Deep Learning-based Fault Detection, Classification, and Locating in Shipboard Power Systems

Soroush Senemmar, *Student Member, IEEE,* and Jie Zhang, *Senior Member, IEEE*
The University of Texas at Dallas
Richardson, TX 75080, USA
Email:{soroush.senemmar, jiezhang}@utdallas.edu

***Abstract*—This paper proposes a new deep learning-based framework for fault detection, classification, and location identification simultaneously in shipboard power systems (SPS). Specifically, three different neural networks based fault detection methods, including deep neural network, gated recurrent unit, and long short-term memory, are developed and compared to detect different faults in SPS. The developed models use real-time line voltages of all SPS buses to detect faults for the entire network. The models are trained and tested based on simulated fault data from an 8-bus SPS. Results show that all the three methods have achieved a more than 99% accuracy in fault detection, classification, and location identification. The impacts of load variations and noisy inputs on the model performance are also investigated, and results show that the models are robust against the load variation and noisy input.**

***Index Terms*—Deep neural network, fault detection, gated recurrent unit, long short-term memory, shipboard power system.**

I. Introduction

There is increasing interest in using unmanned surface vessels for long duration voyages, which is more reliable and economically efficient, and also significantly increase the complexity of the ships [1]. In addition, modern Navy ships have been undergoing rapid growth in both size and complexity, which can lead to challenges in energy management and malfunction. For example, Navy shipboard power systems (SPS) will have more intricate power systems due to new emerging technologies such as the integrated electric propulsion system, pulsed power loads, energy storage, and others. [2]. Thus, it is critically important to develop approaches for predictive maintenance that are needed to manage the ship's machinery and electrical systems when there is no one on board to address it.

The protection of future SPS is one of the most important issues, which is defined as fault detection, fault classification (e.g., single phase to ground, line to line, and three phase), and fault location identification in the SPS. Recently, the application of machine/deep learning to fault detection in power systems has attracted lots of attentions due to its significant benefits. These learning-based methods can detect the fault more accurately [3] and faster compared to traditional methods [4]. Fast and accurate fault detection and classification could help to isolate the fault, while determining the fault location can significantly reduce the maintenance time and improve the system reliability.

This work relates to Department of Navy award (N00014-20-1-2795) issued by the Office of Naval Research. The United States Government has a royalty-free license throughout the world in all copyrightable material contained herein.

Machine/deep learning based fault detection methods have been studied in recent works. For instance, a fault detection and classification method based on support vector machine for transmission networks has been presented in [5], which can detect and classify the fault type using busbar voltages with an accuracy of more than 93%. Chanda and Fu [6] adopted artificial neural network (ANN) for fault classification in distribution networks, and used steady state voltages of all busbars to detect the fault, classify the fault type, and determine the fault location in an IEEE 37-bus network. Yu et al. [3] developed a recurrent neural network (RNN) method to detect and classify the fault in a microgrid, and specifically gated recurrent units (GRU) was used to classify the faults using time-series data with an over 97% accuracy.

Recently, deep learning based fault detection methods in SPS have been studied in the literature [6]–[10]. For instance, Chanda and Fu [6] determined the fault type in a cable of medium voltage DC (MVDC) SPS using a fully connected deep neural network (DNN), and showed that the accuracy of DC fault detection and location identification is more than 95% in all studied scenarios. Ma et al. [7] proposed a DC load monitoring and fault location method applied to a Naval DC-pulsed load. The authors used a long short-term memory (LSTM) method to detect component faults through monitoring time-series current signals of the DC-pulsed load. Li et al. [8] implemented an ANN-based fault detection and classification method in MVDC SPS for each busbar, with a 99% accuracy for classifying fault types at each busbar.

Most of the previous works used the information of one busbar, one component, or steady state voltages to detect the fault, which may not be able to protect the entire network in an efficient manner. In this paper, we propose a holistic deep learning-based framework, including fault detection, classification, and location identification for a medium voltage AC (MVAC) SPS. For the first time, time-series voltages of all busbars are leveraged to monitor the SPS status. To evaluate the effectiveness of the holistic framework, three different deep learning architectures, i.e., a fully connected DNN, an LSTM, and a GRU architecture, are adopted and compared

to detect the fault, classify the fault type, and determine the fault location in SPS. To evaluate the robustness of the proposed method, the dependency and sensitivity of the model on load variations and noisy inputs are also investigated. The contributions of this work are briefly summarized as follows.

