

A Semantic Segmentation Method for Buffer Layer Defect Detection in High Voltage Cable

Zhang Jun^{1*,a}, Duan Xiaoli¹, Xie Yi¹, Duan Jianjia¹, Huang Fuyong¹, Zeng Zeyu¹

¹ State Grid Hunan Electric Power Company Limited Research Institute, Changsha 410007, China

Abstract. A semantic segmentation method based on the fully convolutional network is proposed to detect the buffer layer defect in high voltage cable automatically. One hundred seventy-seven high-resolution X-ray images of cables are collected. FCN-8s and VGG16 backbone are adopted. The results indicated that the FCN-8s achieves the mIoU to 0.67 on the test set, proving to be an efficient way to detect the buffer layer defects.

1 Introduction

High voltage XLPE cable has become one of the core power equipment for large-capacity power transmission in cities due to its advantages such as excellent electrical performance, large transmission capacity, simple manufacturing process, easy installation and maintenance. The rapid growth of high voltage XLPE cable has posed more challenges to cable lines' operation and maintenance. In China, body failure caused by cable buffer layer ablation frequently occurs in recent years [1-2]. Many researchers have studied the buffer layer gap characteristics [3], the structure and performance of the buffer layer [4] and related physical and chemical analysis of the buffer layer ablation phenomenon [5]. However, there is little research on defect detection, and no mature detection technology has been formed.

X-ray imaging detection method is a potential practical technology that can form relatively intuitive defect images. The defect size of buffer layer ablation can be obtained relatively quantitatively at some angles. However, the original X-ray image needs for image processing. Meanwhile, with the growing number of cable acquisitions, manual defect-recognition will inevitably be time-consuming and low efficiency. Therefore, automatic detection for the buffer layer defect in high voltage cables is an up-and-coming technology.

In recent years, the deep learning technique represented by the fully convolutional networks (FCN) [6] is continuously surpassing the previous image semantic segmentation techniques in the pixel level classification task. It is gradually applied in many professions.

In this paper, on-site X-ray image data set collected from a cable tunnel is established. Then, the buffer layer defects are identified by using the technology based on FCN, thus providing an effective way to liberate the workforce.

2 FCN model

The classic convolution neural network (CNN) [7] comprises the input layer, convolution layers and pooling layers, and a fully connected multilayer perceptron classifier. Convolution and pooling operations can greatly simplify model complexity and reduce model parameters. The basic CNN is a feature extractor with the advantage that human engineers do not need to design multiple layers of features. Compared to the standard feed-for-ward neural networks, CNN has better learning and adaptive ability due to its unique designs.

The main function of the convolution layer is to convolve the convolution kernel with the input data of the upper layer, which the convolution kernel in the convolutional layer is applied to feature extraction. It can also reduce the connection between different layers to prevent overfitting and too many parameters. The next layer output can be expressed as follows:

$$y(i, j) = \sum_{m=1}^M \sum_{n=1}^N x(i+m-1, j+n-1)k(m, n) \quad (1)$$

$i = 1, 2, \dots, p, j = 1, 2, \dots, q$

where, $y(i, j)$ is the element of the output matrix y with order $p \times q$; $x(i+m-1, j+n-1)$ is the element of the input matrix x with order $p \times q$; $k(m, n)$ is the element of the convolution kernel k with order $M \times N$.

The pooling layer divides the input data into multiple non-overlapping regions, and take the maximum value (maximum pooling) or average value (average pooling) of each region. It can eliminate non-critical feature samples, thus improving the training efficiency and estimation accuracy, and the pooling formula is:

$$y' = \max_R(y) \text{ or } \text{mean}_R(y) \quad (2)$$

where, R is the pooling region.

^a Corresponding author: sghnzj1988@163.com

The fully connected layer expands the two-dimensional data passing through the convolution layer and pooling layer into one-dimensional data. It can be expressed as follows:

$$o = f(\mathbf{wc}^T + b) \quad (3)$$

where, w is the connection weight vector; c is the expanded one-dimensional data; b is the bias; o is the output; f is an activation function which can enhance the network' nonlinear characteristics. For example, rectified linear unit (ReLU) is a popular activation function which can activate the neurons of the neural network sparsely, and can be expressed as follows:

$$\text{ReLU}(y(i, j)) = \begin{cases} y(i, j), & y(i, j) > 0 \\ 0, & y(i, j) \leq 0 \end{cases} \quad (4)$$

VGG network [8] is a convolution neural network model developed by the Computer Vision Group of Oxford University and Google DeepMind researchers in 2014. VGG network structure is simple and has good generalization performance. The typical network VGG16 is shown in Figure 1.

