Proceedings of the Pakistan Academy of Sciences:
A. Physical and Computational Sciences: 58(S): 37-47 (2021)

ISSN: (Print): 2518-4245; ISSN (Online): 2518-4253
http://doi.org/10.53560/PPASA(58-sp1)731

Pakistan Academy of Sciences

Research Article

Convolutional Neural Network and Long-Short Term Memory based for Identification and Classification of Power System Events

Mauridhi Hery Purnomo[1*], Vincentius Raki Mahindara[2], Rahmat Fabrianto Wijanarko[2], Agustinus Bimo Gumelar[1], Feri Wijayanto[3], and Yanuar Nurdiansyah[4]

[1]Department of Computer Engineering, Institut Teknologi Sepuluh Nopember, Jl. Teknik Kimia, Keputih, Sukolilo, Surabaya 60111, Indonesia
[2]Department of Electrical Engineering, Institut Teknologi Sepuluh Nopember, Jl. Teknik Kimia, Keputih, Sukolilo, Surabaya 60111, Indonesia
[3]Institute for Computing and Information Science, Radbound University Comenisulaan 4, 6525 HP Nijmegen, The Netherlands
[4]Program Study of Information Technology, University of Jember, Jl. Kalimantan 37, Jember 68121, Indonesia

Abstract: In this present era, power system delivery has to be reliable and sustainable. The growth of demands increasing the complexity of the power system operations. An interrupted power supply must not occur for any reason. Hence, the improvement of the controller and protection devices is mandatory. One of the unnecessary interruptions in the power system is a false trip due to the incorrect setting of the protection devices. Therefore, a method to classify the symptom of the power system based on the voltage, current, and frequency measurements is required. However, since there are a ton of maneuver options and fault types, the number of data becomes complex, enormous, and irregular. This is where deep learning takes place. This paper proposed the use of Convolutional Neural Networks (CNN) combined with Long-Short Term Memory (LSTM) to recognize the categorize the type of events in a medium voltage power distribution network. As CNN's models are great at decreasing frequency variation, LSTM is great for temporal modeling, we take benefit of CNN's and LSTM's complementarity in this study by integrating it into a unified architecture. The simulation results indicate that CNN and LSTM can recognize the symptoms in power system operation with accuracy up to 79 % with a total epoch 350.

Keywords: Artificial Intelligence-based Model, Deep Learning Algorithm, Electrical Protection System, Energy Efficiency, Sustainable Power System

1. INTRODUCTION

As far as the many appealing challenges faced by the industry today, the most critical part is to be able to compete in the market by figuring and shaping the new technology revolution. The successful technology revolution requires a support system. One of the most crucial parts is the power system.
It has been a common secret that everyone demands a reliable power system. When the blackout occurs, every second of the power outages leads to an economic loss in every business sector. To tackle this issue, the power system must be fortified with a defensive scheme that considers the possible fault. On the other hand, with the recent trends of distributed energy resources (DER), a lot of intermittent generators penetrate the system that might cause system instability [1–3]. The combination of a bunch of operation schemes and the possibility of distributed generator (DG) penetration equal with no shortage of challenges for the engineers.

A reliable and sustainable power system network means it has a ton of operation maneuvers and a dependable protection system. The operation

Received: February 2021; Accepted: July 2021
*Corresponding Author: Mauridhi Hery Purnomo <hery@ee.its.ac.id>

and protection of the power system must work independently but with a correlated purpose. The false or unsafe operation during power system maneuver that might be caused by DER penetration [4], sudden load injection or rejection, networks topology variation, and a fault condition shall be anticipated by a proper protection system [5]. However, the protection system shall not limit the flexibility of power system operations. For this particular reason, a trade-off point is compulsory to achieve a robust power system [6]. An illustration regarding the correlation between power system operation and protection is presented in Figure 1.

As the electrical protection device, a relay works by comparing the reading value to a specific threshold and delay time before it trips the circuit breaker (CB) [7]. The relay must work to limit the interruption, diminish the damage of the component involved, and minimize the affected area. If the closest relay to the fault location cannot handle the situation, then it must be backup by another relay that has a greater time and more area banned from the system. The coordination of the relays constructs an electrical protection system.

The fault in the power system could be categorized into two types, i.e., series faults such as an open line condition and shunt faults such as the short circuit condition [8, 9]. It is very easy to distinguish the type of fault in a direct observation by witnessing the broken equipment or conductor. Per contra, if the fault is located in a remote area, then the observation-only can be proceeded by examining the electrical parameters, i.e., voltage, current flow, and frequency. Chiefly, it is impossible to witness every single piece of equipment on the system. This is where the protection device takes place.

The study aims to fill the gap of providing timeless power system monitoring to witness the symptoms by planting an artificial intelligence-based model in electrical protection devices (EPD). The EPD has to be able to determine what is happening in the system and decide the precise action to overcome the situation. This paper intends to show the opportunity of implementing a deep learning computation to solve a protection scheme problem in power system operations. It is known that deep learning has an undoubted performance in terms of image processing [10]. In this case, the pattern of data series based on the simulated power system events will be used as a feature to be extracted. Deep learning is based on the convolutional neural networks (CNN) since its powerful ability to recognize the pattern [11]. The objective is to use CNN to perform an event classification based on large data sets gathered in the operation of medium voltage power system distribution networks. Later, the CNN is combined with the long-short term memory (LSTM) technique to increase the accuracy of recognizing and classifying the power system distribution network symptoms.

