

# Enhanced Deep Convolutional Neural Network for Fault Signal Recognition in the Power Distribution System

**Srinivasa Rao T C**

Vasavi College of Engineering,  
Shamshabad, India  
srinivasarao.tc@gmail.com

**Tulasi Ram S S**

J.N.T.U.H.,  
Kukatpally, Hyderabad, Telangana, India

**Subrahmanyam J B V**

C.J.I.T.S.,  
Warangal Dist Telangana, India

**Abstract:** Currently Electrical power system is represented as the complex artificial system throughout the world, as the economic and social advancements based upon consistent, intact, stable and economic models. Because of varied arbitrary causes, unintentional failures happen in electrical power systems. Based on this problem, this work aims to develop EDCNN to detect and classify the fault signals like sag transient and swell in the transmission line. Moreover, the wavelet decomposed fault signals are extracted and identify the fault on the basis of the decomposed signal using the EDCNN approach. Moreover, this paper presents the performance evaluation by deciding the Type I and Type II measures and RMSE measures. In the performance analysis, it evaluates the performance of EDCNN method to several existing methods. The experimentation outcomes examine that the proposed EDCNN method efficiently and classifies recognizes the fault signal in the power distribution system while comparing with the conventional methods such as linear- Support Vector Machine (SVM), quadratic-SVM, and Radial Basis Function (RBF)-NN and Gradient Descent (GD)-NN models.

**Keywords:** Power System; Fault Recognition; Fault Signal; Wavelet Decomposition; EDCNN

## Nomenclature

Abbreviations	Descriptions
EDCNN	Enhanced Deep Convolutional Neural Network
RMSE	Root-Mean-Square-Error
WT	Wavelet Transform
ADN	Active Distribution Network
DS	Distribution System
SVM	Support Vector Machine
LLG	Line-to-Ground Fault
SOGDT	Signal Optimization Based Generalized Demodulation Transform
DN	Distribution Network
ANN	Artificial Neural Network
FSC	Full-wave Signal Construction
ML	Machine Learning
MCA	Mirror-Cycle-Add
MCSA	Motor Current Signal Analysis
SR	Stochastic resonance
DNN	Deep Neural Network
HIF	High Impedance Fault
EMD	Empirical Mode Decomposition
DG	Distributed Generator
MODWPT	Maximum Overlap Discrete Wavelet Packet Transform
PCA	Principal Component Analysis

## 1. Introduction

The detection, identification, and classification, of faulty location and phase of faults on power transmission lines, are the major aims of some protective relaying method. The contemporary power system networks are extremely interrelated, therefore a directional relaying method that gets on with it acts in response to faults in forward direction merely and in the protection zone without intentional time delay is required to guarantee the maximum speed fault clearing and uphold the system steadiness. Faulty phase detection, fault classification, and it is exploited to determine its position, which is necessary to perform the protection of power transmission line and permit single-pole tripping of line circuit breakers in the presence of single-phase to earth faults.

A great number of distributed energy has been linked into the DN that creates the conventional DN change into an ADN. The ADN is helpful for the enhancement of the environment; however, the issues of process stability turn out to be more and more well-known concurrently. ADN has an additional stretchy topological model. Its topology is not a single radial form anymore, however rising towards a meshed form just similar to a transmission network. At the load nodes, a great number of DGs are linked that slowly translate the passive DN to an active power DN. Whether these DGs must be linked to the network is ordered using the power supply and real-time load demand. The precise association positions and capability are not unalterable, so an additional fast optimization algorithm of DGs' sizing and sitting in the DN is essential for dynamic regulation

A number of fault position strategies were used for un-compensated transmission lines [6] that are stated in the feature at this time. In [6], the negative sequence phase angle was exploited to the estimation of the fault position for only single LLG. Other kinds of faults, like LLG, line fault (LLL), Line-To-Line fault (LL) and triple multi-location faults and converted faults are not considered in this work. One more fault position method for LL only exploits matrix inverse lemma was presented in [7] for direct circuit analysis. In [8], Fault position method for a single-circuit line on the basis of the bus-impedance matrix was stated that offer inadequate accurateness exploiting one bus measurements, though, superior accurateness was attained exploiting 2 bus voltage measurements. Various soft computing methods were exploited for fault position estimation, like wavelet-fuzzy combined algorithm [13], ANN [9] [10] [11] and [12] SVM- WT [14], and combined wavelet-neuro-fuzzy combined approach [15].

