Defect Detection Method on Power Distribution Equipment Using Cascaded Network Improved by GS-NMCSO Algorithm

Dongming Zhao, Huimin Yu, Xiang Fang, Lei Tian, and Pengqian Han

***Abstract*—The copper bus-bar in the power distribution cabinet is a device conveying current and connecting electrical components in the circuit. If the fastening bolts on copper bus-bar are loose or falling off, high current can melt the copper bus-bar, causing significant damage. A defect detection method on power distribution equipment based on an improved cascaded network is presented, which fully embodies the advantages of deep-learning on feature extraction, region proposal, small object recognition. This method provides high accuracy, real-time and high-efficiency identification effect for bolts fastening status (normal, falling off and loose). In the paper, Faster-RCNN is improved by adding support vector machine (SVM) after RPN (Region Proposal Network) to improve classification accuracy on the foreground and background images. Only the foreground features including target are sent to the subsequent network for training to make this method take good balance between accuracy and training efficiency. In addition, the gravitational search operator is introduced into the Normal Mutation cat swarm (NMCSO) algorithm to improve the global optimization ability and significantly enhance the classification precision of SVM. The experimental results show that the improved cascaded network is superior to the existing deep-learning algorithms and traditional machine learning methods in accuracy and training time for power distribution equipment detection. The mean average precision (mAP) for image set of this paper is 89.51%, and the training time is the shortest.**

***Index Terms*— Power distribution equipment defect detection, deep learning, cascaded network, cat swarm optimization, gravitational search algorithm**

Manuscript received Feb 22, 2020, revised May 08, 2020.

Dongming Zhao is with the college of information science and Electronic Engineering, Zhejiang University, Hangzhou, 310058, PR China. (E-mail: waitman_840602@163.com).

Huimin Yu, is with the college of information science and Electronic Engineering, Zhejiang University, Hangzhou, 310058, PR China. (E-mail: yhm2005@zju.edu.cn).

Xiang Fang is with Hangzhou Zhijiang Switchgear Co. Ltd., HangShen Group, Hangzhou, 311200, PR China. (E-mail: coolfxcn@aliyun.com).

Lei Tian is with the Artificial Intelligence Laboratory, China Mobile Communication Group Tianjin Co., Ltd., Tianjin, 300020, PR China. (E-mail: 15822591582@139.com).

Pengqian Han is with the Artificial Intelligence Laboratory, China Mobile Communication Group Tianjin Co., Ltd., Tianjin, 300020, PR China. (E-mail: hanpengqian@tj.chinamobile.com).

I. Introduction

Electric power is the foundation and pillar of modern industrial production. Small power transformer room is an important component for electrical enterprises to receive the electric energy from power transmission system. The high-voltage electric energy transmitted from the backbone power grid can be reduced to 380V/220V for ordinary machines and equipment. With an increasing power consumption in enterprise production, the requirements of power distribution room are increasingly strict to ensure the safe operation of power grid. Therefore, the design of efficient, stable and safe electrical equipment defect detection system for power distribution rooms gradually become the current research difficulty.

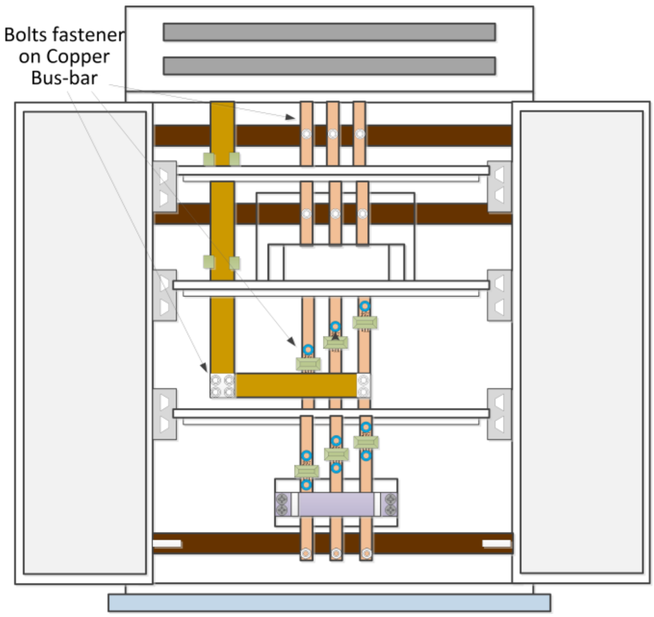

Fig. 1. Structure chart of power distribution cabinet.

The low-voltage power distribution cabinet serving as the hub converter for power exchange is one of the most critical electrical equipment in the distribution room of an enterprise. The switchgear, measuring instruments, protection appliances, and auxiliary equipment are assembled in semi-closed metal cabinet according to electrical wiring requirements. The function is to turn on or turn off the circuit through the automatic switch in normal operation, and cut-off circuit with protection appliance after failure, and trigger alarm. Fig. 1 is the structure chart of power distribution cabinet, copper bus-bar is widely used in the circuit of the low-voltage distribution cabinet to transmit current and connect electrical equipment. Copper bar, also known as copper bus-bar, is made of copper and is a long conductor with a rectangle or rounded rectangle cross-section. With a wide range of switchgear equipment, a large number of distribution circuits and complex wiring, the use of copper bus-bar as a conductive medium to connect widely used equipment is a very good choice, the wiring complexity can be greatly reduced and the load capacity can be significantly increased. The connection in copper bus-bars and in copper bus-bars and equipment is mainly fastened by bolts. If the bolts fixing the copper bus-bar are loose or falling off, high current will melt the copper bar, causing irreparable production losses. Therefore, the application based on artificial intelligence technology to detect the key parts of power distribution equipment is an important research topic, and the focus of this paper.

The core method of defect detection system for power distribution equipment is "object detection + classification". Those object detection algorithms that predate the rise of deep learning can be summed up as image detection window selection, feature extraction and classifier design. The early method for image detection window selection is to traverse all possible boxes in image from left to right, from top to bottom, and scale image size, then the image pyramid is obtained for multi-scale search. There are various methods of feature design. Feature extraction speed on Haar algorithm is very fast, which can express various image edge change information of objects. LBP algorithm can express the object texture information and has a good adaptability to the uniformly changing light. HOG algorithm uses histogram statistics to encode the object edge and has stronger ability for feature expression. The above methods are widely used in object detection, object tracking and object recognition. In addition, classification algorithms mainly include Ada Boost, SVM, decision tree and random forest, etc.

Since 2013, the entire academic and industrial community has gradually applied deep-learning to image detection, such as object classification and object detection, and the detection accuracy has improved dramatically with the powerful feature expression capabilities of deep-learning.

Then, many image detection areas have achieved rapid development, including face recognition, license plate detection, fingerprint recognition and medical image processing. Nowadays, the mainstream object detection algorithms based on deep-learning are SSD, YOLO and RCNN series. Among them, Faster-RCNN is a typical cascade network structure. The RPN network is proposed to replace selective search algorithm so that the detection task can be completed from end-to-end. Faster-RCNN can be understood as the combined set of RPN and Fast-RCNN, which can share image convolution feature, so that it can run at speed of 5fps on a single GPU. In addition, this method has high recognition accuracy especially for small objects in high-resolution images, and is very suitable for complex equipment abnormal state detection in power distribution cabinet. In the paper, Faster-RCNN is improved by adding SVM after RPN to improve classification accuracy on foreground and background images, only the foreground features including target are sent to the subsequent network for training, so the method take good balance between accuracy and training efficiency. Besides, the gravitational search operator is introduced into the Normal Mutation cat swarm optimization (NMCSO) algorithm to improve global optimization ability and significantly enhance classification precision of SVM, which is of great research value.

