JOIV : Int. J. Inform. Visualization, 8(2) - May 2024 812-818

INTERNATIONAL JOURNAL ON INFORMATICS VISUALIZATION

journal homepage : www.joiv.org/index.php/joiv

A Deep Learning-based Fault Detection and Classification in Smart Electrical Power Transmission System

Shihab Hamad Khaleefah [a], Salama A. Mostafa [b,*], Saraswathy Shamini Gunasekaran [c], Umar Farooq Khattak [d], Siti Salwani Yaacob [e] , Alde Alanda [f]

[a] Department of Computer Science, Al Maarif University College, Ramadi, Anbar, Iraq
[b] Faculty of Computer Science and Information Technology, Universiti Tun Hussein Onn Malaysia, Johor, Malaysia
[c] College of Computing & Informatics, Universiti Tenaga Nasional, Selangor, Malaysia
[d] School of Information Technology, UNITAR International University, Petaling Jaya, Selangor, Malaysia
[e] Faculty of Computing, Universiti Malaysia Pahang Al-Sultan Abdullah, Pekan, Pahang, Malaysia
[f] Department Information Technology, Politeknik Negeri Padang, Padang, Indonesia

*Corresponding author: *salama@uthm.edu.my*

Abstract*—Progressively, the energy demands and responsibilities to control the demands have expanded dramatically. Subsequently, various solutions have been introduced, including producing high-capacity electrical generating power plants, and applying the grid concept to synchronize the electrical power plants in geographically scattered grids. Electrical Power Transmission Networks (EPTN) are made of many complex, dynamic, and interrelated components. The transmission lines are essential components of the EPTN, and their fundamental duty is to transport electricity from the source area to the distribution network. These components, among others, are continually prone to electrical disturbance or failure. Hence, the EPTN required fault detection and activation of protective mechanisms in the shortest time possible to preserve stability. This research focuses on using a deep learning approach for early fault detection to improve the stability of the EPTN. Early fault detection swiftly identifies and isolates faults, preventing cascading failures and enabling rapid corrective actions. This ensures the resilience and reliability of the grid, optimizing its operation even in the face of disruptions. The design of the deep learning approach comprises a long-term and short-term memory (LSTM) model. The LSTM model is trained on an electrical fault detection dataset that contains three-phase currents and voltages at one end serving as inputs and fault detection as outputs. The proposed LSTM model has attained an accuracy of 99.65 percent with an error rate of just 1.17 percent and outperforms neural network (NN) and convolutional neural network (CNN) models.

***Keywords*— Electrical power transmission networks; fault detection; classification; neural network; deep learning; CNN; LSTM.**

Manuscript received 11 Oct. 2023; revised 27 Dec. 2023; accepted 21 Feb. 2024. Date of publication 31 May 2024.

I. INTRODUCTION

Transmitting significant quantities of electrical energy from the area where it is created, such as a power plant or power station, to an electrical substation is called electrical power transmission. Alterations are made to the voltage before it is sent to consumers or other substations. These alterations take place at the electrical substation. A "transmission network" is the name given to the system of linked power lines that makes it possible to move electrical energy from one location to another. This term is often used when referring to the system in question. These lines, when combined, make up an Electrical power transmission network (EPTN), more commonly referred to as the power grid [1].

It is made up of a great variety of different components, each of which is unique, intricate, dynamic, and interrelated. Every one of these components is constantly susceptible to disruptions or electrical malfunctions. Power generators and transmission protective devices have to be used to avoid the intermittent operating conditions of the system, such as the increased use of high-capacity electrical power plants and the formation of a grid in which electricity generators and transmission-distribution networks are organized into synchronized electrical power systems and distributed networks. The activation of fast-acting corrective measures that ensure that the power system can function at a persistent state in the smallest possible time intervals is feasible [2]. This was the key factor that inspired the creation of the character, as well as making the electrical system continue operating. The resulting power plant is characterized by nearly all the electricity being produced at 11 kV or 33 kV typical voltages. On the transmission level, it is the same - 100 kV to 700 kV and even a bit higher to handle the power swings. This voltage level is added at the source through the use of transformers and after that is transmitted to the distribution centers more rapidly via transmission lines. Furthermore, when the magnitude of the voltage gets higher, the distance that must be transmitted is even larger. In contrast, a shorter distance should be covered as a result of the lowest voltage magnitude presented [4].

