Research on Automatic Monitoring Technology of Combination of Image Processing and Deep Learning

Zheng Li
Electrical Engineering Faculty, University of new South Wakes, NSW 2052, Australia

Abstract

Recently, the equipment automatically monitors whether the electricity consumption is normal or not is an extremely important link in the system. The safe operation of the system is related to the safety of life and property. This has caused the number of equipment to monitor the status to continue to grow, resulting in shortages of inspection workers and personnel on duty, increased operating costs, and low labor efficiency. The equipment is generally built in areas with more complex environments, and it is not safe for staff to perform inspections. Moreover, with the development of computer software and hardware, deep learning technology has gradually matured. Researchers have used deep learning technology in various fields, including the field of automatic equipment monitoring, so that workers do not need to manually copy the status of equipment on site, and can directly use the front end The device automatically monitors the status of the device. However, the use of deep learning methods cannot simultaneously identify various types of objects on the equipment and determine their status, and when the shooting position of the inspection device is deviated or the environment is bad, the recognition of the object category will be affected, thereby the state of the object The discrimination produces a large error. This paper analyzes the problems encountered by the equipment automatic monitoring technology in view of the above problems, and proposes a method combining image processing and deep learning according to the characteristics of the equipment scene, and uses this method to analyze and process the inspection logo. First, use the deep learning label detection method to identify the category of the object. For this part, this paper proposes a multi-label detection model, which is based on the original YOLOv3 model and optimizes the FPN feature fusion layer and loss function part , And use the K-means clustering method to modify the size of the a priori box of the convolutional neural network to fit the size of the object in the power equipment scene; secondly, according to different categories, different image processing methods are used to determine the state. For this part, this article uses The multi-category comprehensive discrimination method judges the state of each object. This method integrates various image processing methods according to different categories after the oral logo is detected. The problem of judging the state of change by fans. Finally, the state discrimination algorithm is integrated with the deep learning logo detection network to form an automatic equipment monitoring technology that combines image processing and deep learning.

Keywords*: deep learning, image processing, POLO algorithm*

1 Research status at home and abroad

In recent years, automatic equipment monitoring technology has been studied at home and abroad. In China, the country first started the research on this technology. Its subsidiary Shandong Luneng Intelligent Technology[1]. Ltd first started the research on the intelligent monitoring of equipment at the end of the 20th century. At the beginning of the 21st century, Chinese researchers learned about the Technology, and then the country built a laboratory to study automatic monitoring in the field of electric power, this laboratory is to study the automatic monitoring function of robots. In 2004, the laboratory developed the first automatic monitoring system[1]. After that, it has developed a series of equipment intelligent monitoring 2systems. The success of these systems is inseparable from many related national projects and the national "863 projects." support[3]. These systems comprehensively use video monitoring, intelligent cloud platform control, image processing and other technologies, so that the equipment automatic monitoring system can operate in the substation 24 hours a day and automatically

without manual intervention. At the same time, the system also supports the equipment The identification of the status of the upper isolation switch and the indicator light, and the automatic identification of industrial instruments have realized the automatic monitoring system for the first time in the world, making the robot's fully automated inspection further. In 2012, the National Chongqing Branch and Chongqing University cooperated to study automatic monitoring technology[4]. The system was successfully operated in Banan 500kV substation, achieving the standard of automatic control by staff in the background. In 2014, at the Ruian Substation, a substation inspection device developed by Zhejiang Guozi Robot Technology Co., Ltd. was officially launched. Before the mouth, more and more companies have joined the team of research and development equipment automatic monitoring technology, which has made great progress in the unmanned inspection of equipment[5].