In dynamic conditions, a synchronized phasor measurement unit was exploited to calculate the fault position in [13] exploiting two-terminal data. The majority significant requirement of these methods [3–13] was that they need fault classification, i.e., fault type to be recognized initial to start the fault position methods. But that, these methods were exploited for un-compensated lines that might not work accurately for series compensated lines.

In recent times, a lot of intelligent diagnostic algorithms on the basis of the deep learning were presented, to provide an effectual technique to rapidly procedure a great number of fault data and robotically construct precise diagnostic outcomes. Nevertheless, the majority of the current intelligent algorithms were exploited to diagnose faults, which were on the basis of the vibration signals. Because of the characteristic differences among vibration signals and current signals [17], perfect diagnostic outcomes cannot be attained through directly exploiting the spectrum signal as the sample.

The main contribution of this paper is to present the EDCNN model, as the fault diagnosis issue is able to organize as a series of binary classification issue. Additionally, the wavelet decomposed fault signal is extracted, and additionally, this work presents the fault signal identification on the basis of the decomposed signal. Hence, the precise fault diagnosis of the power DS is attained using the wavelet signals descriptors and EDCNN method.

## 2. Literature Review

In 2019, Dezun Zhao et al [1], recommended the fault diagnosis approach in the power DS by exploiting the DBN approach. In this experimentation, the fault signals considered were incorporated sag, transient, and swell signal. At first, wavelet decomposed fault signals were extracted, and an additional DBN method was exploited to identify the fault on the basis of the decomposed signal. To the next of the experimentation, RMSE, Type I, Type II measures were decided to analysis the intention. Moreover, in recognizing fault, the evaluation of the DBN method was analyzed with conventional methods such as quadratic-SVM, linear-SVM, polynomial-SVM, RBF-SVM, LM-NN MLP-SVM, and GD-NN models. Hence, the DBN model was superior to the conventional methods in recognizing the fault signals in the power DS. Here, a new SOGDT was presented in order to roll the bearing nonstationary fault characteristic extraction.

In 2019, Feng Li et al [2], worked on failures in planetary gearboxes that can happen downtime, accidents, and maximum preservation costs. In rotating machinery, MCSA presents a non-intrusive algorithm to detect mechanical faults. In recent times, a lot of intelligent diagnostic algorithms on the basis of the deep learning model were developed, provided that an effectual method to rapidly procedure a great number of fault data and robotically create precise diagnostic outcomes. Nevertheless, the majority of current intelligent algorithms to diagnose faults were on the basis of the vibration signals. Because of the characteristic differences among ideal diagnostic, vibration signals and current signals, outcomes cannot be attained by directly exploiting the spectrum signal as the sample. This work diagnosed faults in planetary gears by preprocessing current signals and exploiting DNN methods.

In 2019, Kuo Chi et al [3], worked on SR to have the ability to improve the weak fault-induced impulses in envelope signal. Nevertheless, the envelope signal was a unipolar half-wave signal that was poorer than the bipolar full-wave signal for the SR effect. To transfer the envelope signal into a bipolar full-wave signal, two FSC algorithms on the basis of the MCA process were proposed named FSCBHCD and FSCBME. Moreover, the FSCBME was to discover the delay time, a significant MCA parameter which increases the energy of the full-wave signal modeled. FSCBHCD was to create the lag time capable of half of the driving signal cycle. Concerning the full-wave signal was modeled as SR input; two new SR algorithms were presented named FSCBHCD SR and FSCBMESR.

In 2019, Debadatta Amaresh Gadanayak and Ranjan Kumar Mallick [4] addressed the challenges in HIF detection since it was always related to nonlinear, asymmetric, low magnitude, and arbitrary fault currents. Here, a novel model for arcing HIF detection by means of the analysis of the inter harmonic content of the current signal in distribution feeders. The source of interharmonics was associated with the arbitrary variation of the electric arc-related with HIF. A new algorithm exploiting MODWPT and a recently presented particular knot on the basis of the EMD was proposed for effectual separation of inter harmonic components.

