A Recognition Method for Time-domain Waveform Images of Electric Traction Network Overvoltage Based on Deep Learning

JIA Junyi*, WU Mingli and WANG Qi
School of Electrical Engineering, Beijing Jiaotong University, Beijing, 100044, China
*Corresponding author’s e-mail: 18121444@bjtu.edu.cn

Abstract. A large amount of time-domain waveform images of overvoltage of electric traction network are collected during the on-board test of EMU. The traditional time-domain or frequency-domain analysis and classification method can’t be directly applied to time-domain waveform images. In order to analyze these images, firstly, image preprocessing is carried out. Then, the data set of time-domain waveform images is established. Finally, the CNN (convolutional neural network) classification model is trained based on data set. The results show that the sensitivity of recognition is above 81%, and this method is suitable for the time-domain waveform image of overvoltage of electric traction network.

1. Introduction
In recent years, the power quality problem of traction network caused by the instability of network-train match is becoming more and more serious, among which overvoltage is the most common phenomenon. The traction network overvoltage often occurs, which causes breakdown of traction converter, fire burn of pantograph and tripping of feeder circuit of traction substation. In serious cases, it may cause the explosion of the voltage transformer, which may affect the safety of electrified railway. Therefore, the analysis and classification of traction network overvoltage is helpful to improve the safety and reliability of traction power supply system.

With regard to the traction network overvoltage analysis method, the traditional methods are mainly time-domain and frequency-domain analysis methods. Time-domain analysis is a direct method to analyze signals in time domain, so time-domain analysis has the advantages of intuition and accuracy; specific methods of frequency-domain analysis include fast Fourier transformation, short-time Fourier transformation and wavelet transformation. Currently variety of instruments used in EMU test results in a variety of data file formats. To analyze the overvoltage, we often export the time domain waveform of the overvoltage to image formats (such as JPEG, IMG and PNG). In the previous EMU tests, a large amount of time-domain waveform images have been collected. However, traditional time-domain and frequency-domain analysis methods cannot be applied to this kind of unstructured data.

Convolutional Neural Networks (CNN) is a kind of feed-forward neural network with deep structure and convolution calculation. It is one of the algorithms of deep learning. A CNN consists of one or more convolutional layers and fully connected layers (corresponding to classical neural networks), as well as the associated weights and pooling layers. This structure enables the CNN to utilize the two-dimensional structure of the input data. Compared with other deep learning structures, CNN can give better results...
in image recognition. The VGG convolutional neural network is a model proposed by Visual Geometry Group at Oxford University in 2014. When this model was proposed, it became the most popular CNN model at that time because of its simplicity and practicability. It performs very well in both image classification and target detection tasks. In summary, based on VGG, this article uses TensorFlow 2.0 (GPU) + Keras framework to supervise training based on the traction network overvoltage time-domain waveform image data set, then uses regularization methods to prevent overfitting, and finally achieves overvoltage type classification.

2. Types of Electric Traction Network Overvoltage
For electric traction network overvoltage, it can be divided into internal overvoltage and external overvoltage. This paper mainly studies several types of common internal overvoltage. The internal overvoltage is divided into temporary overvoltage and operational overvoltage according to its characteristics and causes. Its classification is shown in Figure 1.

![Figure 1. Classification of electric traction network overvoltage](image)

As shown in the figure above, the internal overvoltage of traction network consists of 5 common types: ferromagnetic resonance overvoltage, high-frequency resonance overvoltage, VCB (vacuum circuit breaker) overvoltage and neutral-section overvoltage (including into neutral-section overvoltage and out neutral-section overvoltage).

3. Process of Recognition Based on Deep Learning
3.1. Overvoltage Waveform Collection
The traction drive system of standard EMU consists of two independent traction units, which are mainly composed of PT (potential transformer), on-board traction transformer, traction converter and traction motor. In the on-board EMU test, the test device is connected to the low voltage side of the PT in the electrical cabinet of the EMU. A traction unit diagram of the EMU is shown in Figure 2.

![Figure 2. A traction unit diagram of the EMU](image)

The time-domain waveform images collected in this paper recorded a total of four nominal frequency periods (80 milliseconds). The horizontal coordinate is time (in milliseconds) and the vertical coordinate is network voltage (in kilovolts). Five kinds of traction network overvoltage including ferromagnetic
resonance overvoltage, high-frequency resonance overvoltage, VCB overvoltage and neutral-section overvoltage (including into neutral-section overvoltage and out neutral-section overvoltage). Figures 3~7 show several traction network over-voltage time domain waveform images.