The electrical current's voltage is elevated to such levels to enhance efficiency, thereby minimizing the I2R loss during transmission. This is accomplished by increasing the voltage of the electrical current. When the voltage is raised, the current flowing through the circuit drops in proportion to the voltage to maintain the same amount of power. Because of this, the I2R losses are cut down significantly [5]. Primary transmission involves sending a significant amount of electrical power from the power generators to substations via overhead electrical cables. This process is commonly known as "primary transmission" [6]. In certain circles, this stage is sometimes referred to as "primary transmission." In certain countries, the transfer of information across shorter distances may also take place via the use of underground cables, which are used when the transmission takes place [7].

The electrical power voltage is typically decreased from 66 to 33 kilovolts (kV) on a scale of zero to one hundred when it reaches a receiving station [8]. Electricity is accepted at the receiving area of the station and travels along transmission lines to electric substations close to the "load centers," which may include town, urban, and village areas. Therefore, it is named the secondary transmission. When the train arrives at a substation, a regulator, a step-down transformer, will lower the voltage to around 11kV, which is much closer to its original production level. At this point, the transmission phase gives way to the distribution phase, which is responsible for satisfying the electrical power demands of both primary and secondary customers [9].

It is anticipated that problems will be found in the electrical power system's transmission lines at some point. After that, they are intended to be classed appropriately, and ultimately, they are supposed to be eliminated as soon as it is physically possible. It is feasible to use the protection system for a transmission line to initiate the starting of the other relays that are required to keep the power system protected from disturbances. When a high-quality fault detection system is used, a relaying operation has the potential to be effective, reliable, swift, and risk-free.

Electrical utility businesses face a number of challenges on their path to achieving their objectives of enhanced efficiency and reliability. One of the most significant of these challenges is the management of failures in power systems [10]. To achieve this objective, these companies employ a diverse range of advanced techniques and methodologies that leverage modern advancements in information, communication, and technology. Within the Secondary Distribution Network (SDN), electrical faults represent a segment of the network spanning from the low-voltage transformer to the end-users [11]. The SDN is a network component that reaches all the way from the low-voltage transformer to the end-users. Customers or personnel at utility companies in many developing countries are increasingly responsible for physically and visually identifying and reporting electrical faults. In most instances, diagnostic and repair work on the electrical services will call for the involvement of a maintenance crew, which the customer support representatives will be expected to delegate to them. Finding and reporting errors by phone requires an excessive amount of time due to the nature of the medium. There is an insufficient number of methods for problem-solving, and there are insufficient tools to identify and classify problems. As a result, the overall inefficiency of the process is significantly impacted.

In scientific literature, fault detection models are integrated into power line components. The essential components of these models include filtering (Hough transform, and Gabon filters), clustering, and mathematical conventional pattern recognition algorithms. Song et al. [12] utilized a Canny edge detector in conjunction with a Hough transform to find transmission line spacers that had been damaged. This step was taken to increase the precision of their findings. Initially, a scan window was produced along the conductor's path. Subsequently, candidates' spacers were located in all sliding windows, if they existed during the convolution process. Based on the measurement of the linked components, the shape configuration parameters are used to ascertain whether or not the perceived spacer was compromised. The information that was gathered served as the basis for this conclusion. Insulator extraction was the primary focus of the study carried out by Zhai et al. [13], which made use of a pattern descriptor (variable) that was developed using the Saliency Aggregating Faster Pixel-wise Image (FPISA) method. Based on the color channel that is included inside the Lab color space, the flashover area of the observed insulator was calculated.

The system was able to obtain a detection rate of 92.7 percent after being put through its paces with a hundred photographs of flashover fault insulation. Zhai et al. [14] and Han et al. [15] applied a comparable concept to detect faults associated with the absence of insulator caps, utilizing saliency and adaptive morphology (S-AM), which integrates form and pattern factors. Both of these studies were published in [14] and [15]. The insulator's cap was missing, resulting in these problems. The good features of past investigations play an important role in reaching high levels of accuracy. This should be kept in mind when concentrating on the positive aspects. However, the procedures that were utilized to process the data for the different EPTN components are arduous, involve a substantial amount of time, and need a certain degree of knowledge from the person doing the processing. Moreover, the method falls short in handling individual placement, multi-class classification, and error identification, of complex environments like those of the EPTN.

Wang et al. [16] use a Sparse Autoencoder (SAE) and a dataset collected from the power system dispatching department for categorizing defects in the power system. Because the SAE was constructed with concealed layers of varied dimensions, the influence that such layers have on the accuracy rate of diagnosis may be seen as a direct outcome of its construction. The accuracy of the SAE was measured against that of backpropagation (BP), and it was found to be 71.3% accurate.

