

Tunnel cable defect detection method based on deep learning and Global-Local features

Yang Zhao¹, Yingqiang Shang¹, Tian Guo¹, Zixi Han^{2,*}, Zhaoyi Zhang², Jirong Fu² and Yongrui Zhang²

¹ State Grid Beijing Electric Power Company, China

² Institute of Next Generation Power Systems and International Standards, Wuhan University, China

Abstract

A tunnel cable fault detection method based on deep learning and global-local features is proposed to meet the requirements of tunnel cable inspection and improve the safety of cable operation. First, a method based on color uniqueness and compactness is provided to highlight the salient foreground regions in the inspection image. Then, according to the relationship between the global spatial scenario distribution and local information, a global context model and a local fine detection model are established to deeply and accurately compute the salient features of the image. A cyclic structure network is used to weigh the position of each feature map. Next, a cyclic connection is then established by feeding the output of each block model back to the input. Repeated iteration and noise filtering reduce the influence of background information. Finally, the advantages of the proposed scheme in terms of training time and recognition accuracy were verified by comparing it with four classical target recognition algorithms in the case study.

Keywords

tunnel cable inspection; deep learning; local fine feature

1. Introduction

With the gradual improvement of urban underground distribution networks, the construction and maintenance of cable tunnels play an increasingly important role in the development process of new power systems. However, the natural geographical environment of cable tunnels leads to problems such as water seepage, moisture, cable aging, and equipment corrosion and damage [1,2]. At present, the maintenance of cables and their supporting equipment in tunnels mainly relies on manual work. The large amount of thick smoke, harmful gases, high-voltage cable leakage, and cable tunnel collapse may pose a threat to the safety of workers.

Owing to the rapid development of robotics and image recognition technology, using inspection robots to detect the internal situation of cable tunnels has gradually become an

ICCIC 2024: International Conference on Computer and Intelligent Control, June 29–30, 2024, Kuala Lumpur, Malaysia

* Corresponding author.

✉ 15101691629@163.com (Y. Zhao); 819900413@qq.com (Y. Shang); guotian124com@163.com (T. Guo); 2018302070047@whu.edu.cn (Z. Han); zhangzhaoyi@whu.edu.cn (Z. Zhang); 2020302191016@whu.edu.cn (J. Fu); 2020302191606@whu.edu.cn (Y. Zhang)

 0009-0007-7289-5498 (Z. Han)



© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

effective solution. In recent years, many domestic and foreign companies and research institutions have developed tunnel inspection robots for different application fields. The water tunnel robot developed in [3] can effectively repair the inner wall of the tunnel, improving construction efficiency and work safety. Ref. [4] developed an underground cable detection robot called "Patrol", which can enter the underground and crawl along the cable to find the fault location. At present, most researchers focus on improving the structure of tunnel robots and researching their path planning and obstacle avoidance, with relatively little research on the intelligent functions of robots. Specifically, under the complex and humid internal environment of cable tunnels, a device fault identification scheme for cable tunnel inspection robots not only has significant significance for the construction and maintenance of the tunnel but also plays a good role in ensuring the safety of maintenance personnel inside the tunnel.

In recent years, deep learning has been increasingly widely applied in the field of image recognition. Ref. [5] trains two deep neural networks to calculate pixel features and global recommendations respectively, and comprehensively judges the salient regions of the image. Ref. [6] proposes dense and sparse reconstruction (DSR), which first uses a significance calculation model for reconstruction errors, calculates the entropy and sparsity of each region of the image to reconstruct errors, and uses the K-Means algorithm to obtain the significance of reconstruction error clustering. Ref. [7] obtained more accurate saliency map inference by combining prior knowledge of saliency. Ref. [8] proposed a multi-task deep learning algorithm that applies the Laplacian nonlinear model to a significantly enhanced regression model, which is more commonly used in semantic segmentation and multi-object task detection. Ref. [9] proposes a bidirectional learning framework to aggregate multi-level convolutional features of salient object detection (SOD) models and improve the robustness and accuracy of saliency detection.

Based on the above research, this paper proposes a cable tunnel defect detection method based on deep learning and global-local features, fully considering the relationship between local fine features and global spatial scene distribution, to improve the accuracy of salient feature map detection. A cyclic structure network module was proposed to continuously collect contextual information and iteratively improve convolutional features. The proposed method can effectively identify three common faults of tunnel cable insulation air gaps, insulation scratches, and insulation surface impurities. The implementation of the proposed tunnel cable defect detection method provides a scalable intelligent application model for tunnel inspection robot systems, thus laying the foundation for the further development and implementation of other related applications.

