Core looseness fault identification model based on Mel spectrogram-CNN

Ping He¹, Yong Li¹, Shoulong Chen¹, Hoghua Xu¹, Lei Zhu¹ and Lingyan Wang¹*

¹State Grid Jiangsu Electric Power Co., Ltd. Nanjing power supply branch, Nanjing, Jiangsu Province, China
* Lingyan Wang’s e-mail: 549885860@qq.com

Abstract. In order to realize transformer voiceprint recognition, a transformer voiceprint recognition model based on Mel spectrum convolution neural network is proposed. Firstly, the transformer core looseness fault is simulated by setting different preloads, and the sound signals under different preloads are collected; Secondly, the sound signal is converted into a spectrogram that can be trained by convolutional neural network, and then the dimension is reduced by Mel filter bank to draw Mel spectrogram, which can generate spectrogram data sets under different preloads in batch; Finally, the data set is introduced into convolutional neural network for training, and the transformer voiceprint fault recognition model is obtained. The results show that the training accuracy of the proposed Mel spectrum convolution neural network transformer identification model is 99.91%, which can well identify the core loosening faults.

1. Introduction

With the development of economy and the increasing demand of society for power consumption, as one of the most important power equipment, the number of power transformers in operation is increasing, and the resulting noise also affects the production and life of surrounding residents. Domestic and foreign scholars have conducted in-depth research on the problem of transformer noise, and reduced the noise of transformer by optimizing the structure of transformer, such as changing the material of silicon steel sheet. The noise of transformer contains a lot of information that can reflect the operation state of transformer. When the operation state of transformer changes, the noise will change accordingly. Therefore, the changing features in the transformer sound signal can be extracted, and the transformer operation state can be identified through data mining technology.

The noise signal of transformer contains rich information, so many scholars extract the characteristic parameters that can reflect the operation state of transformer through various data mining technologies, so as to realize the monitoring of transformer fault state. Literature [1] proposes a new scheme to select the optimal gas in each fault diagnosis layer by chi square test, eliminate redundant gas, and classify the selected optimal gas by different machine learning classifiers. The results show that chi square test can effectively extract the characteristic gas, and the effect of hierarchical fault diagnosis with different classifiers is better than that with a single classifier. In literature [3], wavelet packet transform (WPT) is used to extract the characteristics of characteristic gas data in transformer oil and construct the feature vector; Then, Mahalanobis distance (MD) is used to represent the similarity between each vector, and the graph structure is constructed with data features as vertices and similarity as edges; Finally, the graph convolution neural network is used to classify the transformer fault category and accurately identify the transformer fault category. Using various data mining techniques to extract the characteristic quantity
of transformer operation state has been proved to have a good effect, but the training effect of CNN under picture data is better than other machine learning methods.

Firstly, this paper preprocesses the sound signals in various states, draws the idiom spectrum of the original signal, reduces the dimension of the spectrum through Mel filter banks, constructs the Mel spectrum, and makes the data sets in different states. Finally, the data set is imported into the convolutional neural network (CNN) suitable for transformer core preload, and the Mel spectrum CNN transformer core preload voiceprint recognition model is constructed.

2. Sound signal preprocess

2.1. Spectrogram drawing

Spectrogram is an important characteristic frequency spectrogram of sound signal analysis and processing. It can reflect the frequency energy distribution of signal at different times. The relationship between time domain and frequency domain of sound signal is established completely to maximize the sound feature information, which is helpful to the extraction and learning of voiceprint features in the later stage.

The rendering of voiceprint spectrogram includes framing, windowing and discrete Fourier transform. As shown in the figure, a segment of transformer 2s noise is intercepted. Firstly, the intercepted segment is processed by frame. If the frame length is too long, the accuracy of feature quantity will be affected. If the frame length is too short, useful feature quantity will not be extracted. Since the transformer noise is more stable than human voice, the frame length can be appropriately increased to ensure the integrity of voice signal characteristics. In this paper, n = 4096 per frame is taken as 64ms. In order to make the smooth transition between frames, the overlap rate is taken as 50%. Secondly, if the discrete Fourier transform is directly performed on the data after framing, there will be spectrum leakage. Therefore, each frame needs to be windowed first. In this paper, Hamming window is selected to make both ends of the signal smooth and reduce the distortion of the signal.

$$\sigma(n) = \begin{cases} 0.54-0.46\cos\left(\frac{2\pi(m/(N-1))}{N}\right), & 0 \leq n \leq N-1 \\ 0, & \text{else} \end{cases}$$

