Indonesian Journal of Electrical Engineering and Computer Science
Vol. 34, No. 2, May 2024, pp. 1024∼1032
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v34.i2.pp1024-1032 ❒ 1024

Implementing generative adversarial networks for increasing performance of transmission fault classification

Tilottama Goswami[1], Uponika Barman Roy[2], Deepthi Kalavala[3], Mukesh Kumar Tripathi[4]
[1]Department of Information Technology, Vasavi College of Engineering, Hyderabad, India
[2]Department of Electrical and Computer Engineering, University of Western, London, Canada
[3]Department of Computer Science Engineering, Koneru Lakshmaiah Education Foundation, Hyderabad, India
[4]Department of Computer Science and Engineering, Vardhaman College of Engineering, Hyderabad, India

Article Info

Article history:

Received Dec 1, 2023
Revised Dec 14, 2023
Accepted Jan 19, 2024

Keywords:

Classification
GAN
Random forest
Simulink
Transmission fault

ABSTRACT

An electrical power system is a network that facilitates the sourcing, transfer, and distribution of electrical energy. In the traditional power system, there are eleven types of faults that can occur in the system. This paper focuses on the classification of these faults over a stretch of 100 kilometres. The dataset used is synthetic and generated from a simulated model using MATLAB/Simulink software. Data augmentation is carried out during training to improve the accuracy of the classification. An indirect training approach through generative adversarial network (GAN) is used to classify these overhead transmission line faults. The random forest (RF) classification is used as the base learning model on the original dataset and it achieves accuracy of 84%. However, the base learner RF when used on GAN model generated augmented faulty data, it performs exceptionally well achieving 99% accuracy. One of the recent state-of-art methods is compared with this approach.

Corresponding Author:

Tilottama Goswami
Department of Information Technology, Vasavi College of Engineering
Hyderabad, 500031, Telangana, India
Email: tgoswami@staff.vce.ac.in

1. INTRODUCTION

The demand for electricity is growing rapidly due to population growth, resulting in the expansion of electrical power stations. Overhead transmission lines are an essential component of the electrical grid, carrying high voltage electricity to local transformer substations, which then distribute the power to consumers through distribution lines. However, this supply chain faces significant risks at all levels, particularly the overhead transmission lines, which carry high voltages and are vulnerable to faults. A fault is an abnormality in the functioning of a power system that causes a deviation in current and voltage values [1]. These faults can arise from a variety of sources, including bad weather conditions, equipment damage, fire accidents, and human error. When a fault occurs, there is a surge in current flow as voltage approaches zero.

There are two primary categories of faults: shunt faults and series faults. Shunt faults are the most common type and occur when there is a loss of insulation between the phase conductors, causing a heavy current to flow through electrical equipment. Series faults occur due to the loss of a fuse or failure of one or more phase conductors. There are also two other categories of faults: symmetrical and unsymmetrical. A symmetrical fault occurs when there is a fault in all three phases, while an unsymmetrical fault occurs when one or two phases are faulty. For a three-phase overhead electrical network, there are eleven combinations of phases, including the ground wire, making it significantly challenging to identify the occurrence of a fault and which phases are impacted [1], [2]. Different types of faults are described in Table 1.

To study and analyse the electrical fault occurrence in an extensive complicated system like overhead transmission line requires enough data points. Goswami and Roy [2], experimented with 11,000 data points for 11 different combinations of three phase transmission lines across 100 kilometres. The experiment was vast indeed but for any electrical system, the more data is analysed the better the analysis is. It is noteworthy that, the real time data collection of the overhead lines' current and voltages is one of the significant challenges hence the data generation was simulated in [2]. With the set of 11,000 data the best result was obtained as 84%. This paper aims to generate more data points by using extensive data augmentation techniques to load more data into the system and then to perform the machine learning classifications. To perform data augmentation, the authors propose the use of generative adversarial network (GAN) [3] to add more data to the original synthetic fault dataset. A random forest (RF) classifier is used as the base learner to understand the model's base performance. The performance of RF classifier on original data is compared to its performance against the data including original dataset and GAN augmented to bring out uniquely how data augmentation impacts a model's performance in an electrical system.

This paper includes a literature survey of research related to fault classification in section 2 and proposes a methodology for implementing the GAN algorithm on the fault dataset in section 3. Section 4 gives details about GAN architecture and data augmentation used in this paper. Section 5 compares the results of the proposed method with existing implementations, and section 6 concludes the paper and gives a direction for future work. In summary, this paper highlights the significance of accurate fault classification in ensuring the stability and reliability of power systems, and demonstrates the potential of advanced machine learning techniques such as GANs to improve fault classification.