II. CASCADE NETWORK OPTIMIZATION

A. Object detection algorithm

Faster-RCNN is a object detection network improved by Fast-RCNN, which retains the original candidate region sampling, classification and coordinate regression process. In the region selection for object, the RPN network replace time-consuming selective search algorithm, resulting in a significant improvement in both accuracy and recognition speed. The RPN network is to output a set of rectangular region proposal boxes for object with an multi-scale image as input. Each box contains 4 position coordinate variables and 1 score. The process of region proposal is as follows:

--First, the input image is processed through convolution operation to generate the feature map.

--Second, the multi-scale convolution operation is done on the feature map to get the mapping of proposal, which is expressed as 3 scales and 3 length-width ratios at the position of each sliding window. Each original image can be mapped to 9 proposals of different scales. For example, the number of proposal is W×h×9 for shared convolution feature map of W×h.

--Third, the classification layer outputs the results of whether each proposal is an object or not, specified as W×h×9×2 candidate regions with scores and estimated probabilities.

In the process of RPN network training, a binary label is assigned to each region to mark whether an object exists. The criterion is whether the intersection-over-Union (IoU) from the proposal and ground region (GT) exceed threshold value. The anchors of IoU (>0.7) are set as positive sample, and the anchors of IoU (<0.3) are negative sample.

The training process follows multi-task loss principle, and the loss function of image is defined as:

$$L\left(\{p_i\},\{t_i\}\right) = \frac{1}{N_{cls}}\sum_{i} L_{cls}\left(p_i, p_i^*\right) + \lambda\frac{1}{N_{reg}}\sum_{i} p_i^* L_{reg}\left(t_i, t_i^*\right) \tag{1}$$

where i is the index of proposal, and P_i is the proposal probability of class i, t_i is a vector representing the 4 parameter coordinates of the proposaled anchor, t_i^* is the corresponding GT coordinate vector. N_{reg} and N_{cls} are the regression layer function (reg) and classification layer function (cls). The box-classification loss function N_{cls} is the logarithm loss of two categories (object and non-object), expressed as:

$$L_{cls}\left(p_i, p_i^*\right) = -log\left[p_i p_i^* + (1 - p_i^*)(1 - p_i)\right] \quad (2)$$

The box-regression loss function N_{reg} is defined as:

$$L_{reg} = (t_i, t_i^*) = R(t_i - t_i^*) \quad (3)$$

where R is a robust loss function(smooth L_1), expressed as:

$$smooth_{L1}(x) = \begin{cases} 0.5x^2, \text{ if } |x| < 1 \\ |x| - 0.5, \text{ otherwise} \end{cases} \quad (4)$$

B. *Network structure optimization*

The focus in this paper is the detection for bolt fastening state of copper bus-bar in power distribution equipment, which is a "detection+classification" process for objects. The main process is as follows: detecting the position coordinate of object in the image, and then classifying the results and recognizing defects. An ideal object detection network needs to consider running speed and calculation efficiency to generate as few proposals as possible, and it also ensures that almost all objects can be covered. Convolutional neural network (CNN) is suitable for extracting the deep features of images, while RPN network can capture the similarity among different objects, which can be used to generate proposal regions. Although the significant differences between foreground object area and background area exist, such as color features, edge features etc, the actual output of RPN network is the possibility proposal area of the object, which is a probability estimation. The proposal inevitably contains many background areas, and results in computing power waste for final recognition layer.

Recently, some research team has proposed to introduce softmax function after RPN network to improve the regional prediction effect. However, softmax is mainly used for multi classification, and the main purpose of RPN is to distinguish foreground region and background region, which need to solve a binary classification problem. The fitting capability of softmax in binary classification is not optimal, which may lead to poor results. Therefore, an optimized cascaded network structure based on Faster-RCNN is proposed in the paper, where the SVM with radial basis function (RBF) kernel is introduced into RPN to classify the foreground region and background region, and screen out the background image. Only the foreground image containing objects is sent to the final recognition layer. RBF kernel function can map samples to high-dimensional space and its fitting ability is significantly better than many existing regression classification methods such as softmax. This method has the following advantages: the independent SVM network further improves the region proposal quality, filters out the background region without objects, and makes region recommendation more in line with the requirements of object recognition. In addition, the image features on region proposal from different objects are fused with each other, and the information is used in complementary ways, which makes the classification more accurate. As shown in the dotted area of Fig.2, the RPN network learns the texture features of input image, generates a large number of possible areas and inputs them into the subsequent SVM network to screen out the foreground image, so this cascaded network can make up for the deficiency of RPN.

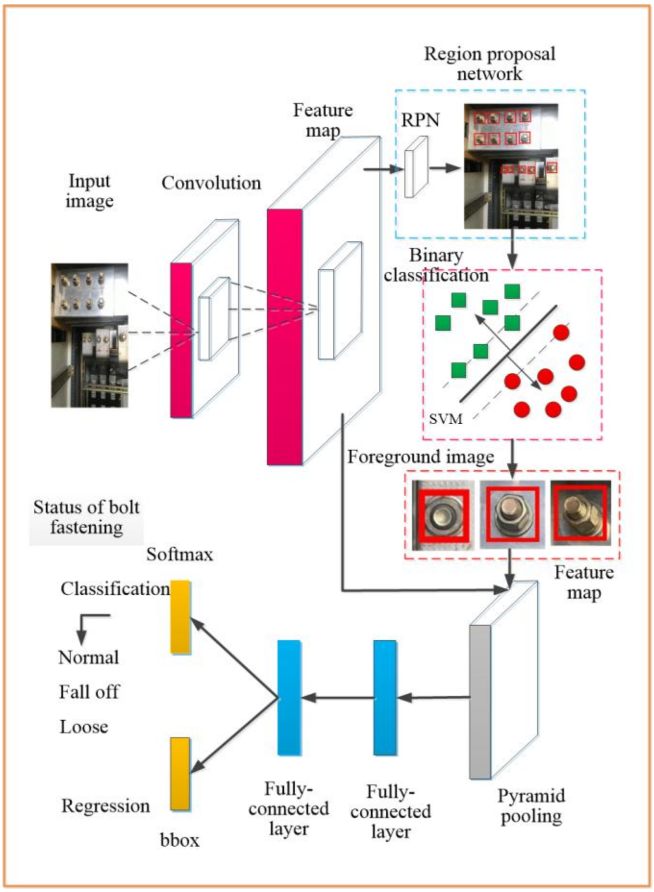

Fig. 2. Optimized cascaded network structure based on SVM.

III. NETWORK PARAMETERS OPTIMIZATION

The detection effect of optimized cascaded network mainly depends on the classification accuracy improvement on SVM for foreground and background images, while the classification ability depends on the penalty parameter C and RBF kernel function λ , in this article, a new swarm intelligence optimization algorithm is presented to optimize this kernel function parameters.

Cat swarm optimization (CSO) algorithm presented by Zhu in 2016 is a new evolutionary computing algorithm that imitates the natural behavior of cats, which has shown superiority in many research fields nowadays. It divides the behavior of cats into two modes: searching and tracking, and calculates the fitness value and iteratively searches for the optimal solution.