James et al. [17] proposed a method using a joint detection technique of WT-based Deep Neural networks (DNN)s for islanding microgrids. The proposed model, for which they gave a result, has produced an average precision accuracy rate of 98.27%. This methodology consists of the definition of the type, the discernment of the phase, and the indication of the location of the fault. The Zhang et al. [18] research focuses on the utilization of LSTM and SVM algorithms that can be used for the prediction of line trips in power systems so that operational dependability and system stability can be further enhanced. The prediction accuracy of the researchers was exaggerated by 97.7 percent by this method. It is evident from the data indicated in [16], [17], and [18] can be deduced that deep learning algorithms excel significantly over other standard methods.

Authors [19] Wang et al. came up with the FF-DNN-based architecture to supervise and detect the gearbox faults of WTs using SCADA acquisition data. This mechanism was installed to localize and feature the causes of the reasons. The framework was developed to help HM detect and monitor the faults of the WT gearbox. The DNN way of generating results gives high-quality predictions that are better than the five data-driven approaches that use them as a benchmark.

Furthermore, aiming at featuring SAE and Deep Belief Network (DBN) and DNN, Cheng et al. [20] also suggest using a rotor current signal for wind turbine gears commissioned into the drivetrain. The same as that presented by Zhang et al. [21], the novel strategy of defect diagnostic testing was given for the systems of solid oxide fuel cells (SOFCs). The SAE algorithm formed the basis of this method. Then the SOFCs, which have many industrial applications, such as auxiliary power units and stationary power generators, were studied using this approach. Raw, unprocessed data from SOFC are fed to the feedforward neural network model, which at this stage achieves an accuracy of 79.94 %. Any data acquired from a real system can be supplied to this model, regardless of going through an intermediary stage.

The development of computer vision is becoming the major milestone in gradually overcoming the restrictions associated with the RGB cameras and the traditional methods for detecting and recognizing the problems within the EPTN used. One of the pioneering steps in this obstacle is the identification of the surface discoloration due to flashover on the insulator which was implemented by means of a convolutional neural network (CNN) classifier employing a pre-trained AlexNet. This research has also been one of the first attempts at implementing deep learning in defect recognition in the field. This research, published by Zhao et al. [22], was among the first of its sort. On a total of 1000 samples, the trials obtained a mean Average Precision (mAP) score of 98.71 percent. The proposed architecture performed better than the manual approach. However, it could only categorize insulator condition inspection photos, necessitating extensive feature engineering.

In addition, Liu et al. [23] introduced the Faster R-CNN method for identifying insulators without caps. Three distinct voltage transmission line levels were used to examine the system, and 1,000 training samples and 500 testing samples were created for each level. The assessment was conducted. The identification of missing cap flaws needed around 120 pictures (80 of which were for training). Due to the tiny size of the dataset, the research additionally emphasizes the possibility of overfitting and applies data augmentation to enlarge the dataset physically.

Jiang et al. [24] developed a new way of retrieving default insulator defects from high-quality images (1920 x 1080 pixels) using a single-shot detector (SSD) as the main network with three perception levels as the meta-architecture: low, medium, and high perception. Ensemble learning is used for integration. This method was proposed to tackle the problems that are of multi-scale level images based on the paper of Liu et al. [23]. The RUIE (Region of Interest (ROI) Union Extraction) image preprocessing technique was employed to generate perception images of medium and high levels. The proposed strategy was tested on the data set which comprised the images with different perception levels and missing cap insulators and achieved an absolute accuracy of 93.69 % and the recall rate of 91.09 %. However, the studies of this kind of work were limited to the contextual features of a particular fault inspection in the transmission line that were related to the insulator component. They only focused on the other problems that occurred at the same time and they did not take into account other technical concerns involved. In this aspect, the research has a significant flaw. The bulk of the time, the characteristics formed by these techniques may not accurately depict the insulators, and the imaging utilized in these processes may need modifications.

II. Materials and Methods

Fault detection on a smart grid will not only ward drops but will result in the effective and reliable operation of the EPTN. Fault detection plays a crucial role in improving the stability of an electrical power transmission system in a smart grid for several reasons:

1) Fast response time: Traditional fault detection methods are usually dependent on human inspection or a tedious, slow process of automation. But in smart grids, the use of advanced sensing systems dealing with real-time data analytics enables the identification of faults and the response to them within a very short time. The sooner the system identifies and takes action, the higher the likelihood of avoiding a disruption in the grid.