1.1 Research status of target detection algorithms

In fact, the mouth mark detection algorithm is to distinguish the categories of all objects in an image or a frame of video, and it is necessary to detect the position coordinates of each object. In daily life, oral label detection technology provides us with a lot of convenience[6]. After the rapid development of deep learning, slogan detection technology has also developed. Most of the deep learning techniques in front of the mouth use convolutional neural networks, which can be used to quickly find the objects of interest in the video images. The found objects need to be recognized for their type and position coordinates. This technology is It is applied to many fields, including: security, medical, industry, automobile, etc. Significant changes have taken place in the oral logo detection technology this year. Alex Krizhevs and his teacher researched and designed the AlexNet network model this year, and used GPU graphics cards for training in optimizing the hardware[7]. The previous neural network parameters were too many and learned The difficult problem is solved. After that, other researchers optimized and improved the network based on this network, and gave birth to many oral logo detection network frameworks with a high detection rate. Before this, the feature extraction part of the oral logo detection technology was extracted by human experience. The robustness of the network becomes very poor, and the detection results are not ideal when detecting some objects with unobvious features. The research on this technology before the mouth can be divided into two directions according to its basic principles: one direction is called the two-step method, which is based on the regional suggestion frame, mainly including R-CNN, SPP-Net, and the framework of these networks. Improve other oral logo detection algorithms obtained on the basis of. These algorithms consist of two steps. The first step is to find the possible bounding box of the object in the measured range[8]. The second step is to use the mouth mark detection algorithm object classifier to determine the possible bounding box of the object in the measured range in the first step. What is the specific category of the object, and the regression algorithm is used to clarify the specific location of the bounding box; the other direction is called the one-step method, which is based on end-to-end, mainly including POLO series algorithms, SSD algorithms, etc., these algorithms are directly used The algorithm of regression thinking performs detection and obtains the object category and position coordinates at the same time. Before the mouth, these two logo detection algorithms are also used in the field of automatic monitoring of power equipment. However, these two methods have their own obvious advantages and disadvantages. For example, the detection accuracy of the two-step method is higher than that of the one-step method. The detection puzzle is faster than the two-step method[9-11].

1.2 Research status of digital image processing technology

We can simply understand digital image processing technology as processing methods such as image acquisition, reconstruction, enhancement, and restoration by a computer. Since the 21st century, computer theory and technology have been comprehensively developed and improved in various fields, from the earliest medical field to the current artificial intelligence, such as: traffic monitoring on the road, collection of daily weather conditions, and coding of various commodities , Bank’s equipment and government departments’ alarm systems, etc. In modern society, the amount of information that images can contain is increasing day by day, so we have high requirements for digital image processing technology[12]. Due to the advantages of this technology, such as high processing accuracy, strong adaptability, and wide application range, we can integrate this technology into all walks of life. Although the advantages of image processing technology are obvious, there are also many aspects that affect its

development. For example, this technology is very dependent on the computer's operational fascination and the amount of data stored by the computer. Therefore, some processing procedures will be very complicated. These factors are the reasons for the stagnation of image processing technology. At this stage, image processing technology will also provide great help for deep learning logo detection. Therefore, the development of image processing technology also affects the rapid development of deep learning.

1.3 Main research content and organizational structure

This paper mainly studies the use of image processing and deep learning technology to classify and recognize objects on power equipment, and provide technical support for the automatic monitoring of power equipment. The main ports to be monitored on power equipment are three categories: indicator lights, dashboards and isolating switches. The later chapters are based on the research on the principles of deep learning port label detection and how to modify the model to improve the detection accuracy. The Faster R-CNN model and the YOLOv3 model have carried out in-depth study and research. The ideas of these models have been improved to improve the detection accuracy[13]. At the same time, various digital image processing technologies are investigated and studied, and different researches are integrated. The object designs different state discrimination methods. Finally, this article combines image processing technology and deep learning logo detection algorithm to make the automatic monitoring technology of power equipment can be realized in engineering applications.

2 Deep learning target detection theory and related models

2.1 Overview of Deep Learning

Deep learning is a learning in which researchers use the structure and characteristics of the human brain to imitate the human brain. Deep learning combines the low-level semantics in the network and combines them into high-level semantics through different forms. These features are used to represent certain attributes, and the features are expressed in the distribution of these features. The working mechanism of deep learning and the working mechanism of the human brain It is closely related. The principle is shown in Figure 2-1. Starting from the human eye retina, the low-level V1 area is used to extract the contour features of the received external information. After the contour feature is extracted, it passes through the V2 area[14]. It is the partial information of the mouth mark formed by the previous features, and then the V4 area is used to determine what the mouth mark is, and then the PFC layer is used to classify the objects that are seen. Through this process, it is discovered that the combination of low-level features is high-level features. As the levels increase, the information expressed becomes more and more abstract, and the working mechanism of deep learning is to slow down the low-level features of the object to be recognized. Slow the process of forming high-level abstract features.