2. COMMON PROBLEM ON MEDIUM VOLTAGE POWER SYSTEM DISTRIBUTION NETWORK

2.1 Fault and Fault-like Event in Power System

An event in the electrical system such as ground fault, phase fault, start or cease of electrical machines, and network topology changes has its symptoms. These symptoms are categorized based on the electrical parameter, i.e., frequency, current flow, and voltage. As an illustration, during the event of a short circuit fault, the closest source feeder relay to the fault location will experience a drop voltage

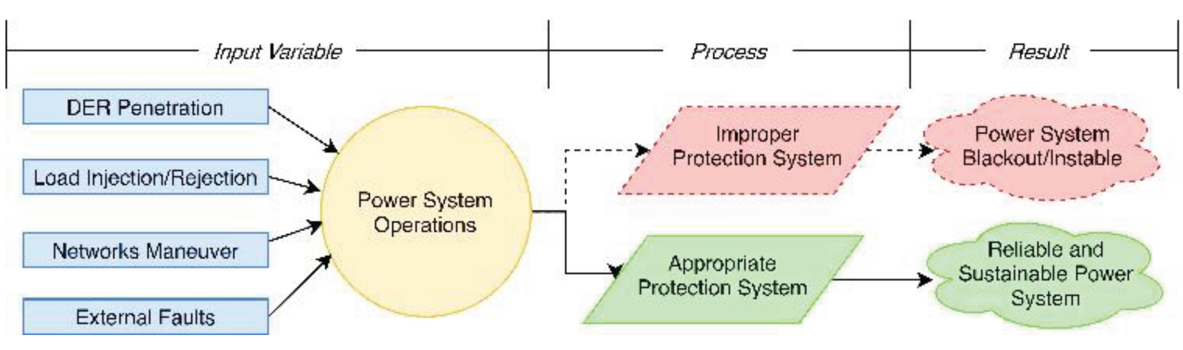


Fig. 1. Reliable and sustainable power system is a result of an appropriate protection system and proper power system operations

followed by increasing current flow and a higher frequency rate at the source site. Meanwhile, in the event of motor starting, the feeder may sense a dip voltage, surge current, and frequency drop. These two events are different even though the symptom of voltage and current is similar.

The fault conditions require CB to trip while it is prohibited to trip during motor starting events. Somehow, due to an improper setting, there is a record of CB trip during the motor starting. This happens quite often in most petrochemical or oil-and-gas companies [12]. The root cause is because the multi-function relay was installed with a default setting and not considering the in-rush condition. As a containment action, the engineer usually broader the setting or disabling to anticipate the drop voltage or the surge current during motor starting. This action comes with a risk that the relay may do not trip when the short circuit occurs during the start process. To validate this hypothesis, time-series data is gathered using a commercial software called ETAP in a transient stability domain.

To set the simulation, the sequence of events is divided into three domains, which are: i) PRE-event, ii) ON-event, and iii) POST-event [13]. To visualize these time domains, Figure 2 represents the voltage, frequency, and current measurements of a feeder that suffering a short circuit condition at t = +1 s.

As shown in Figure 2, PRE-event denoted as the state where there is no switching condition. meaning to say, this domain happens before t = 1 s and might be stated as the normal or basic condition. Meanwhile, ON-event just after the switching. In this domain, the event is happening without any controller react to overcome the events. The ON-event occurs at t = +1 s until any reactions or following events. Regarding that, the EPD must take action or decision during this time domain. The POST-event is any event that happens with regard to the previous events. It could be an opening of CB to clear the fault. However, in this POST event, there is a possibility that the system might go unstable or being stable. This condition is most likely depending on how long the time required of the CB to react.

2.2 Power System Network and The Protection System

The simple system consists of two generators where one of them is considered as a distributed generator (DG). The main transfer bus of this system is an 11 kV that is connected into four different feeders. Every feeder on this system is equipped with a relay that able to observe voltage (V), frequency (f), and current flow (I) in rms. The relay ID is namely A until F as shown in the single line diagram (SLD) Figure 3.

There are several loads connected to the main bus of 11 kV. However, this paper is focused on the 11 kV events only. Several situations that taken into consideration are i) normal condition, ii) three-phase symmetrical (LLL) fault of phase a-b-c in 11 kV, iii) line to line (LL) fault of phase a-b in 11 kV, iv) line to ground (LG) fault of phase a-b in 11 kV, v) the event of DG outage, and vi) direct online (DOL) motor starting of 3 000 kW induction

