Research on Fault Line Selection Method Based on Convolutional Neural Networks

Jinfeng Gao and Yufeng Wang

College of Electrical Engineering, Zhengzhou University, Zhengzhou, Henan, China
Email: 810429024@qq.com

Abstract. In order to improve the reliability and accuracy of Single-Phase-to-Ground Fault line selection for small current grounding systems, this paper proposes a method named Deep Convolutional Neural Networks with Wide First-layer Kernels (WDCNN). The proposed method uses the original current signal as input and uses a wide kernel in the first convolutional layer to extract features and suppress high-frequency noise. The following convolutional layers with small convolution kernels are used for multilayer nonlinear mapping. Batch normalization layers and dropout layers are carried out for improving the generalization ability of the model. The simulation results show that the recognition rate of the proposed WDCNN can reach 99%, and it is not affected by factors such as grounding impedance, fault location and fault time.

Keywords. Single-Phase-to-Ground Fault; deep learning; convolutional neural networks; fault identification.

1. Introduction

Most of medium and low voltage distribution networks use small current grounding in China [1]. After a single-phase ground fault occurs in a system with such a grounding method, the inductive current generated by the arc suppression coil can compensate for the capacitive ground current of the system, so that the fault current is very small, the protection device will not immediately trip, and it can still run with electricity 2 hours. Although the low-current grounding system improves the reliability of electric energy, because non-fault relative ground voltage increases, this voltage is likely to break through the weak points of the insulation in the system and pose a threat to the safety of electrical equipment.

After a single-phase ground fault occurs, the amplitude of the fault current in the small current grounding system is relatively small compared to the normal operating current, which makes it harder to select the fault line. The line selection method is divided into two types: one is the external signal approach and the other is the fault signal approach. The external signal method requires external injection equipment, and the use of fault signals for feeder protection is the most in-depth research [2-6]. The fault signal method is to extract the transient, steady state or full information features of the fault signal, directly set the threshold to form a criterion, or obtain a classifier with strong generalization ability through sample training for line selection. This type of machine learning method requires artificial feature extraction, and the final recognition effect is determined by the ability of the feature to represent the original signal [7-8]. As a further development of machine learning, deep learning methods can eliminate deviations caused by manually extracting features, allowing the network itself to learn the most representative features in classification tasks. Existing research shows that the deep learning method can train a large number of samples to extract features from complex systems with weak fault features, and realize fault diagnosis and classification.
Essentially, fault line selection is a multivariate time series classification problem. For time series classification, convolutional neural networks can learn features that do not vary with space from the original input time series, and thus are considered to be the most widely used architecture in time series classification problems [9-10]. Convolutional neural networks have been widely used in time series classification. In the context of power big data, it has broad prospects for its application in power system fault line selection.

2. Sample Acquisition and Analysis
Because deep learning requires massive data as support, enough training and iterations are needed to generate the ideal model. However, in actual production, it is hard to acquire a large number of fault samples, so this article uses a simulation model and batch simulation to obtain a large number of samples. In order to facilitate research, this article uses the 4 lines model as the simulation object. The transformer capacity is 250MVA, and the overhead line is used. The parameters are shown in table 1, and the topology is shown in figure 1.

| Line type | Phase sequence | Resistance (Ω/km) | Inductance (mH/km) | Capacitance (μF/km) | Length of each line(km) |
|-----------|----------------|-------------------|-------------------|---------------------|------------------------|
| Overhead line | Positive | 0.450 | 1.260 | 0.061 | 10, 20, |
| Zero | 0.700 | 3.640 | 0.038 | 30, 40 |

According to the line number of the fault, the sample can be divided into 4 categories. Different faults include different fault phase angles, different fault distances, different grounding impedances, and different load conditions. The frequency for sampling is 12.8kHz, the quantity of sampling points is 2560 and the sample specification is $4 \times 2560$. The phase angle of the fault is randomly selected from $0^\circ$ to $360^\circ$, the fault location is randomly selected on the entire feeder, and the grounding impedances are $0.1\Omega$, $100\Omega$, $500\Omega$, and $1000\Omega$, respectively. In this paper, 24,000 samples are obtained through automated simulation, of which 15,000 are applied as the training set and 9000 are applied as the test set. Take the metal single-phase fault grounding of the outgoing line 4 as an example. Figure 2 shows the zero sequence current waveform of each line. It can be seen from the current that the amplitude of the zero-sequence current of each line will change abruptly after a fault occurs, and the polarity of the first half wave of the faulted line and the non-faulted line is different.