(1)

Finally, the discrete Fourier transform is performed on each frame data after windowing. The formula is to obtain the time-frequency matrix and draw the spectrogram.

$$X(k) = \sum_{n=0}^{N-1} x(n)e^{-j\frac{2\pi kn}{N}}, (k = 0,1,2,...N-1)$$

(2)

The abscissa of the spectrogram represents the number of frames (time) after framing, and the ordinate represents the frequency. The color represents the size at this frequency at this time, which is the power spectral density. The spectrogram generated by short-time Fourier transform is shown in the figure:
2.2. Mel spectrogram

The spectrogram obtained through the above preprocessed can describe the relationship between time domain and frequency domain during transformer operation. Although the dimension is compressed to 62 columns in time, there are 2049 lines in frequency dimension, which has a great impact on the training and recognition speed of subsequent convolutional neural network. Therefore, the frequency dimension must be compressed. Therefore, Mel filter is considered to reduce the dimension of time-frequency matrix.

Mel filter makes nonlinear processing for linear frequency band according to the structural characteristics of human ear, increases the weight of low-frequency part and decreases the weight of high-frequency part, highlighting the role of effective frequency band. The conversion relationship between Mel frequency and actual frequency is as follows:

$$f_{mel}(f) = 2595 \cdot \log \left(1 + \frac{f}{700\text{Hz}}\right)$$

(3)

Where, $f_{mel}(f)$ is the frequency under mel scale, and $f$ is the actual frequency.

Mel filter is a filter bank composed of M triangular filters ($M = 40$). The center frequency is $f(m)$. On the Mel frequency scale, the distance between filters is equal. The transfer function of the filter bank is:

$$H_m(k) = \begin{cases} 0, k < f(m-1) \\ \frac{k - f(m-1)}{f(m) - f(m-1)}, f(m-1) \leq k \leq f(m) \\ \frac{f(m+1) - k}{f(m+1) - f(m)}, f(m) < k \leq f(m+1) \\ 0, k > f(m+1) \end{cases}$$

(4)

Where, is defined as:

$$f(m) = \frac{N}{f_s} \left( F^{-1}_{mel} \left( f_h \right) + m \frac{F_{mel}(f_i) - F_{mel}(f_h)}{M+1} \right)$$

(5)

Where, and $f(h)$, $f(l)$ is the upper and lower limits of filter frequency, $f_s$ is the sampling frequency of transformer voiceprint sampling ($f_s = 64\text{kHz}$), and $N$ is the frame length when short-time Fourier change is performed.
3. Model recognition based on convolutional neural network

3.1. Convolutional neural network

Convolutional neural network has outstanding ability in image recognition. It is an artificial neural network which is good at processing plane two-dimensional data. It is composed of multiple groups of convolution layer, pooling layer and full connection layer. It is usually used to classify the pixel values in the image. Compared with the traditional depth learning method, convolutional neural network avoids the complex process of feature extraction and data reconstruction. The network can extract image features by itself, including colour, texture, shape and image topology. In dealing with two-dimensional images, especially displacement recognition. It has good robustness and computational efficiency in the application of scaling and other forms of distortion invariance. Its basic structure is shown in the figure.

\[
y_{xy} = f \left( \sum_i \sum_j \sigma_y v_{(x+i)(y+j)} + b \right)
\]

Where, \( y_{xy} \) represents the eigenvalue \((x, y)\) at the position of the original graph, \( \sigma_y \) represents the value of the convolution kernel, \( v_{(x+i)(y+j)} \) represents the input at the position, and \((x+i, y+i)\) is the deviation value:
3.2. Dataset production

Take the Mel spectrogram drawn above as the input of CNN, and the pixels of the input image are \([62 \times 40 \times 3]\) Data (width \(\times\) high \(\times\) Depth). Among them, the width 62 represents the time component after framing of the sound signal, the height 40 represents the Mel frequency component, and the depth 3 represents the colour channel. Since the generated spectrogram is colour, there are three RGB channels. The Mel spectrograms generated in each state of the transformer are input into CNN network. The network trains, learns and extracts the features of the Mel spectrograms generated in each state, and finally learns the spectrograms that can identify various states of the transformer, so as to complete the transformer core state recognition based on convolutional neural network.