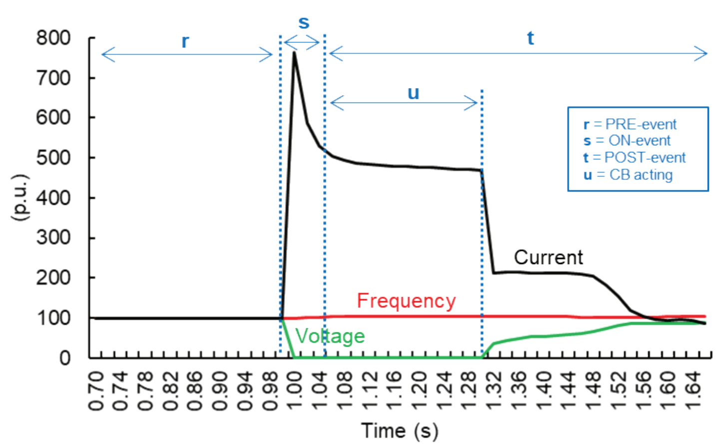


Fig. 2. The sequence of events in the power system on the perspective of voltage, frequency, and current flows

machine. Those six situations will be simulated under several circumstances, such as the penetration of DG and the total load connected.

2.3 Event Classification Based on Convolutional Neural Network (CNN)

2.3.1 CNN Architecture

Recently, Artificial Intelligence (AI) has witnessed a massive development that could bridge the divide between human and machine capacities. This field aims to allow computers to see the environment as people do, to interpret it in the same way, and even to use understanding for many functions such as the identification of images and videos, image analyses and classification, recommendation systems, linguist processing, etc. In power system generation with intelligent electricity meters increasing and the widespread use of power generation technology such as solar panels, we have much information on the use of electricity. This information is a multivariate time sequence of power that could be used in turn for modeling and even predicting future energy usage. The progress made in Artificial Intelligence (AI) by Deep Learning, mainly through one specific algorithm – Convolutional Neural Network (CNN), has been built and improved over the moment. CNN can learn characteristics from sequence information, assist multiple-variate information automatically, and able to immediately generate a multi-step forecasting vector [10]. A CNN is a deep learning algorithm capable of capturing input data, assigning significance (learnable weights and biases) to multiple aspects/objects, and being prepared to distinguish between them [14] Figure 4 depicts the typical CNN block scheme that later will be used on the power system operations.

The fundamental objective in implementing deep learning is to eliminate the complicated and eventually restricting choice of features. The role of a convolutional layer can be articulated merely: it uses a three-dimensional quantity of data to generate a fresh three-dimensional quantity of data.

As shown Figure 5, illustrates the 3D convolution process used in CNNs, and input features are used for the convolution operation. The first convolution layer uses low-level characteristics like edges, rows, and angles. The input is $N \times N \times D$ and is converted with H kernels, each of which is separately $k \times k \times D$. Convolution of with one kernel generates an output function and separately generates features map with kernels. Each kernel passes one element at a time from the top-left corner. One element will be moved down from the kernel and one item will be passed over again from left to right. This method is performed until the kernel hits the bottom-right corner. If input height and width are equal to 32 and the kernel value is 5, there will be 32 distinctive places from left to correct and 32 distinctive places from top to upper that the kernel can hold. According to these locations, each function in the output will comprise 32×32 $((N-k+1) \times (N-k+1))$ components. To be able to achieve a single component in a single output, multiplier activities between two components are input=$(k \times k \times D)$ and kernel=$(k \times k \times D)$ are needed for each place in the sliding window method.

In Figure 6, the same color links are limited to having the same weight at all times. This can be achieved by initializing all the connections within a group with the same weights and by always averaging a group's weight updates before applying them at the end of each backpropagation iteration. This filter produces the output layer of the function map. A neuron is triggered in the feature map where a filter is identified at the respective place in the prior layer to contribute to its activities.

3. RESULTS AND DISCUSSION

3.1 Data Gathering

As stated in Section 2.1. The most crucial part of the electrical circumstances is the ON-event. Therefore, by considering that the relay sampling is one cycle (equal to 20 ms for 50 Hz system) and the relay must react within five cycles, the data consist of five samplings construct by the PRE-event and ON-event as shown in Figure 7.

Says there are five rows of data to be recognized by the CNN, then the first two data consist of the PRE-event data or the normal (basis) condition. When the event is occurring, there will be a deviation between the PRE and ON events. Using the SLD, as shown in Figure 3, the whole relay on the system is assumed able to measure the current, voltage the frequency. Each phase on every feeder will be measured regarding the voltage and current flows in rms value. While the frequency is

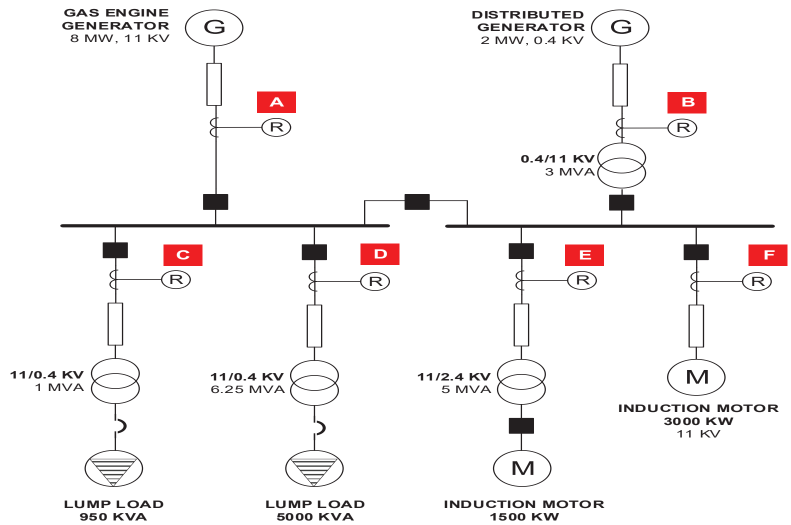


Fig. 3. Single Line Diagram (SLD) being used for the simulation and data gathering.