2) Prevention of cascading failures: In an EPTN, when a fault occurs at one end of the grid, there is a possibility of system-wide failures through a cascading effect, meaning the fault propagates across the grid to other parts of the network and causes serious power outages. Quick and accurate detection of defects, followed by their immediate isolation, allows the grid operators to prevent or alleviate further failures within the system and lower the risk of being impacted by these particular failures.

3) Isolation of faulty components: Location-finding intelligent fault detectors can identify the exact source of a fault in the grid almost instantaneously. By tracking down the fault within the grid, the operators could isolate the affected parts and limit the impact on the remaining system, and the unchanged areas of the grid can be monitored to ensure the stability of the power system.

4) Integration with control systems: Modern smart grids consist of state-of-the-art control systems that can operate automatically reacting to fault detection. These control systems are flexible, and they can perform functions such as reconfiguration of the grid, adjusting various generation and heavy load levels, and many other actions to maintain grid stability in the event of a fault.

5) Optimized grid operation: The operators in the power grid are able to get insights into the overall health and performance of the system by constantly monitoring the system for any fault. The data gathered could be used for grid operation optimization, for re-routing of power flows, and even for preventive maintenance. All of the above factors combined ensure better stability of the system.

6) Enhanced resilience: Smart grids not only identify flaws quickly, but with fast remedial actions, smart grids are also more resistant to disturbances like equipment failure, extreme weather, and cyber-attacks. A higher resilience level that is obtained as a result eventually leads to a stable and reliable EPTN.

A. Electrical Fault Dataset

Jamil et al. [25] developed a dataset that was used in this investigation of electrical faults. MATLAB was used to do the modeling and simulation of it. The constructed method that is based on ANNs has been simulated, developed, and put into reality using a system that is indicative of a conventional 400 103 V three-phase transmission line. The system also contains generators positioned at both ends of the transmission line. Each end of the transmission line is home to one of the system's two generators, each of which produces a voltage of 400 103 V. These power generators have been strategically placed in these particular locations so that any issues that may arise with the transmission line may be seen, documented, and investigated.

The transmission line has been referred to as a distributed type of property, which makes the proposed approach fit for the application of very long transmission lines with high accuracy, which consequently produces more credible results. This is facilitated by the modeling that can be done of transmission lines of significant distances. The simulated system model is implemented using the SimPowerSystems toolbox, an internal part of the Simulink instruction set which is a tool of MATLAB [26]. The experiment is done by employing Simulink. ZP and ZQ depict the impedances of the sources of active power of generators on the left and right sides, respectively. By using SimPowerSystem's three-phase V–I measurement block for terminal A, the toolbox allows us to get crucial three-phase voltages and currents at the start of the simulation. Simulation of the transmission line of 300km with models that describe different faults that occur at different locations on the line by considering the difference of the fault resistance values. In this study, the radiation of 50Hz is taken into consideration.

B. Long Short-Term Memory

A detailed examination of the problem of vanishing gradients may be found in [27]. When the gradient of the neural network's error function is sent back through a unit of the neural network, that unit scales it by a certain factor before sending it on to the next unit. This takes place every time the gradient is sent back through an individual unit of a neural network. In essentially all of the practically meaningful possibilities that may be evaluated, this factor is either more than one or less than one. Because of this, the gradient rapidly increases or decreases over time in a recurrent neural network. (From the perspective of language modeling, time steps are analogous to different places that words occupy in a phrase.) Therefore, the gradient either comes to dominate the subsequent stage in the weight adaptation process or essentially disappears. To circumvent this scaling effect, the authors of the study redesigned the unit that makes up a neural network (NN) in such a manner that the scaling factor that corresponds to it is always set to one. The new unit type that is produced as a result of achieving this design aim has fairly restricted capabilities when it comes to learning [28]. As a consequence, the unit was improved by adding several gating units. The completed device can be shown in Figure 1, where we have included two changes to the LSTM unit that were first presented in [29] and [30].