A. *Support vector machine*

For the classification task on deep-learning network, the classification algorithms commonly used in the output layer include polynomial logic regression, softmax and SVM, among which, SVM is the most commonly used algorithm for two-category classification problem, and needs only a few training samples to obtain the global optimal solution, so it is widely used in object detection, image classification, nonlinear regression, pattern recognition and other related fields.

SVM uses linear model to introduce nonlinear input vector into multi-dimensional space and generate nonlinear class boundary, so a maximum margin between two classes is established in multi-dimensional space to classify accurately. The classification training sample set is expressed as (x_i , y_i) where $y_i \in \{-1,1\}$, $x_i \in R^d$, i=1,2,3,......,n. The input sample space is mapped to the high-dimensional feature space by nonlinear mapping function φ , and then the linear classification function is constructed in the high-dimensional space, the equation of discriminant function is:

$$f(x) = \text{sgn}(\omega\varphi(x) + b) \quad (5)$$

Where ω is the weight vector and b is the offset vector. According to the structural risk minimization principle, the optimization problems to be solved are as follows:

$$\min f(w) = \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}\xi_i \quad (6)$$

Where C is the error penalty factor, which can control the punishment degree of right and wrong classification for samples, enables effective tradeoff between wrong samples proportion and algorithm complexity, ξ_i is the relax matrix.

For SVM model, RBF as a nonlinear function can reduce the computational complexity in training process and is used as the kernel function.

The expression equation of RBF is:

$$K(x_i, x_j) = \exp\left(-\lambda\left\|x_i - x_j\right\|^2\right), \lambda > 0 \quad (7)$$

In equation(6) and equation(7), C and λ are the core parameters that need to be adjusted precisely. C is the error penalty factor to control the penalty degree of error samples, while the kernel function parameter λ mainly affects the sample distribution complexity in space. The parameter set C and λ determine learning and generalization ability of SVM, which ensures the edge maximization effect of optimized objective function. After the C and λ are optimized to the best value, SVM achieves the highest accuracy for foreground and background image classification.

B. Normal mutation cat swarm optimization with gravitational search operator

In this section, gravitational search operator is introduced to solve the global optimization problem of NMCSO, and a novel normal mutation cat swarm optimization with gravitational search operator (GS-NMCSO) is proposed and used to optimize the penalty parameters C and λ of SVM. Each cat represents a vector (a set of C and λ) with two dimensions. The objective of GS-NMCSO is to find the best individual cat, that is, to find the best value of C and λ .

Normal mutation cat swarm optimization (NMCSO)

The normal mutation cat swarm algorithm (NMCSO) is proposed by Pappula and Ghosh in 2018, the mutation coefficient is optimized by the normal distribution curve in seeking mode so as to enhance the optimization accuracy. The NMCSO algorithm uses the mixed ratio (MR) to determine the distribution for the cat swarm pattern and adjust whether the cat's pattern is seeking or tracing.

Seeking Mode (SM): Cats continue to look around and search to determine the best position, initially in a static but alert state. The smaller mutations in the parent neighborhood is more advantageous to the NMCSO to make a more accurate local search and faster local convergence, and is more consistent with the actual situation. For individuals in a cat swarm, the perception range for prey is gradually reduced from the current static position. The NMCSO uses a normal distribution curve model to update the position of parental cats in the seeking mode, rather than variation of parental position based on SRD percentage approach. The mutation probability near the parent location is smaller, but far away from the parent location is higher. This normal mutation process is very benefit for the parent casts to make a precise search compared with the mutation process in classical CSO. The normal distribution density function with mean (μ) and standard deviation (σ) is:

$$f_{normal}(x;\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (8)$$

The normal random numbers (N) can be expressed as:

$$N(\mu,\sigma^2) = \mu + \sigma N(0,1) \quad (9)$$

Where, $N(0,1)$ is a random number of standard normal distribution with zero mean (μ =0) and standard deviation of one (σ =1). The standard deviation of zero mean and variable normal distribution drive cats to produce normal distribution random variation, thus guides the cats to update position and ensure that the cats are always randomly distributed within the perceptual range of the parent.

The position x_i^m of an individual after normal mutation is:

$$x_i^m = x_i + N(0,\sigma^2) = x_i + \sigma N(0,1) \quad (10)$$

Tracing Mode (TM): Cats rapidly change position to track prey continuously. The position vector of each cat at the next moment depends on speed and global optimal position of cats. We define the ith cat position $x_{i,d}$ and the velocity $v_{i,d}$ as $x_{i,d} = (x_{i,1}, x_{i,2}, x_{i,3}..., x_{i,D})$, $v_{i,d} = (v_{i,1}, v_{i,2}, v_{i,3}..., v_{i,D})$. The global best position $x_{g,d}$ of the cat swarm is represented as $x_{g,d} = (x_{g,1}, x_{g,2}, x_{g,3}..., x_{g,D})$ .The velocity of the ith cat in

d-dimension is calculated as follows:

$$v_{i,d} = w * v_{i,d} + c * r * (x_{gd} - x_{i,d}) \tag{11}$$

Where w is inertia weight, c is acceleration constant and r is a random number in [0,1]. The ith cat position in d-dimension is:

$$x_{i,d}{}^{n} = x_{i,d} + v_{i,d} \tag{12}$$

Evaluate every cat based on fitness function. The position of the cat with best fitness is set to be the optimum solution and replace current position. The SM and TM steps are repeated continuously until the global optimum solution is obtained.

Gravitational search operator

Because the special consideration for the accurate local search in NMCSO, the potential risk 'premature convergence' always exists, which leads to the algorithm falling into local optimum. High complexity also increases the operation time. In this section, the gravitational search operator is added to the operation process of NMCSO, which can change the dimension information of cats by applying the displacement operation based on the gravitational forces among cats.

The gravitational and inertial masses are expressed by the following:

$$m_{i,d} = \frac{f(x_{i,d}) - \max_{x_{j,d} \in w(t)} f(x_{j,d})}{\min_{x_{j,d} \in w(t)} f(x_{j,d}) - \max_{x_{j,d} \in w(t)} f(x_{j,d})} \tag{13}$$

$$M_{i,d} = \frac{m_{i,d}}{\sum_{x_{j,d} \in w(t)} m_{j,d}} \tag{14}$$

Gravitational and inertial masses are simply calculated using the fitness evaluation. A heavier mass means a better cat which is more attractive and moves more slowly. $f[X(t)]$ represents the fitness value of cat i at time t, and $W(t)$ is the aggregation of N cats $X(t)$ at time t. The aggregation of all cats is N_a, and the i, j cat position is $x_{i,d}$, $x_{j,d}$ ($1<d<D$).