2. Composition of robot system

Power cables are increasingly favored by governments and power departments in urban power lines, therefore, the application of laying power cables as a power supply method in power construction is becoming more and more common. But high-voltage power cables are installed in underground passages (pipe galleries or tunnels), and cable tunnels refer to corridors or tunnel like structures used to accommodate a large number of cables, which can provide sufficient protection for the cables. China's cable tunnel technology began in the

1980s and has undergone decades of development. Currently, complex urban cable tunnel networks are built in major cities in China. Many cable tunnels face problems such as collapse, seepage, water accumulation, fire, cable smoldering, aging and rusting of power supporting facilities, animal invasion, and accumulation of toxic and combustible gases on the front. To ensure the safety of cables inside the tunnel and maintain the stable operation of the entire power grid, it is necessary for staff to regularly inspect the cable tunnel. However, due to the long history of construction and the harsh environment inside old-fashioned cable tunnels, inspectors not only need to face the harsh environment of high temperature and humidity, but also constantly face the threat of landslides and toxic gases during the inspection process. All of these factors pose a serious threat to the life safety of cable tunnel inspectors. At present, there are many online monitoring devices or track type inspection robots installed in cable tunnels. The online monitoring devices in tunnels mostly use wired communication connections, which makes construction difficult before or during the technical transformation process of the tunnel. For example, distributed fiber optic temperature measurement requires laying fiber optic cables along the entire length of the cable body, while other monitoring methods only set monitoring equipment at the attachment, which is low in cost-effectiveness and difficult to ensure monitoring accuracy; Orbital robots need to carry out track construction inside tunnels, which is costly, requires a long construction period, poses a great threat to personnel in harsh environments, and is prone to damage to tunnel structures.

The working system of the cable tunnel inspection robot includes an algorithm workstation, a ground station, a cable tunnel inspection robot, and a camera and sensor mounted on it, as shown in Fig. 1. The equipment carried by the robot includes visible light cameras, thermal imaging cameras, gas sensors, relay boards, WiFi boards, gimbals, and lithium batteries. To eliminate the problem of limited wireless signal transmission in the complex environment of cable tunnels, this system converts the local area network at the bottom of the cable tunnel into high-power 5.8GHz WiFi to communicate with ground stations.

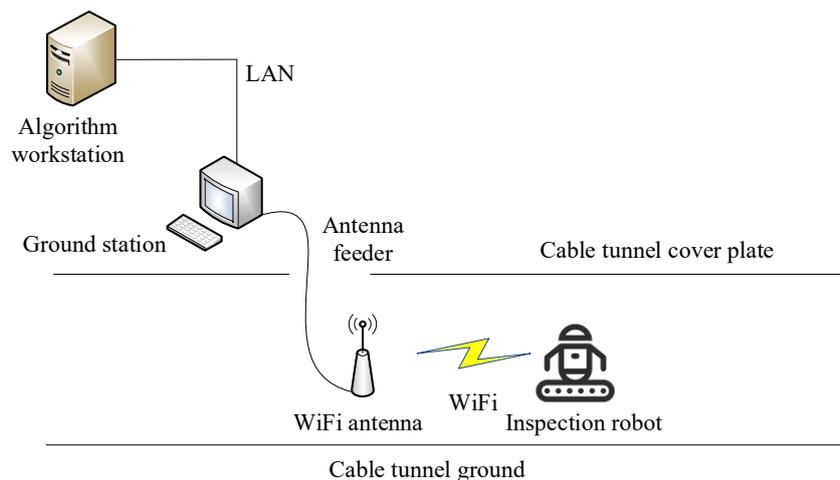


Figure 1: Composition of cable inspection robot system.

3. Cable defect detection combining deep learning and global-local features

This article first uses color uniqueness and color compactness to calculate the salient foreground region of an image, to preserve the structural information of salient regions and the contours of salient objects. Then, based on a convolutional neural network (CNN), a saliency detection network is proposed within the same deep learning framework fully considering the relationship between local fine features and global spatial scene distribution. Using a module with contextual relationships to weigh the position of each feature map, the proposed method detects local features of salient objects and improves the salient feature map, based on the relationship between each pixel and adjacent regions. Finally, a cyclic structure network module is provided by feeding the output of each block back to the input to establish a cyclic connection, and iteratively filtering out background noise to obtain salient objects.