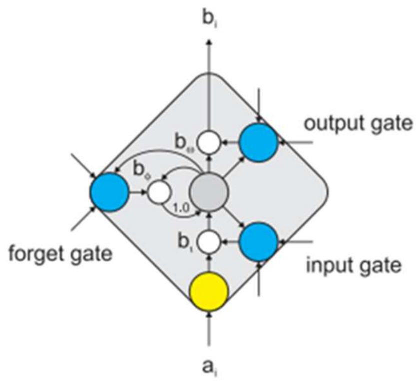


Fig. 1 LSTM memory cell with gating units [31].

$$\begin{aligned} i_t &= \sigma(w_i[h_{t-1}, x_t] + b_i \\ f_t &= \sigma(w_f[h_{t-1}, x_t] + b_f \\ o_t &= \sigma(w_o[h_{t-1}, x_t] + b_o \end{aligned} \tag{1}$$

To further explain the formula, $\boldsymbol{i_t}$ stands for the input gate**,** f_t stands for the forget gate, o_t stands for the output gate, σ stands for the sigmoid function, w_x stands for the weight of the representative x gate, h_{t-1} stands for the output of the previous LSTM block, x_t stands for input at a current timestamp, and b_x stands for the basis of the respective gate x.

When a $tanh$ activation function is utilized, a CNN unit I only has two components: the input activation ai and the output activation bi. These two components are connected in the following ways:

$$b_i = \tanh(a_i) \tag{2}$$

The LSTM unit includes the following additional intermediate steps: After the application of the activation function to a_i, the resulting value is then multiplied by a factor named b. Then, being self-connected, the inner activation value at the previous step influences the summation which is a result of bφ multiplication. After that, the output is multiplied by b and then input into yet another activation function, which ultimately produces bi. The extra units displayed as blue circles and referred to as input, output, and forget gate, respectively, are responsible for controlling the components $b\iota, b\varphi, b\omega \in (0, 1)$, which are represented by the little white circles. The gating units sum up the activations from the previous hidden layer and the activations from the current layer at the previous time step, along with the inner activation of the LSTM unit, before incorporating them into the overall output. After applying a logistic sigmoid function, the final number is squashed, and the function's output is either $b\iota, b\varphi, or\ b\omega$, depending on the case. For brevity's sake, we will not go into the relatively lengthy equations that describe the LSTM network. These examples can be found, in [32]. Figure 2 shows our proposed model.

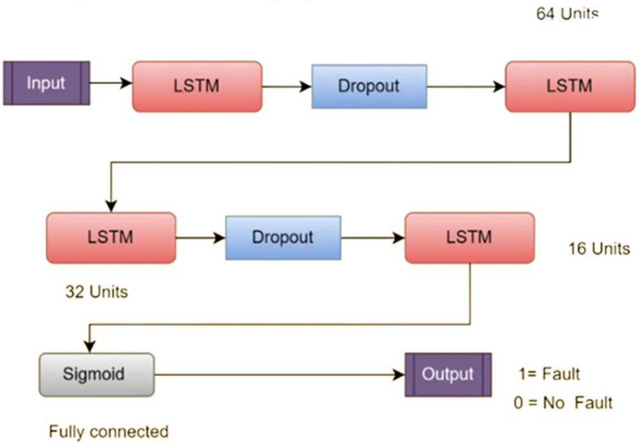


Fig. 2 Proposed LSTM architecture

One possible interpretation of the whole LSTM unit, including the gating units, is that it is a differentiable kind of computer memory. Because of this, LSTM units are often referred to as LSTM memory cells in addition to their more common name. The vanishing gradient issue can be solved using the LSTM design for just a minor increase in the computing costs, and this is true regardless of whether or not one accepts the suggested interpretation of the gating units. In addition to this, it has the advantageous quality of being able to accommodate regular recurrent NN units as a special case.

III. Result and Discussion

Task accomplishment of LSTM has relied on the sequence or fed used to input and output the system. To attain an effective performance, the perfect training set is required that consists of 10 different kinds of typical faults occurring respectively at 10 usual positions of the transmission line to be analyzed. This is how a simple outcome can be accomplished. Due to all those concerns, there were created by the team as many as a hundred single incident scenarios for the construction of both the 10 primary and less typical groups of failures. The current inserted to define fault behavior now starts with a number of potential faults across multiple real-time scenarios. It entails identifying the type of problem that occurs at the defective phases, and the range of the problem along the EPTN.

In essence, this experiment aimed to determine the many types of electrical faults that may occur inside a power transmission system. The LSTM model that is being used has undergone training by making use of a dataset with four sensors that have been strategically positioned to capture the data from the current and the voltage of the power transmission system. As shown in Table 1, the collected findings demonstrate that LSTM has surpassed all of the other classifiers discussed in the section under Related Work. Table 1 shows our obtained results using the same dataset used in [17] with ANN. We can see clearly how the proposed LSTM model had far outperformed the ANN and CNN models.