In 2019, Muhammad Sarwar et al [5], developed a precise HIF isolation and detection method in a power distribution network. From current and voltage sensors the developed method uses the data available. The model uses multiple techniques consisting of Fisher Discriminant Analysis, PCA, Binary and Multiclass SVM for identification and detection of the HIF. These data-driven models experimented on IEEE 13-node DN for the identification and detection of HIF with the unbroken and a broken conductor. Moreover, the robustness of ML approaches was been evaluated.

### 3. Signal Decomposition and Power Distribution System

#### 3.1 Proposed Methodology

Fig. 1 demonstrates the proposed method for fault classification and diagnosis of transmission signals in the power system. Moreover, by exploiting the wavelet descriptors, the extraction of the features of the signal; and utilizing the EDCNN model, the signal classification and diagnosis are performed. Moreover, the specimen bus system presents the training signals, which can be exploited for the EDCNN model. In reality, the specimen bus system consists of the normal signal and the equivalent fault signals like transient, sag, and swells signals. Ahead of exploiting the signals to the EDCNN model, the wavelet descriptors of every signal are recognized. Moreover, the fault signal is diagnosed, and the EDCNN method classifies the categories of signals.

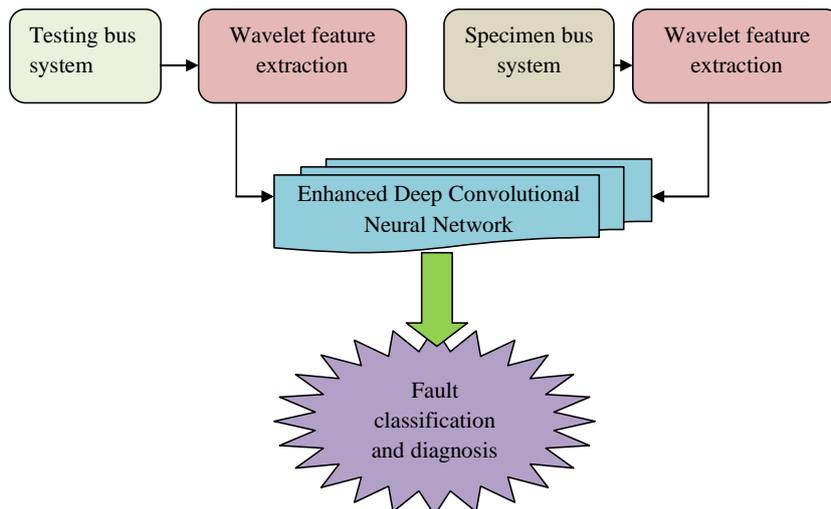


Fig. 1. Schematic illustration of the proposed methodology for fault diagnosis

### 3.2 Signal Decomposition

In a power bus system, the number of transmission lines is diverse besides its length. Consider a bus system consist of transmission lines with three different distances such as  $L_1$  (in km),  $L_2$  (in km) and  $L_3$  (in km). Additionally, the position of the fault can vary as each transmission line has its own load. Hence, previous to localizing fault, the measurements of every transmission line must be integrated with each other. Nevertheless, it a difficult job to confine the fault incidence in such conditions.

Consider  $x = [u_1; u_2; u_3; \dots; u_N]$  as the integrated voltage measurements, whereas the signal descriptors are indicated as  $u_1; u_2; u_3; \dots; u_N$  and denote the number of voltage sources. Consider  $I$  be the number of observation instances, and  $N_O$  as the number of observation records, in order that the size of the integrated voltage sources is allocated as  $N_o \times I$ . In, eq. (1) indicates the wavelet decomposition of the integrated voltage measurements, whereas  $\psi_{p,q}(t)$  denotes the wavelet coefficient (depicted in eq. (2)), that exploits two integers like  $p$  and  $q$  for its explanation, whereas  $p$  and  $q$  show the scaling and sampling numbers, correspondingly. In eq. (2) the phrase  $(c_0, d_0)$  indicates the constant which is fixed as  $(2, 1)$ .

$$D(x) = \int_{-\infty}^{\infty} f(t) \chi_{p,q}^*(t) dt \quad (1)$$

$$\psi_{p,q}(t) = \frac{1}{\sqrt{c_0^p}} \chi\left(\frac{t - nd_0 c_0^p}{c_0^m}\right) \quad (2)$$

In addition,  $|D(x)|^R \times N_o \times I$  indicates the size of the decomposed signal, whereas  $|y|^R$  indicates the number of rows in a matrix  $y$ . As the higher dimension pacts to raise the calculation complexity, the dimension  $D(x)$  is important to minimize. Hence, to minimize the dimension of  $D(x)$ , PCA is utilized.