3.1 Detection for significant foreground features

According to the basic principles of the human visual attention system, using visual cues such as color uniqueness and color compactness can better perform saliency detection. Color uniqueness, as a measure of contrast, mainly describes the differences in color and brightness between different regions. Color compactness, as a spatial distribution feature, typically results in a compact spatial distribution of salient objects, while the background area is more widely distributed throughout the entire image.

This article combines color uniqueness and color compactness to calculate foreground salient features. Firstly, given image I , a unique mapping S_U is obtained based on center priors through positional information. The compactness feature can be decomposed into images using a Gaussian mixture model (GMM), and then the saliency map S_C can be obtained through feature clustering calculation. Unlike the method of selecting the best function from uniqueness [10,11], this paper uses a simple multiplication of S_C and S_U to obtain a composite salient foreground region.

$$S_{FG} = S_C \cdot S_U \quad (1)$$

3.2 Global Context Detection Model

Firstly, the $256 \times 256 \times 3$ sized image after superpixel segmentation is taken as input, where the three dimensions represent width, height, and channel, respectively. The algorithm in this article consists of 5 convolutional layers and 2 fully connected layers. *Conv* represents the convolutional layer, the *pool* represents the pooling layer, and *fc* represents the fully connected layer. The *conv* layer and the *fc* layer are composed of linear transformations, and the *relu* layer has a non-linear correction unit function. Only the *conv* layer and the *fc* layer have learnable parameters. For the transformation layer, the size of the feature map is defined as width \times height \times depth, where the first two dimensions describe the spatial size and depth refers to the number of channels. The last layer of the network structure has two

neurons, which use the Softmax function as the output to represent the probability that the detected superpixels belong to the background or belong to salient objects.

3.3 Local fine detection model

The upper branch calculates the approximate salient region based on global contextual relationships, while the lower branch is used to detect local details of salient objects. The local fine detection model adopts inputs similar to the global context detection model, which are then normalized to $227 \times 227 \times 3$ sized images. The upper and lower branch models share the same deep structure and have independent parameters [12]. By estimating the significance probability score, the output window is predicted. The calculation formula is:

$$S_{\text{score}}(x_{\text{gc}}, x_{\text{lc}}) = P(y=1 | x_{\text{gc}}, x_{\text{lc}}; \theta_1) \quad (2)$$

where x_{gc} and x_{lc} represent the output of the global context detection model and the second to last layer of the local fine detection model. y represents the significance prediction of superpixels, where $y=1$ when superpixels are salient regions and $y=0$ when superpixels are background.

This article uses the minimum result and the maximum loss between segmentation labels in the last network layer to classify salient and background features, as shown in (3).

$$L(\theta; \{x_{\text{gc}}^i, x_{\text{lc}}^i, y\}_{i=1}^m) = -\frac{1}{m} \sum_{\substack{i \in \{1, \dots, m\} \\ j \in \{0, 1\}}} \{y^{(i)} = j\} \log P(y^{(i)}) \quad (3)$$

$$y^{(i)} = j | x_{\text{gc}}^i, x_{\text{lc}}^i, \theta_j$$

The parameters in the algorithm can be decomposed into several parts. In Eq. (2), $\theta_j = \{\mathbf{w}_{\text{gc},j}, \mathbf{w}_{\text{lc},j}, \alpha, \beta\}$. $\mathbf{w}_{\text{gc},j}$ and $\mathbf{w}_{\text{lc},j}$ represent the last layer parameters of the global context and local fine detection model, respectively. α, β is the proportion of global and local modeling weights. Label probability is calculated via (4).

$$P(y=j | x_{\text{gc}}^i, x_{\text{lc}}^i, \theta_j) \propto \mu(x_{\text{gc}}, x_{\text{lc}}, \theta_j^\mu), \quad (4)$$

$$\theta_j^\lambda = \mathbf{w}_{\text{gc},j}, \theta_j^\mu = \{\mathbf{w}_{\text{gc},j}, \mathbf{w}_{\text{lc},j}, \alpha, \beta\}, j \in \{0, 1\}$$

where μ represents the significant probability of modeling the global background and local environment, i.e., $\mu(x_{\text{gc}}, x_{\text{lc}}, \theta_j^\mu) \propto e^{f(\alpha \mathbf{w}_{\text{gc},j}^T x_{\text{gc}} + \beta) \mathbf{w}_{\text{lc},j}^T x_{\text{lc}}}$.

