Detection Method of Conductor Strand Defects Based on Multi-modal Data Fusion

Jiehui Wu[1], Jianrong Zhang[1], Liang Fan[2], Lei Zhang[2*], Huafeng Su[1], Jinduo Zhou[1], Guanke Liu[1], Zhongyu Li[1], Jianzhong Li[1], Zhibin He[1]

[1]Dongguan power supply bureau, Guangdong Power Grid Corporation, Dongguan, China
[2]department of Data, Guangzhou zhongke Intelligent Inspection technology Co., Ltd, Guangzhou, China

*zhanglei@igiai.com

Abstract. Combined with the advantages of infrared image and visible image, this paper proposes a multi-modal data fusion method for detecting the wire strand defects. The segmentation performance of UNet model in images with poor illumination conditions is improved; it solves the problem that the contrast between the wire and the background in the infrared image is low and it is difficult to distinguish, and reduces the false alarm rate of the loose detection. The experimental results show that this method can be used in UAV patrol images under different lighting conditions, which is more conducive to embedded in UAV for all-weather intelligent patrol, and has the advantages of high recall rate and low false detection rate, compared with the reference method using only infrared or visible light images.

Keywords: UAV object detection; UNet image segmentation; multi-modal data fusion; conductor strand defects detection.

1 Introduction

The wire is the core equipment in the whole power grid line, responsible for conveying the electricity in the power grid, and its safe operation is related to the stability of the whole power system. The defect of wire scattered stock is easy to cause a sharp increase of power line power, resulting in line short circuit or wire broken stock phenomenon. Therefore, in the early stage to accurately detect the wire dispersion defects and solve them, can effectively prevent further damage to the wire, so as to ensure the safety and stability of the power line transmission. After the rise of UAV inspection technology in recent years, UAV inspection is widely used in power line operation and maintenance with its advantages of high efficiency and reliability, low cost and not affected by terrain factors. At present, UAV image acquisition technology is basically used for defect detection, to collect data from the key positions of power lines, and manual reading or image processing algorithm is used for detection in the later stage.

H. Wu et al. (Eds.): CSIEDE 2022, ACSR 103, pp. 755–766, 2022.
https://doi.org/10.2991/978-94-6463-108-1_84

With the rapid development of deep learning technology, more and more deep learning-based object detection and object tracking algorithms have emerged. Deep convolutional neural networks have powerful feature extraction and model fitting capabilities, and can use the target data sets to autonomously learn the target to be detected, and improve their own models. Current widely used target detection algorithms based on deep learning can be roughly divided into two categories: (1) Two-stage target detection algorithm, Such as Fast R-CNN 1 (Region-Conventional Neural Network), Faster R-CNN 2, Mask R-CNN 3, etc. These algorithms all divide the target detection into 2 stages. First, the regional candidate network 2 (RPN) is used to extract the information of the candidate targets, Using the detection network to complete the prediction and identification of the location and category of the candidate targets; (2) Single-step target detection algorithm, Such as SSD 4 (Single Shot Multibox Detector), YOLO 5 (You Only Look Once), YOLO 9000 6, YOLO V4 7, etc. Such algorithms do not need to use the RPN. Instead, going directly through the network to generate location and category information about the target, It is an end-to-end object detection algorithm. Therefore, the single-step object detection algorithm has a faster detection speed.

At present, the automatic conductor strand defects detection technology can be roughly divided into two categories. One uses the computer vision algorithm to process the visible image and directly detects; the other uses the resistance heating characteristics of the area to judge the temperature map of the infrared image to predict the situation of the wire. In the above methods, the image detection algorithm used in the former is difficult to detect the slight dispersion of the wire, and the recall rate of the algorithm is not good; while in the latter, due to the complex background of infrared image, low contrast and high noise, it is easy to judge a large number of misjudgment.

Infrared images and visible light images using different band acquisition, two kinds of images show low correlation in terms of characteristics, has different advantages of detecting wire strands defects. Based on that, this paper proposes a combination of infrared and visible light image multi-modal data fusion detection method, in order to achieve high precision and high recall of wire defect detection.

2 Related work

2.1 UNet model

The UNet model 8 consists of an encoding network and a decoding network, and the model structure is shown in the following figure.1 and figure.2. UNet first performance a convolution and gradually pooling pictures, and then pool the features of the feature sampling, each sampling and the corresponding size before pooling features on the channel, and then after the splicing feature sampling, continue to splicing and convolution, and finally get the same prediction as the input image size. The entire network was subsampled and upsampled. Each subsampling involves halving the size of the feature map, doubles the number of channels, and the feature map becomes smaller and compact, more three-dimensional, each upsampling doubles the size of

the feature map, halving the number of channels, and the feature map finally returns to the size of the input.