Different from the CNN in convolution with a full connection layer to get a fixed-length feature vector for classification, FCN can accept any size of the input image. The deconvolution layers are applied to upsampling the feature map of the final convolution layer and make the feature map back to the same size of the input image. Thus, the network can have a prediction on each pixel, while retaining the original input space information of images. Figure 2 is the architecture of FCN adopted in this paper. Its main ideas include:

1. Adopting the end-to-end structure, which makes the network easy to train;
2. Cancel the full connection layer;
3. When the feature map of the image is down-sampling to a certain extent, reverse up-sampling is performed to match the semantic segmentation annotation map of the image;
4. Up-sampling will lose some information. Therefore, FCN takes into account the response of the shallow layer in the network to better predict the details in the image. As shown in Figure 3, Pool4 and Pool3 responses are considered as outputs of model FCN-16s and FCN-8s, respectively, and combined with the original output of FCN-32s for the final semantic segmentation prediction.

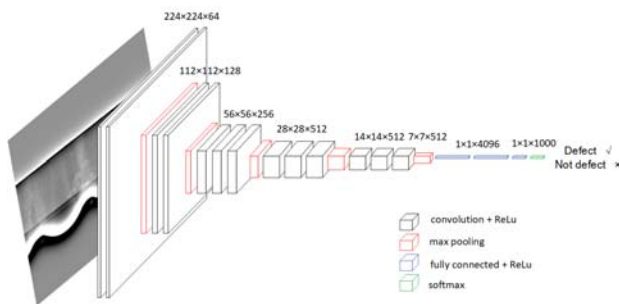


Figure 1. VGG16 network

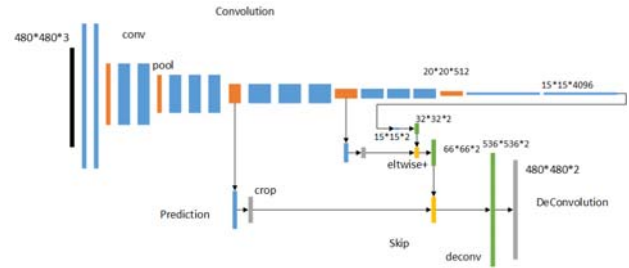


Figure 2. Architecture of FCN-8s, in which the backbone is VGG16.

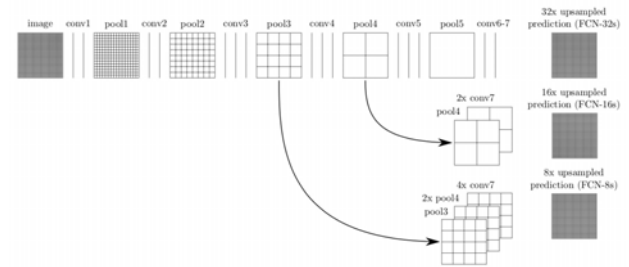


Figure 3. Skip Layer of FCN^[6]

3 The semantic segmentation method flow chart

The flowchart of our semantic segmentation method is shown in Figure 4 and demonstrated as follows:

- Step 1. Image acquisition. Use a CCD or flat panel detector to obtain an X-ray digital grayscale image.
- Step 2. Image preprocessing. Use contrast adjustment, edge sharpening, window width and window level adjustment for grayscale images to form an image suitable for human eyes to judge defects. Then, the image is classified and labelled with or without flaws, and the image with defects can be processed with image enhancement.
- Step 3. Model training. The FCN model is used for training, and the hyperparameters of the model are set to obtain the parameters of the trained model.
- Step 4. Model evaluation. Use mIoU to evaluate model performance.
- Step 5. Model saving. Save the trained model.
- Step 6. Model testing. Test the new data set using the trained model.

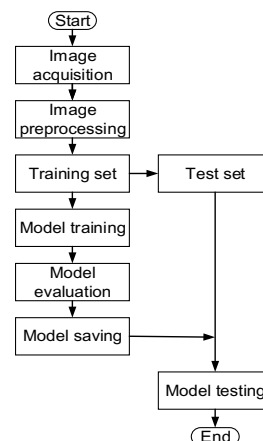


Figure 4. Flowchart of our semantic segmentation method.

4 Example

We collected 177 original X-ray images in this example from a cable tunnel. For the original high-resolution image, the sliding window method is used to cut the image into local images, and the local images with defects are labelled at pixel-level. In this example, the original image resolution is 2176×1972 and scaled to 896*896, and the local image resolution is 448×448. The stride of the sliding window is 224 and each local image marked with defect pixels is rotated by 180 degrees to enhance the generalization capability. Finally, we get 196 local images with defects, in which 155 local defect images are used for training, and the rest of the local images are used for testing. The FCN-8s and the backbone VGG16 are adopted. The deep learning library TensorFlow is used for model training.

The cross-entropy loss function is used as follows:

$$\text{Loss} = -y \log y' - (1 - y) \log(1 - y') \quad (5)$$

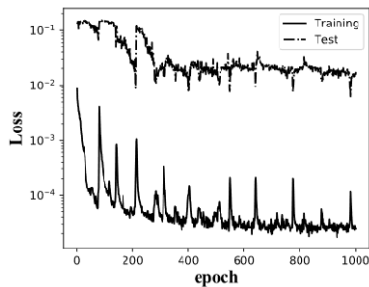
where, y denotes the positive and negative label; y' denotes the output probability.