Figure 3. Ferromagnetic resonance overvoltage  
Figure 4. High-frequency resonance overvoltage

Figure 5. VCB overvoltage  
Figure 6. Into neutral-section overvoltage

Figure 7. Out neutral-section overvoltage

3.2. Image Preprocessing
Time-domain waveform image data must be preprocessed before it can be trained in the neural network. The main purpose of image preprocessing is to eliminate irrelevant information in the image, restore useful and real information, and enhance the detectability of related information, so as to improve the reliability of feature extraction and recognition. Image preprocessing mainly includes gray transformation, geometry correction, image enhancement and so on. This paper uses OpenCV, an image processing library in Python, to extract images and preprocess them. The preprocessing operations
include geometry correction, grayscale, binarization and image scaling.

3.2.1. Geometry Correction
Geometry correction includes image segmentation and affine transformation, where affine transformation includes translation, rotation, and scaling. First of all, the image is segmented according to the characteristics of the image. The purpose is to remove the irrelevant information such as edge blank area, redundant identification, and get the image containing the traction network overvoltage waveform. Then, the image affine transformation is carried out to ensure that the horizontal and vertical coordinate axes are along.

3.2.2. Grayscale
In the task of classifying time-domain waveform images, the color of the picture is not important, so we transform the color image into a single-channel grayscale image, which is called grayscale. For RGB images, there are three grayscale methods: maximum, mean and weighted average, which are expressed as (1), (2) and (3):

\[ I^{\text{max}}(R, G, B) \]  
\[ I^{\text{mean}} = \frac{1}{3}(R + G + B) \]  
\[ I = 0.59 \times R + 0.11 \times B + 0.30 \times R \]

In the above expressions, I is the brightness value after graying a pixel point, and R, G, B are the brightness values on different channels before graying the point. I, R, G and B range from 0 to 255. This paper directly uses the cvtcolor function in OpenCV library to achieve grayscale.

3.2.3. Binarization
Image binarization is the process of rendering the whole image with a distinct black and white effect, which greatly reduces the amount of data in the image and highlights the outline of the target. The conversion formula for image binarization is:

\[ I_2(x, y) = \begin{cases} 1, & I_1(x, y) \geq T \\ 0, & I_2(x, y) < T \end{cases} \]  

In formula (4), \( I_1(x, y) \) is the brightness value of the pixel point before the conversion, \( I_2(x, y) \) is the brightness value of the pixel point after conversion. The threshold T selected in this paper is 200.

3.2.4. Scaling
Since the size of the two-dimensional tensor of the input layer of the neural network used in this paper is 700*525, the size of the image should be adjusted at the last step of image preprocessing to meet the requirements of the neural network for the size of the input layer.

3.3. CNN Training
Convolution neural network is a feed-forward neural network that consists of input layer, convolutional layer, pooling layer, fully connected layer and output layer. The training of convolution neural network consists of a forward process that computes the loss function through the convolutional layer, the pooling layer and the fully connected layer in turn, and a backward propagation process that updates the network parameters layer by layer using the gradient descent algorithm.

The convolutional layer is composed of convolutional kernels, which act as filters. Convolutional kernel obtains global information by sliding over the input feature map. The mathematical model of the convolution layer is as follows.

\[ X^l_j = f \left( \text{sum} \left( X^{l-1}_{M_j} \odot w^l \right) + b^l \right) \]  

In the formula (5), l represents the number of layers; sum ( ) represents the algebraic sum of all elements in the matrix; X is a two-dimensional tensor; \( \odot \) represents the Hadamard product; w is the weight matrix corresponding to the convolution kernel; j represents the elements in the two-dimensional tensor; \( M_j \) denotes the area associated with the j element in the previous layer; b is the offset term; f is
the activation function, and the rectified linear unit (ReLU) function is chosen in this article. The mathematical equation of ReLU is:

\[ RReLU = \begin{cases} x, & x > 0 \\ 0, & x \leq 0 \end{cases} \]  
(6)

The main role of the pooling layer is to reduce dimensions and parameters. Pooling layer is a form of non-linear down-sampling. Maximum pooling and average pooling are the most common types of pooling. In this paper, maximum pooling, which selects the maximum value in the area as the pooled value, is used.

