inferred as an accuracy obtained of 99.5 percent and classification class of 97. The six PQDs present being voltage flicker, voltage notch, impulse transient of voltage, oscillatory transient of voltage, voltage harmonics and voltage swell.

It is inferred that, as compared to reported work [9] and [12], the average classification accuracy is over 99% for 96 types of PQDs combinations with an SNR of 20dB which is on upper side. To obtain the optimal parameters for the BI-LSTM based deep learning model, 50 number of epochs are repeated for the model training. Loss and accuracy curves obtained during the training of the proposed model

Table 4 Comparison of the proposed PQ assessment framework with other methods

Type of PQD (20DB Noise)	[11] CNN	2-layer LSTM	BI-LSTM with Dense	BI-LSTM with Attention layer	Proposed method
D1	99.5	99.8	99.7	99.8	99.99
D2	98	99.3	99.5	99.7	99.7
D3	98.5	99.3	99.4	99.6	99.8
D4	100	99.1	99.5	99.7	99.9
D5	100	99.2	99.6	99.7	99.8
D6	100	99.8	99.1	99.3	99.5
D7	100	98.7	99.2	99.6	99.8
D8	100	98.5	98.9	99.2	99.4
D9	98	98.4	98.9	99.4	99.7
Overall accuracy	99.33	99.01	99.31	99.55	99.73

All entries in %

are presented in Fig. 8 and Fig. 9 respectively. The confusion matrix obtained during training and testing of the proposed model is illustrated in Fig. 10 and Fig. 11 respectively. It is found from the results that the classification accuracy of the training confusion matrix is 99.73% for nine types of PQD signals with SNR of 20 dB. The classification accuracy for the proposed model is also presented for nine different basic PQD signals [15] with SNR of 0 dB (noiseless), 20dB and 40dB of noise. Table 3 shows the classification accuracy of nine

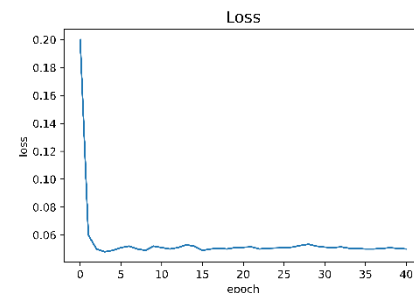


Fig. 8: Loss curve

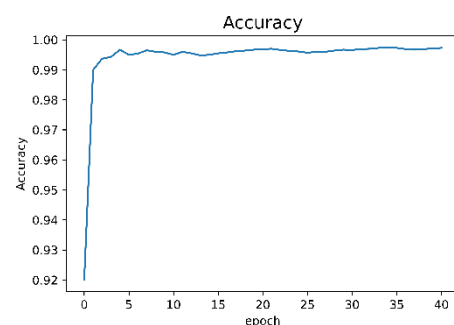


Fig. 9: Accuracy curve

PQD signals including normal voltage signal obtained through the proposed model. Overall classification accuracy of 99.87%, 99.73% and 99.50% was obtained for nine classes of PQD for SNR of 0 dB, 20 dB and 40 dB respectively. Additionally, a comparison of the proposed PQ assessment framework with other methods is illustrated in Table 4, five methods including the proposed methods with SNR of 20 dB are compared with each other for average accuracy and overall accuracy of nine types of PQD signals. It is observed that the proposed method achieves higher accuracy as compared to other