The corresponding non-standardized significance prediction score function is represented as:

$$f(x_{\text{gc}}, x_{\text{lc}}, \theta_j^\mu) = \mathbf{w}_{\text{gc},1}^T x_{\text{gc}} + f_u(\alpha \mathbf{w}_{\text{gc},j}^T x_{\text{gc}} + \beta) \mathbf{w}_{\text{lc},j}^T x_{\text{lc}} \quad (5)$$

where $f_u(t) = \begin{cases} t, & 0 \leq t \leq 1 \\ 0, & \text{otherwise} \end{cases}$. $f_u(\alpha \mathbf{w}_{gc,j}^T \mathbf{x}_{gc} + \beta)$ is the range of saliency detection

results obtained from multiple global context detection models. The range of $\alpha \mathbf{w}_{gc,j}^T \mathbf{x}_{gc} + \beta$ in (5) is $(-\infty, 0] \cup [1, +\infty)$. $f_u(\alpha \mathbf{w}_{gc,j}^T \mathbf{x}_{gc} + \beta)$ has high confidence in saliency in the global model. It is assumed that in the local fine detection model, Eq. (6) contains non-zero processing.

$$f_u(\alpha \mathbf{w}_{gc,j}^T \mathbf{x}_{gc} + \beta) \mathbf{w}_{lc,j}^T \mathbf{x}_{lc} \quad (6)$$

3.4 Circular network design

CNN can detect from low-level visual features to high-level features, so the simple combination of convolutional features can allow noise to propagate to the prediction layer without limitation, or may lead to the loss of some information during the information propagation process. Convolutional recurrent structures can better explore and process environmental information [13,14].

This article proposes a reweighting mechanism based on the initial architecture to adjust the transmitted features, using a cyclic structure to perceive situational information and connecting the output of each block to the input. By using the same cyclic connection and sharing weights multiple times in each layer, the new architecture can increase the depth of traditional CNN layers without significantly increasing the total parameters. Connect the calculated salient region to the first input module in the loop structure, and the downsampling layer after the feature map f^k is generated by the k -th block. A convolutional layer of kernel size m is used to slide on the local feature space window of $m \times m$. Normalize the L_2 layer and finally connect the activation layer to form features. The detailed steps are as follows.

Firstly, to calculate the weighted mapping map M^k , convolution operations need to be performed on the subsequent output channels, where W represents the kernel, and b represents the bias parameter. Each value in the obtained weighted mapping determines the importance of each spatial position.

Then, the Softmax operation is applied spatially to M^k to obtain the final weighted mapping via (7).

$$\lambda^k(x, y) = \frac{\exp(M^k(x, y))}{\sum_{(x', y')} \exp(M^k(x', y'))} \quad (7)$$

where $\lambda^k(x, y)$ represents the normalized response value at (x, y) . k is the number of blocks. If pixel i is significant at position (x, y) , then assign higher significance values to the relevant regions of that pixel in the weighted map.

Finally, the weighted mapping is upsampled to obtain the features $f_i^k(c) = \lambda_v^k \times f_{i-1}^k(c)$ where c represents the number of feature channels. λ_v^k is shared across all channels in f_i^k .

4. Simulation analysis

To verify the advantages of the proposed scheme in terms of training speed and recognition accuracy, four classic object detection and recognition algorithms, i.e., the classic CNN algorithm [15], SVM algorithm [16], BPNN algorithm [17], and ELM algorithm [18], are provided for comparison. These methods are respectively applied to defect recognition of internal equipment in cable tunnels in a city in China. The main comparative indicators of the simulation include network training time and algorithm recognition accuracy.

4.1 Network training time

Train the images collected in cable tunnels using 5 different algorithms mentioned above. To minimize the impact of network parameters on the training speed of the five methods, the small batch gradient descent method was used during the training process, and the learning rate was uniformly set to 0.1. Meanwhile, the number and resolution of images in all training datasets remain consistent. Specifically, due to the instability during the training process, which may result in a deviation in the required time for each training session, the average of 10 training sessions for 5 algorithms is taken as the final result. The specific training time comparison is shown in Fig. 2.