### 3.3 Conventional DCNN

The basic model of DCNN [16] comprises of few steps that might emerge numerous times in an arranged manner in an exacting network. The major three layers of a DCNN are the convolution layer, pooling layer, and fully connected layer. An exacting deal of these layers begins with the input layer and terminates with the output layer. These layers are explained in a brief manner as stated below:

#### i) Input Layer

In every DCNN, this is the initial layer. Here, usually, the input is considered as the pixel values attained from any image. Occasionally raw pixel values are supplied to the network; however few forms of pre-processing can be also performed for superior outcomes. For grayscale images, the shape of the input would be (height of the image)  $\times$  (width of the image)  $\times$  (Number of Channels = 1). In the scenario of true or RGB color images the input shape will be (height of the image)  $\times$  (width of the image)  $\times$  (Number of Channels = 3). Nevertheless, DCNN can as well obtain 3D data or 1D data volume and procedure these with a small adaptation.

#### ii) Convolutional Layer

The convolutional layer is the central element of DCNN. The convolution process is the construction block of this layer. Few filters or kernels are needed which can discover exact structures from a presented image data for the convolution process.

One more term exploited in DCNN is a pace that indicates the number of cells to be transferred. The technique filters work is that the filter is positioned at the topmost corner of the input image subsequently a part-wise production is done among the values of the filter and equivalent values of the image pixels subsequently; the generated outcomes are summed to obtain a value. The summation increased is subsequently positioned at one cell of the output. Subsequently, the filter is enthused to the right one stride and similar computations are frequent. Numerous filters can be exploited on a similar layer. When exploiting more than one filter, the convolution process is performed with all these filters independently and the outputs are subsequently stacked together that are exploited as input to other layers. Here, a noteworthy point is so as to the depth of the filter in a layer should be equal to the depth of the layer channel. Hence, for RGB images filter of shape  $f \times f \times 3$  is to be exploited for the first convolution layer as the input image encompass the shape height  $\times$  width  $\times 3$  (here,  $f$  is the size of filter).

Convolution process can be of two kinds, such as the same convolution and valid convolution. For instance, the valid convolution, the output shape is minimized. In the scenario of a similar convolution

process, the input image is 0 padded in order that the output encompasses a similar width and height as the input. The formerly stated instance exhibits a predefined filter being exploited. Nevertheless, normally the filters to be exploited are initialized with arbitrary numbers. By exploiting the backward propagation and feed-forward, these arbitrarily selected numbers are optimized to robust for any exacting issue. One iteration of backward propagation and feed-forward is known as an Epoch.

An activation function is needed to append non-linearity for increasing a widespread solution. The convolution layer output is augmented with a bias value (part-wise) and subsequently, it is conceded during an activation function. Various kinds of activation functions can be exploited such as tanh, sigmoid, leaky ReLU, ReLU, function and so on. In this paper, ReLU activation function is exploited widely, which is practical.

### iii) Pooling Layer

The pooling layer extracts standard features from the input and minimizes the width and height of the input. This layer is usually exploited subsequently to one or more convolution layers. Also, by minimizing the dimension, this layer maximizes computational performance using minimizing the number of parameters these layers reduce the possibility of in excess of the fitting. Pooling can be performed exploiting numerous techniques, such as max-pooling (whereas the utmost value is occupied from an exacting shape of filter), average pooling (the average value from a definite shape of filter) and so on.

### iv) Fully Connected Layers

Subsequent to transient the input data by few pooling and convolution layers, the output from the final layer is flattened that is reshaped into a linear array and presented as input to the nodes of a fully connected layer. Subsequent to fetching the outputs of all preceding layers into a single linear array a matrix multiplication with the weights is performed followed using an adding with a bias value.

### v) Output Layer

Subsequent to a fully connected layer, there is an output layer. These final two layers can be evaluated with the final layers of common NN. Nevertheless, a softmax output unit is commonly exploited to acquire the output. A softmax unit presents a probability to each of the output classes that augment up to 1. The class comprising the highest value will be the accurate class.