Finally, after several convolutional and max pooling layers, the high-level reasoning in the neural network is done via fully connected layers. Neurons in a fully connected layer have connections to all activations in the previous layer, as seen in regular (non-convolutional) artificial neural networks. Its mathematical equation is:

\[ y = f (w \cdot x + b) \]  
(7)

In the formula (7), y is the output tensor, x is the input tensor, w is the weight, and f is the activation function.

The convolutional neural network model used in this paper is a VGG neural network. The structure of VGG is very simple. By repeatedly stacking small convolutional kernels of 3*3 and maximum pooling layers of 2*2, VGG successfully constructs a convolution neural network with a depth of 16~19 layers. The structure of the VGG used in this paper is shown in Table 1.

**Table 1. Structure of VGG-16 neural network**

| Module     | Content                                                  |
|------------|----------------------------------------------------------|
| Input      | 2-dimensional tensor                                     |
| Module1    | Convolutional layer1: 64 3*3 convolutional kernel        |
|            | Convolutional layer2: 64 3*3 convolutional kernel        |
|            | Pooling layer1: 2*2 filter                               |
| Module2    | Convolutional layer3: 128 3*3 convolutional kernel       |
|            | Convolutional layer4: 128 3*3 convolutional kernel       |
|            | Pooling layer2: 2*2 filter                               |
| Module3    | Convolutional layer5: 256 3*3 convolutional kernel       |
|            | Convolutional layer6: 256 3*3 convolutional kernel       |
|            | Convolutional layer7: 256 3*3 convolutional kernel       |
|            | Pooling layer3: 2*2 filter                               |
| Module4    | Convolutional layer8: 512 3*3 convolutional kernel       |
|            | Convolutional layer9: 512 3*3 convolutional kernel       |
|            | Convolutional layer10: 512 3*3 convolutional kernel      |
|            | Pooling layer4: 2*2 filter                               |
| Module5    | Convolutional layer11: 512 3*3 convolutional kernel      |
|            | Convolutional layer12: 512 3*3 convolutional kernel      |
|            | Convolutional layer13: 512 3*3 convolutional kernel      |
|            | Pooling layer5: 2*2 filter                               |
| Module6    | Fully connected layer1                                   |
|            | Fully connected layer2                                   |
|            | Fully connected layer3                                   |
| Output     | Classification result                                    |

As shown in the table above, the network contains 13 convolutional layers, 5 pooling layers and 3 fully connected layers. Since weight updating does not happen in pooling layer, 16 layers participate in weight updating, so the network is referred to as VGG-16 network for short. An image of 700 x 525 is input to VGG and the output is the type of overvoltage to which the image belongs.

4. Case Analysis

4.1. Process of experiment

This paper uses VGG-16 convolutional neural network to classify 5 kinds of time-domain waveform
images of traction network overvoltage. The 5 types of overvoltage are: ferromagnetic resonance overvoltage, high-frequency resonance overvoltage, VCB overvoltage, into neutral-section overvoltage and out neutral-section overvoltage. 400 time-domain waveform images are selected for each type of overvoltage, and preprocessed as experimental samples. Then these 2000 samples are divided into training set, validation set and test set according to 7:1:2 ratio, and then the training of VGG-16 neural network is started using TensorFlow 2.0 (GPU) +Keras framework. The process is shown in Figure 8.

4.2. Effects of regularization methods
In statistics, overfitting is "the production of an analysis that corresponds too closely or exactly to a particular set of data, and may therefore fail to fit additional data or predict future observations reliably". In deep learning, if the validation error increases while the training error steadily decreases, then overfitting may have occurred.

In the training process of the neural network, when the model is too complex and the number of iterations is too high, overfitting may happen. Regularization prevents overfitting and reduces generalization errors. In this paper, two regularization methods are used: dropout and stopping training in time.

4.2.1. Dropout
Dropout is a method to optimize neural network with deep structure. In the training process, the regularization of the neural network is achieved by reducing the interdependence of nodes by randomly zeroing some weights or outputs of hidden layers. Dropout is implemented differently according to the different structure of the neural network. The dropout used in this paper is only for the last fully connected layer, that is, the fully connected layer 1 and the fully connected layer 2 in Table 1. The keep rate of dropout in this paper is 0.5, that is, 50% of the weight is kept. As shown in Figure 9, the validation
set error and the training set error change with the number of iterations before and after dropout.