3. WDCNN Structure
This section proposes a model of “first layer wide convolution kernel deep convolution neural network” (WDCNN) for the characteristics of power system current signals. The network except the first layer convolution kernel is a large convolution kernel the convolution kernels of other layers are all small convolution kernels. In order to enhance the expressive power of WDCNN, the convolution kernel size
of the remaining convolutional layers of this model is $5 \times 1$. This design can effectively reduce the number of parameters, so it can deepen the number of neural network layers to enhance the expressive ability of the model. After the convolution pooling operation, in order to reduce the impact of data distribution on the model, the output data needs to undergo batch normalization (BN). The structure of the entire model is shown in figure 3.

![Figure 3. WDCNN structure.](image)

4. Experiment
The server selected for the experiment has a CPU model of E5-2630 v4, a server memory size of 64G, and a GPU model of NVIDIA TITAN Xp. The keras framework developed by Google was selected for the construction of the algorithm model. The specific parameters are shown in table 2. The data processing and model training script language is Python, and Seaborn is used for visual analysis.

4.1. Training Process and Results
WDCNN is shown in figure 4 during the training process. The relationship between the accuracy and the value of the objective function and the training algebra are shown in the blue lines in figures 4a-4b, respectively. In order to explore the role of batch normalization (BN) in the entire training process, this section performs ablation testing on the model, removes the BN layer and trains the model again, and the results are shown in Figures 4a-4b In the orange line. It can be seen from Figure 4a that the model with BN layer can achieve a recognition rate of 99% after about 10 trainings. After removing the BN layer, the recognition rate rises not only slowly but also with great volatility. According to the picture 4b, during the training process of the model with BN layer, the target loss function drops smoothly and eventually stabilizes within 0.1. After the model is retrained except for the BN layer, it is not difficult to find that the value of the objective function decreases slowly and has certain volatility. Compared with the BN-layer model, although the accuracy rate is basically the same, the loss function value is doubled. In summary, the BN layer not only reduces the training difficulty, but also increases the accuracy of fault line selection, which is an important part of the model.

4.2. Anti-noise Performance Analysis
In the current fault line selection algorithm based on machine learning, the common idea is to extract the fault feature of the zero-sequence current signal through wavelet analysis (WPT), and then use common machine learning algorithm models such as support vector machine (SVM), multilayer perceptron (MLP) and Gradient Boosting Decision Tree (GBDT) to classify the extracted features [4]. This section compares the performance differences between the WDCNN model and the common WPT-SVM, WPT-GBDT, and WPT-DNN models using deep learning. Among them, SVM uses radial basis
function, the number of GBDT weak learners is 100, and the number of neurons in each layer of DNN are 100, 32 and 4.

![Validation accuracy](image1)

**Figure 4.** Visualization of the training process.

| No | Layer type  | Kernel Size/Strid | Kernel Number | Output Size (Width × Depth) | Padding |
|----|-------------|-------------------|---------------|----------------------------|---------|
| 1  | input       | —                 | —             | 2560×4                      | —       |
| 2  | Convolution1| 16×1/16×1         | 128           | 2560×128                    | yes     |
| 3  | Pooling 1   | 2×1/2×1           | 128           | 1280×128                    | no      |
| 4  | Convolution2| 5×1/2×1           | 256           | 1280×256                    | yes     |
| 5  | Pooling 2   | 2×1/2×1           | 256           | 640×256                     | no      |
| 6  | Convolution 3| 5×1/2×1          | 128           | 640×128                     | yes     |
| 7  | Pooling 3   | 2×1/2×1           | 128           | 320×128                     | no      |
| 8  | Convolution 4| 5×1/2×1          | 128           | 320×128                     | yes     |
| 9  | Pooling 4   | 2×1/2×1           | 128           | 160×128                     | no      |
| 10 | Convolution 5| 5×1/2×1          | 128           | 156×128                     | no      |
| 11 | Pooling 5   | 2×1/2×1           | 128           | 78×128                      | no      |
| 12 | Fully- connected | 512              | 1             | 256×1                       | —       |
| 13 | Softmax     | 4                 | 1             | 4                           | —       |

The experimental results are shown in figure 5. From the figure, it can be seen that the WPT-DNN algorithm is superior to WPT-SVM and WPT-GBDT. Although all features are artificially extracted, the use of deep learning models is still better than traditional machine learning models. These algorithms that rely on artificial feature extraction are not very good in anti-noise ability. For example, in the SNR = 0dB environment with strong noise, the accuracy of these three is less than 85%. WDCNN’s accuracy rate is more than 90%.