The operator selects the first $2*N$ cats of greatest mass from N_a (small fitness value) to form the superior cat swarm aggregation (R), which is used as the attractive element to exert a gravitational effect on $W(t)$. The gravity is defined as follows:

$$F_{i,d} = \sum_{j \in R, j \neq i} rand(0,1) * F_{i,j,d} \tag{15}$$

$$F_{i,j,d} = G \frac{M_i * M_j}{r_{i,j} + \varepsilon} (x_{j,d} - x_{i,d}) \tag{16}$$

Where $F_{i,d}$ is the sum of the forces on the ith cat in the D-dimensional space, G is the gravitational constant, M_i is the passive inertial mass related to the ith cat, $r_{i,j}$ is the Euclidean distance between the ith and jth cats, the new location $x_{i,d}$ is:

$$x'_{i,d} = x_{i,d} + \frac{F_{i,d}}{M_i} \tag{17}$$

The optimized location information is shared in cats to ensure that every cat improves a poor dimension value using global location information to continuously improve population information before next iteration. Thus, the GS-NMCSO avoids premature convergence and local optimal problems, and fundamentally improves population diversity. Besides, all population positions are updated by cats with better values as the initial positions of next iteration, which avoids invalid search and effectively improves the optimization rate. The algorithm flowchart of GS-NMCSO is shown in Fig. 3.

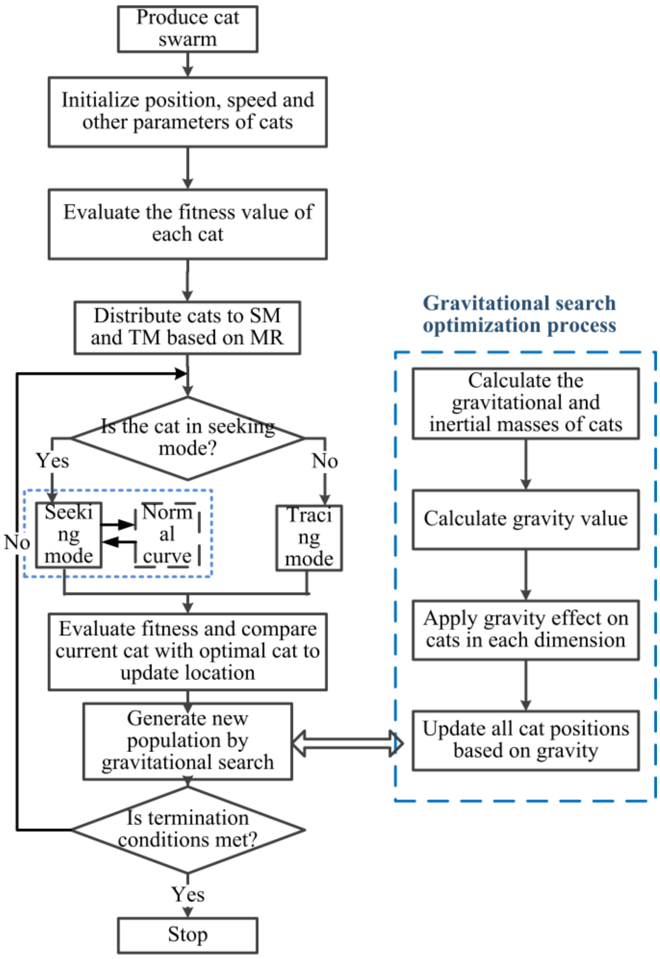

Fig. 3. The GS-NMCSO operation process flowchart.

C. The optimization process of GS-NMCSO

According to the general rule of SVM, the classification accuracy will be improved synchronously with the increase of penalty coefficient C and kernel radius parameter λ. However, if the values of parameter C and λ are too large, SVM will enter the over fitting state. Otherwise, if the value of C and λ is too small, the under fitting problem may occur, which will greatly affect classifier performance.

The classification performance of SVM is greatly affected by parameter C and λ, and the best values are hard to define. GS-NMCSO algorithm is considered to optimize SVM, which is used to find the optimal parameter values of C and λ. The fitness value of GS-NMCSO is calculated by mean square error (MSE) to measure the deviation degree between variable and mean value. The *fitness* expression is as follows:

$$fitness=\sum_{i=1}^{n}\frac{1}{n}\left(f\left(x_i\right)-\bar{f}\right)^2 \tag{18}$$

$$\bar{f}=\sum_{i=1}^{n} f\left(x_i\right)/n \tag{19}$$

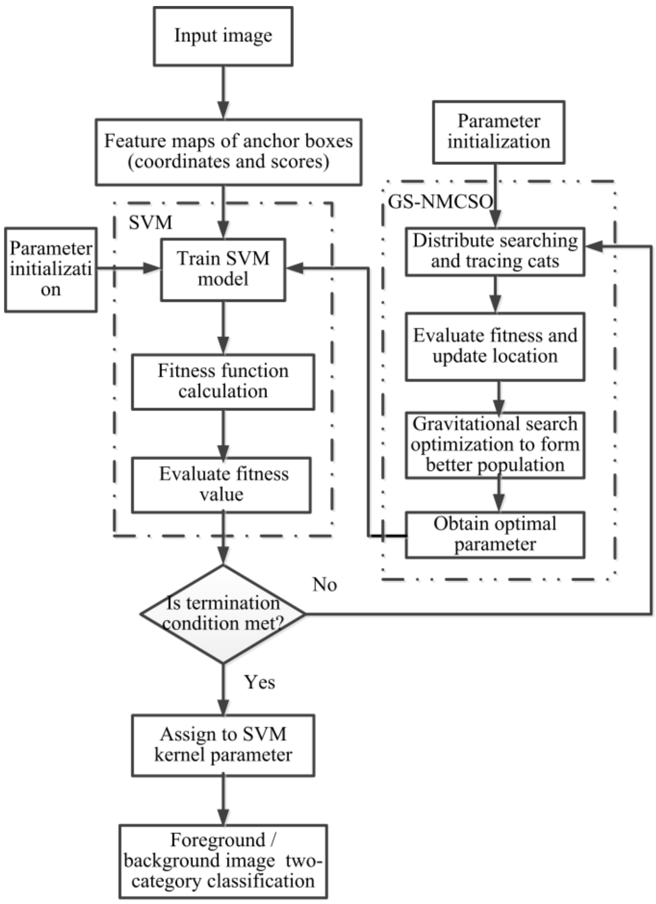

Fig. 4. Parameter optimization process of SVM based on GS-NMCSO.

The parameter optimization process of SVM as in Fig. 4 are as follows:

Step1. Initialization algorithm parameters.

Initialization parameters of algorithm include population size, maximum iteration number, PMO (Probability of Mutation Operation), CDC (Count of Dimensions to Change), and SMP (Seeking Memory Pool), initial velocity and position of cats, etc.

Step2. Define parameters of SVM.

Each cat means a set of solution for C and λ, and has two dimensions which represent two parameters. The initial position of a cat is transmitted to the parameter C and λ. The range of C and λ are [0.1, 20000] and [0.01, 50] respectively.

Step3. Prepare training data for RPN network.

RPN network randomly selects anchor box from input image to train the SVM model, and transforms images into feature maps, that is, the set of position coordinates and scores. Simultaneously, the proportion of positive samples and negative samples is close to 1:1 to ensure data balance. Because the number of negative samples is always more than positive samples, thereafter part of positive samples is padded with some negative samples.

Step4. Algorithm operation.

Seeking mode. The mutation coefficient is optimized by the normal distribution curve to enhance optimization accuracy. Through Normal mutation operation, the *i*th cat's position in every dimension (x_{id}^{m}), that is, the *i*th set of parameters C and λ, is obtained.

Tracing mode. The position $x_{i,d}$ and velocity $v_{i,d}$ of the *i*th cat in the *d*-dimension are updated continuously to search the optimal position $x_{g,d}$. The fitness of all cats is evaluated, the best cat is set to replace the current position.

Step5. Repeat optimization process.