Table 1. Types of faults in electric system (used in this work)

Type of fault	Cause of the fault	Probability of occurrence	Corresponding class labels
Line-to-Ground	One of the three transmission lines short with the ground	0.7	AG, BG, CG
Line-to-Line	A phase comes in contact with the other due to heavy wind	0.15	AB, BC, AC
Double-Line-to-Ground	Two phases come in contact with the ground potential	0.1	ABG, BCG, ACG
Three Phase	The entire system gets disrupted due to fatal failure of the towers, rupture of the transmission line and heavy wind flow	0.05	ABC, ABCG

2. LITERATURE REVIEW

The field of power system research is vast, with potential for new discoveries that can significantly impact electrical networks. One area of focus is fault identification and classification, which has prompted numerous investigations around the world. This paper explores some promising analyses in this area. Goswami and Roy [2], classify faults in three-phase overhead transmission lines that span 100 kilometers. Using a Simulink model, they generate faulty data for 11 categories of faults. The authors employ supervised machine learning classification models, including decision tree (DT), K-nearest neighbor (kNN), and spport vector machine (SVM). The 11-class data is split into a train-test set with an 80-20 ratio. It is found that SVM with a radial basis kernel function performs the best with an accuracy of 91.06%. Anika *et al.* [3] discovered that neural network is one of the promising fields to research on electrical power system. Anika *et al.* [3] focuses on identifying and classifying faults in power system bus bars using artificial neural networks. The dataset consists of four fault types for a nine-bus system obtained from the IEEE 9 bus system architecture. With an input matrix data of 7,776×54, the authors allocate 70% of the data for training and divide the remaining 30% into validation and test sets of 15% each. They use the neural network pattern recognition toolbox from MATLAB to incorporate a two-layer feedforward network structure on the train set data. The authors report a maximum accuracy of 80% for detecting and classifying faults.

Pan *et al.* [4] presents a fault classification mechanism for the distributed generation system, which is a wing of power electronics. The authors not only use convolutional neural networks (CNN) to classify faults in this system but also successfully introduced image based deep learning model to train on an electrical system. They preprocess voltage and current data with a MATLAB tool-wavelet transform, converting them into images as scalogram that are further trained with CNN. The data is split into training, validation, and testing sets, with the training set consuming 70% of the data. The CNN model trains the data in 13 layers with a 0.001 learning rate, optimizer as Adam, 50 maximum epochs, and 15 minimum batch size. The validation frequency is maintained at 10, and this model results in an average testing accuracy of 98.9%. A wavelet neural network approach is used in another study [5] to classify faults in a hybrid network of distributed generators for wind and photovoltaic systems. The data is generated from MATLAB/Simulink for a range of 100 kilometres distance and 30 kilovolts of voltage. The data is split equally for the training, validation, and testing sets. They train the input data of size 560×420 for faults located every 10 kilometres using the wavelet neural network. The model achieves a classification accuracy of 99.4% at 2,520 epochs. The authors contributed to the highest possible performance of an electrical system using simulated data, which is like the proposed methodology, for a stretch of 100 kilometers.

All these literature surveys helped the authors to realize that the electrical system fault classification is being researched with powerful deep learning models like neural network and CNN. The supervised models are good to train on the data as they are labelled but the authors discovered a scope to experiment a unsupervised models to explore how impactful it is to identify the patterns of the data and then to classify. GAN is primarily known for image augmentation but the usage of CNN in [4] on the MATLAB simulated numerical data also explained a potential use of GAN for data augmentation as well as to introduce unsupervised technique to classify the faults in this paper.

3. METHOD

The proposed method uses GAN [6] for data augmentation and thereby improving the classification model's performance. RF classifier is used as a classifier to classify 11 different types of faults in a three-phase overhead transmission line. Figure 1 represents the approach how input data goes through different steps to classify the faults.