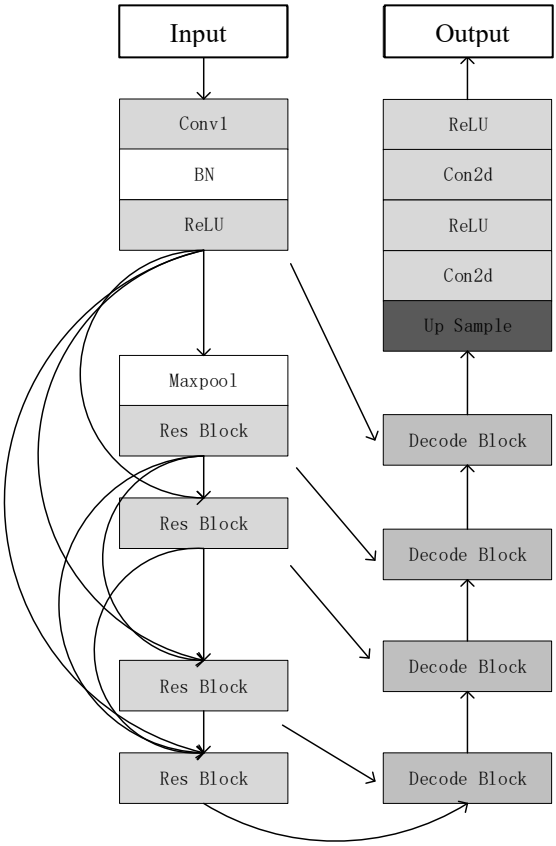

Fig. 1. more detail about UNet structure

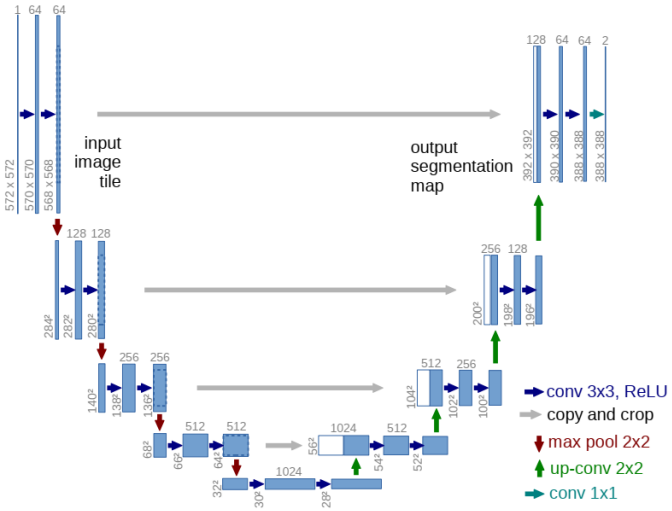

Fig. 2. UNet global structure

2.2 The ResNet50 encoding network

The Resnet50 network contains 49 convolutional layers, a fully connected layer. As shown in the figure below, the 9 structure of the Resnet50 network can be divided into seven parts. The first part does not contain the residual blocks, and mainly calculates the convolution, regularization, activation function, and maximum pooling of the inputs. The second, third, fourth, and five parts all contain the residual blocks. The block in the figure does not change the size of the residual block, and is only used to change the dimension of the residual block. In the Resnet50 network structure, the residual blocks have three layers of convolution, and that network has a total of 1 + 3× (3 + 4 + 6 + 3) =49 convolutional layers, plus the final full connection layer is a total of 50 layers, which is also the origin of the Resnet50 name. The input of the network is 224×224×3. After the convolution calculation of the first five parts, the output is 7×7×2048. The pooling layer will convert it into a feature vector, and finally the classifier will calculate the feature vector and output the category probability.