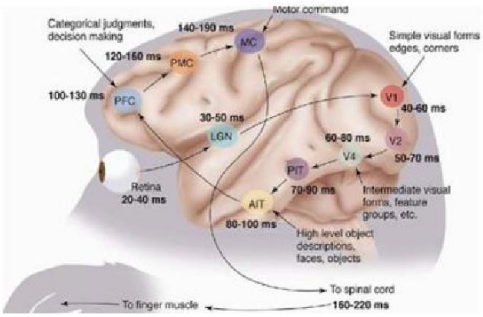

Figure 1: Human brain structure diagram

Because there is only an output layer and an input layer in a layer of perceptron, people think that if a layer is

added between the two layers to form a multi-layer, so that the samples can be classified by the convex domain, this layer is called the hidden layer , This constitutes a multi-layer perceptron, which is the original deep learning network model. As we have seen, this is the earliest model of deep learning, divided into three layers: input layer, hidden layer and output layer. Due to the number of intermediate layers, the neural network can be divided into shallow and deep layers. In 1969, Marvin Minsky and Simon Pipette pointed out that theoretical analysis cannot prove whether the upgrade from a single-layer perceptron to a multi-layer perceptron is of practical significance, which makes experts in many fields feel that the research on perceptrons difficulties, did not continue to study on this subject[15].

2.2 Convolutional Neural Network

In 1989, Yann Lecun of New York University developed a new neural network model called a convolutional neural network. It is essentially a multi-layer perceptron, in which each layer is essentially formed by a combination of many neurons. These neurons will form a feature map, and you need to know what features these feature maps represent, so you need to use Filter to achieve the function of extracting features, these feature maps constitute each layer of the convolutional neural network. It includes the following parts.

Convolutional layer: The foundation of the deep learning convolutional neural network is this layer, because it determines the depth of the lower layer matrix. The input of each node in this layer is not all of the upper layer, but the upper layer three times three or five Multiply a piece of five, the depth of the matrix will become deeper after such processing[16].

Pooling layer: The function of this layer is to reduce the resolution of the picture, because the length and width of the matrix will become smaller after passing through this layer, but the depth will not change, so this will reduce the number of neurons in the fully connected layer , This will reduce the training parameters and optimize the entire network.

Fully connected layer: This layer is the last one to two layers of the convolutional neural network. Its mouth is to classify the features extracted after passing through the convolutional layer and the pooling layer, because we look at each category To determine the specific category, the softmax activation function is more suitable for this layer. This function can handle multi-classification tasks.

2.3 Target detection algorithm based on convolutional neural network

Deep learning has become the focus of attention by researchers in terms of oral logo detection technology[17]. It started with the convolutional neural network achieving a high detection rate in the ImageNet dataset. After that, researchers made improvements in all aspects of the convolutional neural network. In many public A high detection accuracy rate is obtained on the data set. Traditional classification algorithms have been unable to catch up with the pace of convolutional neural networks in classification algorithms, and the improvement of classification effect has also led to the rapid development of deep learning slogan detection technology[18].

3 Research on target detection algorithm

Before the mouth, there are one-step oral labeling detection method and two-step oral labeling detection method. The two-step method includes the regional convolutional neural network R-CNN and its optimized algorithm. The difference between the two-step method and the one-step method is that the two-step method first generates the regional candidate frame of the oral mark, and then uses the SVM and other algorithms to extract the features. Classification, the final output is the type of object to be detected and the specific location coordinates. The one-

step method includes SSD, POLO, and optimized algorithms YOLOv2 and YOLOv3.

3.1 Faster R-CNN network

Before introducing Faster R-CNN, let's introduce R-CNN and its optimized network. The main steps of R-CNN logo detection network are shown in Figure 3-1: input image, select region candidate frame, feature extraction and region classification[19].