The SM and TM process repeat until the optimum solution is obtained. After each iteration, the gravitational search operator updates all cat positions and generates the new cat positions of the entire population for next iteration.

After the optimization process completion, we select the best individual to update $x_{g,d}$.

Step6. Check termination condition and output results.

If maximum iteration number has been reached, the optimization process is stopped, and the optimum solution is output. Otherwise, steps 4-6 repeat to update the cats' positions.

Step6. Assign optimum value to SVM

By assigning the global optimal solution of GS-NMCSO to parameter values C and λ, the foreground and background anchor by RPN network are classified precisely.

IV. EXPERIMENT AND ANALYSIS

In order to test the performance of GS-NMCSO algorithm and cascaded network structure, the distribution equipment defect detection experiment is designed. Considering the preciseness and clarity requirement, the optimization experiment for GS-NMCSO and classification experiment for SVM optimized by GS-NMCSO are verified separately.

A. Image Dataset construction

The image Dataset selected in the experiment is a set of power distribution cabinet pictures with a specific resolution taken in a large electrical equipment manufacturing enterprise. The pictures include copper bus-bars, bolts, low -voltage circuit-breakers and other components. A total of 8,400 pictures were taken from 232 switchgear units in 27 groups of distribution rooms, containing 753 rectangular or rounded rectangular bus-bars and approximately 29,000 fastening bolts. These pictures were taken by the high-definition camera equipped with the laser SLAM inspection robot (EAI smart type). The Dataset includes pictures of different environments in multiple time periods (shadow, dark light, front view, squint, etc.) to ensure that the proposed method has strong generalization ability and can adapt to different conditions.

In order to build a training set for the bolt tightening status identification model for copper bus-bar in distribution cabinet, 4590 images in the collected distribution cabinet units were manually tagged in this paper. The marking task consisted of marking the bolt location with a regression box and the bolt tightening state (normal, loose, falling off) corresponding to this box, with 3500 randomly selected in the marker images into the training set and 1090 into the verification set. The training process of the improved cascaded network based on Faster-RCNN follows the multi-task loss principle. After the model training, the accuracy of verification set is used to evaluate the model reliability. That is, at every iteration interval, the accuracy of the current training model on verification set is calculated and the model is saved. After all iterations are completed, the model with the highest accuracy is obtained by observing the accuracy change curve on validation set and used for subsequent test. A testing dataset is generated to evaluate the proposed method and test the model performance for defect detection of power distribution equipment. The test image is selected from 38 power distribution cabinets responsible for UPS power conversion. Each image acquisition task is divided into four periods, 20 rounds in total, and 1187 copper bus-bar images of power distribution cabinets are obtained.

B. Evaluation standard

In this experiment, the proposed method is tested and evaluated from two aspects of average precision and processing time (Frames Per Second, FPS). The test objects include GS-NMCSO algorithm, object detection framework and the improved cascade structure. True positive (TP), false negative (FN) and false positive (FP) are important indexes of image detection. Through these indexes, precision rate (P) and recall rate (R) can be calculated.

The average precision (AP) of a single category and the mean average precision (mAP) of all categories are calculated based on the P-R relationship between precision rate and recall rate. TP refers to the set of labels correctly detected, FP refers to the set of wrong labels in the detection results, and FN refers to the set of labels not detected. The expression of the above evaluation indexes is as follows:

$$P = \frac{TP}{TP + FP} \times 100\% \tag{20}$$

$$R = \frac{TP}{TP + FN} \times 100\% \tag{21}$$

$$AP = \int_0^1 P(R)d(R) \tag{22}$$

$$mAP = \frac{1}{m}\sum_{k=1}^{m}(AP)_k \tag{23}$$

C. Initialize parameters setting

Considering the configuring convenience on environment parameters, Caffe is used to train and verify the improved cascaded network in this experiment. The experimental PC was configured as: Core i7 CPU, 3.6 GHz basic frequency, 32GB RAM, and the operating system is Ubuntu 14.04. The specific training environment parameters are shown in Table I. In addition, the initialization parameters of GS-NMCSO algorithm are shown in Table II, the initialization parameters of Faster-RCNN are shown in Table III.

D. Experimental results and discussion

The model trained in this paper shows excellent results on the bolt fastening state detection of positioning copper bus-bar. Fig. 5 is the visual result of images defect detection.

TABLE I

EXPERIMENTAL ENVIRONMENT PARAMETERS

Definition	Parameters
operating system	Ubuntu 14.04
deep-learning framework	Caffe
language	Python, C++
CPU	Intel Core i7-9700k 3.6GHz*8
GPU	NVIDIA RTX2060（8GB RAM）
RAM	32.00GB

TABLE II

INITIALIZATION PARAMETERS OF GS-NMCSO ALGORITHM

Definition	Parameter values
population	100
Dimension (*D*)	2
SMP	2.5
SRD	0.15
CDC	75%
PMO	0.3
w	0.1~0.9
c	2
r	[0,1]
ε	0.01
G	100* e^{-20}

TABLE III

INITIALIZATION PARAMETERS OF FASTER-RCNN

Definition	Explain	Parameter values
base_lr	initial learning rate	0.01
lr_policy	drop the learning rate in steps by a facter of gamma	Step
Gama	facter of droped learning rate of *lr*	0.1
Stepsize	attenuation step of *lr*	5000
test_interval	maximum iteration number to test	1000
batchsize	number of images per iteration	256
IoU_FG_RPN	IoU value of RPN foreground proposals	[0.6,1]
IoU_BG_RPN	IoU value of RPN background proposals	[0,0.3]
NMS_IoU	threshold value of IoU for NMS method	0.7
momentum	previous update weight	0.9
Weight_decay	facter of regularization	0.0005
Max_iter	number of execution steps	10000

GS-NMCSO Algorithm Verification

Aiming at the optimization effect of GS-NMCSO on SVM kernel function parameters, several swarm intelligence optimization algorithms are applied to compare and verify the iteration speed and accuracy. The fitness function is set as the mean square error of individual output, the minimum fitness value represents that swarm have obtained the global optimal solution.

Several frequently-used algorithms have been introduced for comparison, including classic cat swarm optimization (CSO), quantum particle swarm optimization (QPSO) and Cauchy mutation cat swarm optimization (CMCSO).

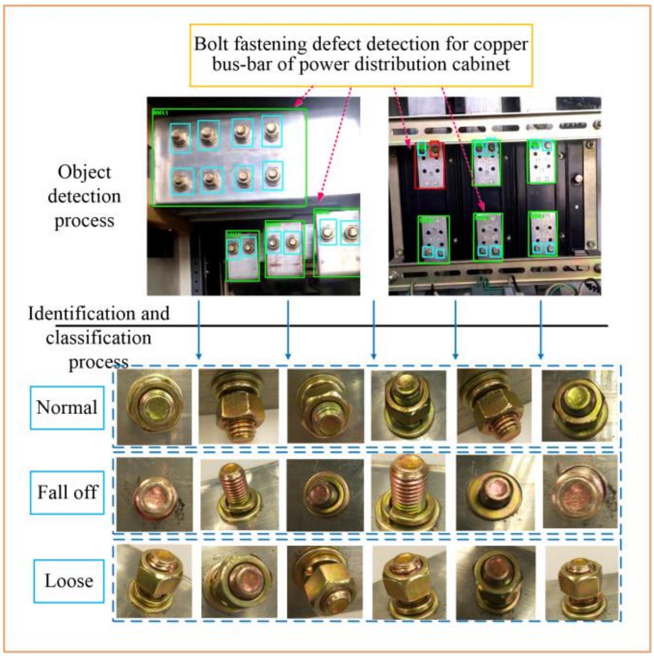

Fig. 5. Visual diagram of bolt fastening defect detection for copper bus-bar of power distribution cabinet.