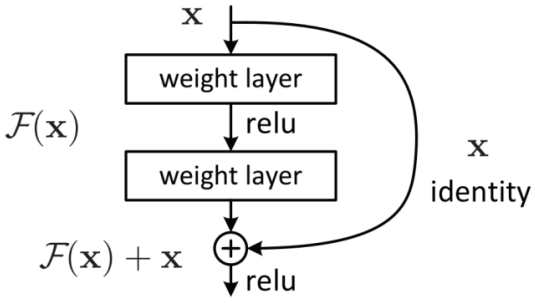

Fig. 3. Residual Network unit module

2.3 Data fusion

Data fusion 10 is a technology that yields more consistent, informative, and accurate information from uncertain, imprecise, inconsistent, conflicting, and similar raw data. Current data fusion has been applied in many fields, including wireless sensor networks, radar systems, object tracking, object detection and identification, intrusion detection, etc. Data fusion belongs to a kind of attribute fusion, which is the intelligent synthesis of multi-source image data in the same region to produce more accurate, complete and reliable estimation and judgment than a single information source.

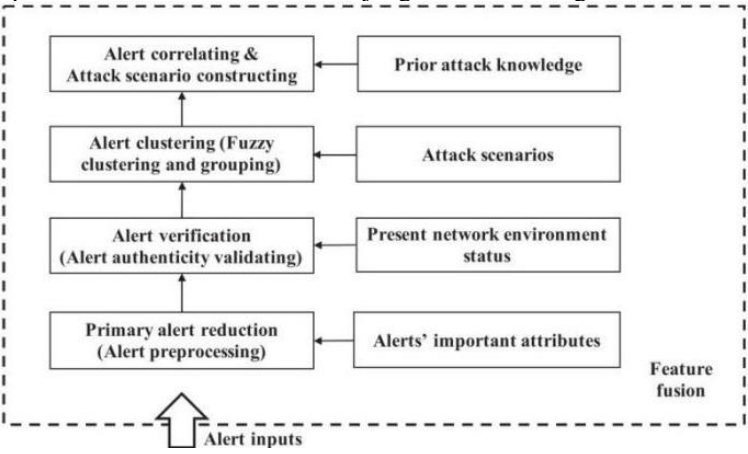

Fig. 4. Data fusion technique

3 Materials and Methods

3.1 The segmentation model based on the UNet

Traditional image processing method requires specific background conditions or light conditions, with poor real-time monitoring and weak generalization ability. The features extracted by shallow machine learning algorithms are not adequately discriminative and representative for complex situations. The features extracted by the deep learning method are highly representative, and the model has high recognition accuracy and generalization ability, so this paper uses the UNet model based on the deep convolutional neural network to segment the wire and the background in the images.

1) Build an encoding network.

This paper adopts ResNet50 as a UNet coding network, because the wire semantics is relatively simple, the edge detail information is difficult to capture, so the image of deep semantic information and shallow feature details are very important. So based on the ResNet50 added six cross-layer connection, through the maximum pooling operation unified dimension, we form a dense network structure. This structure can reuse shallow features, enhance the depth of feature extension. The coding network includes

four residual modules (Res Block), the residual module structure is shown in the figure.5 below, each module contains two convolution layers, Rectified Linear Unit, ReLU) and two batch normalized layers (Batch Normalization, BN). The convolution kernel is 3 * 3, and the jump connection can prevent the gradient from disappearing during backpropagation.

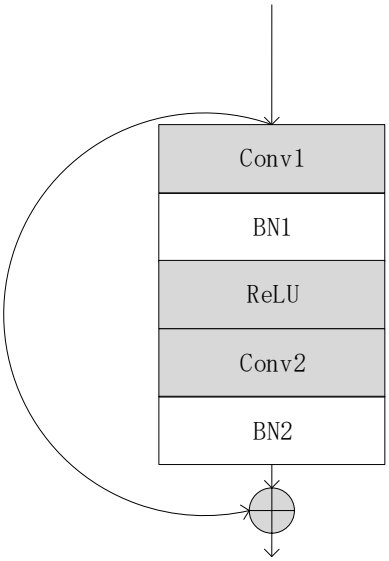

Fig. 5. Encoding network structure

2) Build a decoding network.

UNet decoding network includes 5 decoding modules, decoding module structure is shown in the following figure, each module includes upsampling, two convolution layer Conv2d, two activation function layer ReLU, upsampling method uses the nearest neighbor interpolation method, used to double the feature graph size, convolution kernel is 3 * 3 size, step size is 1, for fusion features. There are 4 jump connections between the coding and decoding network to splicing the same-dimensional features, realizing the multiplexing of shallow features.

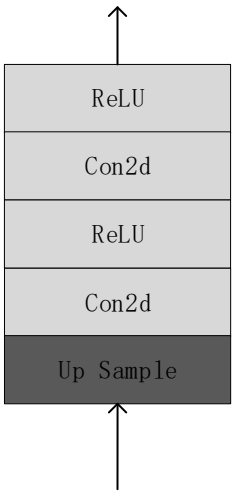

Fig. 6. Decoding network structure

3) Define the loss function.