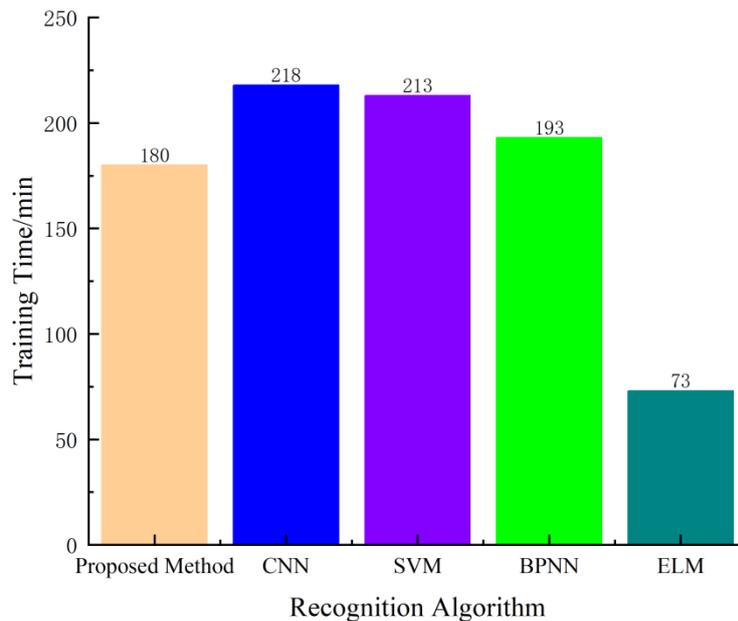


Figure 2: Comparison of network training time for 5 methods

As shown in Fig. 2, the proposed algorithm has a significant reduction in network training time compared to CNN, SVM, and BPNN algorithms. This is because the convolutional neural network has obtained ideal weights and biases through pre-training, making it easier for the network to converge during secondary training. The ELM method has the shortest training time because it only needs to update the weights between the

hidden layer and the output layer. The entire training process is the process of solving the MP generalized inverse matrix, which is easier to implement compared to the network training process in other methods.

4.2 Algorithm recognition accuracy

Using 5 algorithms to identify defects in 100 images collected in cable tunnels, the recognition results include 4 possibilities, i.e., defects in the image that were correctly identified, defects in the image that were not detected, defects in the image that were not detected, and defects in the image that were not detected but were mistakenly identified. The first two situations indicate accurate recognition, while the latter two situations indicate the presence of misidentification. The mathematical expression for recognition accuracy is the quotient of the number of correctly recognized images and the total number of images, as shown in Table 1.

It is worth noting that due to the uncertainty of all five methods mentioned above, the results of identification accuracy in Table 1 are the average of 20 experiments. Based on the statistical data in the table, the five methods have higher recognition accuracy for insulation scratches, and the specific reason may be related to the collected internal images of the cable tunnel itself. Because the calibrated training and testing image dataset contains more cable equipment images, the characteristics of cable insulation scratch images are easier to detect and collect. In addition, compared to the four traditional recognition algorithms, the algorithm proposed in this paper shows higher recognition accuracy. This is because the method fully considers the relationship between local fine features and global spatial scene distribution, achieving accurate calculation of significant image features, and thus achieving higher recognition accuracy in the final cable tunnel defect recognition application.

In summary, although the proposed algorithm is not as fast as the ELM algorithm in network training, its recognition accuracy is the highest. Moreover, after the network training is completed, the network can be encapsulated, and the subsequent recognition process can be carried out through the trained network encapsulation. Therefore, the speed of network training has little impact on the subsequent recognition process.

Table 1
Comparison of recognition accuracy of 5 algorithms

Fault Type	Recognition accuracy/%				
	proposed method	CNN	SVM	BPNN	ELM
Insulation air gap	95.6	94.7	91.7	93.2	88.2
Scratches on insulation	97.3	95.5	92.3	94.3	87.3
Insulation surface impurities	94.8	92.6	88.9	89.7	85.9
Total	95.9	94.3	91.0	92.4	87.1

5. Conclusion

This article proposes a cable tunnel defect detection method based on deep learning and global-local features, which fully considers the relationship between local fine features and global spatial scene distribution, and improves the accuracy of salient feature map detection. A cyclic structure network module was proposed to continuously collect context information and iteratively improve convolutional features. Compared to the classic recognition algorithms, the proposed method outperforms CNN, SVM, and BPNN in terms of network training time. Compared to ELM algorithm, although the network training speed of the proposed method is lower than that of ELM, the final recognition rate (95.9%) is higher than ELM algorithm (87.1%). In addition, the trained network can be encapsulated and directly used for defect recognition in cable tunnel inspection robots, greatly improving the engineering applicability.