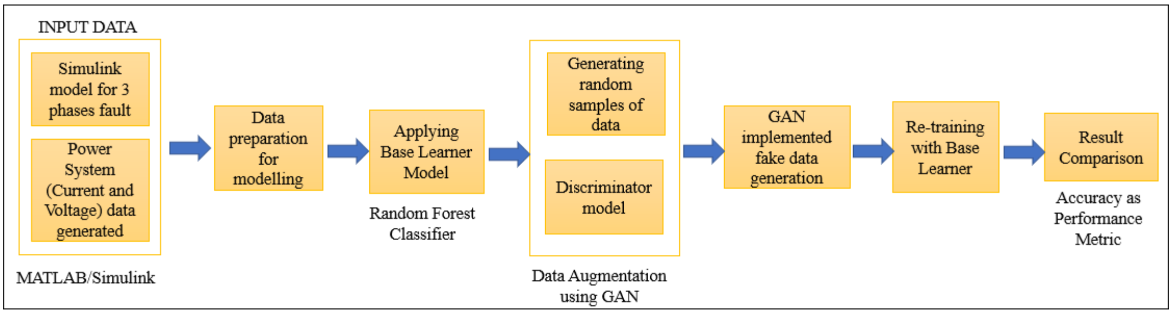

Figure 1. Block diagram of process workflow

3.1. Data creation using MATLAB/Simulink

The authors chose to simulate the data using the MATLAB/Simulink model with various blocks creates a power system transmission line. The real time data collection of high voltage transmission lines is dangerous and thereby the current and voltage values are simulated in a similar power system simulated environment. They include source block, impedance block, Transmission lines, bus bar block and 3-phase fault block. More details about the data collection can be found in [2]. The 3-phase fault block uses breaker blocks to program various kinds of faults like phase-to-phase, phase-to-ground, or a combination. The fault block records the voltages and currents under various fault conditions. Three-phase voltages and currents acquired are initially transformed into root mean square (RMS) values. The Simulink model used is shown in Figure 2. For each of the 11 fault types 1,000 cases are generated, and considering 300 cases as healthy data the overall dataset consists of 11,300 instances. The first 300 data points of healthy data are ignored as it does not align with the research objective to classify among the 11 fault types. This data is illustrated in Table 2, where the "ia", "ib", and "ic" columns represent the respective current values for each phase, and the "va", "vb", and "vc" columns represent the corresponding voltage values. The "class" column indicates the type of fault. Overall, the dataset is of size 11,300×7. The simulated data was recorded in a Microsoft Excel sheet and underwent further preprocessing.

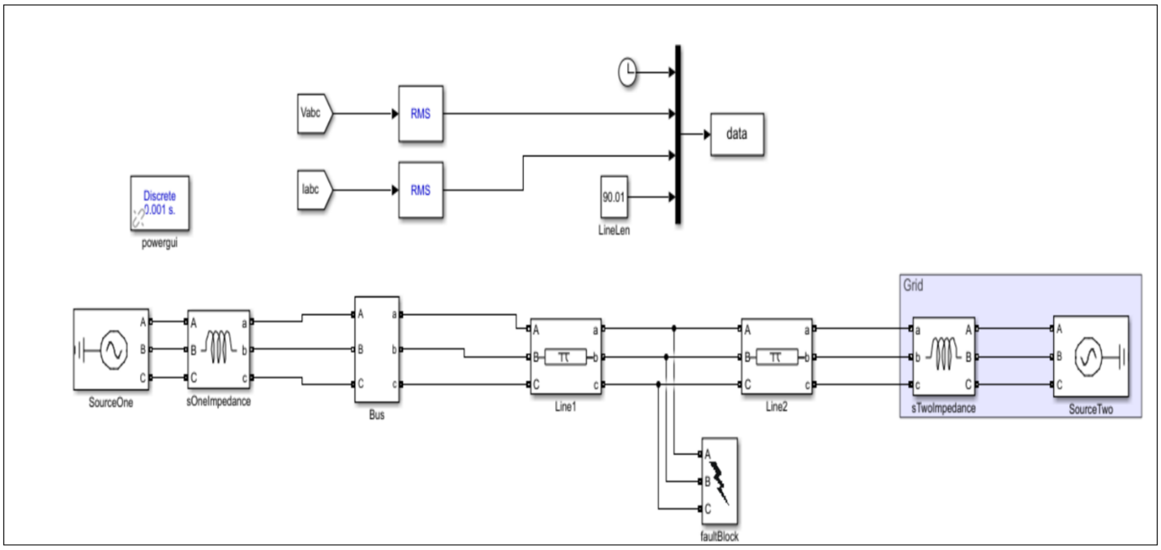

Figure 2. Simulink model for data acquisition using simulation

Table 2. Power system fault dataset

ia	ib	ic	va	vb	vc	Class
133337.2	133338.7	133338	2638.917	2638.801	2638.881	ABCG
133337.2	133338.7	133338	2638.917	2638.801	2638.881	ABC
133337	133338.9	133336.9	2638.843	2638.784	2638.878	ABCG
133337	133338.9	133336.9	2638.843	2638.784	2638.878	ABC
133340.4	133338.8	133336.9	2638.854	2638.794	2638.86	ABCG
133340.4	133338.8	133336.9	2638.854	2638.794	2638.86	ABC
133339.3	133338.7	133339.7	2638.811	2638.786	2638.858	ABCG
133339.3	133338.7	133339.7	2638.811	2638.786	2638.858	ABC
133337.4	133338.1	133340.1	2638.804	2638.791	2638.843	ABCG
133337.4	133338.1	133340.1	2638.804	2638.791	2638.843	ABC