TABLE I
THE RESULTS OF THE TESTED MODELS

Model	Accuracy	Loss	Precision	Recall	F1-score
ANN	93.55	-	50.27	44.28	46.81
CNN	94.60	-	95.13	94.77	94.95
LSTM	99.65	1.17	99.63	99.65	99.646

An alternative way of determining how well a trained NN performed is to plot the confusion matrices for the various kinds of errors that were made by the neural network. This step is done as part of the evaluation process. Figure 3 presents the confusion matrix for the three phases of the LSTM classifier's development: training, testing, and validation.

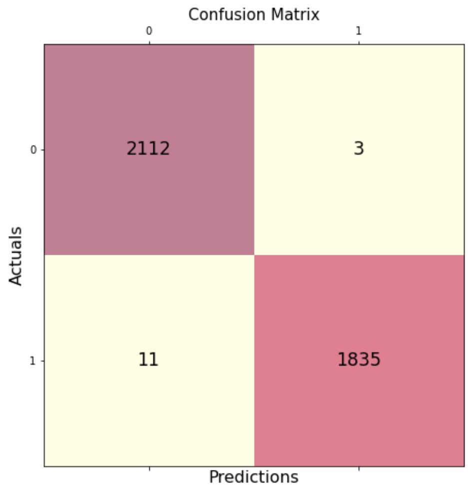
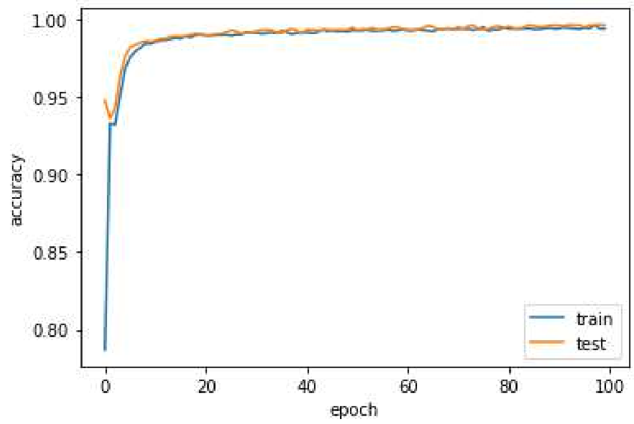
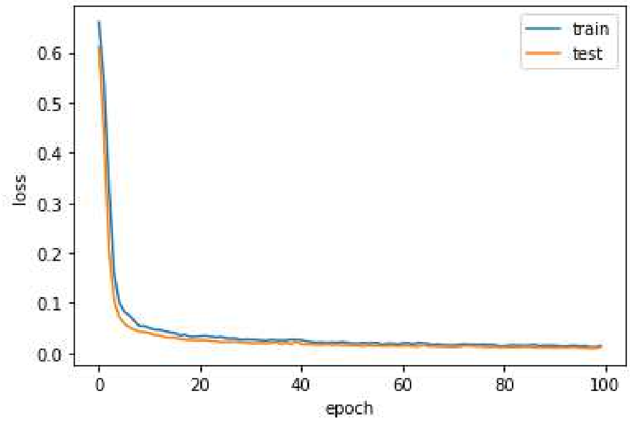


Fig. 3 Confusion Matrix

The confusion matrix compares results to predicted labels, with 0 indicating no faults and 1 denoting faults. The NN picked has almost one hundred percent accuracy in identifying faults. The graphs for accuracy and loss during testing and training are shown in Figure 4. As can be seen from the plots in Figure 4, the selected LSTM has performed quite well, as seen by the fact that there is almost no overfitting, which is a sign of successful outcomes. Figure 4 (a) resamples the accuracy curve for training and testing over one hundred epochs to show that the accuracy went up to near 95 percent in just the 10th epoch and kept a steady and stable performance over the next 90 epochs after that. This graph is a resampling of the original graph that was shown on the left. The error rate of the LSTM has decreased from 60 percent to 5 percent in the 15th epoch and has continued to go lower to reach 1.17 percent over the next 85 epochs. The plot of Figure 4 (b) illustrates the loss or errors of the model, which are the rows that were classified incorrectly.



(a) accuracy



(b) loss

Fig. 4 The result plots of the LSTM model

This shows that the LSTM error rate has decreased significantly. Our study has made one step toward developing LSTM-based fault detection in power transmission systems. However, the largest are still needed for improvement of the accuracy and efficiency of these approaches for ensuring the stability and reliability of the power grid. These future engagements may provide alongside the improvement of the system's overall efficiency and dependability, which can, in the end, be beneficial for both utility providers and the end-user.