- Develop a holistic deep learning-based framework for fault detection, classification, and location identification in SPS, and
- Real-time monitor and detect faults for the entire SPS network.

The rest of this paper is organized as follows. Section II formulates the SPS fault detection problem and describes the proposed learning-based model. Section III presents and discusses simulation results of an 8-bus SPS, followed by the conclusion in Section IV.

II. Problem Description

A. Shipboard power system

Figure 1 shows a synthetic shipboard power system (SPS), which is a standard 3-phase MVAC 8-bus SPS, working under the 13.8 kV voltage and 50 Hz frequency [11]. Busbar 1 (BB1) is considered as the reference busbar in the SPS. The SPS has four generators including two main generators (G1 and G3) and two auxiliary generators (G2 and G4). All the loads are aggregated in load points numbered 1 to 6. Cables are modeled based on MVAC cable characteristics. The two main generators have the rated power of 36 MW with a 0.9 power factor (PF). The two auxiliary generators have the rated power of 4 MW with a 0.8 PF. The rated voltage of generators is 13.8 kV. The voltage control mode for G1 is based on the voltage, and for other generators is based on PF. Each of the L2 and L5 nodes has 30 MW load with a 0.9 PF; each of the L1, L3, L4, and L6 nodes has 2 MW load with a 0.9 PF.

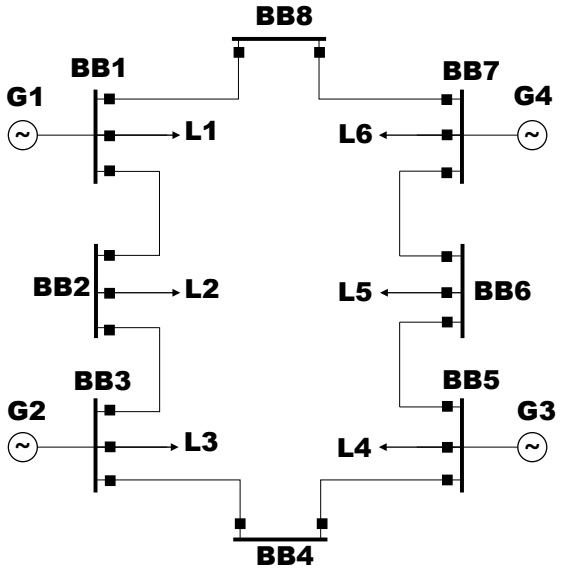

Fig. 1: An 8-bus shipboard power system.

B. Power flow simulations and fault scenarios

To develop and evaluate the learning-based fault detection method, a large number of fault scenarios are simulated and generated. The proposed 8-bus model in Fig. 1 is simulated in the DIgSILENT PowerFactory software [12]. In addition, three busbars, i.e., BB1, BB4, and BB6, are considered for fault simulation scenarios. At each busbar, three types of faults are simulated, including single-line to ground (LG), line-to-line (LL), and three-phase (3L) faults. In each simulation, voltages of all busbars are sampled in a 5 seconds time-frame, where faults occur at second 0.5. The frequency of sampling is 200 samples per cycle. The simulations are conducted based on time-series dynamic simulations in the software.

C. Deep learning models

Three different deep learning architectures, i.e., a fully connected DNN, an LSTM, and a GRU architecture, are adopted and compared to detect the fault, classify the fault type, and determine the fault location in SPS.

Recurrent neural networks (RNN) have been widely used in power systems with different time-series problems. The main benefit of RNN is the ability to create a history through time-series data based on the past and present information of the system. RNN methods may have the problem of vanishing gradient during the training. The LSTM and GRU [13] methods can manage the weights and biases of each cell to prevent the exploding of the gradient. LSTM has more functions in each cell while GRU is faster due to a lower complexity. The mathematical formulations and block diagram of GRU and LSTM methods are presented as follows.