3.2. Data preparation for modelling

The Simulink raw data needs to be prepared to suit the input requirements of a machine learning classification model. From the data collected in the subsection 3.1, three voltage and three current values are taken, making the number of input features as six. The class label for each of the instances is represented using categorical type. The machine language does not comprehend a non-numeric value so it is important to transform the classes into numerical categories. Label encoding [7] technique is used to convert categorical labels into numeric form. The data prepared thus, is of size 11,300×7, with 11,300 instances, six input features and one output feature, which is the class label representing one of the eleven types of faults.

3.3. Applying base learner model

Goswami and Roy [2], experimented with DT as the machine learning model. DT is non-parametric supervised learning model, which classifies data based on decision rules inferred from features of the corresponding dataset. RF is an ensemble of DT classifiers which uses the concept of bootstrap aggregation of bagging. This model bootstraps samples of data and let each decision tree model to train independently on those data samples and then aggregates the results using majority voting principal. This not only improves performance of the weak learner models but us powerful enough to classify on larger dataset.

The data prepared in subsection 3.2 is given to a RF classifier, which is used as base learner in our experiments. The fault dataset is split into training data and testing data. RF classifier is trained with training data with 100 n estimators and prediction is done using test data. Accuracy score, is used as a performance measure, one of the most popularly used measures in classification task [8]. The model is trained and tested to yield an accuracy of 0.84 and precision of 0.85.

3.4. Data augmentation using GAN

Collecting labelled data during the occurrence of faults in a complex power system is challenging in the real-world scenario. Overhead transmission lines carry high voltage current for hundreds of kilometres, making it difficult to label data during any fault occurrence. The performance of a classifier depends upon number of instances in the training data. Having a greater number of instances may greatly improve the performance of classifier. In order to achieve our goal of generating a synthetic version of current and voltage tabular data, we have chosen to use a GAN. GANs are unsupervised models that are primarily used for image data augmentation [9]. However, in this case, we have utilized this powerful tool to synthesize our current and voltage tabular data. With the help of GANs, we can acquire realistic data that would otherwise be difficult to obtain. The GAN architecture used in the paper and method of creating augmented data are explained in detail in section 4.

3.5. Retraining with base learner

Number of instances in the fault data increases with data augmentation. RF classifier is implemented on the augmented dataset, expecting improved performance. The augmented data is again split into training and testing data. Classification is performed as explained in subsection 3.3.

3.6. Result comparison

The results obtained in 3.3 is by training the base learner on the actual dataset. And the results obtained in 3.5 is by training the augmented data (explained in 3.4) with the base learner. Hence the result comparison extensively provides an idea how the data augmentation using GAN impact the training process of the base algorithm. The GAN architecture used for the data augmentation is discussed in detail in the next section.

4. GENERATIVE ADVERSARIAL NETWORKS FOR DATA AUGMENTATION

In situations where more data can improve prediction performance, labelling a vast dataset can be laborious [10]. GAN has the potential to generate synthetic data from labelled samples for prediction purposes [11]. For this, it is important to understand the pattern of the data and identify the distribution pattern of the data in the dataset. Unsupervised algorithms play a crucial role in this scenario as they identify patterns from the distribution of data and perform their analysis without labelled data [12]. However, the unique functionality of GAN makes it a preferred choice. Although GAN is an unsupervised algorithm, it employs supervised learning to operate. This section first defines the GAN in first part and its implementation details in the second part of the section.

GANs have versatile applications falling into distinct categories. Within image processing, GANs excel in fake image detection [13], image conversions [14], and improving the quality of images in automation systems [15]. They play a pivotal role in identifying manipulated images, transforming images between domains, and ensuring high-fidelity visual inputs for automation tasks. GANs enhance the quality of audio signals by reducing noise, improving clarity, and aiding in speech recognition tasks [16]. In cybersecurity, GANs are instrumental in both fake image detection [13] and intrusion detection [17]. GANs contribute to identifying manipulated images and generating synthetic data representing potential security threats, thereby aiding in the development of robust intrusion detection systems. Moreover, in the context of the internet of things (IoT), GANs enhance security by training adversaries [18]. From literature, GANs find application in agriculture [19]. They contribute to the identification and classification of diseases in various parts of the plants. In high performance computing clusters, GANs optimize cluster performance by generating synthetic data for testing and benchmarking purposes [20]. They contribute to tasks such as image denoising, reconstruction, and synthesis, ultimately advancing medical diagnostics. Additionally, GANs play a role in healthcare by generating synthetic medical images [21], simulating patient data for research, and improving the efficiency of medical imaging processes [22]. In space applications [23], GANs are utilized for tasks such as generating synthetic data for simulating space environments and aiding in the analysis of astronomical data.