Acknowledgements

This work was supported by the self funded project "Key Technology Research on Collaborative Inspection of Quadruped Robots in Power Tunnels" by State Grid Beijing Electric Power Company.

References

- [1] W. Ding, S. Yuan, X. Gao, et al., "Research on construction disturbance characteristics caused by super large diameter pipe jacking in electric power tunnel," *Rock and Soil Mechanics*, vol. 31, no. 9, pp. 2901–2906, 2010.
- [2] J. Li, X. Huang, X. Yang, "Numerical simulation and settlement prediction of subway locating under cable tunnels," *Journal of Railway Science and Engineering*, vol. 5, no. 1, pp. 68–71, 2008.
- [3] S. J. Julier, J. K. Uhlmann, "Unscented filtering and nonlinear estimation," *Proceedings of the IEEE*, vol. 92, no. 3, pp. 401–422, 2004.
- [4] B. Jiang, A. P. Sample, R. M. Wistort, et al., "Autonomous robotic monitoring of underground cable systems," *International Conference on Icar, IEEE*, pp. 673-679, 2005.
- [5] M. H. Yang, R. Xiang, H. Lu, et al., "Deep networks for saliency detection via local estimation and global search," *Proceedings of the IEEE Conference on Computer Vision, Los Alamitos: IEEE Computer Society Press*, pp. 3183-3192, 2015.
- [6] X. H. Li, H. H. Lu, L. H. Zhang, et al., "Saliency detection via dense and sparse reconstruction," *Proceedings of the IEEE International on Conference on Computer Vision, Los Alamitos: IEEE Computer Society Press*, pp. 2976-2983, 2013.
- [7] L. Z. Wang, L. J. Wang, H. C. Lu, et al., "Saliency detection with recurrent fully convolutional networks," *Proceedings of the European Conference on Computer Vision, Heidelberg: Springer*, pp. 825-841, 2016.
- [8] X. Li, L. M. Zhao, L. N. Wei, et al., "Deep Saliency: multi-task deep neural network model for salient object detection," *IEEE Transactions on Image Processing*, vol. 25, no. 8, pp. 3919-3930, 2016.

- [9] P. Zhang, D. Wang, H. Lu, et al., "Amulet: aggregating multi-level convolutional features for salient object detection," Proceedings of the IEEE International Conference on Computer Vision, Los Alamitos: IEEE Computer Society Press, vol. 2017, pp. 202-211, 2017.
- [10] X. Chen, H. Ma, X. Wang, et al., "Improving object proposals with multi-thresholding straddling expansion," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Los Alamitos: IEEE Computer Society Press, pp. 2587-2595, 2015.
- [11] R. Achanta, S. Hemami, F. Estrada, et al., "Frequency-tuned salient region detection," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Los Alamitos: IEEE Computer Society Press, vol. 22, no. 9/10, pp. 1597-1604, 2009.
- [12] D. W. Zhang, H. Z. Fu, J. W. Han, et al., "A review of co-saliency detection algorithms: fundamentals, applications, and challenges," ACM Transactions on Intelligent Systems and Technology, vol. 9, no. 4, pp. 1-31, 2018.
- [13] X. Wang, H. M. Ma, X. Z. Chen, et al., "Edge preserving and multi- Scale contextual neural network for salient object detection," IEEE Transactions on Image Processing, vol. 27, no. 1, pp. 121-134, 2017.
- [14] W. Zhang, Y. Han, Q. Huang, et al., "The fast multi-scale convolutional sparse coding based super-resolution for infrared image," Journal of Computer-Aided Design & Computer Graphics, vol. 30, no. 10, pp. 1935- 1942, 2018.
- [15] Z. H. Zhao, S. P. Yang, Z. Q. Ma, "License plate character recognition based on convolutional neural network LeNet-5," Journal of System Simulation, vol. 22, no. 3, pp. 638-641, 2010.
- [16] A. Tsoupos, V. M. Khadkikar, "A novel SVM technique with enhanced output voltage quality for indirect matrix converters," IEEE Transactions on Industrial Electronics, vol. 66, no. 2, pp. 832-841, 2019.
- [17] Z. Xiao, S. J. Ye, B. Zhong, et al., "BP neural network with rough set for short term load forecasting," Expert Systems with Applications, vol. 36, no. 1, pp. 273-279, 2009.
- [18] G. B. Huang, Q. Y. Zhu, C. K. Siew, "Extreme learning machine: theory and application," Neurocomputing, vol. 70, no. 1-3, pp. 489-501, 2006.