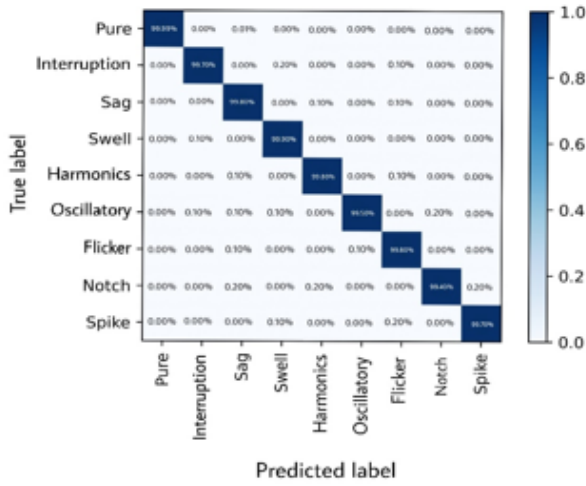


Fig. 10: Confusion matrix training

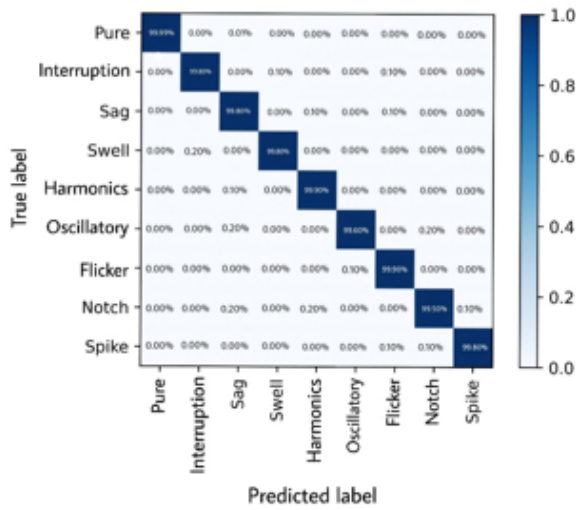


Fig. 11: Confusion matrix test

methods which could be contributed to the fact that attention mechanism and attention gate mechanism focus on the important feature of the PQR. Fig. 12 and Fig. 13 shows the comparison of overall accuracy of the models and average accuracy of the models for each class of PQR respectively.

For comparison of the proposed BI-LSTM method with attention gate and attention layer mechanism, baseline models based on RNN and CNN are considered. For CNN based model, reference [11] is considered and for RNN based methods, self-implemented models using two-layer LSTM model (hidden layers 100), BI-LSTM model (hidden layers 100) with dense layer, BI-LSTM model with attention layer mechanism and the proposed model with BI-LSTM with attention gate applied in the LSTM cell itself with attention layer are considered. The models are tested and trained for PQR signals with 20 dB SNR. It can be inferred from the results, CNN based method gives good performance over LSTM and BI-LSTM based models but addition of attention layer to the BI-LSTM model improves the classification accuracy over the CNN based method and further addition of attention gate to the BI-LSTM model (with attention layer) results in better classification accuracy. It is concluded that the addition of attention mechanism improves the results of classification because of its capability of selecting the most salient part or features of the PQR signal.

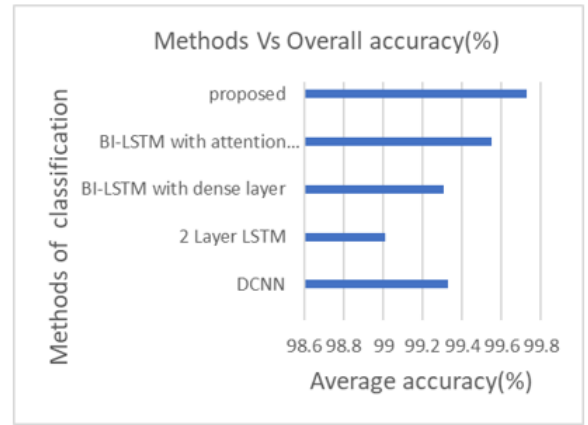


Fig. 12: Comparison of overall accuracy of each different class of PQR.

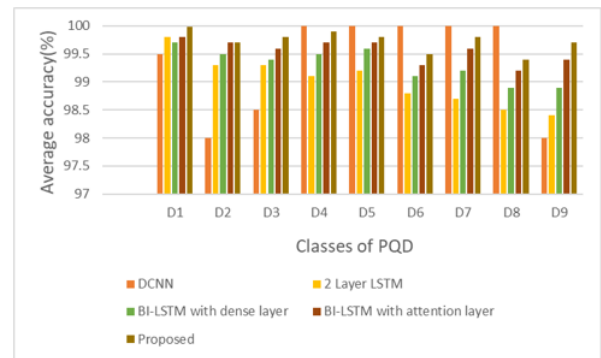


Fig. 13: Gives comparison average accuracy of different models for each class of PQR respectively.

4.1 Classification accuracy with phase shift addition

The PQR voltage signals considered for testing are generally pure signals with same phase. Generally, the methods of PQR classification are trained with PQR voltage signals with same phase or zero initial phase. In this work, signals are also generated with random phase difference with any angle ranging from $-\pi$ to π radians and testing of the proposed model is done with these signals. Table 5 gives the classification accuracy of the proposed model tested with phase shifted PQR signals and signals without phase shifted PQR signals with SNR of 0 dB, 20 dB and 40 dB. It could be seen from the results that addition of phase shift to the PQR signals decreases the classification accuracy to around one percent which is majorly due to the fact that the addition of phase shift to the PQR signal increases the randomness of the signal. In addition to the training of the model with PQR signals without phase shift, the model is also trained with PQR signals with phase shift to obtain good level of accuracy of classification.