![Figure 9. Error in function of number of iterations](image)

It can be seen from Figure 9 that before using dropout, when the number of iterations is greater than 90, the verification set error increases with the increase of training rounds, and overfitting happens; after using dropout, the increase of verification set error is suppressed and overfitting is eliminated.

### 4.2.2. Stopping Training in time

When it is found that the verification set error increases, it’s important to stop training in time to prevent overfitting. As shown in Figure 9, when the number of iterations $n \geq 100$, the verification set error (using dropout) almost does not change. When $n = 100$, the verification set error (using dropout) is 14%, so the training can be stopped when the $n = 100$.

### 4.3. Confusion Matrix and Sensibility

After the training of neural network, its performance is analyzed on the test set. There are 80 samples for each type of overvoltage in the test set. Table 2 is the confusion matrix of the VGG classification model, and table 3 is the recognition sensitivity of each overvoltage type.

**Table 2. Confusion matrix of VGG classification model.**

| predict | 1 | 2 | 3 | 4 | 5 |
|---------|---|---|---|---|---|
| 1       | 65| 11| 2 | 0 | 2 |
| 2       | 8 | 66| 1 | 3 | 2 |
| 3       | 1 | 0 | 72| 2 | 5 |
| 4       | 2 | 1 | 1 | 73| 3 |
| 5       | 1 | 4 | 2 | 3 | 70 |
Table 3. Recognition sensibility of classification model

| Type                        | Sensibility (%) |
|-----------------------------|-----------------|
| Into neutral-section overvoltage | 81.25           |
| Out neutral-section overvoltage | 82.5            |
| Ferromagnetic resonance overvoltage | 90              |
| VCB overvoltage | 91.25           |
| High-frequency resonance overvoltage | 87.5            |

It can be seen from table 2 that for ferromagnetic resonance overvoltage, VCB overvoltage and high-frequency resonance overvoltage, the sensitivity of the VGG-16 classification model is above 87%. For the in and out neutral-section overvoltage, the sensitivities are 81.25% and 82.5% respectively, and it can also be observed from table 3 that these two kinds of overvoltage are easy to be confused.

5. Conclusion

This paper presents a recognition method for the time domain waveform image of traction network overvoltage based on deep learning. Firstly, through the on-board EMU test, five kinds of traction network overvoltage time-domain waveform images are collected and sorted out; then the waveform images are preprocessed to get the sample set; then the VGG-16 neural network is trained to obtain the classification model. The results show that:

Choosing appropriate regularization methods, such as dropout and stopping training in time, can avoid overfitting of neural network and improve the generalization ability of neural network.

The method proposed in this paper is for time-domain waveform image. The time-domain waveform image can be exported from a variety of file formats. Therefore, the deep learning classification method for time-domain waveform image can not only save the steps of format transformation between different files, but also simplify the process of feature extraction. With the development of image processing technology and the improvement of computer operation performance, time-domain waveform image recognition method based on deep learning has practical significance.

References

[1] Simonyan, K., Zisserman, A. (2015) Very Deep Convolutional Networks for Large-Scale Image Recognition[A]. International Conference on Learning Representations, 10: 18-30.
[2] Liu, C., He, J., Zhang, C., Cao, M. (2019) Pattern Recognition Method for Time-domain Waveform Images of GIS Partial Discharge[J]. Proceedings of the CSU-EPSA, 31(10): 24-30.
[3] Wang, C., Lu, X. (2020) A Motor Bearing Fault Diagnosis Based on Deep Learning Method[J]. Journal of Lanzhou Jiaotong University, 39(02): 42-50.
[4] Zhang, Q., Wang, J., Li, W. (2019) Insulator State Detection of Convolutional Neural Networks Based on Feedback Mechanism[J]. Transactions of China Electrotechnical Society, 34(16):3311-3321.
[5] Deng, R., Liu, B., Song, W. (2019) Low-Frequency-Oscillation Analysis and Suppression of the Coupling System between Traction Network and Multi-Locomotives[J]. Transactions of China Electrotechnical Society, 34(S1):327-335.
[6] Ma, T., Xu, Y. (2015) Influence of High-voltage Transformer on Electric Locomotive Passing the Neutral Section[J]. Electric Drive for Locomotives, 2015(02):20-25.