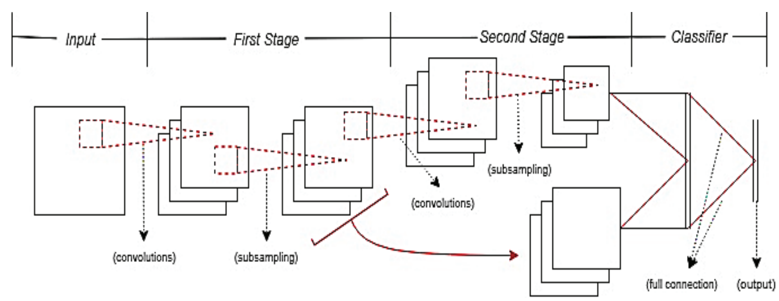


Fig. 4. A typical CNN block scheme [10]

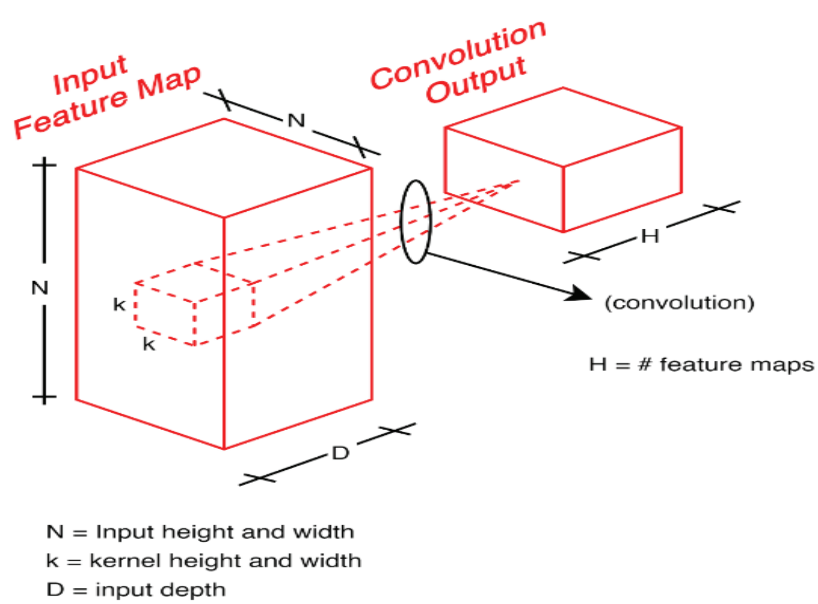


Fig. 5. Convolutional process representation [15]

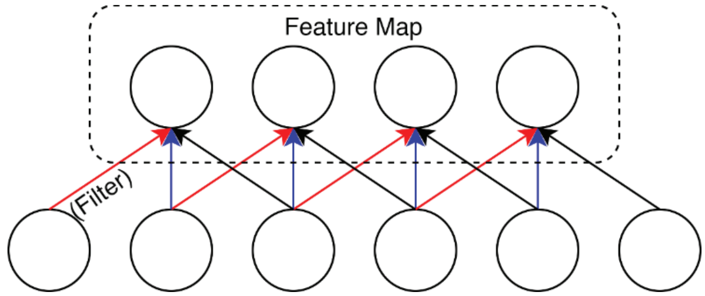


Fig. 6. The filters and maps are represented as neurons in a convolutional layer [15]

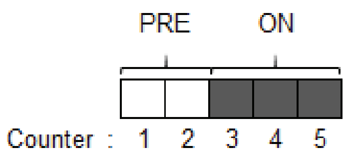


Fig. 7. Visualization of the data required to be trained in ANN

measured only in a single phase. The illustration of data measurement is shown in Figure 8.

Generally speaking, there are two types of feeders in Figure 3. The upper side of the 11 kV bus might be called a source feeder while the lower side is the load feeder. Figure 9 shows the illustration of the current reading for the whole relay with a variety of conditions, as stated in Section 2.2. Moreover, Figure 10 shows the illustration voltage and frequency reading in accordance with the data counter. In total, there are 193 symptoms constructed based on the combination of the power system operations and events.

3.1.1 Current flow analysis

Relay A and B could be grouped as the source feeder relay. For this type of relay, it might be known from Figure 9a. and Figure 9b. that there is a tremendous current spike when the LLL fault is occurring. The other current surge occurs when there is a motor starting. Especially for Relay A, the minor current increase happens when the DG is falling to outage condition. For sure, on the other perspective, Relay B will sense the current drop, as shown in Figure 9b. When the LG fault happens, both Relay A and B only sense a tiny deviation in terms of current reading.