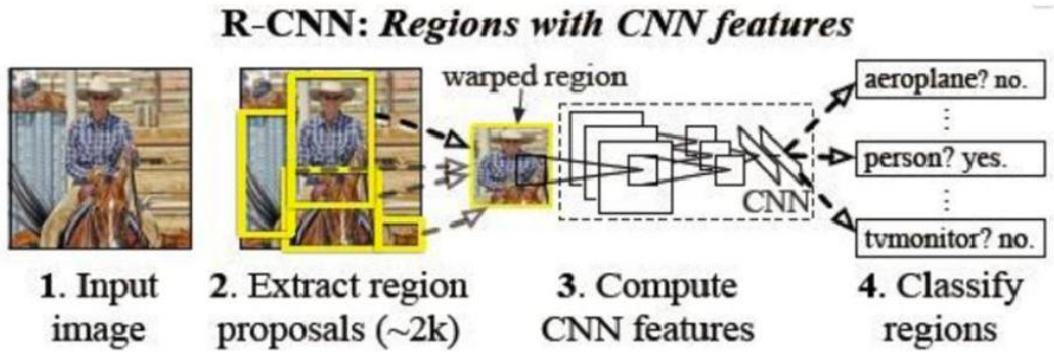

Figure 2: R-CNN target detection network steps

First of all, it can be seen from the above figure that the first step is the input of the image to be detected; the second step is to use selective search and other similar algorithms to propose about two thousand regional suggestion boxes; the third step is to carry out all the regional suggestion boxes Cut to a uniform size; finally, the features are sent to each classifier for classification. However, most of the candidate frames of each region extracted by R-CNN overlap, and each candidate frame will extract features from it. In this way, many repeated features will be extracted and forward propagation must be carried out. It will increase the amount of calculation; and under a certain convolutional neural network model, the input size of the network must be determined in advance. In order to solve the above problems, the spatial pyramid pooling method is used, and SPP-net is proposed[20].

SPP-net first divides the area suggestion box mapping area on the feature map of the convolutional layer into three windows of 1X1,2X2, 4X4 size; secondly, the maximum pooling operation is performed in these three windows, so as to pass through the spatial pyramid The feature map generated by each convolution kernel of the pooling will get a (1X1+2X2+4X4)-dimensional feature vector; the dimension of the final generated feature vector is determined, and we can use this feature vector as an input parameter to input The fully connected layer is classified. There are two advantages of this method: 1. The way to get the feature map is to directly use the initial image calculation, instead of the convolutional neural network calculation for each region suggestion box in R-CNN, which will reduce the amount of calculation [21]. 2. SPP-net does not need to crop the suggestion box, because the generation of the region suggestion box in SPP-net also requires the use of selective search algorithms on the input image to obtain the corresponding area in the feature map through mapping, and The region suggestion box is pooled in a spatial pyramid in the corresponding region in the feature map.

The function of spatial pyramid pooling is that the proposed area of any size will generate a fixed-size feature vector through SPP. Although SPP-net is faster than the R-CNN network in detecting mystery, the problems in R-CNN will also exist in SPP-net. The training process of the convolutional neural network and the regression training process of the slogan are carried out separately. This kind of staged training will cause the parameter optimization process to not be integrated, which will make it impossible for the network to achieve a high detection rate. To this end, the original author designed the Fast R-CNN network.

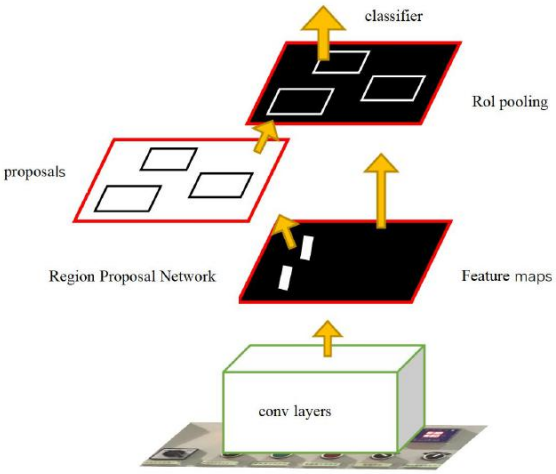

Figure 3: Faster R-CNN structure

The method of Faster R-CNN to generate suggested regions is to use the RPN network, which is different from Fast R-CNN, and then use the ROI pooling layer to map the candidate region data into uniform size candidate region feature vectors, and then input them into the classification Detector and regressor for image detection processing.