### 4.3. Visual Analysis

In this section, the neuron expression of the zero sequence current signal in all convolutional layers and fully connected layers will be visually analyzed, as shown in figure 6. Figure 6a shows a single-phase ground fault at line 1, and figure 6b shows a single-phase ground fault at line 3. Dark purple indicates that the neuron is not activated, and yellow indicates the neuron that is the most activated. It can be seen from the figure that after the convolution operation, the location and the number of neurons in the first convolutional activation layer with different types of fault signals have certain differences. Among them, the expression of the transient part is mainly yellow, which indicates that the number of activated
neurons is large, and the activation of the neuron in the steady state is periodic. As the convolutional layer increases, the differences in the signals of different fault types also increase. After passing through the fully connected layer, the difference between the two can be clearly identified. It can be seen from the positions marked by the red boxes in figures 6a-6b that the characteristics of the transient part in the zero-sequence current are retained in the neural network and always appear yellow, that is, the part has been in the activation state shows that this part of the characteristics plays an important role in the fault line selection.

**Figure 5.** Comparison of classification accuracy under different noisy environment.

**Figure 6.** Visualization of all Layers in WDCNN.
5. Conclusion
This paper gives an account of a line selection algorithm for small current grounding systems. It does not require zero-sequence current and avoids complex artificial feature extraction of the original signal. The recognition rate can reach 99% under fault conditions. This method has good anti-noise ability, and the accuracy rate is still around 95% in the environment of SNR = 5dB. Compared to methods that rely on manual feature extraction, they have better performance. In addition, visualization techniques are used to demonstrate WDCNN’s fault line selection process. The results show that transient signal characteristics play an important role in fault line selection. Since the proposed fault line selection method is based on simulation data, there are still many problems that need further discussion. Considering that the built model and the actual collected data have a high degree of similarity in features, how to use the actual data to further fine-tune on the network that has been trained with simulation data to achieve transfer learning is the next direction worth studying.

References
[1] Zeng X, Yu K and Wang Y 2016 A novel single phase grounding fault protection scheme without threshold setting for neutral ineffectively earthed power systems CSEE Journal of Power and Energy Systems 2 73-81.
[2] Lianjie H, Changkai S and Zhuo Y 2017 A fault section location method for small current neutral grounding system based on energy relative entropy of generalized S-transform Transactions of China Electrotechnical Society 32 274-80.
[3] Tailin X, Junlang H and Jianxin Z 2018 Fault line detection in low current grounding system based on GA optimized BP neural network Electrical Automation 40 66-9.
[4] Feng Z and Dechang Y 2013 Fault line selection based on wavelet packet and improved BP neural network for power distribution grid Automation Instrumentation 34 4-8.
[5] Moslem D, Mohammad H and Taher N 2016 Fast fault detection and classification based on a combination of wavelet singular entropy theory and fuzzy logic in distribution lines in the presence of distributed generations Int. J. Electr. Power Energy Syst. 78 455-62.
[6] Ren Z and Zhang Y 2015 Fault line selection of distribution network based on improved Hilbert-Huang transform and identification confidence degree Power Syst. Prot. Control 43 8-13.
[7] Lecun Y, Bengio Y and Hinton G 2015 Deep learning Nature 521 436.
[8] Esteva A, Kuprel B and Novoa R 2017 Dermatologist-level classification of skin cancer with deep neural networks Nature 542 115-18.
[9] Janssens O, Slavkovikj V and Vervisch B 2016 Convolutional neural network based fault detection for rotating machinery Sound Vib. 377 331-45.
[10] Sheng G, Tao Y and Wei G 2018 A novel fault diagnosis method for rotating machinery based on a convolutional neural network Sensors 18 1429.