Meanwhile, for Relay C and D that operate as a feeder type relay, the immense current drop occurs when the LLL fault is happening. This is due to there is no source feeding to the load during that occasion. A bit of current decrease happens during the motor starting. During the DG outage and LG fault, there is no current variation on the feeder line. Relay E and F belong to an induction machine that operates on 2.4 and 11 kV. During the LLL fault,

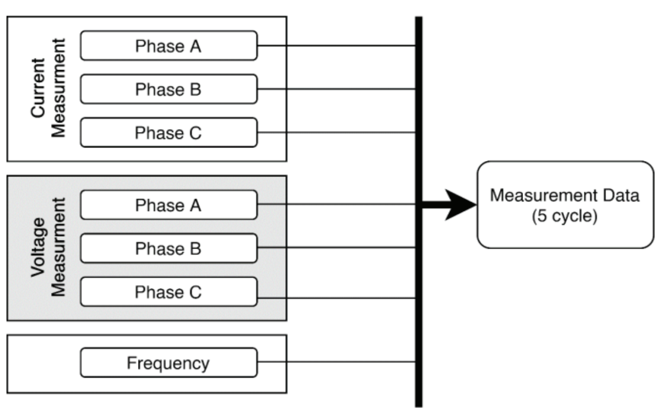
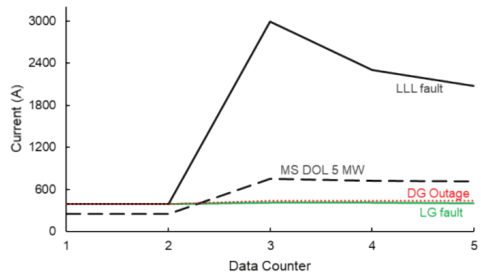
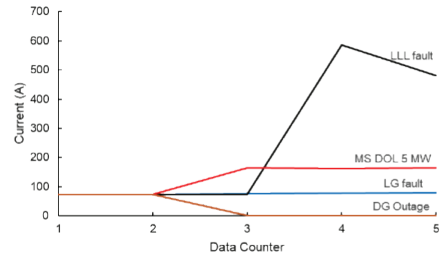


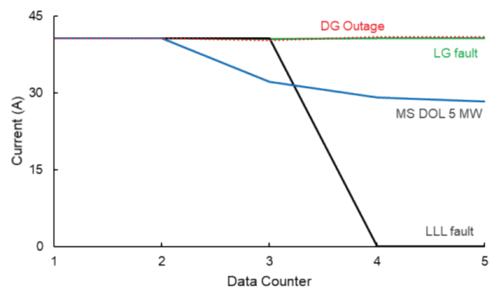
Fig. 8. Practical measurement data from the simulation



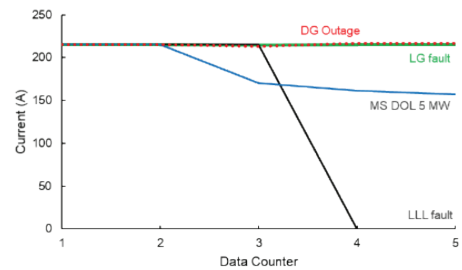
(a) Current reading on Relay A



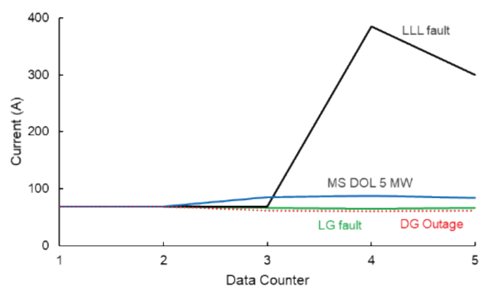
(b) Current reading on Relay B



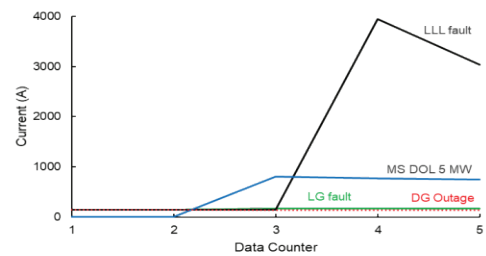
(c) Current reading on Relay C



(d) Current reading on relay D



(e) Current reading on Relay E



(f) Current reading on Relay F

Fig. 9. Voltage (a) and Frequency (b) reading on each relay with a varied event in 11 kV bus: LG fault, LLL fault, DG Outage, and Motor Starting (MS).

both relays sense a huge current spike since the motor is acting as a generator that feeds the current to the fault location. A minor current spike happens during the motor starting while there is such a big difference in terms of current reading during DG outage and LG fault.

3.1.2 Voltage analysis

Since all the relay is connected in parallel to the 11-kV bus, the voltage reading of the all-six relays must be the same. It might be known from Figure 10.a. that the voltage collapse is immensely happening during the LLL fault. There is also a voltage collapse condition that happens during the motor starting, but not as severe as the LLL fault. In the event of a DG outage, the voltage reading shows if there is no deviation compared to the PRE-event condition. There is a minor increment of voltage during the LF fault.

3.1.3 Frequency analysis

Similarly, to the voltage measurement, the frequency reading only needs a one-perspective measurement since the system is fully synchronized. Figure 10.b. shows that the major frequency collapse happens when the DG outage. A minor frequency collapse also happening when there are a motor starting and the LG fault. Meanwhile, during the LLL fault, the frequency is increasing dramatically.