## 3.4 Proposed Enhanced Deep Convolutional Neural Network

By LeNet-5 architecture, the DCNN model exploited is enthused [22]. This model comprises seven layers. Meanwhile, the input images are feed to the network enchanting 50 images as a batch. All images are re-sized to contain 32 x 32 dimensions and all these images are true to color or RGB images. For optimization Adam (Adaptive Moment Estimation) optimizer is exploited. Adam optimizer takes into account the preceding values of the parameter for creating the learning. The updating rule for Adam is stated in (3), (4) and (5).

$$u_t = \alpha_1 u_{(t-1)} + (1 - \alpha_1) d\omega \quad (3)$$

$$r_t = \alpha_2 r_{(t-1)} + (1 - \alpha_2) d\omega^2 \quad (4)$$

$$W = W - \beta \frac{u_t}{\sqrt{r_t + \epsilon}} \quad (5)$$

In eq. (3), (4) and (5),  $\alpha_1$  and  $\alpha_2$  indicates the two hyperparameters. Additionally,  $\alpha$  indicates a hyperparameter that states the learning rate.  $d\omega$  indicates the gradient of the learning parameter and  $\epsilon$  indicates a constant value added with the denominator to evade division by zero. For performance  $\alpha_1 = 0.9$ ,  $\alpha_2 = 0.99$  and  $\beta = 0.001$  were exploited that are exploited conservatively. The parameter and hyperparameter (size of filter, stride, and padding number so on) settings for each layer are stated below

Subsequent to the fully connected layer a softmax output unit assumes the groups. All 10 output classes are assigned a probability value which sums up to 1. The class comprising the greatest probability is the selected calculation. Initially, all the parameters that are the values of the bias values and filters for every layer are arbitrarily initialized for training the model. Subsequent to one pass of feedforward, the Adam optimizer exploits the derivative of the loss function regarding parameters to optimize all the parameters to robust the model exploiting eq. (3). Here, the loss function is a measurement of how better the model is performing that is computed using enchanting as seem at the predicted class and the actual class. Numerous kinds of loss functions can be exploited. In this work, definite cross-entropy is exploited as a loss function. This function is stated in eq. (6).

$$G_p = -\sum_i p_i \log(x_i) \quad (6)$$

In eq. (6),  $p$  indicates the predicted class and  $x$  indicates the true class.  $G_p$  indicates the loss function of  $p$  regarding  $x$ . Adam optimizer discovers a parameter set which creates the value of this loss function lowest. Experiment data exploited for this model are exploited to confirm the model and, as a result, that the model is widespread appropriately.

## 4. Results and Discussions

### 4.1 Simulation Set up

The simulation with respect to the fault signal recognition in the distributed power system was experimented in MATLAB and also shows the experimentation results. Here, the proposed technique was compared with the conventional techniques with respect to the swell, sag, and transient signals. Here, the performance measures such as accuracy, sensitivity, specificity, and precision, False Positive Rate (FPR), False Negative Rate (FNR), Negative Predictive Value (NPV), False Discovery Value (FDR), F1-Score, G-Mean and Matthews Correlation Coefficient (MCC) for validating the proposed algorithm.

### 4.2 Performance Analysis

In this section, the performance evaluation of the proposed model with the conventional model for swell, sag, and transient signals. Fig 2, 3, 4 and 5 show the performance evaluation of the proposed technique with existing techniques for normal signals, swell, sag and transient signals. In Fig 2, the accuracy of the proposed method is 35% better than the Linear SVM, 12% better than the Quadratic SVM, 34% better than the RBF-SVM, and 11% better than the GDNN algorithms. The accuracy of the proposed method is 22% better than the Linear SVM, 21% better than the Quadratic SVM, 14% better than the RBF-SVM, and 19% better than the GDNN algorithms and it is shown in Fig 3. In Fig 4, the accuracy of the proposed method is 16% better than the Linear SVM, 18% better than the Quadratic SVM, 14% better than the RBF-SVM, and 21% better than the GDNN algorithms.