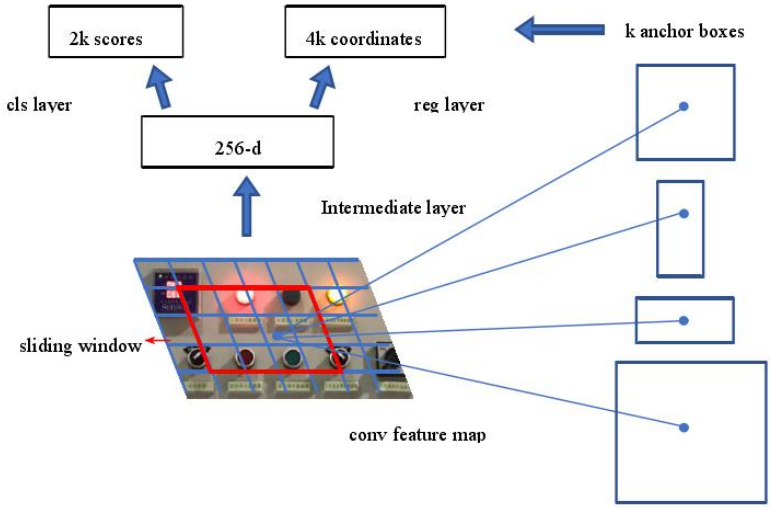

Figure 4: RPN network

The Faster R-CNN training method is simpler than the traditional R-CNN, because the entire network model training does not need to separate the classification function and the regression function, so that the network training is a whole. In short, as far as the overall network structure is concerned, a network model consists of two parts: feature extraction and generation of suggested area boxes. Faster R-CNN combines the two parts together, which saves the time of generating candidate boxes in batches; in terms of features In terms of extraction, in order to extract more features of the object to be detected, we use a deeper network, but the deepening of the network will cause the problem of gradient disappearance, so the shared convolutional layer is used to extract the features, which can improve the detection accuracy; In terms of candidate region generation, using the RPN method to generate suggestion boxes can reduce redundancy, omit unnecessary suggestion boxes, and reduce the amount of calculation; as far as the suggested region size is concerned, because feature maps of different sizes are sent to the final detection The classifier will cause the detection accuracy to decrease, so when performing ROI pooling, map

the proposed region data to feature maps of the same size[22].

3.2 POLO network

Due to the slow detection of the mouth mark detection network and the difficulty of network optimization, the POLO network is proposed, which regards the mouth mark detection task as a regression problem, and divides the image to be detected into N*N grids, as shown in Figure 2- As shown in 8, each grid needs to complete a prediction of the category and position. If the grid detects the center point of the object to be detected, it will start to detect n candidate frame objects. The candidate frame consists of five parts. Composition, which includes the coordinates of the center, the width and height of the candidate frame, and the confidence level. The POLO model process first adjusts the image size, then uses the convolutional neural network model, and finally uses the non-maximum suppression method to make the detection result more accurate.

The process of POLO's one-step detection network is actually very simple. There is no complicated detection process. Just feed the picture at the input of the convolutional neural network and the result can be directly output. Therefore, the detection degree is compared with the R-CNN series[23]. The algorithm is very fast, and compared to other algorithms with the same degree of confusion, the accuracy of POLO is more than doubled. The POLO algorithm can make full use of the information on the entire image, so as to avoid the erroneous information caused by background interference. This aspect is different from other verbal mark detection algorithms, which are obtained by the classifier used by other verbal mark detection algorithms. It is only partial image information, because they use regional candidate frames, so the full text information cannot be obtained well.

3.3 YOLOv3 network

Under the original network framework of YOLOv2, researchers have optimized the network so that the detection accuracy is improved while the detection level is also maintained at a high level. At the same time, the detection accuracy of small marks in dense scenes is also improved. This is YOLOv3 network, the network has been improved from these places:

In terms of multi-scale features, YOLOv3 uses 3 feature maps of different scales. Combined with the YOLOv3 network structure diagram, after the 79th layer of the entire network, through the following convolutional layers, a feature map that is twice as small as the original image size will be obtained, because after such down sampling, the feature map will have With a larger receptive field, it is more convenient to detect the larger size of the mouth mark. Similarly, after the 79th layer, the up-sampling on the right is performed to fuse the features with the 61st layer[24]. When the 91st layer with higher semantics can be obtained, a more suitable feature map for small object detection will be obtained. Similar to the above process, there is a convolutional layer under the 91st layer. Through this convolutional layer, a feature map that is 16 times smaller than the original image will be obtained. This receptive field is moderate on the feature map and is suitable for objects of moderate size. The detection effect is good[25]. After the 91st layer, the feature map continues to be used and merged with the previous 36th layer feature map to obtain a feature map that is 8 times smaller than the original image. This feature map has a smaller receptive field and is more suitable for detecting small-sized targets.