4.1. GAN architecture

The fundamental concept behind GAN is that the generative model attempts to create fake data that can deceive the discriminator [10], [24]. Like other neural networks, the discriminator can also be misclassified by noisy data. Therefore, the discriminator constantly checks new data to distinguish between fake and real data and prevent itself from being fooled. The feedback from the discriminator improves the quality of synthetic data generated by the generative model and enhances its similarity to the real data. This process creates a trade-off between the quality of real and fake data, ultimately improving the quality of the generated data. The working of generator and discriminator is shown in Figure 3.

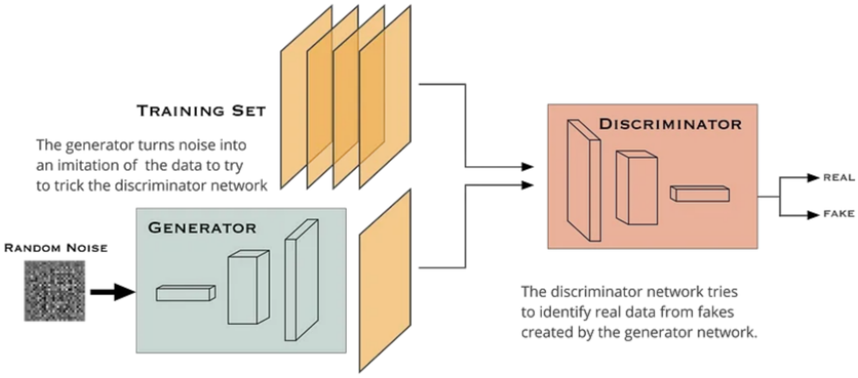

Figure 3. GAN-generator and discriminator [25]

4.2. Implementation of GAN

In this section, the implementation of the GAN model is explained in detail. The generator model is defined in 4.2.1 followed by the definition of the discriminator model in 4.2.2. The first one is responsible for generating the fake instances of the data to augment the data and the later is to discriminate between the fake and real data. The next step is to train the combined generator-discriminator model as discussed in 4.2.3 followed by the re-training the base algorithm on the augmented dataset as mentioned in 4.2.4.

4.2.1. Defining the generator model

Generative model is created to generate artificial data instances using random noise, which is reshaped to match the shape of the original input data. To generate fake instances of data, a latent space needs to be created with dimensions equal to the input data. A three-layered sequential model is used as the generator, which utilizes the Keras library. The first two layers are activated using the rectified linear unit (Relu) function, while the last layer is activated using the linear function [26]. The output layer is designed to have the same shape as the dataset. The generator model summary can be found in Table 3.

4.2.2. Defining the discriminator model

The next step is to create a discriminator model, which is similar to the generator model, except for the last layer that outputs a binary classification value indicating whether the input is real or fake, and is activated using the sigmoid function. The role of discriminator is to classify between fake and real data. The model is compiled with binary cross-entropy loss as it classifies binary result like fake or real and the Adam optimizer [26]. Adam is a well-known optimizer to converge fast when the learning rate is fixed and thereby it earns a great scope to be used as an optimizer in this experiment. Table 4 summarizes the discriminator model.

Table 3. Generator model

Layer (type)	Output shape	Param #
dense_12(Dense)	(None, 15)	165
dense_13(Dense)	(None, 30)	480
dense_14(Dense)	(None, 7)	217
Total params: 862		
Trainable params: 862		
Non-trainable params: 0		

Table 4. Discriminator model

Layer (type)	Output shape	Param #
dense_15(Dense)	(None, 25)	200
dense_16(Dense)	(None, 50)	1300
dense_17(Dense)	(None, 1)	51
Total params: 1,551		
Trainable params: 1,551		
Non-trainable params: 0		

4.2.3. Train with the combined generator-discriminator model

The current approach involves combining two models, the generator and the discriminator. The generator and discriminator models compete for a specified number of epochs to create a model that is competent enough to distinguish between real and fake data. During each epoch, the model considers a batch of data, consisting of both real and fake data. Specifically, half of the batch is comprised of real data, while the other half is made up of fake data generated by the generator. The model calculates the average loss for this batch. The generator's parameters are then updated based on the error calculated by the discriminator. This process is repeated for a fixed number of epochs, unless the error from the generator reaches a minimum. The model is observed to reach its lowest possible loss at 500 epochs indicating that the highest accuracy is obtained.