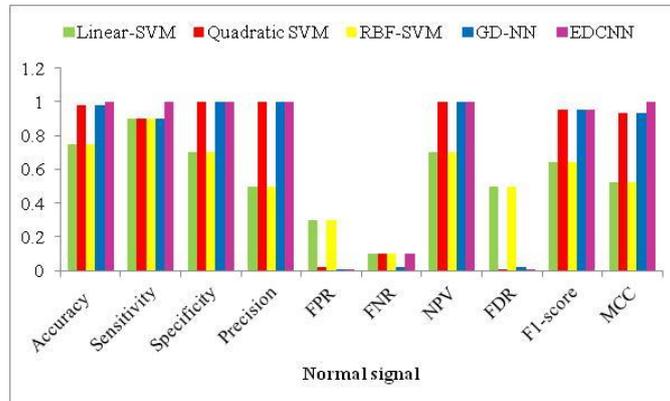


Fig. 2. Analysis of the proposed technique with conventional models for normal signal

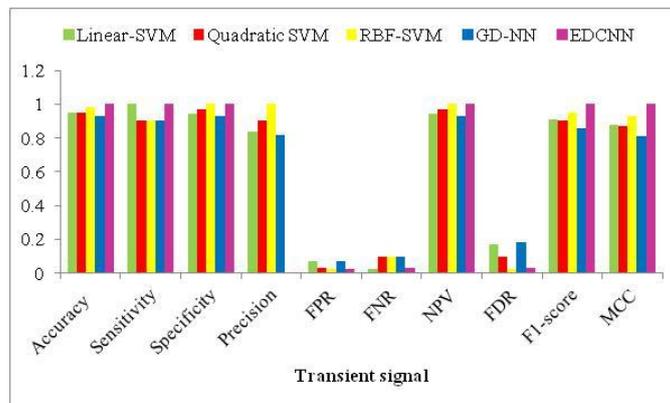
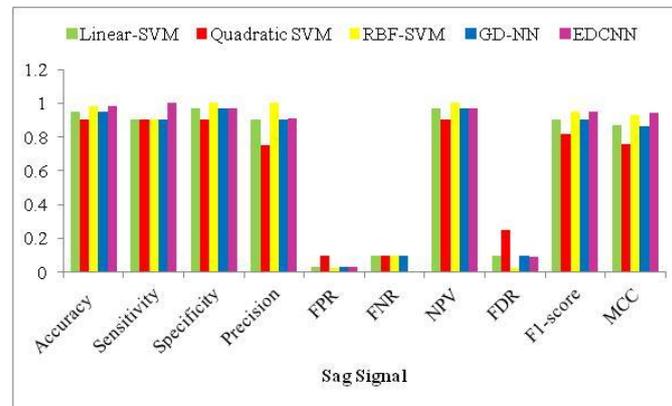
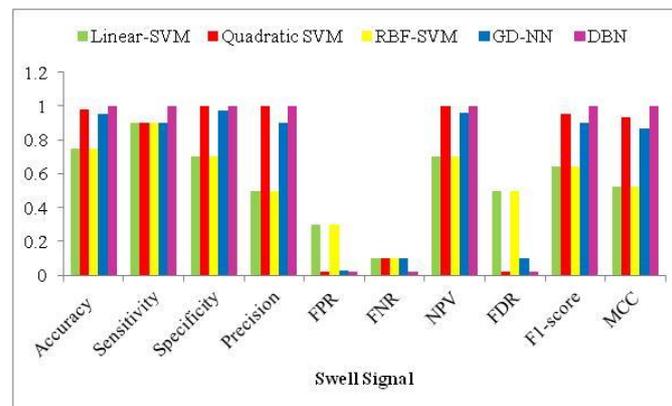


Fig. 3. Analysis of the proposed model with the conventional models for Transient signal



**Fig. 4.** Analysis of the proposed technique with conventional models for Sag signal



**Fig. 5.** Analysis of the proposed technique with the conventional models for Swell signal

## 5. Conclusion

In this paper, the fault diagnosis methodology was introduced in the power distribution system by exploiting an Enhanced Deep Convolutional Neural Network (EDCNN) method. Moreover, the fault signals have considered in this work, which incorporated such as sag, transient, and swell signal. At first, the extraction was done on the basis of the wavelet decomposed fault signals, and additionally, EDCNN model was exploited to diagnose the fault on the basis of the decomposed signal. After the implementation, Type I, Type II was determined for the evaluation purpose. At last, the analysis of the EDCNN model in recognizing fault was analyzed with conventional models such as linear-SVM, quadratic-SVM, and RBF-NN and GD-NN models. Finally, the performance of the EDCNN model was compared with the conventional models in diagnosing the fault signals in the power distribution system.