IV. Conclusion

The study explored both defect detection and classification in a three-phase transmission lines system as its primary focus, with recurrent neural networks and its typical Long Short-Term Memory (LSTM) networks being the tool of choice. We simulated three-phase power grid faults as either changes in voltage or current, which were then combined into the LSTM network and normalized using the pre-developed method to accurately detect line-to-ground faults, which is crucial in the power transmission systems. The proposed LSTM model has attained an accuracy of 99.65 percent with an error rate of just 1.17 percent and outperforms NN and CNN models. On the other hand, the findings of our study also indicate the need for the extension of the research to include the other fault types like line-line faults, double-line-to-ground faults, and symmetrical three-phase faults which would eventually enable the future study and development of the energy sector. In the scope of the research, the LSTM model proved to be a promising structure because of the ability to simultaneously store and exploit large volumes of data while the accuracy of the localization was the outstanding feature of the model. The support for this claim is the results of comparing the method with the other methods illustrated in the related literature, where our approach consistently performed better than the uncompetitive techniques. In the future, research could be directed towards expanding the scope of LSTM models to accommodate a wider spectrum of fault types. Thus, the system's efficiency would be significantly enhanced, and the system could be integrated with other existing technologies. Along with this, the works could be prioritized on clarifying the normalization mechanism and tuning the model parameters that are used for detection to increase the accuracy and speed. Besides, thorough information fusion technologies like integrating data from multiple sensors or sources would become the crucial element for superior fault performance and practicability in real-world scenarios. Moreover, assessing the feasibility of linking the LSTM-based fault detection systems with current technologies, including edge computing and Internet devices of Things (IoTs), may lead to the development of real-time monitoring and decision-making systems that will also enhance the reliability and resilience of the grid. Another aspect of technology transfer is an industry that put together field testing and validation trials that would inform on the practicality of the steps and the possibility of scaling up the process.

Acknowledgment

This work was supported by Dato' Low Tuck Kwong International Energy Transition Grant under the project code 202202003ETG.

References

[1] A. Thaeer Hammid, O. I. Awad, M. H. Sulaiman, S. S. Gunasekaran, S. A. Mostafa, N. Manoj Kumar, B. A. Khalaf, Y. A. Al-Jawhar, and R. A. Abdulhasan, (2020). A review of optimization algorithms in solving hydro generation scheduling problems. *Energies*, *13*(11), 2787.

[2] D. A. G. Pedder, A. D. Brown, and J. A. Skinner, "A contactless electrical energy transmission system," *IEEE Trans. Ind. Electron.*, vol. 46, no. 1, pp. 23–30, 1999.

[3] J. D. Glover, M. S. Sarma, and T. Overbye, *Power system analysis & design, SI version*. Cengage Learning, 2012.

[4] T. Ahmed, K. Nishida, and M. Nakaoka, "The commercial advancement of 16 MW offshore wave power generation technologies in the southwest of the UK," in *8th International Conference on Power Electronics-ECCE Asia*, 2011, pp. 1476–1483.

[5] M. H. Haque, "Capacitor placement in radial distribution systems for loss reduction," *IEE Proceedings-Generation, Transm. Distrib.*, vol. 146, no. 5, pp. 501–505, 1999.

[6] Idoniboyeobu, D.C., Bala, T.K. and Blue-Jack, K.I., 2017. Performance Evaluation of the 132KV Sub-Transmission Lines in the Nigeria Power Network: A Case Study of Port Harcourt Sub-Region, Rivers State. International Journal of Research in Engineering and Science, 5(12), pp.28-40.

[7] Arrillaga, J., Watson, N.R. and Liu, Y.H., 2007. Flexible power transmission: the HVDC options. John Wiley & Sons.

[8] J. Casazza and F. Delea, "Understanding Electric Power Systems: An Overview of the Technology, the Marketplace, and Government Regulations," 2011.

[9] M. J. Heathcote, The J & P transformer book: a practical technology of the power transformer. Newnes, 1998.