4.2.4. Re-training the base learner on the augmented dataset

After successfully training a generator model to produce synthetic data, evaluate the quality of the dataset. To do this, the generated instances are appended to the original dataset. The effectiveness of the proposed GAN based model is determined by using the same supervised model as the base learner in the first step. So, RF algorithm is retrained using the augmented dataset. The GAN model has augmented fault data, increasing the number of instances from 11,300 to 66,300 instances, by synthesizing 5,000 new instances per class. The classifier performs very well on this new dataset, achieving an accuracy of 0.99. The experiment is carried out using Python language in Anaconda's Jupyter notebook environment, utilizing various libraries such as scikit-learn (sklearn), pandas, Keras-TensorFlow, NumPy, and Matplotlib [27].

5. RESULTS AND DISCUSSION

The paper discusses the effectiveness of a GAN-based model for creating new synthetic data which has resulted in significant improvements in the performance of supervised algorithms such as RF. The study conducted in the paper focuses on using GAN for data augmentation and evaluating the impact of the augmented data on the model's performance. Although the authors in [2] have used algorithms such as DT, KNN, and SVM, this study aims to investigate the efficiency of GAN rather than comparing algorithm performance. SVM is time-consuming when training data is large, while KNN may suffer due to potential outliers that could negatively affect model performance. Hence, the study chooses RF as the classifier to implement multiple decision trees on augmented data, which overcomes the issues of time complexity and outlier anomalies. The comparisons between RF model, and other supervised models with proposed GAN based model are shown in Table 5. It can be observed that supervised models like DT, KNN, SVM, and RF perform with an accuracy of 84.2%, 86.15%, 96.06%, and 84% respectively. When GAN is used for data augmentation, the size of the training set increases that helps in improving the classification accuracy of RF learning model to 99%.

Table 5. Comparative result of supervised model versus GAN-based model

	Supervised model	GAN based model
Objective	Classification of fault in the overhead transmission line	Classification of fault in the overhead transmission line
Dataset	11 types of faults; synthetically generated using Simulink	11 types of faults; synthetically generated using Simulink and GAN
Model category	Supervised	Unsupervised/Supervised
Language	Python	Python
Working library	Scikit Learn	Scikit Learn, Tensorflow-Keras
Algorithms	Decision tree, kNN, SVM	RF, GAN
Preformance (Accuracy)	RF (with original synthetic dataset): 84% DT: 84.2% KNN: 86.15% SVM: 96.06%	RF (with GAN augmented dataset): 99%

6. CONCLUSION AND FUTURE WORK

The paper emphasizes the potential of power system fault analysis as an area of research to optimize computation for operational purposes. The proposed methodology reduces the need for human intervention, avoiding redundancy in efforts. The system employs deep learning algorithm to detect fault occurrences in the three-phase medium transmission line by recognizing patterns of current and voltages in the input data.

The GAN algorithm is particularly powerful in creating new data points which are very similar to original data points thus making the classifier more powerful. The experiment demonstrates that GAN is not limited to image data but also can handle tabular data. Proposed model provides reliable classification of fault types in medium-length transmission systems transmitting voltages. The system's intelligence is sufficient to classify fault types for high-voltage transmission lines up to a distance of 100 kilometres. The RF classification is used as the base learning model on the original dataset and it achieves accuracy of 84%. However, the base learner RF when used on GAN generated augmented faulty data, it performs exceptionally well achieving 99% accuracy. This paper successfully and strongly contributes to a unique horizon of practical implementation of unsupervised model like GAN to recognize the patterns in a complicated three phase electrical transmission system, to augment the data and thereby resulting to perform outstandingly with the best reliable accuracy in any electrical system.

ACKNOWLEDGEMENTS

The authors would like to acknowledge Vasavi College of Engineering, Hyderabad for supporting the work.

REFERENCES

[1] C. L. Wadhwa, *Electrical power systems*, 7th ed. New Academic Science, 2017.

[2] T. Goswami and U. B. Roy, "Predictive model for classification of power system faults using machine learning," in *IEEE Region 10 Annual International Conference, Proceedings/TENCON*, 2019, vol. 2019-October, pp. 1881–1885, doi: 10.1109/TENCON.2019.8929264.

[3] A. Anika, M. Junaed-Al-Hossain, S. H. Alam, and N.-Al-Masood, "Fault detection and classification of power system busbar using artificial neural network," in *2019 IEEE International Conference on Power, Electrical, and Electronics and Industrial Applications, PEEIACON 2019*, 2019, pp. 31–34, doi: 10.1109/PEEIACON48840.2019.9071944.

[4] P. Pan, R. K. Mandal, and M. M. R. R. Akanda, "Fault classification with convolutional neural networks for microgrid systems," *International Transactions on Electrical Energy Systems*, vol. 2022, pp. 1–21, 2022, doi: 10.1155/2022/8431450.