4 Equipment state discrimination algorithm

4.1 Discrimination algorithm of circular indicator status based on color space conversion

The state of the indicator light is judged based on color characteristics. It is mainly divided into three steps: 1. Convert the image color to the HSV model; 2. Extract the specific color of the indicator light according to the interval of different colors; 3. Determine the status of the indicator light in the HSV color space[26].

4.2 Spatial color conversion

In front of the mouth, in most cases, the camera is used to collect image information, and the image information is in RGB mode

According to the depth of the color, each color has a different level, with a range of 0255, so that through the combination of these three primary colors, tens of millions of colors can be formed, which can completely represent the various colors of the natural world. But in this mode, the correlation of these three variables is very high. If you want to distinguish these three colors well, you can do it in HSV mode. Hue: the attribute of the color, saturation: the purity of the color, lightness: the composition of the brightness, compared with the RGB model, the HSV model is more convenient for color judgment and processing[27]

As can be seen from the above figure, the angle around the central axis can represent the hue, the distance from the center of the cone to the central axis can represent the saturation, and the value from the black at the bottom to the white axis at the top can represent the lightness. The conversion formulas for the two models are as follows Show:

$$H = \begin{cases} 0^\circ \quad , \text{if } \mathbf{M} = \mathbf{m} \\ 60^\circ \times \frac{\mathbf{g}-\mathbf{b}}{M-\mathbf{m}} + 0^\circ \quad , \text{if } \mathbf{M} = \mathbf{r} \\ 60^\circ \times \frac{\mathbf{b}-\mathbf{r}}{M-\mathbf{m}} + 120^\circ \quad , \text{if } \mathbf{M} = \mathbf{g} \\ 60^\circ \times \frac{\mathbf{r}-\mathbf{g}}{M-\mathbf{m}} + 240^\circ \quad , \text{if } \mathbf{M} = \mathbf{b} \end{cases} \tag{1}$$

In the HSV color space, the distribution values of various colors are as follows. According to the specific distribution of colors, the H channel stores the color information of the image, so you can set the threshold according to the color you want to recognize, and extract different color information; the image of the V channel can be seen from the image of the V channel in the figure below. The light and fire of the indicator can be judged[28]

4.3 Color discrimination

From the HSV model in the previous section, we can see that the color can be extracted in the H channel. The following takes the green indicator as an example. It can be seen that the H value of the green information is in the range of 35-90, so first extract the green area in the HSV image As shown in the figure below, it can be seen from the figure below that the outline of the green indicator light has been extracted, with white information surrounding the outline, and the amount of irrelevant information has been reduced after extraction.

Figure5: Green area

4.4 Judgment of brightness

It can be seen from the previous section that V represents the brightness of the image, but in the on and off state, the distribution interval of the V value of the indicator light is different, so by identifying the V value of the area

where the indicator light is located, the V value can be determined. The on and off state of the indicator light. Because in the first step, the deep learning technology has been used to identify the position information of the indicator in the image to be identified. At this time, only the coordinates of the center of the identified indicator are calculated, and the V of the neighborhood of the center of the indicator is extracted. The value can be used to determine the magnitude of the value. Experiments show that if this value is greater than 210, the indicator light is in the on state, on the contrary, if the value is less than 210, the indicator light is in the off state. Finally, mark the indicator light in the bright state in the original picture, and get the recognition result as shown in the figure. The status of the indicator light represents the power running status monitored by the power equipment[28].The detected result is the operating state of the power equipment where it is located, and then the identified result is saved and uploaded to the background to complete the preservation of the inspection data.

4.5 Distinguishing switch state discrimination algorithm based on template matching

Since the isolation switch on the device is divided into two states, which are on or off, and the styles of the isolation switches of these two states are quite different, the template matching method is used to identify the specific state of the switch.