The evaluation index of the model is:

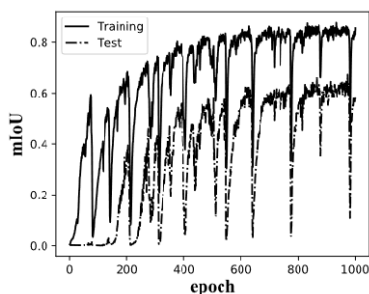
$$\text{mIoU} = \frac{1}{M} \sum_{m=1}^M \frac{\text{TP}}{\text{TP} + \text{FP} + \text{FN}} \quad (6)$$

where, M denotes the patches number, which is computable. TP denotes the true positive predictions; FP means false positive predictions; FN denotes the false-negative predictions.

The model training process is as follows:



(a) Loss



(b) mIoU

Figure 5. The training process of FCN-8s

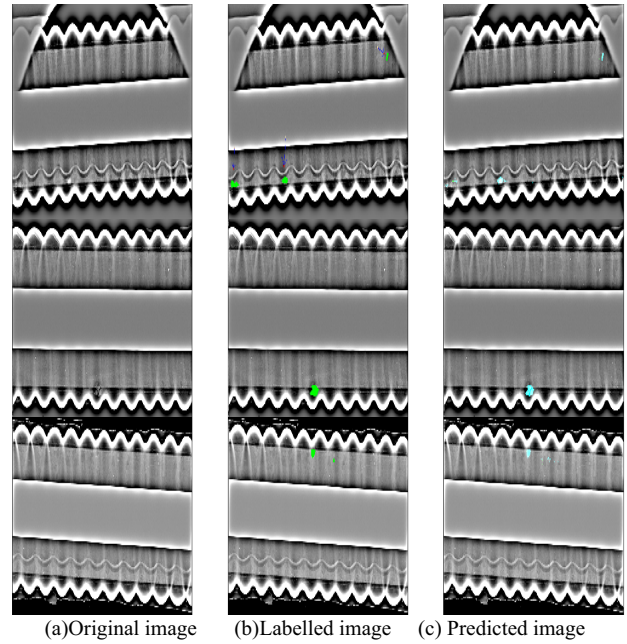


Figure 6. FCN-8s for X-ray image semantic segmentation of high voltage XLPE cable

For the FCN-8s, the initial learning rates are set as 0.001, and the SGD updater is used. The weight decay to use for regularizing the model is set 10e-4 to prevent overfitting. The FCN-8s is trained for 1000 epochs, and the batch size is set as 10.

As can be seen from Figure 5, the Loss and mIoU converge after about 600 epochs on both the training and test sets except the test Loss, which converges after about 250 epochs. However, there are some unstable jump points in the training process. The mIoU finally achieve 0.67 on the test set. Meanwhile, Figure 6 shows some semantic segmentation results on the test set, which indicates that the FCN-8s is an efficient way to detect the buffer layer defects.

5 Conclusion

This study adopts a semantic segmentation method for high voltage XLPE cable buffer layer defect detection. The FCN is developed from classic CNN and can achieve pixel-level prediction. The details through the pixel-level information, such as the area of buffer layer defect, can be automatically acquired. The on-site cable X-ray images are used to verify the FCN-8s, which is finally proved to be useful. In the future, higher performance models and more data should be used to improve the defect detection effectiveness.

Acknowledgments

This work is supported by the State Grid Hunan Electric Power Company Limited Research Institute science project (No. B316A5200017). The authors would like to thank the Cao Xianhui and Liu Weike for providing the verification data.

References

1. Li Lin, Yan Bo, Zhou Zhangbin, et al. Analysis of the fault cause of 110kV cable body. *Electric World*, 60(05):11-15. (2019)
2. Li Chenying, Li Hongze, Chen Jie, et al. Analysis of High Voltage XLPE Power Cable Buffer Layer Discharge Problem. *Electric Power Engineering Technology*, 37(02):61-66. (2016)
3. Yang Fan, Zhu Ningxi, Liu Xiaodong, et al. Experiment and Analysis of Evaluating Buffer Layer Gap Condition of Cable Based on Impedance. *Guangdong Electric Power*, 31(12):93-98. (2012)
4. Deng Shenghua, Jiang Fuzhang, Liu Heping, et al. Research on Materials and Structural Characteristics of Buffer Layers for High Voltage Cable. *Electric Wire & Cable*, 02:19-27. (2019)
5. Zhang Weiwei, YAN Youxiang, WANG Lei, et al. Physical and Chemical Analysis and Simulation Study on Discharge Burn Defect of Buffer Layer in 110 kV XLPE Cable. *High Voltage Apparatus*, 56(05):155-162. (2020)
6. Long, J., Shelhamer, E., Darrell, T. Fully convolutional networks for semantic segmentation. *Computer vision and pattern recognition*. 3431-3440. (2015)
7. Lecun, Y., Bottou, L. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11), 2278-2324. (1998)
8. Simonyan, K. and Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. *Computer Science*, arXiv:1409.1556. (2014)