1) Gated Recurrent Unit: A GRU scheme with the data flow is illustrated in Fig. 2. In the block, h_t and h_{t-1} are the outputs of the present and previous blocks, respectively, and the update gate decides about passing or blocking it for the present block, as follows.

$$h_t = (1 - z_t) \times h_{t-1} + z_t \times \tilde{h}_t \tag{1}$$

where z_t is an update gate that stores the information of updates of each unit, which is the most important gate in GRU.

$$z_t = \sigma(W_z.[h_{t-1}, x_t]) \tag{2}$$

where σ is a differentiable and smooth activation function. In addition, $\tilde{h}_t$ is a variable that combines the information from previous outputs with the present input, given as follows.

$$\tilde{h}_t = tanh(W.[r_t \times h_{t-1}, x_t]) \tag{3}$$

where r_t is a reset gate and defined as follows.

$$r_t = \sigma(W_r.[h_{t-1}, x_t]) \tag{4}$$

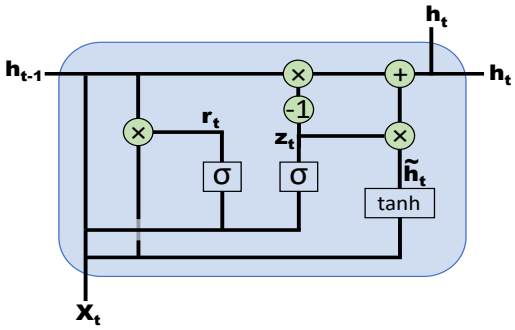

Fig. 2: The block diagram of GRU.

2) Long Short-Term Memory: LSTM has two more gates, i.e., forget and output gates, compared to GRU. Figure 3 shows the block diagram of LSTM, and the mathematical formulations for LSTM are briefly described as follows.

$$f_t = \sigma(W_f.[h_{t-1}, x_t] + b_f) \tag{5}$$

where f_t is a forget gate and uses a "sigmoid" function to determine to keep or throw away the cell state.

$$C_t = f_t \times C_{t-1} + i_t \times \tilde{C}_t \tag{6}$$

where C_t and C_{t-1} are the present and previous cell states, respectively. In addition, $\tilde{C}_t$ is a new candidate value which is obtained from a $tanh$ function on a combination of input data and previous block output.

$$\tilde{C}_t = tanh(W_c.[h_{t-1}, x_t] + b_c) \tag{7}$$

$$i_t = \sigma(W_i.[h_{t-1}, x_t] + b_i) \tag{8}$$

$$O_t = \sigma(W_o.[h_{t-1}, x_t] + b_o) \tag{9}$$

where i_t and O_t are the input and output gates, respectively.

$$h_t = O_t \times tanh(C_t) \tag{10}$$

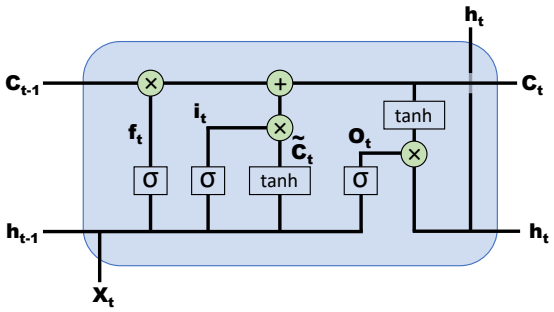

Fig. 3: The block diagram of LSTM.

3) Deep neural network: The DNN architecture is defined as an ANN method with multiple hidden layers between input and output. These multiple hidden layers can significantly improve the model performance in extracting features from inputs and connect them to the proper outputs. In this paper, the fully-connected DNN architecture is adopted, in which each node in hidden layers is connected to all the nodes from previous and next layers. The fully connected DNN is suitable for modeling complex non-linear systems. Since the fully connected DNN connects the input through linear algebraic models and simple activation functions, the computational burden is low, making it more efficient for real-time processes. However, fully-connected DNN cannot work with time-series data, which cannot keep any history from the previous entered data into the model.

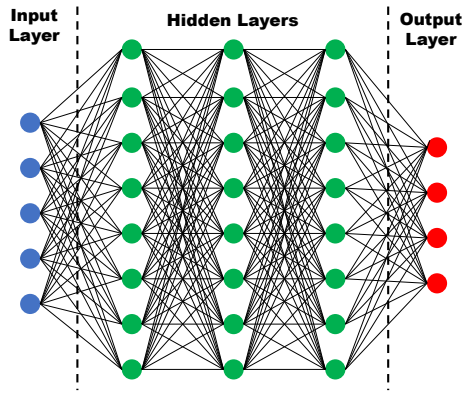

Fig. 4: A fully connected DNN architecture.