## References

- [1] Dezun Zhao, Tianyang Wang, Robert X. Gao, Fulei Chu, "Signal optimization based generalized demodulation transform for rolling bearing nonstationary fault characteristic extraction", *Mechanical Systems and Signal Processing*, Volume 134, 1 December 2019.
- [2] Feng Li, Xinyu Pang, Zhaojian Yang, "Motor current signal analysis using deep neural networks for planetary gear fault diagnosis", *Measurement*, Volume 145, Pages 45-54, October 2019.
- [3] Kuo Chi, Jianshe Kang, Rusmir Bajric, Xinghui Zhang, "Bearing fault diagnosis through stochastic resonance by full-wave signal construction with half-cycle delay", *Measurement*, Volume 148, December 2019.
- [4] Debadatta Amaresh Gadanayak, Ranjan Kumar Mallick, "Interharmonics based high impedance fault detection in distribution systems using maximum overlap wavelet packet transform and a modified empirical mode decomposition", *International Journal of Electrical Power & Energy Systems*, Volume 112, Pages 282-293, November 2019.
- [5] Muhammad Sarwar, Faisal Mehmood, Muhammad Abid, Abdul Qayyum Khan, Adil Sarwar Khan, "High impedance fault detection and isolation in power distribution networks using support vector machines", *Journal of King Saud University - Engineering Sciences*, In press, corrected proof, Available online 10 July 2019.
- [6] Wang, B., Dong, X., Lan, L., Xu, F.: 'Novel location algorithm for single-line-to-ground faults in transmission line with distributed parameters', *IET Gener. Transm. Distrib.*, 2013, 7, (6), pp. 560-566
- [7] Choi, M.S., Lee, S.J., Lim, S.I., Lee, D.S., Yang, X.: 'A direct three-phase circuit analysis-based fault location for line-to-line fault', *IEEE Trans. Power Deliv.*, 2007, 22, pp. 2541-2547

- [8] Liao, Y.: 'Fault location for single-circuit line based on bus-impedance matrix utilizing voltage measurements', *IEEE Trans. Power Deliv.*, 2008, 23, pp. 609–617
- [9] Jain, A., Thoke, A.S., Modi, P.K., Patel, R.N.: 'Classification and location of single line to ground faults in double circuit transmission lines using artificial neural networks', *Int. J. Power Energy Convers.*, 2010, 2, (2), pp. 109–125
- [10] Yadav, A., Thoke, A.S.: 'Transmission line fault distance and direction estimation using artificial neural network', *Int. J. Eng. Sci. Technol.*, 2011, 3, (8), pp. 110–121
- [11] Yadav, A., Warlyani, P., Thoke, A.S.: 'Fault classification, distance location and faulty section identification in teed transmission circuits using artificial neural network', *Int. J. Comput. Appl.*, 2012, 47, (15), pp. 14–21
- [12] Jain, A.: 'Artificial neural network-based fault distance locator for double-circuit transmission lines', *Hindawi-Adv. Artif. Intell.*, 2013, 2013, ID 271865, pp. 1–12
- [13] Reddy, M.J., Mohanta, D.K.: 'A wavelet-fuzzy combined approach for classification and location of transmission line faults', *Int. J. Elect. Power Energy Syst.*, 2007, 29, pp. 669–678
- [14] Ekici, S.: 'Support vector machines for classification and locating faults on transmission lines', *Applied Soft Comput.*, 2012, 12, pp. 1650–1658
- [15] Reddy, M.J., Mohanta, D.K.: 'Adaptive-neuro-fuzzy inference system approach for transmission line fault classification and location incorporating effects of power swings', *IET Gener. Transm. Distrib.*, 2008, 2, (2), pp. 235–244.
- [16] Yiping Du, "Bayesian Deep Convolutional Neural Network for SMS Quality Analysis", *Journal of Visual Communication and Image Representation*, In press, journal pre-proof, Available online 10 October 2019.
- [17] T C Srinivasa Rao, S S Tulasi Ram and J B V Subrahmanyam, "Fault Recognition and Diagnosis Based on a Decision Tree for Power Distribution Systems", *Journal of Computational Mechanics, Power System and Control (JCMPS)*, Volume 2, Issue 2, April 2019.