In the power line image, the number of background negative samples is much larger than the number of wire positive samples, so this paper adopts Dice function 11 as a loss function. Dice loss function is commonly used to calculate the similarity between samples, it focuses on the prospect image classification is correct, do not pay attention to the background pixels, can effectively alleviate the prospect, background sample imbalance, the calculation formula is as follows:

$$L_{dice}(P,T) = 1 - \frac{2\times\sum_{i=1}^{W}\sum_{j=1}^{H}|t_{ij}\cap p_{ij}|}{\sum_{i=1}^{W}\sum_{j=1}^{H}(|t_{ij}|+|p_{ij}|)} \quad (1)$$

4) Model iterative training.

The UNet model is trained on the physical computer platform Windows. The model training uses Adam (adaptive moment estimation optimizer) as the network training optimization strategy for 100 epochs. The learning rate is dynamically adjusted from 0.001, and after each epoch of updates, the learning rate is multiplied by 0.9. Observe the change of the training loss LOSS, when the LOSS value does not drop for 5 consecutive epochs, we stop the training, and obtain the model with the best convergence effect (the lowest final Loss value).

3.2 Data fusion step

The fusion image obtained in step 2 is input into the training UNet model to obtain the resulting mask image. The white part of the mask image is the wire area and the black part is the background. The process of generating the mask image is shown in the following figure 7. below.

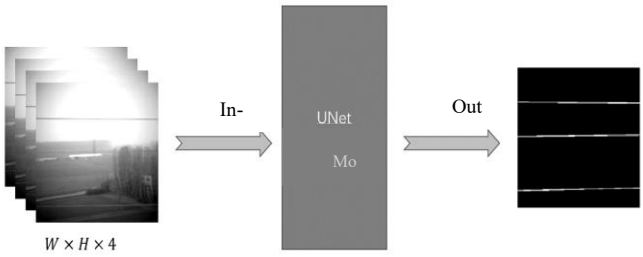

Fig. 7. Mask image generation process

3.3 Algorithm detection strategy

The mask is applied to the infrared image to divide the wire area and background from the infrared image, and to determine the temperature threshold of the heat map mapping to the infrared image. Then obtain the abnormal temperature threshold by Otsu 12 according to the interclass temperature threshold, and then split the wire dispersion area from the wire area according to the threshold. The Otsu algorithm counts the mean and variance of the image gray scale of the wire area, and divides the wire area into normal and abnormal categories according to the optimization criterion of maximizing the variance between classes. the boundary between the two types of pixels is the anomalous temperature threshold of segmentation. The specific detection process of wire distribution area is as follows: The total number of pixels in the wire area is N, the gray range is between 0 and L-1, the number of pixels in the corresponding gray level i is n i, and the odds can be calculated as:

$$p_i = n_i/N, i = 0,1,2 \cdots, L-1 \tag{2}$$

$$\boldsymbol{Table}\ 1. \sum_{i=0}^{L-1} p_i = 1 \tag{3}$$

Pixels in the wire area are divided into two categories C_0 and C_1 by threshold value T, C_0 consists of pixels with gray values between [0, T], C_1 consists of pixels with gray values between [T+1, L-1], and the image mean is:

$$u_T = \sum_{i=0}^{L-1} ip_i \tag{4}$$

The mean values of C0 and C1 are, respectively:

$$u_0 = \sum_{i=0}^{T} ip_i/\omega_0 \tag{5}$$

$$u_1 = \sum_{i=T+1}^{L-1} ip_i/\omega_1 \tag{6}$$

Where $\omega_0 = \sum_{i=0}^{T} p_i$, $\omega_1 = 1 - \omega_0$ then:

$$u_T = \omega_0 u_0 + \omega_1 u_1 \tag{7}$$

Within-class variance is defined as:

$$\sigma_B^2 = \omega_0(u_0 - u_T)^2 + \omega_1(u_1 - u_T)^2$$
$$= \omega_1\omega_0(u_0 - u_1)^2 \quad (8)$$

The abnormal temperature threshold T is made successively taken between [0, L-1], so that the maximum T value of σ_B^2 is the optimal threshold. The area in the image greater than the threshold T is the wire dispersion area.