3.2 Training and Classification using CNN

3.2.1 Implementation of CNN

The network input is a 135 × 35 matrix that is considered as a single data channel. The first layer, C1, conducts four input conversions with 2 × 2 kernels, creating four-function maps of size 133 × 33. In the second layer, C2, execute 2 × 2 kernels. The information is rescaled to suit the Keras sequential model's three-dimensional input criteria. For a simple univariate model, the input form is 35-time stages with one feature. In the convolutional layer, the sequence was not divided into several subsequences, but rather 32-time stages. Six neurons in the thick layer generate six outputs.
One of the popular Recurring Neural Network (RNN) models is LSTM. As shown in, the schematic of LSTM and σ denotes a sigmoid feature. There are three basic gates on LSTM: i) an input, ii) output and iii) a forget gate. The procedure operation between these three doors allows LSTM able to fix long-term dependencies which are not learned by particular RNNs.

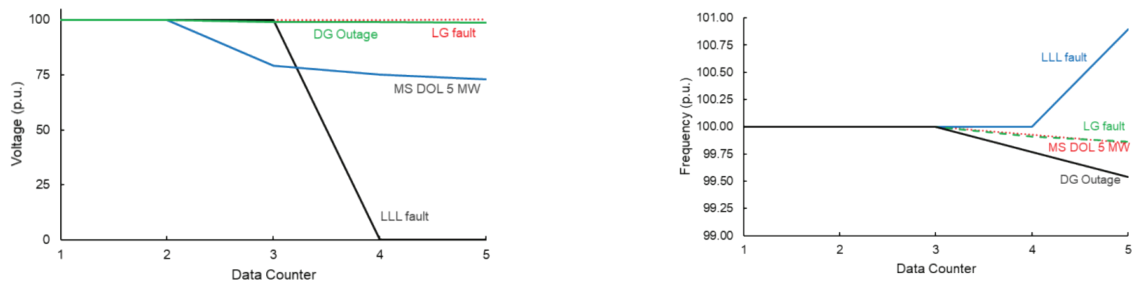
Both in Keras, a CNN-LSTM can be defined for training. A CNN LSTM can be described by bringing the first layer is CNN layers, and the second layer is followed by LSTM layers with a Dense layer at the output layer. It is helpful to think of this architecture as defining two sub-models. The CNN Model for features extraction and the LSTM Model for analysis of features over time phases. Figure 12 depicts the illustration of the feeding process from CNN to LSTM while the architecture based on CNN is summarized in Table 1. Figure 13 shows the Keras sequential model.

3.2.2 Experimental Result

Figure 14 shows the experimental result of the proposed CNN to recognize the symptoms of power system operations. It might be known that the model with 200 epochs might reach 69 % accuracy, as shown in Figure 14.a. An increasing epoch up to 350 shows that there is an accuracy spike that reaches 79 % at epoch 275 as depicts in Figure 14.b. This result later is confirmed by the heatmap of the dataset of features 1-35 (x-axis) with the value of each feature per row (y-axis) as shown in Figure 14.c. This indicates that CNN and LSTM can generalize and not just memorizing the pattern. Furthermore, this might be preliminary proof that CNN and LSTM might be helpful to supervise the power system operations.

Table 1. CNN based architecture

Layer	Kernel/Pooling Window	Layer Size
Input	-	1@135x35
Lstm (C1)	[32@2x2]	32@133x33
Lstm (C2)	[32@2x2]	32@122x31
Lstm (C3)	[32@2x2]	32@110x9
Dense (P2)	[6@2x2]	6@58
Full Out (F1)	-	58



a) Voltage reading on 11 kV bus

b) Frequency reading on 11 kV bus

Fig. 10. Voltage (a) and Frequency (b) reading on each relay with a varied event in 11 kV bus: LG fault, LLL fault, DG Outage, and Motor Starting (MS).