III. Simulation Results

A. Data sampling and pre-processing

It was mentioned before that the SPS model is simulated in DIgSILENT PowerFactory software and line voltages for each line at each busbar are sampled with the rate of 200 samples per cycle. The goal of the developed deep learning methods is to detect, classify, and determine the location of the fault in the shortest time possible, by monitoring the line voltages at each sample-step. Since the time-frame for simulations is 5 seconds, there are 50,000 samples per line voltage and each busbar has three line voltage (AB, BC, and AC); thus the vector of input data is 50,000×24, where the rows represent sample-steps and columns represent line voltages of all busbars. Figures 5-7 show the sampled line voltages from all busbars when the LG fault (Fig. 5), LL fault (Fig. 6), or 3L fault (Fig. 7) occurs on BB1. It is shown that the voltages are in steady-state before second 0.5 (i.e. when the fault occurs), and begin to decrease after the fault.

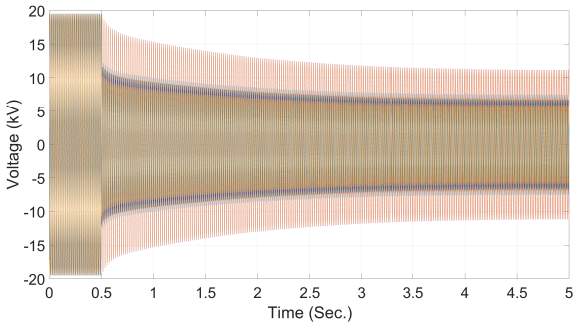

Fig. 5: Busbar voltages with LG fault on BB1.

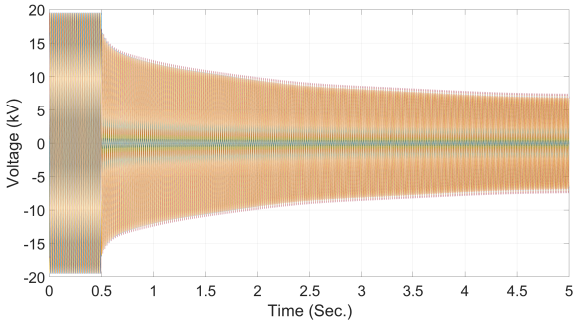

Fig. 6: Busbar voltages with LL fault on BB1.

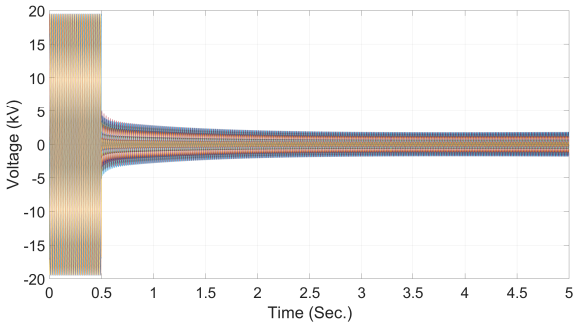

Fig. 7: Busbar voltages with 3L fault on BB1.

1) Fully connected DNN: Three busbars are candidates for fault scenarios, where three types of fault are simulated at each candidate busbar. To this end, nine vectors of input data are created. All the vectors are placed in a row to form a 450,000×24 input matrix. The output of the model is the type of fault and its location. To this end, the non-fault condition is shown with 0 in the output vector, while the fault conditions are labeled as in Table I. Depending on the non-fault condition or faulty condition, the corresponding output value of each sample step is determined. When the input and output vectors are generated, they are divided into training and testing datasets. In this work, 80% of the simulated data is considered as training data, and the rest is considered as a testing dataset. Thus, a vector of 360,000×24 is used to train the fully connected DNN. The flowchart of the fully connected DNN scheme is shown in Fig. 8(a). In this DNN, we use 3 dense layers where each layer has 32 units. The first dense layer maps the input vector into higher order dimensions to extract the features in the input data. The next two layers are utilized to extract more features from the input data. The activation function is the rectified linear unit (ReLU) function for all hidden layers. The last layer (i.e., output) converts extracted features into a 1-D vector which is the output vector. The activation function for the output layer is the softmax function. Hence, the output of the DNN network shows the probability of mapping the input vector to each element of the output vector, and the highest probability is selected as the predicted output for that specific input vector.