The individual evaluation dimension is 2, which represents a set of parameters (C, λ), the search scope of C and λ are [0.1, 20000] and [0.01, 50] respectively. The maximum iteration number is 100. Due to the algorithm randomness, the four algorithms perform 20 independent operations respectively to optimize the SVM kernel function parameters, take the optimal value and the mean value for comparison.

The optimization curve of GS-NMCSO, CMCSO, CSO and QPSO algorithm is shown in Fig. 6. Table IV shows the optimal values and averages obtained after 20 separate experiments, cumulatively generating 2000 sets of SVM kernel function parameters (C, λ) data. The experimental analysis is mainly carried out from the aspects of iteration speed, optimal value and corresponding fitness value (MSE). The x-axis represents the algorithm iteration number, and the y-axis represents the value of MSE. It can be seen from Fig. 6 that as the number of iterations increases, the GS-NMCSO and CMCSO algorithm curves have a steep downward trend, rapidly decline after the 5th iteration, and gradually converge after the 10th iteration. Besides, the GS-NMCSO curve achieves global optimization at 13th iterations, slightly faster than CMCSO at 15th. In addition, the CSO curve converges significantly slower and shows a slow decreasing trend as the iterative process continues, and is always differentiated from the GS-NMCSO and CMCSO curves. The global optimal solution of CSO was not found until about the 70th iteration. The QPSO curve begin to decline after the 5th iteration, but the trend was always relatively gentle at this stage, converges to the global optimization at the 28th iteration. However, the optimal fitness value is significantly greater than the other three algorithms, and QPSO is the lowest in terms of optimization accuracy.

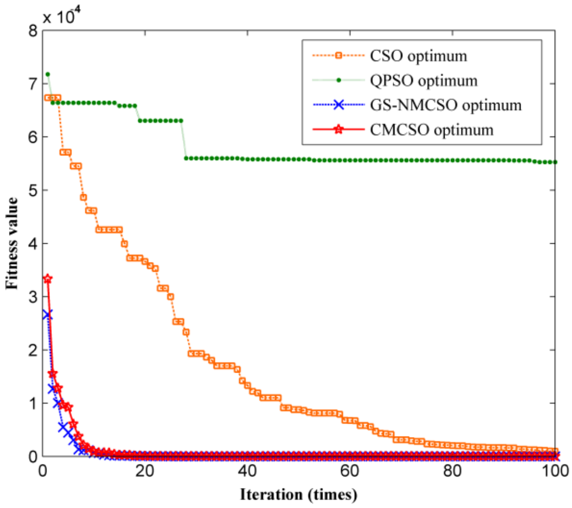

Fig. 6. The fitness curve based on CSO, QPSO, CMCSO and GS-NMCSO algorithm.

TABLE IV

OPTIMAL FITNESS VALUES OF SEVERAL ALGORITHMS

Algorithm	Optimum value	Mean value
GS-NMCSO	0.0245E-04	0.1284E-04
CMCSO	0.0473E-04	0.2057E-04
CSO	0.2822E-04	0.9183E-04
QPSO	5.4928E-04	6.0416E-04

As shown in Table IV, the optimal and mean fitness values of GS-NMCSO algorithm are smaller than those of CMCSO, CSO and QPSO algorithm, with the fastest convergence speed and the highest accuracy, so the proposed algorithm is better than CMCSO, CSO and QPSO in terms of speed and accuracy, and is optimal for SVM kernel function parameter optimization. The reason why GS-NMCSO has excellent optimization effect is that it makes use of the advantages of normal distribution curve to ensure that the updated position is always distributed in the perception range of parent cats in the searching stage so as to improve the search accuracy. After a round of iteration, through the set of cats with better fitness value, the initial position of the cats is updated before the next iteration on gravitational effect of better population to improves the global optimization ability, meanwhile, GS-NMCSO operation process is more complex and requires higher computing resource compared with other several algorithms, because the algorithm includes especial normal mutation and gravitation optimization process.

Object detection performance verification

In this section, the performance of detect detection for power distribution equipment is verified and analyzed. The Faster-RCNN based on region proposal and SSD based on regression are compared to prove the detection capability of the improved cascaded network. In addition, YOLO is also an important end-to-end detection network, which has very fast operation speed and requires low computing resources, but the number of default output boxes is relatively few and the accuracy of the model is relatively low, so it is not used as the contrast object for the small target detection experiment in this paper. The section also uses two traditional methods of manual feature extraction + machine learning classification algorithm, that is, HOG feature+Ada Boost classifier and Deformable Part Model (DPM), to verify the results on the identification of detection of power distribution cabinet. The power distribution cabinet is cube equipment. This experiment need to ensure that the captured image contains the detection target component as much as possible. The original image size is set as 4096×2160 pixels, and then unified scaling to 409×216 pixels to reduce the computational complexity. The component to be detected is bolt, which is large in number and small in size, mostly in the range of 40×40 to 90×90 pixels. In this paper, the improved cascaded structure, in a global to local way, extracts the foreground image feature from the RPN region proposal, and then locates the bolt position and judges the fastening state.

The P-R curve in training stage is drawn to show the detection effect on different object detection algorithms. As shown in Fig. 7 and Table V, the detection precision of improved cascaded network by GS-NMCSO, Faster-RCNN and SSD networks are significantly higher than that of DPM and HOG+Ada Boost methods. The mAP values of these three deep-learning methods are over 70%, which is much higher than 45.89% of DPM and 10.82% of HOG+Ada Boost. The improved cascaded network adds SVM after RPN network, which significantly improves the proportion of foreground image in region proposal process. Thus the method can improve the detection precision and save the training time. The mAP is up to 89.51%, which is higher than 80.73% of basic Faster-RCNN and 70.28% of SSD network. The detection speed of the several methods is also important indicator. SSD is the fastest among all the algorithms in this experiment because of the end-to-end training advantage, reaching 26 frame/s, the application of multi-scale feature map, convolution detection, priori anchor box setting for SSD also ensure high accuracy, and the detection speed of improved cascaded network and basic Faster-RCNN is 5.2 frame/s and 6.8 frame/s respectively, which are very close to each other and can meet the requirements of normal video image detection. Besides, the detection speed of DPM and

HOG+Ada Boost is relatively slower, only 0.7 frame/s and 1.34 frame/s.

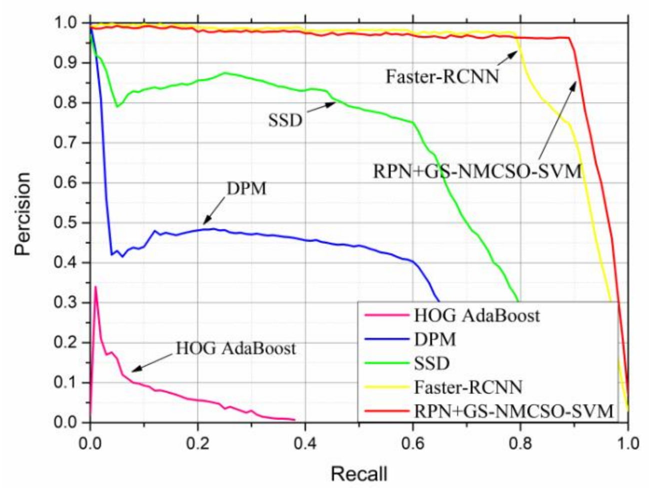

Fig. 7. The P-R curve comparison of several object detection algorithms.