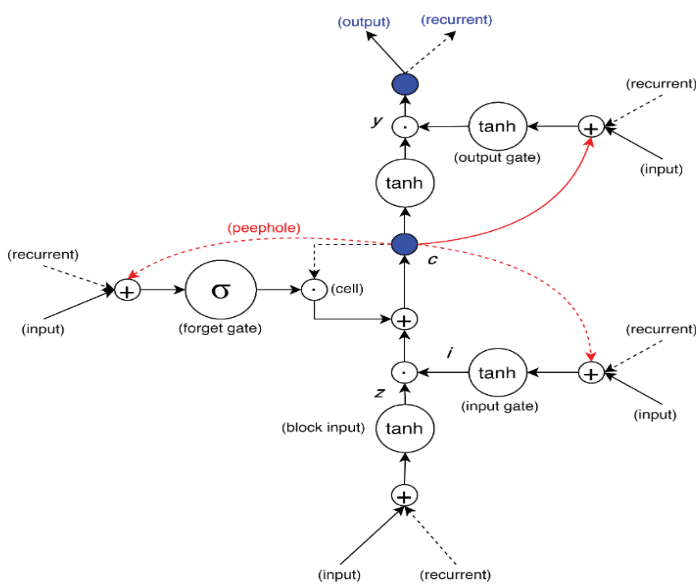


Fig. 11. LSTM Block

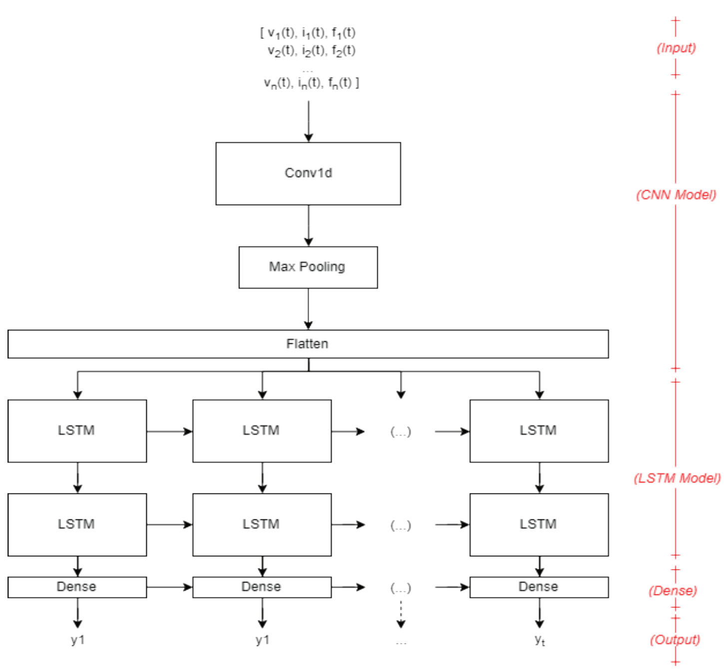
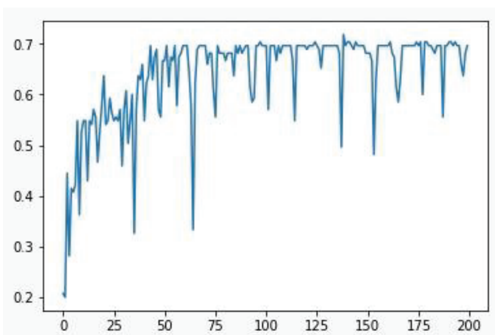


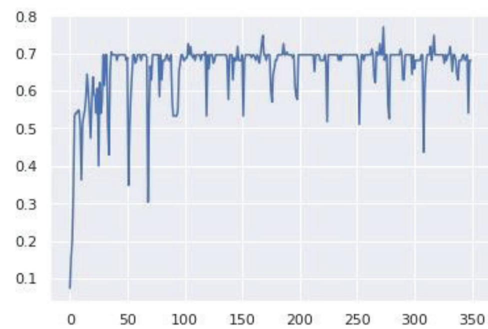
Fig. 12. Feeding into LSTM layers

```
from keras.layers import (
    Activation, Flatten, Dropout, LSTM,
    Dense,


model.add(LSTM(32, return_sequences=True,
               input_shape=(35, 1)))
model.add(LSTM(32, return_sequences=True))
model.add(LSTM(32))
model.add(Dense(6, activation='softmax'))

model.compile(loss='categorical_crossentropy',optimizer='adam',metrics=['acc'])
model.summary()
```

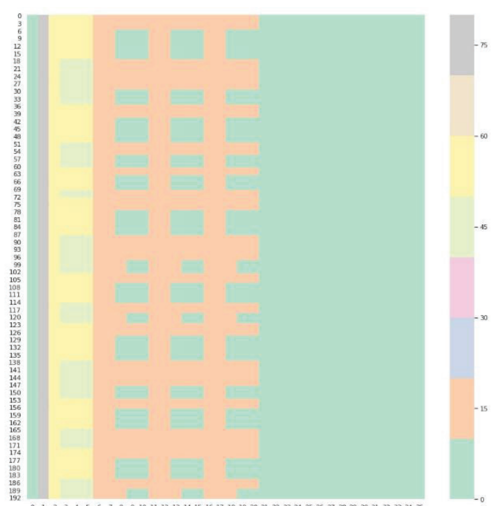
Fig. 13. Keras sequential model



(a)



(b)



(c)

Fig. 14. Trial result of the CNN to recognize the symptoms of power system operations: (a) 200 epoch (b) 275 epoch, (c) heatmap dataset

4. CONCLUSION

The learning time would also improve proportionately if the conventional Neural Network (NN) were to be equal to a CNN, as the number of parameters would be much greater. With a significant reduction in parameters in a CNN, the learning period is reduced proportionately. In an ideal practice, a conventional NN might be built with the same performance as a CNN. A standard CNN-equivalent neural network would have more parameters that add to greater noise during the training stage. Therefore, the output of conventional CNN is always less efficient. In the future, it is expected that the expenditure of more time in a customized CNN architecture can result in an ideal model for the future task. Ensemble designs have shown that other classification activities are achieving stronger outcomes. Given that a model is better at rating specific event types, while its overall accuracy is low, multiple models can increase the classification performance. Lastly, the research was restricted to six events in power system operation with several kinds of circumstances of the case. A further kind of events and symptoms might be achieved by working with a greater number of the bus that could assist to develop stronger CNN models and produce precious fresh findings.