[10] M. A. Mahmoud, N. R. Md Nasir, M. Gurunathan, P. Raj, and S. A. Mostafa, "The current state of the art in research on predictive maintenance in smart grid distribution network: Fault's types, causes,

and prediction methods—a systematic review," *Energies*, *14*(16), 5078, 2021.
[11] J. P. Avilés Arévalo, "Optimization of distribution networks using evolutionary algorithms," 2018.
[12] Y. Song et al., "A vision-based method for the broken spacer detection," in 2015 IEEE International Conference on Cyber Technology in Automation, Control, and Intelligent Systems (CYBER), 2015, pp. 715–719.
[13] Y. Zhai, H. Cheng, R. Chen, Q. Yang, and X. Li, "Multi-saliency aggregation-based approach for insulator flashover fault detection using aerial images," *Energies*, vol. 11, no. 2, p. 340, 2018.
[14] Y. Zhai, D. Wang, M. Zhang, J. Wang, and F. Guo, "Fault detection of insulator based on saliency and adaptive morphology," *Multimed. Tools Appl.*, vol. 76, no. 9, pp. 12051–12064, 2017.
[15] J. Han *et al.*, "A method of insulator faults detection in aerial images for high-voltage transmission lines inspection," *Appl. Sci.*, vol. 9, no. 10, p. 2009, 2019.
[16] Y. Wang, M. Liu, and Z. Bao, "Deep learning neural network for power system fault diagnosis," in *2016 35th Chinese control conference (CCC)*, 2016, pp. 6678–6683.
[17] J. Q. James, Y. Hou, A. Y. S. Lam, and V. O. K. Li, "Intelligent fault detection scheme for microgrids with wavelet-based deep neural networks," *IEEE Trans. Smart Grid*, vol. 10, no. 2, pp. 1694–1703, 2017.
[18] S. Zhang, Y. Wang, M. Liu, and Z. Bao, "Data-based line trip fault prediction in power systems using LSTM networks and SVM," *Ieee Access*, vol. 6, pp. 7675–7686, 2017.
[19] L. Wang, Z. Zhang, H. Long, J. Xu, and R. Liu, "Wind turbine gearbox failure identification with deep neural networks," *IEEE Trans. Ind. Informatics*, vol. 13, no. 3, pp. 1360–1368, 2016.
[20] F. Cheng, J. Wang, L. Qu, and W. Qiao, "Rotor-current-based fault diagnosis for DFIG wind turbine drivetrain gearboxes using frequency analysis and a deep classifier," *IEEE Trans. Ind. Appl.*, vol. 54, no. 2, pp. 1062–1071, 2017.
[21] Z. Zhang, S. Li, Y. Xiao, and Y. Yang, "Intelligent simultaneous fault diagnosis for solid oxide fuel cell system based on deep learning," *Appl. Energy*, vol. 233, pp. 930–942, 2019.
[22] Z. Zhao, G. Xu, Y. Qi, N. Liu, and T. Zhang, "Multi-patch deep features for power line insulator status classification from aerial images," in *2016 International Joint Conference on Neural Networks (IJCNN)*, 2016, pp. 3187–3194.
[23] X. Liu, H. Jiang, J. Chen, J. Chen, S. Zhuang, and X. Miao, "Insulator detection in aerial images based on faster regions with convolutional neural network," in *2018 IEEE 14th International Conference on Control and Automation (ICCA)*, 2018, pp. 1082–1086.
[24] H. Jiang, X. Qiu, J. Chen, X. Liu, X. Miao, and S. Zhuang, "Insulator fault detection in aerial images based on ensemble learning with multi-level perception," *IEEE Access*, vol. 7, pp. 61797–61810, 2019.
[25] R. Jamil, M., Sharma, S.K. and Singh, "Fault detection and classification in electrical power transmission system using artificial neural network," *Springerplus*, vol. 4, no. 1, pp. 1–13, 2015.
[26] Demuth, H.B., Beale, M.H., De Jess, O. and Hagan, M.T, "Neural network design," *Martin Hagan*, 2014.
[27] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.
[28] A. A. Hezam, S. A. Mostafa, Z. Baharum, A. Alanda, and M. Z. Salikon, "Combining Deep Learning Models for Enhancing the Detection of Botnet Attacks in Multiple Sensors Internet of Things Networks," *JOIV: International Journal on Informatics Visualization*, *5*(4), 380-387, 2021.
[29] F. A. Gers, J. Schmidhuber, and F. Cummins, "Learning to forget: Continual prediction with LSTM," *Neural Comput.*, vol. 12, no. 10, pp. 2451–2471, 2000.
[30] F. A. Gers, N. N. Schraudolph, and J. Schmidhuber, "Learning precise timing with LSTM recurrent networks," *J. Mach. Learn. Res.*, vol. 3, no. Aug, pp. 115–143, 2002.
[31] M. Sundermeyer, R. Schlüter, and H. Ney, "LSTM neural networks for language modeling," 2012.
[32] J. Graves, A. and Schmidhuber, "Framewise phoneme classification with bidirectional LSTM and other neural network architectures," *Neural networks*, vol. 18, no. 5–6, pp. 602–610, 2005.