TABLE V

COMPARISON ON PRECISION, FPS AND TRAINING TIME FOR SEVERAL OBJECT DETECTION ALGORITHMS

Algorithm	Detection precision mAP	Detection speed FPS (frame/s)	Training time (hour)
Improved cascaded network by GS-NMCSO	89.51	5.2	63
Faster-RCNN	80.73	6.8	128
SSD	70.28	26	104
DPM	45.89	0.7	161
HOG+Ada Boost	10.82	1.34	77

In addition, the training time of algorithm is also important evaluation indicator, which represents the requirements for basic computing resources. The improved cascaded network takes the shortest time, only 63 hours per training, which is significantly ahead of other methods. This is due to the addition of GS-NMCSO algorithm-optimized SVM network after RPN, which classifies the output features in foreground and background, and only feeds the image containing the target into the subsequent process, saving arithmetic power and facilitating rapid convergence of the network, resulting in the shortest training time. In general, the problem to be solved is small target recognition in power distribution equipment, that is, to detect the 40×40 pixel bolts in the whole image of power distribution cabinet and accurately classify the normal, loose and falling off states for bolt fastening. This application requires the highest detection precision, and a shorter training time to form a daily operation schedule. The cascaded network on GS-NMCSO has the highest mAP (89.51%), the shortest training time (63 hours), and the detection speed (5.2 frame/s) can meet basic needs. It has a good performance in the speed and accuracy of defect detection of distribution equipment. Compared with other methods, the improved cascaded network performs better in defect detection of power distribution equipment.

Through the above analysis, this section verifies the performance of the improved cascaded network for defect detection on bolt fastening state, and then proceeds to the testing process. The 20 independent experiments in total were carried out on the test Dataset (1187 images), in order to ensure the missed rate for defect detection, that is, recall rate is as high as possible. The threshold value of score is 0.75, and the best mAP is 90.17%, recall rate is 91.22%, False positives rate is 9.83%, and missed detection rate is 8.78%.

Cascaded network structure evaluation

For validating the advantages of RPN+SVM-GS-NMCSO network, this improved cascaded structure is compared with two models of basic Faster-RCNN and RPN+softmax classifiers to test the defect detection precision of power distribution cabinet. In order to verify the complex target detection ability of several network structures, 1200 power distribution cabinet images containing more than 40 bolt components are selected in this section for testing Dataset.

The precision comparison of the three network structures is shown in Fig. 7. The basic Faster-RCNN is to locate and classify target directly in feature maps, and its mean average precision (mAP) is the lowest among the three structures, only 78.79%. The mAP value of RPN+SVM cascaded network proposed in the paper is 83.89%, and that of RPN + softmax cascaded network is 80.25%. The reason for the difference in precision is that the object detection framework is difficult to directly locate each small bolt target with 40×40 pixels in the original image of 4096×2160 pixels. The region proposals of RPN network inevitably contain much background images, this phenomenon will cause background interference to subsequent model training. Moreover, the bolt component has no obvious feature expression in complex background image, even if cascaded structure is used, it is difficult to correctly detect such small target as bolts. Besides, the original high-definition image is scaled to 409×216 pixels and then imported into the network to save computing resources, which also lose some precise features of bolt image and reduce the average precision of object detection network.

According to the above experimental results, the cascaded structure on RPN+classifier proposed in this paper is very necessary to execute defect detection for small bolts from the high-definition image of power distribution cabinet. Next, the effect of introducing SVM or softmax after RPN continues to be compared. It can be seen from Fig. 8 that the precision of RPN+SVM-GS-NMCSO cascaded network is higher than that of RPN+softmax cascaded network in terms of several detection states, normal, loose and fall off. Softmax is a multi-classification function, which is suitable for the output of multi-task learning problems, and RPN output only needs to distinguish between foreground and background, which is a typical two-category problem, so SVM is more suitable for foreground selection of RPN output image. Meanwhile, the kernel function parameters of SVM is optimized by GS-NMCSO, the RPN+SVM cascaded network has stronger learning ability and better effect on the bolt fastening state detection for copper bus-bar of power distribution cabinet.

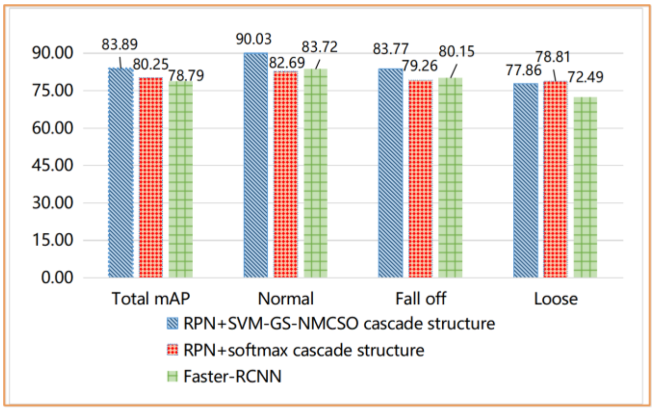

Fig. 8. Precision comparison between RPN+SVM-GS-NMCSO cascaded structure, RPN+softmax cascaded structure and basic Faster-RCNN.

For the three cases of normal, falling off and loose, the RPN+SVM-GS-NMCSO cascaded network has a high recall rate on the bolt-fasteners for various types and sizes of copper bus-bars. However, some of the normal fastening bolts are also detected as loose, and the model has a certain false positive rate. The precision of the three networks is the highest when the bolts are in normal tightening state, while the precision of the loose state is the lowest. This is because normal fastening and loose are very similar in front images, and the feature differences can only be reflected with a certain angle. The loose feature is the least obvious as the pre-state of defect occurrence, so the recognition precision for loose is lowest. As the defect detection system for power distribution cabinet requires to reduce the missing detection rate as much as possible, a certain proportion of false positive is acceptable.

V. Conclusion

Based on the actual needs of defect detection on power distribution cabinet, this paper presents a method to improve the recognition precision and operating efficiency for bolt fastening status of copper bus-bar as falling off, loose and other defect states, and optimize the deep-learning framework Faster-RCNN. A defect detection method on power distribution equipment using RPN+SVM cascaded network improved by GS-NMCSO algorithm is proposed. The penalty parameters C and λ of SVM are optimized by introducing a novel GS-NMCSO algorithm. The foreground/ background two-category classification of RPN output image is realized by SVM-GS-NMCSO, and only the foreground image is sent to the subsequent network for training, so as to improve the precision of small object detection and save training time and computing resource. The proposed approach shows a promising application and accuracy in the defect recognition of power distribution cabinet. The reduced time consumption and improved detection precision makes it feasible to automatically detect electrical equipment in distribution room using intelligent inspection robot platform.