[5] A. Bhuyan, B. K. Panigrahi, and S. Pati, "Fault classification for DG integrated hybrid power system using wavelet neural network approach," in *1st Odisha International Conference on Electrical Power Engineering, Communication and Computing Technology, ODICON 2021*, 2021, pp. 1–5, doi: 10.1109/ODICON50556.2021.9428944.

[6] Z. Wang, Q. She, and T. E. Ward, "Generative adversarial networks in computer vision: a survey and taxonomy," *ACM Computing Surveys*, vol. 54, no. 2, p. 38, Mar. 2021, doi: 10.1145/3439723.

[7] J. T. Hancock and T. M. Khoshgoftaar, "Survey on categorical data for neural networks," *Journal of Big Data*, vol. 7, no. 1, p. 41, 2020, doi: 10.1186/s40537-020-00305-w.

[8] T. M. Mitchell, *Machine learning*. New York: McGraw-Hill Education, 1997.

[9] Z. Zhao, R. Birke, and L. Y. Chen, "CTAB-GAN: effective table data synthesizing," in *Proceedings of Machine Learning Research*, 2021, pp. 97–112.

[10] J. Gui, Z. Sun, Y. Wen, D. Tao, J. Ye, "A review on generative adversarial networks: algorithms, theory, and applications," *IEEE transactions on knowledge and data engineering*, vol. 35, no. 4, pp. 3313–3332, 2021, doi: 10.1109/TKDE.2021.3130191.

[11] S. Shao, P. Wang, and R. Yan, "Generative adversarial networks for data augmentation in machine fault diagnosis," *Computers in Industry*, vol. 106, pp. 85–93, 2019, doi: 10.1016/j.compind.2019.01.001.

[12] K. R. Dalal, "Analysing the role of supervised and unsupervised machine learning in IoT," in *Proceedings of the International Conference on Electronics and Sustainable Communication Systems, ICESC 2020*, 2020, pp. 75–79, doi: 10.1109/ICESC48915.2020.9155761.

[13] M. Berrahal, M. Boukabous, M. Yandouzi, M. Grari, and I. Idrissi, "Investigating the effectiveness of deep learning approaches for deep fake detection," *Bulletin of Electrical Engineering and Informatics*, vol. 12, no. 6, pp. 3853–3860, 2023, doi: 10.11591/eei.v12i6.6221.

[14] P. N. Murthy and S. K. Y. Hanumanthaiah, "Novel hybrid generative adversarial network for synthesizing image from sketch," *International Journal of Electrical and Computer Engineering*, vol. 13, no. 6, pp. 6293–6301, 2023, doi: 10.11591/ijece.v13i6.pp6293-6301.

[15] D. Islam, T. Mahmud, and T. Chowdhury, "An efficient automated vehicle license plate recognition system under image processing," *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS)*, vol. 29, no. 2, pp. 1055–1062, 2023, doi: 10.11591/ijeecs.v29.i2.pp1055-1062.

[16] J. C. Setyono and A. Zahra, "Data augmentation and enhancement for multimodal speech emotion recognition," *Bulletin of Electrical Engineering and Informatics*, vol. 12, no. 5, pp. 3008–3015, 2023, doi: 10.11591/eei.v12i5.5031.

[17] H. Abdulameer, I. Musa, and N. S. Al-Sultani, "Three level intrusion detection system based on conditional generative adversarial network," *International Journal of Electrical and Computer Engineering*, vol. 13, no. 2, pp. 2240–2258, 2023, doi: 10.11591/ijece.v13i2.pp2240-2258.

[18] I. Idrissi, M. Azizi, and O. Moussaoui, "An unsupervised generative adversarial network based-host intrusion detection system for internet of things devices," *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS)*, vol. 25, no. 2, pp. 1140–1150, 2022, doi: 10.11591/ijeecs.v25.i2.pp1140-1150.

[19] C. Kim, H. Lee, and H. Jung, "Fruit tree disease classification system using generative adversarial networks," *International Journal of Electrical and Computer Engineering*, vol. 11, no. 3, pp. 2508–2515, 2021, doi: 10.11591/ijece.v11i3.pp2508-2515.

[20] A. Karthik, J. Shetty, G. Shobha, and R. Dev, "Implementation of generative adversarial networks in HPCC systems using GNN bundle," *IAES International Journal of Artificial Intelligence*, vol. 10, no. 2, pp. 374–381, 2021, doi: 10.11591/ijai.v10.i2.pp374-381.