TABLE I: The label of fault type and location

Fault type	BB1	BB4	BB6
No fault	0	0	0
LG	1	4	7
LL	2	5	8
3L	3	6	9

2) LSTM and GRU: The input vector for LSTM and GRU is more complex, since the input vector should consider a time-series of data. In the LSTM and GRU models, the input vector is a 3-D vector, where the number of rows is equal to the number of samples (here 450,000), the number of column is equal to the time-step (i.e., 5 in this work, meaning the model has the memory of 5 previous samples), and the 24 layers represent the number of variables (i.e., 8 busbars × 3 phases). Each row in the input vector has the values of line voltages at the current sample-step plus four previous samples, creating a time-series. Thus, the input vector dimension is 450,000×5×24. It is worth to note that 80% of the data is considered as training data and 20% as test data. Therefore, the size of input vector for training the LSTM/GRU model is 360,000×5×24.

The flowchart of the LSTM/GRU scheme is shown in Fig 8(b). Four layers of RNN are considered to generate the network, and each layer has 50 units of LSTM/GRU. It was mentioned before that vanishing gradient may happen in RNN methods. To this end, the dropout technique is applied to prevent overfitting, and some nodes are removed randomly with all its inputs and outputs connections. At each layer, 20% dropout is considered. The output layer is a dense layer with 10 units and the softmax activation function to map the LSTM/GRU output into a 1-D readable vector for users.

B. Results and discussion

The fault detection results of the DNN, LSTM, and GRU are presented and compared under three scenarios: (i) with perfect input from simulations, (ii) with load variations, and (iii) with noisy input data. All simulations are conducted on a computer with Core i7-1.5 GHz CPU, 16 GB RAM, and NVIDIA GeForce MX230 GPU. All the power flow and fault simulations are conducted in DIgSILENT PowerFactory and the three learning models are implemented with TensorFlow.

TABLE II: The accuracy of fault detection with different methods

Model	Fault detection		Fault classification		Fault location		Loss
	Training	Testing	Training	Testing	Training	Testing	
Fully connected DNN	99.80%	99.77%	99.67%	99.62%	99.22%	99.18%	0.0238
LSTM	99.74%	99.86%	99.74%	99.84%	99.41%	99.41%	0.0202
GRU	99.87%	99.92%	99.84%	99.91%	99.84%	99.91%	0.0093

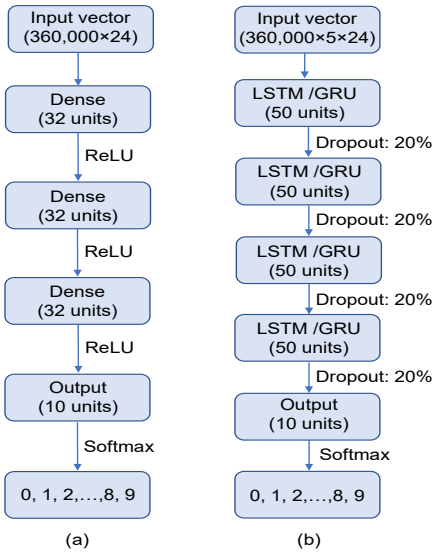

Fig. 8: The flowcharts of (a) Fully connected DNN, and (b) LSTM/GRU

To train the three learning models, 100 epochs for the fully connected DNN and 20 epochs for LSTM & GRU are considered in this work. In addition, we use the "adam" optimizer to determine the weights and biases for each model. Since the output vector has the values of 0-9, the fault detection is a multi-class classification problem, and the categorical crossentropy loss function is used.

Table II shows the accuracy of the three models with perfect simulated data in fault detection, classification, and location identification, including the validation & testing accuracy, and the loss. The results show that all the three models can detect the fault correctly with a more than 99% accuracy. In addition, the GRU and LSTM models perform slightly better than the fully connected DNN model, due to the consideration of time-series. This high accuracy could play a key role in monitoring future Navy ships by detecting the fault quickly and reliably.