5 Conclusion

A deep learning-based model using modified BI-LSTM is successfully implemented in this paper for segmentation and classification of single and multiple (combined) PQRs. Attention mechanism is applied in two different ways to extract important data and knowledge from the voltage signal containing PQR, namely attention gate mechanism and attention layer mechanism. Attention gate is applied to the LSTM cell so that operation of forget gate is limited up to past data and not current data, saving computation time and reducing the training parameters. Attention layer helps in bringing out the most decisive and important information from the output of the BI-LSTM

Table 5 Av. accuracy of classification with and without phase shift in PQD signals

PQD signal	Average accuracy with phase shift	Average accuracy without phase shift
0 dB	98.53	99.87
20 dB	98.32	99.73
40 dB (Noiseless)	98.11	99.5
Average (Av.) Accuracy is marked in %		

layers. Following important conclusions could be drawn from the experimental results:

- 1) Performance of the proposed modified BI-LSTM model is presented and compared with different baseline models in terms of accuracy and effectiveness. The proposed model outperforms the other models in terms of performance.
- 2) Multiple PQD are classified with very high accuracy of over 99 % with SNR of 20 dB. Segmentation of the PQD signal (single or multiple types) is achieved by the proposed model with high accuracy, saving computation time and resources.
- 3) 96 different combinations of PQDs with maximum six type of PQDs with noise are classified with very high accuracy requiring less computational time.
- 4) Classification accuracy is also obtained for PQDs with phase shift showing that phase shift in the PQD signal with SNR of 20 dB brings down the accuracy by around 1%.

As a result, it is concluded that the proposed technique has the potential to detect any combination of PQDs from voltage signal quickly and accurately.

6 Acknowledgments

The authors, Pierluigi Siano and Mohammed AL-Numay acknowledge financial support from the Distinguished Scientist Fellowship Program, King Saud University, Riyadh, Saudi Arabia.

References

- [1] Poras Khetarpal, Madan Mohan Tripathi: 'A critical and comprehensive review on power quality disturbance detection and classification', *Sustainable Computing: Informatics and Systems*, Volume 28, 2020, 100417, ISSN 2210-5379, <https://doi.org/10.1016/j.suscom.2020.100417>.
- [2] M. Markovska, D. Taskovski, Z. Kokolanski, V. Dimchev and B. Velkovski: 'Real-Time Implementation of Optimized Power Quality Events Classifier', *IEEE Transactions on Industry Applications*, vol. 56, no. 4, pp. 3431-3442, 2020, doi: 10.1109/TIA.2020.2991950.
- [3] Beleiu, H.G.; Beleiu, I.N.; Pavel, S.G.; Darab, C.P.: 'Management of Power Quality Issues from an Economic Point of View', *Sustainability* 2018, 10, 2326. <https://doi.org/10.3390/su10072326>
- [4] R. J. Haddad, B. Guha, Y. Kalaani and A. El-Shahat, "Smart Distributed Generation Systems Using Artificial Neural Network-Based Event Classification," in *IEEE Power and Energy Technology Systems Journal*, vol. 5, no. 2, pp. 18-26, June 2018, doi: 10.1109/JPETS.2018.2805894.
- [5] IEEE Recommended Practice for Monitoring Electric Power Quality, Nov. 1995. IEEE Std. 1159-1995.
- [6] S. Santoso, E. J. Powers, and W. M. Grady, "Power quality disturbance data compression using wavelet transform methods," *IEEE Trans. On Power Delivery*, vol. 12, no. 3, pp. 1250-1256, July 1997
- [7] M. Sahani and P. K. Dash, "Automatic Power Quality Events Recognition Based on Hilbert Huang Transform and Weighted Bidirectional Extreme Learning Machine," in *IEEE Transactions on Industrial Informatics*, vol. 14, no. 9, pp. 3849-3858, Sept. 2018.
- [8] P. K. Dash, B. K. Panigrahi and G. Panda: "Power quality analysis using S-transform," in *IEEE Transactions on Power Delivery*, vol. 18, no. 2, pp. 406-411, April 2003.
- [9] M. Valtierra-Rodriguez, R. de Jesus Romero-Troncoso, R. A. Osornio-Rios and A. Garcia-Perez: "Detection and Classification of Single and Combined Power Quality Disturbances Using Neural

Networks', *IEEE Transactions on Industrial Electronics*, vol. 61, no. 5, pp. 2473-2482, May 2014.