[21] Z. Rguibi, A. Hajami, D. Zitouni, Y. Maleh, and A. Elqaraoui, "Medical variational autoencoder and generative adversarial network for medical imaging," *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS)*, vol. 32, no. 1, pp. 494–505, 2023, doi: 10.11591/ijeecs.v32.i1.pp494-505.

[22] Z. Iklima, T. M. Kadarina, and E. Ihsanto, "Realistic image synthesis of COVID-19 chest X-rays using depthwise boundary equilibrium generative adversarial networks," *International Journal of Electrical and Computer Engineering*, vol. 12, no. 5, pp. 5444–5454, 2022, doi: 10.11591/ijece.v12i5.pp5444-5454.

[23] M. A. Zaytar and C. El Amrani, "Satellite image inpainting with deep generative adversarial neural networks," *IAES International Journal of Artificial Intelligence*, vol. 10, no. 1, pp. 121–130, 2021, doi: 10.11591/ijai.v10.i1.pp121-130.

[24] L. Gonog and Y. Zhou, "A review: generative adversarial networks," in *Proceedings of the 14th IEEE Conference on Industrial Electronics and Applications, ICIEA 2019*, 2019, pp. 505–510, doi: 10.1109/ICIEA.2019.8833686.

[25] I. Goodfellow *et al.*, "Generative adversarial networks," *Communications of the ACM*, vol. 63, no. 11, pp. 139–144, Oct. 2020, doi: 10.1145/3422622.

[26] A. Shrestha and A. Mahmood, "Review of deep learning algorithms and architectures," *IEEE Access*, vol. 7, pp. 53040–53065, 2019, doi: 10.1109/ACCESS.2019.2912200.

[27] P. Fabian *et al.*, "Scikit-learn: machine learning in Python," *Journal of Machine Learning Research*, vol. 12, no. 85, pp. 2825–2830, 2011.

BIOGRAPHIES OF AUTHORS

Tilottama Goswami received her B.E. (Honors) in Computer Science Engineering from NIT Durgapur; MS in Computer Science (Distinction) from Rivier University, NH, USA; Ph.D. in Computer Science from University of Hyderabad, India. She is Professor in the Department of Information Technology, Vasavi College of Engineering, Hyderabad, having 24 years of experience in academia, research, and industry. Her research interests are computer vision, machine learning, and image processing. She has been granted Australian and German patents, conferred with Distinguished Scientist Award by IJIEMR-Elsevier SSRN; emerged as finalist in Global Awards 2023 by WomenTech Network, UGC-Basic Scientific Research Fellowship (Government of India). She is an IEEE-SM, ACM, ISOC member and also served as Chair of ACM Hyd Deccan Chapter. She can be contacted at email: goswami.tilottama@gmail.com.

Uponika Barman Roy received her B. Tech. in Electrical and Electronics Engineering from BVRIT Hyderabad; M.S. in Software Engineering from Western University, London, Canada. Her areas of interest are data science and machine learning. She has publications in reputed journals and conferences. She worked as Data Scientist in TCS, Bangalore for three years. She is currently employed as a software developer intern in Carfax, Canada. Throughout the period of her career in industry, she has been an active researcher and academic contributor. She closely worked on various projects in the machine learning domain. She contributed to different IEEE and ACM workshops as a resource person for machine learning and data science. She can be contacted at email: ubarmanr29@gmail.com.

Deepthi Kalavala has received B.Tech. in Computer Science Engineering and M.Tech. in Computer Science Engineering from JNT University, Hyderabad and JNTU College of Engineering, Hyderabad respectively. She has submitted her Ph.D. in computer science to the University of Hyderabad in Machine Learning domain. She is currently associated with IT industry as a Data Scientist to fulfil her passion. She has 13 years of teaching experience and 6 years of research experience. Her specialization is machine learning, deep learning and algorithms. She has 7 journal publications and has presented papers in 4 international conferences. She has reviewed about 20 articles for various Elsevier journals and IEEE conferences. She is a life member of ISTE and IAENG. She can be contacted at email: kalvaladeepus@gmail.com.

Dr. Mukesh Kumar Tripathi has received a Ph.D. degree in Computer Science and Engineering from Visvesvaraya Technological University (VTU), Belagavi. He also received a B.E. degree in Information Technology from Guru Nanak Dev Engineering College, Bidar, India. He is working as Assistant Professor with the Department of Computer science and Engineering, Vardhaman College of Engineering, Hyderabad, India. He has supervised and co-supervised more than five masters and 20 B.E. students. He is authored or co-authored more than ten publications and more than 250 citations. His research interests include soft computing, machine learning, intelligent systems, image processing, and Hyperspectral. He can be contacted at email: mukeshtripathi016@gmail.com.