To compare the proposed methods with published methods in the literature, Table III summarizes the detection accuracy from the references. In Refs. [6] and [8], the fault detection accuracy was not reported. With the hypothesis that the fault detection accuracy is higher than the classification accuracy, we use the "$\geq$ classification accuracy" to represent the fault detection accuracy. It is important to note that the accuracies of other models reported here are directly taken from the references, which are not the fault detection accuracy on our 8-bus SPS model.

TABLE III: The accuracy of different fault detection models

Model	Accuracy		
	Detection	Classification	Locating
Proposed GRU Model	99.92%	99.91%	99.91%
Proposed LSTM Model	99.86%	99.84%	99.41%
Proposed DNN Model	99.77%	99.62%	99.18%
Decision Tree [14]	97%	85%	-
K-nearest neighbors [15]	90.4%	90.4%	-
Fully connected DNN [6]	$\geq 95\%$	95%	-
Fully connected DNN [8]	$\geq 99.58\%$	99.58%	-
Differential relay [14]	96%	-	-

C. Load dependency

The reported results in Table II are based on the load values in Section II. Hereinafter, we call those load values as rated loads. However, the SPS load can vary during the normal operation. Thus, it is essential to evaluate the model accuracy when the loads are different than rated values. Figure 9 compares the accuracy of the three proposed models with $\pm 20\%$ and $\pm 40\%$ load variations. The results show that all the three models are robust against the load variation. In the worst scenario in load variations, the accuracy may reduce by 0.1%, which is acceptable.

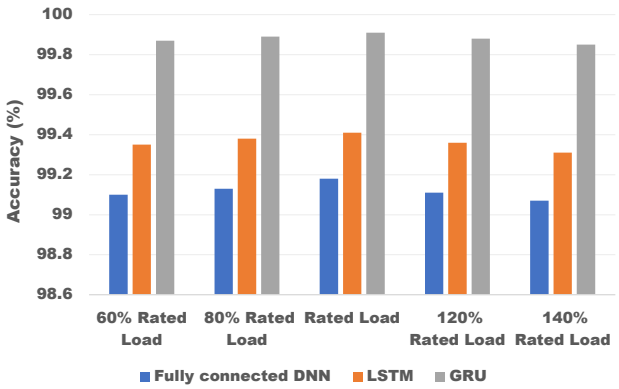

Fig. 9: Accuracy of models under load variations.

D. Sensitivity analysis with noisy input

The impact of noisy input signals on fault detection is also investigated. In real-world systems, the sampled data and transferred signals from sensors inevitably become noisy. In this work, we use the white Gaussian noise to model the impact of noisy input signal. The intensity of noise is usually shown by a signal-to-noise ratio (SNR). Both 40 dB and 30 dB SNR are considered in this study to investigate the performance of the proposed models under noisy input signals. Table IV shows the accuracy of the models under noisy input

signals. The accuracy of fully connected DNN in the worst scenario decreases by approximately 7%. However, the DNN model is still effective in detecting the faults. Both LSTM and GRU models are more robust against noisy input with less than 5% reduction in accuracy with the 40 dB noise. It should be mentioned that the impact of noisy signal could be reduced with some signal processing techniques, which will be investigated in our future work.

TABLE IV: Accuracy of models with noisy input data

Method	SNR	Accuracy		
		Detection	Classification	Locating
DNN	Original Signal	99.77%	99.62%	99.18%
	40 dB	97.92%	97.64%	96.23%
	30 dB	94.32%	93.84%	92.13%
LSTM	Original Signal	99.86%	99.84%	99.41%
	40 dB	98.13%	97.86%	96.82%
	30 dB	94.67%	94.35%	94.03%
GRU	Original Signal	99.92%	99.91%	99.91%
	40 dB	98.22%	97.94%	97.27%
	30 dB	95.35%	95.10%	95.10%

IV. Conclusion

In this study, three deep learning models, including the fully connected DNN, LSTM, and GRU, were developed for fault detection, classification, and location identification in shipboard power systems. The developed models used real-time line voltages of all SPS busbars to detect faults for the entire network. The results showed that all three methods achieved a more than 99% accuracy, which is promising in comparison with other fault detection methods in the literature. More specifically, the GRU-based fault detection model performed the best, followed by the LSTM-based model, and the DNN model. In addition, the methods were investigated based on load generation and noisy input data. The simulation results indicated that the models' performance was only decreased by approximately 0.1% in sharp load changes and 7% with 40 dB noisy input signals, which is acceptable in real-world systems.