When simulating the influence of iron core looseness on transformer noise signal, set the pre tightening force of iron core on the test platform, which is realized by changing the bolt pre tightening force to 0.6 \(F_N\), 0.8 \(F_N\), \(F_N\) and 1.2 \(F_N\) respectively (\(F_N\) is the rated pre tightening force of transformer iron core).

Refer to GB / T 1094.10-2003 to arrange the measuring points of transformer noise, as shown in the figure. The sampling frequency is 64kHz, and the sampling time of each group of data is 10s. In addition, the test site is open and conducted late at night to reduce the impact of external noise on the experimental data.

Firstly, the sound signals under different preloads are transformed into Mel spectrograms in batches, and these pictures are made into training sets and verification sets in 8:2. The convolution neural network is used to identify the operation status of the transformer under different preloads. In order to ensure the effectiveness of learning, after corresponding the data with the label, the data is disrupted and 80% is randomly selected as the training set and input into the network, and the remaining 20% is used to verify the effectiveness of the network.

3.3. Network structure

For the CNN network used in the recognition model, three convolution layers and three pooling layers are overlapped, two full connection layers are used, and the activation function relu is used to improve the training speed of the network.

| Layer(type)   | Output shape          | Param |
|--------------|-----------------------|-------|
| Rescaling_1  | (None,400,600,3)      | 0     |
| Conv2D       | (None,400,600,16)     | 448   |
| Max_pooling2D| (None,200,300,16)     | 0     |
| Conv2D_1     | (None,200,300,32)     | 4640  |
| Max_pooling2D_1 | (None,100,150,32) | 0     |
| Conv2D_2     | (None,100,150,64)     | 18496 |
| Max_pooling2D_2 | (None,50,75,64)  | 0     |
| Flatten      | (None,240000)         | 0     |
| Dense        | (None,128)            | 30720128 |
| Dense_1      | (None,4)              | 516   |

3.4. Training results

It can be seen from the training result chart that it shows good results in both the training set and the verification set. After 60 iterations, the accuracy has reached 99.91% and the loss is close to 0. Therefore, the training effect of the network is very good.
4. Conclusion

By simulating the loose fault of transformer iron core, set the pre tightening force of transformer iron core as respectively, collect the sound signals under different pre tightening force of iron core, make them into Mel spectrogram data set, and send them to convolution neural network for training. The training results show that the accuracy of training set and verification set has reached 99.91%, and the training performance is very good.

Acknowledgments

Key science and technology projects of Jiangsu Electric Power Co., Ltd(J2021053)

References

[1] Zi Zi Jian, Qin Yurui, Li Jingli. Transformer hierarchical fault diagnosis using machine learning [J/OL]. Journal of power system and automation: 1-6 [2021-10-23] https://doi.org/10.19635/j.cnki.csu-epsa.000849.
[2] Wan Zhouli, Liu Hui. Research on Transformer Fault Diagnosis Based on improved GWO optimized DBN network [J]. Modern electronic technology, 2021,44 (19): 163-168
[3] Fan Xiaodan, Fu Weiping, Zhao Zhilong, Zu Shutao, Zhang lishuo, Hu Weitao. Research on fault diagnosis of oil immersed transformer based on long-term and short-term memory network [J]. Transformer, 2021,58 (09): 27-32
[4] Liu Hua, Liu Jiangyong. Power transformer fault diagnosis based on graph convolution neural network [J]. Journal of Hunan University of science and Technology (NATURAL SCIENCE EDITION), 2021,36 (03): 75-81
[5] Zhang Jiusi, Ma Hongzhong, Li Yong, Xu Honghua, Zhu Hao. Analysis and diagnosis of transformer winding looseness fault based on VMD [J]. High voltage apparatus, 2021,57 (08): 198-208
[6] Zhang Zhiheng. Research on Transformer Fault Diagnosis Based on probabilistic neural network optimized by intelligent algorithm [D]. Xi'an University of technology, 2021
[7] Ye Jianhua, Yang Li. Transformer fault diagnosis based on harmonic search optimization support vector machine [J]. Transformer, 2021,58 (06): 33-37
[8] Wang Kai. Research on transformer fault prediction method based on Ensemble Learning [D]. Shandong University, 2021
[9] Zhang Shimin, Feng Yao. Transformer core and winding fault discrimination method based on vibration signal [J]. Electrical measurement and instrumentation, 2021, 58 (06): 161-166
[10] Guo Jian. Research on transformer condition evaluation and fault diagnosis model [D]. Shandong University, 2021