The template matching method is to use a template picture to find the similar part in another picture, use the template picture to scan and search each position on the picture to be detected when matching, and then use the template picture and the picture on the picture to be detected. The objects are compared, the similarity is calculated, and the calculated similarity is compared with the set value. If it is greater than the value, it can be judged that the object has been found. Therefore, template matching technology can also be used in the state of the isolation switch. Before the detection, select the isolation switch pictures of different states as the comparison template, and then use the picture to match the collected picture to be tested. If the match is successful, it is complete. Judgment of the state of the isolating switch. Before speaking, the template matching method can be divided into: square difference matching method, normalized square difference matching method, correlation matching method and coefficient matching method[29].

4.6 Experimental process and results

In the template matching function, set a threshold directly at the input of the function. When the matching value is greater than this threshold, it is considered the same object. The technology of discriminating the state of the knife switch by template matching method is not perfect before the mouth. Since the same template can only be matched with its similar mouth mark, if the appearance of the mouth mark is different, then a single template cannot be correct. Matching results, so this article collects as many as possible the same kind of samples with different appearances to be tested, so that multiple templates can be used to match the objects detected in the first step to improve the accuracy of state discrimination. In addition, the template matching method relies heavily on the size of the image. If the size of the result detected by the oral label in the first step is very different from the template, the detection accuracy will decrease. Therefore, in the subsequent research, these two The problem is the focus of research.

5 Summary and Outlook

5.1 Research summary

This article has done an in-depth research on equipment automatic monitoring technology. The core technology in this research is deep learning logo detection technology and digital image processing technology. This article explores the problems existing in the automatic monitoring technology of power equipment from the two aspects of specific algorithm research and industrial application.

Regarding the oral mark detection technology, researchers continue to study in depth, and there have been various oral mark detection frameworks based on convolutional neural networks. This article studies the advantages of the two-step R-CNN series framework to the one-step POLO series framework. Disadvantages, finally decided to use the YOLOv3 framework as the basic network of the power equipment multi-label detection algorithm, and on this basis, improve the characteristics of the label to be detected in the power equipment scenario, so that it can meet the needs of automatic monitoring. Among them, because the public neural network training set lacks samples for objects on power equipment, in order to solve this problem, a combination of self-made data sets and public data sets is used to construct a special data set for automatic monitoring of power equipment. The data set contains four categories of objects: circular indicator lights, isolation switches, pointer instruments and digital instruments. Followed by the first from the YOLOv3 network

Start with the verification frame, loss function and feature fusion network to improve the detection accuracy and detection rate of the network, and meet the requirements of the application.

Regarding image processing technology, since there is no algorithm in front of the mouth that can simultaneously recognize the status of objects on different power equipment, after the first step of the mouth mark detection to determine the type and location of the object, this article uses the different types of objects separately Different image processing methods have designed different state discrimination algorithms, and finally the different algorithms and the first step verbal mark detection algorithm are written in the same state discrimination framework. It solves the problem of low efficiency and low accuracy after dividing object classification and state recognition into two frames.

5.2 Research Outlook

Based on the idea of fusion of image processing and deep learning, this paper studies equipment automatic monitoring technology, optimizes the existing mouth mark detection network and integrates image processing related technologies, realizes equipment automatic monitoring function, and improves recognition accuracy and efficiency. Because of the rapid development in the field of deep learning, more efficient logo detection networks have emerged, and because the environment in which power equipment is located is more complex, the requirements for equipment monitoring are also high. Once monitoring problems occur, it will lead to serious consequences. Therefore, there is still a lot of room for optimization in the research of the oral logo detection algorithm in this technology. Try to increase the accuracy of the algorithm by 100%. The factors that affect the detection accuracy include some physical interference factors in the camera shooting process and power equipment. There will be more and more objects on the Internet, and the generalization ability of the network cannot be guaranteed. These problems will also have a great impact on the results of object state discrimination, and the above problems need to be further optimized and studied.

Referneces

[1] K-means algorithm for optimizing the selection of initial clustering centers[J]. Yang Yifan, He Guoxian, Li Yongding. Computer Knowledge and Technology. 2021(05)
[2] Improved SSD algorithm for multi-target detection[J]. Ma Yuandong, Luo Zijiang, Ni Zhaofeng, Xu Bin, Wu Fengjiao, Sun Shouyu, Yang Xiuzhang. Computer Engineering and Applications. 2020(23)
[3] Research on interface standardization of unattended auxiliary monitoring system in traction substation [J]. Zhao Chaopeng, Zhang Li. Electrified Railway. 2019(S1)
[4] Power depth vision: basic concepts, key technologies and application scenarios [J]. Wang Bo, Ma Fuqi, Dong Xuzhu, Wang Peng, Ma Hengrui, Wang Hongxia. Guangdong Electric Power. 2019(09)
[5] YOLOv3 network based on improved loss function [J]. Lu Shuo, Cai Xuan, Feng Rui. Computer System Application. 2019(02)
[6] A review of research on target detection algorithms [J]. Fang Luping, He Hangjiang, Zhou Guomin. Computer Engineering and Applications. 2018(13)