4 Results & Discussion

4.1 Experiment environment and configuration

Table 1. Test environment and configuration

name	Specific parameters
CPU	Intel XEON E5-2620V4 @ 2.10GHz, 8C/16T
Graphics card	NVIDIA Titan xp 12GB
operat	Redhat Centos 7.7.1908
CUDA version	9.0
CUDNN	7.5.0.56

4.2 The Defect Detection Dataset

There is a large amount of power pole and tower data on the distribution network lines managed by Dongguan Power Supply Bureau of Guangdong Power Grid Company. Due to the aging of some lines, this experiment collected sufficient training and test positive sample data by screening the image data of regular line UAV inspection.

4.3 Experimental results and comparative analysis

Table 2. Caparison of performance between baseline and our method

Method	Image amount	Iterations	Batch-size	Recall	Precision
Base-line	1246	100000	128	75.21%	82.55%
With Fusion (Ours)	1864	100000	128	85.98%	90.87%

As shown in the following figure 8.: we detect the image of the wire broken stock.

Fig. 8. Algorithmic detection diagram

5 Conclusions

The proposed algorithm in this paper combines the advantages of infrared and visible images to realize the defect detection of multi-modal data fusion. On the one hand, the visible images are used to supply some of the wire features, which improves the segmentation performance of the UNet model in the images with poor illumination conditions; on the other hand, the wire mask output by the segmentation model is applied to the infrared images to solve the problem of low discrimination between the wire and background contrast in the infrared image and reduce the false positive rate of scattered stock detection. Compared with the previous manual reading, the algorithm greatly improves the operation efficiency; compared with the image processing method using only infrared or visible light images, the method can be used in the UAV inspection images with different light conditions, which is more conducive to be embedded in the UAV for all-weather intelligent inspection, and has the advantages of high recall rate and low false detection rate.

The training set, validation set and test set used in this paper are obtained from UAV field acquisition. A series of comparative experiments prove that the multi-modal data fusion detection method combining infrared and visible light images can effectively detect the fault of the wire, which has certain reference value for the patrol inspection of the wire. In the future, we will further combine transfer learning, to improve the detection accuracy and speed up under the case of limited hardware con-

ditions or target data. By doing this ,we would realize more rapid and accurate real-time wire dispersion defect detection in different complex environments, and improve the robustness and practicability of the detection model.

Acknowledgment

We would particularly like to acknowledge all the team members, and the sponsor department China Southern Power Grid Co., LTD. for their wonderful collaboration and patient support.

Fund project: Technology project of China Southern Power Grid Co., LTD. "Research and development and Demonstration Application of UAV Grid Inspection-Topic 2: Research on Intelligent Identification and Analysis Technology of Multi-modal Data of Distribution Network Equipment" (Project No.: [031900KK52210006 (GDKJXM20210897)]))

References

1. Girshick R. Fast r-cnn[C]//Proceedings of the IEEE international conference on computer vision. 2015: 1440-1448.
2. Ren S, He K, Girshick R, et al. Faster r-cnn: Towards real-time object detection with region proposal networks[J]. Advances in neural information processing systems, 2015, 28.
3. He K, Gkioxari G, Dollár P, et al. Mask r-cnn[C]//Proceedings of the IEEE international conference on computer vision. 2017: 2961-2969.
4. Liu W, Anguelov D, Erhan D, et al. Ssd: Single shot multibox detector[C]//European conference on computer vision. Springer, Cham, 2016: 21-37.
5. Redmon J, Divvala S, Girshick R, et al. You only look once: Unified, real-time object detection[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 779-788.
6. Redmon J, Farhadi A. YOLO9000: better, faster, stronger[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 7263-7271.
7. Bochkovskiy A, Wang C Y, Liao H Y M. Yolov4: Optimal speed and accuracy of object detection[J]. arXiv preprint arXiv:2004.10934, 2020.
8. Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation[C]//International Conference on Medical image computing and computer-assisted intervention. Springer, Cham, 2015: 234-241.
9. He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 770-778.
10. Meng T, Jing X, Yan Z, et al. A survey on machine learning for data fusion[J]. Information Fusion, 2020, 57: 115-129.
11. Milletari F, Navab N, Ahmadi S A. V-net: Fully convolutional neural networks for volumetric medical image segmentation[C]//2016 fourth international conference on 3D vision (3DV). IEEE, 2016: 565-571.
12. Otsu N. A threshold selection method from gray-level histograms[J]. IEEE transactions on systems, man, and cybernetics, 1979, 9(1): 62-66.