- [10] W. G. Morsi and M. E. El-Hawary: 'A New Fuzzy-Based Representative Quality Power Factor for Nonsinusoidal Situations', in *IEEE Transactions on Power Delivery*, vol. 23, no. 2, pp. 930-936, April 2008.
- [11] 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.
- [12] Yaping Deng, Lu Wang, Hao Jia, Xiangqian Tong, Feng Li: 'A Sequence-to-Sequence Deep Learning Architecture Based on Bidirectional GRU for Type Recognition and Time Location of Combined Power Quality Disturbance', *IEEE Trans. Industrial Informatics*, vol. 15, no. 8, pp. 4481-4493, 2019.
- [13] S. Mishra, C. N. Bhende and B. K. Panigrahi: 'Detection and Classification of Power Quality Disturbances Using S-Transform and Probabilistic Neural Network', *IEEE Transactions on Power Delivery*, vol. 23, no. 1, pp.280-287, Jan.2008.
- [14] J. Li, Z. Teng, Q. Tang and J. Song: 'Detection and Classification of Power Quality Disturbances Using Double Resolution S-Transform and DAG-SVMs', *IEEE Transactions on Instrumentation and Measurement*, vol.65, no. 10, pp. 2302-2312, Oct.2016.
- [15] M. V. Chilukuri and P. K. Dash: 'Closure on "Multiresolution S-transform-based fuzzy recognition system for power quality events"', *IEEE Transactions on Power Delivery*, vol. 20, no. 1, pp. 540-, Jan. 2005.
- [16] J. Huang, Z. Jiang, L. Rylands and M. Negnevitsky: 'SVM-based PQ disturbance recognition system', *IET Generation, Transmission & Distribution*, vol. 12, no. 2, pp. 328-334, 30 1 2018
- [17] X. Shi, H. Yang, Z. Xu, X. Zhang and M. R. Farahani: 'An Independent Component Analysis Classification for Complex Power Quality Disturbances With Sparse Auto Encoder Features', *IEEE Access*, vol. 7, pp. 20961-20966, 2019
- [18] D. De Yong, S. Bhowmik and F. Magnago: 'An effective Power Quality classifier using Wavelet Transform and Support Vector Machines', *Expert Systems with Applications*, vol. 42, pp. 6075-608, September 2015
- [19] F. Renna, J. H. Oliveira, and M. T. Coimbra: 'Deep convolutional neural networks for heart sound segmentation', *IEEE journal of biomedical and health informatics*, 2019
- [20] Y. Xie, R. Liang, Z. Liang, C. Huang, C. Zou and B. Schuller: 'Speech Emotion Classification Using Attention-Based LSTM', *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 27, no. 11, pp. 1675-1685, Nov. 2019, doi: 10.1109/TASLP.2019.2925934.
- [21] S. O. Sahin and S. S. Kozat: 'Nonuniformly Sampled Data Processing Using LSTM Networks', *IEEE Transactions on Neural Networks and Learning Systems*, vol. 30, no. 5, pp. 1452-1461, May 2019, doi: 10.1109/TNNLS.2018.2869822.
- [22] K. Greff, R. K. Srivastava, J. Koutník, B. R. Steunebrink and J. Schmidhuber: 'LSTM: A Search Space Odyssey', *IEEE Transactions on Neural Networks and Learning Systems*, vol. 28, no. 10, pp. 2222-2232, Oct. 2017, doi: 10.1109/TNNLS.2016.2582924.
- [23] Z. Xueqing, Z. Zhansong and Z. Chaomo: 'Bi-LSTM Deep Neural Network Reservoir Classification Model Based on the Innovative Input of Logging Curve Response Sequences', *IEEE Access*, vol. 9, pp. 19902-19915, 2021, doi: 10.1109/ACCESS.2021.3053289.
- [24] A. Pandey and D. Wang: 'Dense CNN With Self-Attention for Time-Domain Speech Enhancement', *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 29, pp. 1270-1279, 2021, doi: 10.1109/TASLP.2021.3064421.
- [25] Granados-Lieberman, D.; Valtierra-Rodriguez, M.; Morales-Hernandez, L.A.; Romero-Troncoso, R.J.; Osornio-Rios, R.A.: 'A Hilbert Transform-Based Smart Sensor for Detection, Classification, and Quantification of Power Quality Disturbances', *Sensors* 2013, 13, 5507-5527. <https://doi.org/10.3390/s130505507>