[7] A facial expression recognition method based on isolation loss function [J]. Zeng Yiqi, Guan Shengxiao. Information Technology and Network Security. 2018(06)
[8] Research on the recognition method of pointer instrument based on machine vision[J]. Tong Weiyuan, Ge Yisu, Yang Chengguang, Jin Yiming, Gao Fei. Computer Measurement and Control. 2018(03)
[9] Improved target recognition algorithm based on template matching [J]. Ding Xiaoling, Zhao Qiang, Li Yibin, Ma Xin. Journal of Shandong University (Engineering Science Edition). 2018(02)
[10] Car body color recognition algorithm based on HSV color space [J]. Hu Zhuoyuan, Cao Yudong, Li Yang. Journal of Liaoning University of Technology (Natural Science Edition). 2017(01)
[11] A branch image segmentation method based on RGB color components[J]. Li Yanwen,Hu Haobo,Ding Yi,Chen Ou,Jia Yuhang. Journal of Yanshan University. 2016(05)
[12] Overview of the research status of substation inspection robots [J]. Yang Xudong, Huang Yuzhu, Li Jigang, Li Li, Li Beidou. Shandong Electric Power Technology. 2015 (01)
[13] Yesterday, today and tomorrow of deep learning [J]. Yu Kai, Jia Lei, Chen Yuqiang, Xu Wei. Computer Research and Development. 2013(09)
[14] Robot-based image recognition method of substation switch status [J]. Chen Anwei, Le Quanming, Zhang Zongyi, Sun Yong. Automation of Electric Power Systems. 2012(06)
[15] Support vector machine and its application in mechanical fault diagnosis [J]. Yuan Shengfa, Chu Fulei. Vibration and Shock. 2007(11)
[16] Target recognition, location and detection based on CNN integration[D]. Jianghao Rao. University of Chinese Academy of Sciences (Institute of Optoelectronic Technology, Chinese Academy of Sciences) 2018
[17] Research on Intelligent Railway Image Recognition Algorithm Based on Convolutional Neural Network [D]. Wang Cong. Beijing Jiaotong University 2018
[18] Meter value and symbol recognition based on multi-channel mechanism [D]. Wu Caiyun. Fujian Normal University 2018
[19] Research on indoor positioning algorithm based on improved convolutional neural network [D]. Zhang Jin. Harbin Engineering University 2018
[20] Background subtraction method of Gaussian mixture model based on HSV [D]. Hu Haoran. Beijing University of Chemical Technology 2017
[21] The automatic identification system of the state of the high-voltage isolating switch of the intelligent inspection robot in the substation [D]. Li Hui. Tianjin University 2017
[22] Research and design of intelligent inspection system for substation [D]. Wang Shaoya. North China Electric Power University 2015
[23] Research on automatic verification system of digital meters based on image recognition [D]. You Xiaojun. Anhui University of Science and Technology 2013
[24] Research on the detection algorithm of sea surface infrared ship target based on mathematical morphology [D]. Li Jijun. Shandong University 2012
[25] Research on rail surface defect detection technology based on computer vision [D]. Zhen Li. Nanjing University of Aeronautics and Astronautics 2012
[26] Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks.[J] . Ren Shaoqing,He Kaiming,Girshick Ross,Sun Jian. IEEE transactions on pattern analysis and machine intelligence . 2017 (6)
[27] Guest Editorial: Deep Learning[J] . Marc'Aurelio Ranzato,Geoffrey Hinton,Yann LeCun. International Journal of Computer Vision . 2015 (1)
[28] The Canny Edge Detector Revisited[J] . William McIlhagga. International Journal of Computer Vision . 2011 (3)
[29] Overview of the research status of substation inspection robots. Yang Xudong, Huang Yuzhu, Li Jigang, Li Li, Li Beidou. Shandong Electric Power Technology. 2015