Acknowledgment

This material is based upon work sponsored by the Department of the Navy, Office of Naval Research under ONR award number N00014-20-1-2795. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Office of Naval Research.

References

[1] R. R. B. Powers, "Automation as a manpower reduction strategy in navy ships," Ph.D. dissertation, Massachusetts Institute of Technology, 2016. [Online]. Available: https://dspace.mit.edu/handle/1721.1/104386

[2] K. Satpathi, A. Ukil, and J. Pou, "Short-circuit fault management in dc electric ship propulsion system: Protection requirements, review of existing technologies and future research trends," *IEEE Transactions on Transportation Electrification*, vol. 4, no. 1, pp. 272–291, 2018.

[3] J. J. Q. Yu, Y. Hou, A. Y. S. Lam, and V. O. K. Li, "Intelligent fault detection scheme for microgrids with wavelet-based deep neural networks," *IEEE Transactions on Smart Grid*, vol. 10, no. 2, pp. 1694–1703, 2019.

[4] G. L. Kusic, "State estimation and fast fault detection for ship electrical systems," in *2007 IEEE Electric Ship Technologies Symposium*, 2007, pp. 209–214.

[5] H. Livani and C. Y. Evrenosoğlu, "A fault classification method in power systems using dwt and svm classifier," in *PES TD 2012*, 2012, pp. 1–5.

[6] N. K. Chanda and Yong Fu, "Ann-based fault classification and location in mvdc shipboard power systems," in *2011 North American Power Symposium*, 2011, pp. 1–7.

[7] Y. Ma, D. Oslebo, A. Maqsood, and K. Corzine, "Dc fault detection and pulsed load monitoring using wavelet transform-fed lstm autoencoders," *IEEE Journal of Emerging and Selected Topics in Power Electronics*, pp. 1–1, 2020.

[8] W. Li, A. Monti, and F. Ponci, "Fault detection and classification in medium voltage dc shipboard power systems with wavelets and artificial neural networks," *IEEE Transactions on Instrumentation and Measurement*, vol. 63, no. 11, pp. 2651–2665, 2014.

[9] Q. Liu, T. Liang, and V. Dinavahi, "Deep learning for hardware-based real-time fault detection and localization of all electric ship mvdc power system," *IEEE Open Journal of Industry Applications*, vol. 1, pp. 194–204, 2020.

[10] Y. Ma, D. Oslebo, A. Maqsood, and K. Corzine, "Pulsed-power load monitoring for an all-electric ship: Utilizing the fourier transform data-driven deep learning approach," *IEEE Electrification Magazine*, vol. 9, no. 1, pp. 25–35, 2021.

[11] F. Shariatzadeh, N. Kumar, and A. K. Srivastava, "Optimal control algorithms for reconfiguration of shipboard microgrid distribution system using intelligent techniques," *IEEE Transactions on Industry Applications*, vol. 53, no. 1, pp. 474–482, 2017.

[12] "PowerFactory - DIgSILENT," Mar 2021, [Online; accessed 16. Mar. 2021]. [Online]. Available: https://www.digsilent.de/en/powerfactory.html

[13] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Empirical evaluation of gated recurrent neural networks on sequence modeling," *arXiv preprint arXiv:1412.3555*, 2014.

[14] D. P. Mishra, S. R. Samantaray, and G. Joos, "A combined wavelet and data-mining based intelligent protection scheme for microgrid," *IEEE Transactions on Smart Grid*, vol. 7, no. 5, pp. 2295–2304, 2016.

[15] T. S. Abdelgayed, W. G. Morsi, and T. S. Sidhu, "A new approach for fault classification in microgrids using optimal wavelet functions matching pursuit," *IEEE Transactions on Smart Grid*, vol. 9, no. 5, pp. 4838–4846, 2018.