The main contributions of this paper include:

--First, a RPN+SVM-GS-NMCSO cascaded structure is proposed based on the Faster-RCNN model, a two-category classification model SVM is cascaded after RPN network to identify and eliminate the interference of background image, which makes the learned features more accurate and precisely identifies small objects such as bolts. In the 8400 images Dataset of power distribution cabinet constructed in this paper, the mAP reaches 89.51%, which is significantly higher than basic Faster-RCNN, SSD, DPM and HOG+Ada Boost, the detection speed is up to 6 frame/s, and training time is only 66 hour. The recognition accuracy is ensured and the training efficiency is improved at the same time. Compared with basic Faster-RCNN and RPN+softmax models, the mAP of RPN+SVM-GS-NMCSO cascaded network in the image Dataset of multi-component power distribution cabinet reaches 83.89%, and the precision is the best for the bolt recognition in several states of the normal, falling off and loose.

--Second, The GS-NMCSO algorithm is presented and used to find the optimal parameter C and λ values of SVM. The fitness value of GS-NMCSO is calculated by mean square error (MSE) to measure the deviation degree between variable and mean value. Taking the MSE of SVM model as the fitness function, the optimal and average fitness values of GS-NMCSO are better than those of CSO, QPSO and CMCSO algorithms with the fastest convergence speed (only 13 iterations to achieve convergence). The GS-NMCSO algorithm have the best effect on SVM kernel function parameter optimization.

In the future, we will apply the improved cascaded network structure and GS-NMCSO algorithm to more components on defect detection in the power distribution cabinet to further improve the object detection ability .

Acknowledgment

No potential conflict of interest was reported by the authors. The authors would like to thank the editors and the anonymous reviewers for helpful comments and suggestions.

References

[1] G. Wang, J. Guo, Y. Chen, Y. Li and Q. Xu, "A PSO and BFO-Based Learning Strategy Applied to Faster R-CNN for Object Detection in Autonomous Driving," *IEEE Access*, vol. 7, no. 8, pp. 18840 -18859, 2019.

[2] J. Chen, Z. Liu, H. Wang, A. Núñez and Z. Han, "Automatic Defect Detection of Fasteners on the Catenary Support Device Using Deep Convolutional Neural Network," *IEEE Transactions on Instrumentation and Measurement*, vol. 67, no. 2, pp. 257-269, 2018.

[3] Da Silva, Giovanni Lucca, et al. "Convolutional neural network-based PSO for lung nodule false positive reduction on CT images." *Computer Methods and Programs in Biomedicine*, vol. 162, no. 10, pp. 109-118, 2018.

[4] Chen T , Zhang L , Lin H , et al. "Quantitative analysis of chromium in pork by PSO-SVM chemometrics based on laser induced breakdown spectroscopy." *Journal of Analytical Atomic Spectrometry*, vol. 34, no. 5, pp. 884-890, 2019.

[5] B. Wang, Y. Sun, B. Xue and M. Zhang, "Evolving Deep Convolutional Neural Networks by Variable-Length Particle Swarm Optimization for Image Classification," 2018 IEEE Congress on Evolutionary Computation (CEC), 2018, Rio de Janeiro, Brazil, pp. 1-8.

[6] Y. Chin, Y. Hsieh, M. Su, S. Lee, M. Chen and J. Wang, "Music emotion recognition using PSO-based fuzzy hyper-rectangular composite neural networks," *IET Signal Processing*, vol. 11, no. 7, pp. 884-891, 2017.

[7] He Xia, Tangyi Ping, Yuangong Pin, et al. "Study on Bayonet Recognition Engine Based on Cascade Multitask Deep Learning," *Computer Science*, vol. 46, no. 1, pp. 310-315, 2019.

[8] Krizhevsky, Alex , I. Sutskever , and G. Hinton . "ImageNet Classification with Deep Convolutional Neural Networks." NIPS'12: Proceedings of the 25th International Conference on Neural Information Processing Systems, 2012, New York, USA, pp. 1097-1105.

[9] S. Ren, K. He, R. Girshick and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1137-1149, 2017.

[10] R. Girshick, "Fast R-CNN," 2015 IEEE International Conference on Computer Vision (ICCV), 2015, Santiago, USA, pp. 1440-1448.

[11] J. Redmon, S. Divvala, R. Girshick and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, Las Vegas, USA, pp. 779-788.

[12] Liu W. et al. "SSD: Single Shot MultiBox Detector," Computer Vision-ECCV 2016. 2016, Amsterdam, Netherlands, pp. 21-37.

[13] Pappula, Lakshman, and D. Ghosh. "Linear antenna array synthesis using cat swarm optimization." *AEU - International Journal of Electronics and Communications*, vol. 68, no. 6, pp. 540-549, 2014.

[14] Pappula, Lakshman, and D. Ghosh. "Synthesis of linear aperiodic array using Cauchy mutated cat swarm optimization." *AEU - International Journal of Electronics and Communications*, vol. 72, no. 3, pp. 52-64, 2017.

[15] Crawford B, Soto R, Berríos N, et al. "A Binary Cat Swarm Optimization Algorithm for the Non-Unicost Set Covering Problem", *Mathematical Problems in Engineering*, vol. 10, no. 1, pp. 1-9, 2015.

[16] B. Kumar, M. Kalra and P. Singh, "Discrete binary cat swarm optimization for scheduling workflow applications in cloud systems," 2017 3rd International Conference on Computational Intelligence & Communication Technology (CICT), 2017, Ghaziabad, India, pp. 1-6.

[17] Rashedi E, Rashedi E, Nezamabadi-Pour H. "A comprehensive survey on gravitational search algorithm," *Swarm & Evolutionary Computation*, vol. 41, no. 3, pp. 141-158, 2018.

[18] Dongming Zhao, Kewen Xia, Huijuan Liu, et al. "Application of GS-MOCSO Algorithm to Slow-wave Structure Design of STWT," *IAENG International Journal of Computer Science*, vol. 44, no. 4, pp. 575-586, 2017.

[19] Dongming Zhao, Kewen Xia, Huijuan Liu, Xiaoxu Shi. "A pitch distribution in slow-wave structure of STWT using Cauchy mutated cat swarm optimization with gravitational search operator," *Journal of the Chinese Institute of Engineers*, vol. 44, no. 4, pp. 297-307, 2018.

[20] Jia Y, Shelhamer E, Donahue J, et al. "Caffe: Convolutional Architecture for Fast Feature Embedding," MM '14: Proceedings of the 22nd ACM international conference on Multimedia. 2014, Orlando, USA, pp. 532-539.

Dongming Zhao received a Ph.D. degree in electronic science and technology from Hebei University of Technology in 2018 and is studying in communication and information engineering for postdoctoral fellow in Zhejiang University, His main research interests are computer vision, intelligent information processing.

Huimin Yu, professor, doctoral supervisor, is currently serving as the research leader in state key lab of CAD&CG of Zhejiang University, and is active in the study of image/video intelligent processing and analysis, computer vision and multimedia information processing.

Xiang Fang is currently serving as the assistant chief engineer in Hangzhou zhijiang switchgear Co. Ltd. of HangShen Group, her main research interests are electrical engineering and electrical equipment manufacturinge.

Lei Tian received a master's degree in Data Science Engineering from Peking University in 2007, he is a technical expert of big data application in China Mobile Communication Group Tianjin Co., Ltd, her main research interests are artificial intelligence and bigdata.

Pengqian Han received a master's degree in Aeronautical Engineering from Beihang University in 2017, he is a communication network engineer in China Mobile Communication Group Tianjin Co., Ltd, her main research interests are control technology of precooling hypersonic aeroengine.