6. ACKNOWLEDGEMENTS

This research was supported by the MSIT (Ministry of Science, ICT), Korea, under the ITRC (Information Technology Research Center) support program (IITP-2019-2014-1-00639) supervised by the IITP (Institute for Information & communications Technology Planning & Evaluation) and Priority Research Centers Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2018R1A6A1A03024003).

6. CONFLICT OF INTEREST

The authors declare no conflict of interest.

5. REFERENCES

1 D. Yang., J. Kim., Y.C. Kang., E. Muljadi., N. Zhang., J. Hong., and T. Zheng, Temporary Frequency Support of a DFIG for High Wind Power Penetration. *IEEE Transactions on Power Systems* 33(3): 3428–3437 (2018). DOI: 10.1109/TPWRS.2018.2810841

2 E. Muljadi., N. Samaan., V. Gevorgian., J. Li., and S. Pasupulati. Different factors affecting short circuit behavior of a wind power plant. *IEEE Transactions on Industry Applications,* 49(1): 284–292 (2013). DOI: 10.1109/TIA.2012.2228831

3 J. Kim., E. Muljadi., V. Gevorgian., and A.F. Hoke. Dynamic capabilities of an energy storage-embedded DFIG system. *IEEE Transactions on Industry Applications* 55(4): 4124–4134 (2019). DOI: 10.1109/TIA.2019.2904932

4 J.P. Holguin., D.C. Rodriguez, and G. Ramos. Reverse Power Flow (RPF) detection and impact on protection coordination of distribution systems. *IEEE Transactions on Industry Applications.* 56(3): 2393–2401 (2020). DOI: 10.1109/TIA.2020.2969640

5 V.R. Mahindara., D.F.C. Rodriguez., M. Pujiantara, A. Priyadi., M. H. Purnomo., and E. Muljadi. Modern concerns and challenges of overcurrent protection coordination in distribution systems. Presented at the *2020 IEEE/IAS 56th Industrial and Commercial Power Systems Technical Conference (I&CPS)*, Virtual Conference, Apr. 2020.

6 A. Hooshyar, and R. Iravani. Microgrid Protection. *Proceedings of the IEEE* 105(7): 1332–1353 (2017), DOI: 10.1109/JPROC.2017.2669342

7 V.R. Mahindara., M.G. Istiqlal., M. Pujiantara., D.A. Asfani., A. Priyadi., and M. H. Purnomo. Obtaining the setting of inverse-curve overcurrent relay using serial computing modified particle swarm optimization in real system applications. In: *2018 International Seminar on Intelligent Technology and Its Applications* (ISITIA), Aug., 187–192 (2018). DOI: 10.1109/ISITIA.2018.8710868

8 A. Prasad., J. Belwin Edward., and K. Ravi, A review on fault classification power transmission systems: Part—I, *Journal of Electrical Systems and Information Technology* 5(1): 48–60 (2018). DOI: 10.1016/j.jesit.2017.01.004

9 A. Prasad., J. Belwin Edward., and K. Ravi, A review on fault classification methodologies in power transmission systems: Part-II, *Journal of Electrical Systems and Information Technology* 5(1): 61–67, (2018). DOI: 10.1016/j.jesit.2016.10.003

10 Y. LeCun., Y. Bengio., and G. Hinton. Deep learning. *Nature* 521(7553): 436–444 (2015). DOI: 10.1038/nature14539

11 H. Chiroma, U.A. Abdullahi, S.M.Abdulhamid, A.A. Alarood, L.A. Gabralla, I.A.T. Hashem, D.E.

Gbenga, A.I. Abubakar, A.M. Zeki, and T. Herawan. Progress on artificial neural networks for big data analytics: A survey, *IEEE Access* 7: 70535–70551 (2019). DOI: 10.1109/ACCESS.2018.2880694

12 R.M. Vincentius, A.W. Nugraha., P.S. Talitha., P. Margo., P. Ardyono., and H.P. Mauridhi. Recognition of electric machines boundary as the constraint of over current relay coordination in real industrial application with serial firefly algorithm optimization. In *2019 IEEE 12th International Symposium on Diagnostics for Electrical Machines, Power Electronics and Drives* (SDEMPED), Aug. 153–159 (2019). DOI: 10.1109/DEMPED.2019.8864828

13. F.W. Rahmat., M. Pujiantara., V. Lystianingrum., V.R. Mahindara., and T.P. Sari. Self-Classification of multifunction relay based on neural network for industrial scale," in 2020 International Seminar on Intelligent Technology and Its Applications (ISITIA), Jul. 13–18 (2020), DOI: 10.1109/ISITIA49792.2020.9163719

14. N. Buduma. *Fundamentals of Deep Learning.* O'Reilly Media, Inc. Boston, MA (2017).

15. K. Ovtcharov., O. Ruwase., J. Kim., J. Fowers., K. Strauss., and E.S. Chung. Toward accelerating deep learning at scale using specialized hardware in the datacentre. *In 2015 IEEE Hot Chips 27 Symposium (HCS), Aug*. (2015), 1–38. DOI: 10.1109/HOTCHIPS.2015.7477459