

Reflection Interference Elimination of Protection Platen Status Recognition in UHV DC Converter Station

Min Xie, Wei Wang, Qingzhu Shao, Bin Yu, Yang Yu*

State Grid Anhui Electric Power Co., Ltd, Anhui, China

*Corresponding Author.

Abstract

In response to the reflection interference problem in the pressure plate state monitoring of the extra high voltage DC converter screen, a network based on multi-stage connection and adaptive region attention is proposed. First, an interactive multi-level connection structure is proposed to give different ratios to different levels of elements so as to solve the background detail problem of the multi-scale features and restore the reflection images. Secondly, after the channel-level focal layer is introduced to the multi-stage connection, useful network channels are focused to solve adaptive adjustment problems of channel features in multi-stage connection process. Finally, to solve the problems such as the non-natural reflected light, dispersed spot distribution and large light intensity in the monitor room of the station, a wide regional non-local block is proposed, in which the feature map is divided into wide meshes so that the reflected light is evenly distributed to capture long-term spatial dependencies between long-range pixels. After experimental verification is carried out against the public data set and the monitored image data set of extra high voltage DC converter station, the results prove that the network can efficiently remove reflection interference in the images, greatly enhancing the effect of the pressure plate state monitoring.

Keywords: UHV DC converter station, Pressure plate state, De-reflection, Multi-stage connection, Wide regional non-local block

I. INTRODUCTION

Video surveillance has widely been applied to extra high voltage DC converter stations to monitor the operating status of various equipment. Traditionally, information processing of images detected is mainly completed by manual, which reduces the overall efficiency and makes the accuracy of monitoring the operational status of the equipment vary from person to person. In recent years, advancement of artificial intelligence and computer vision technology has allowed machine vision-based

panoramic monitoring & research of the equipment in extra high voltage DC converter stations. This achieves efficient and accurate collection and identification of equipment operating status characteristic information, and helps effectively carry out video inspection, thus saving a lot of manpower and improving inspection efficiency [1]. Currently, however, during current video inspection of the equipment in converter stations, affected by reflected light, large-area reflection of the collected images occurs in specific scenes, which is unfavorable for image processing. Although this problem can be solved in some scenes by adjusting direction of the camera, a problem that the reflection of the panel cabinet surface glass makes it difficult to monitor part of the captured image always arises during panel cabinet monitoring of secondary equipment (as shown in Fig 1). This greatly affects target detection, semantic segmentation and other subsequent machine vision tasks. In some severe cases, the entire area to be detected is interfered by the reflected light, leading to failure to implement or suspension of the tasks. This is also an unresolved problem arising from panoramic monitoring and research of the equipment in extra high voltage DC converter stations.

At present, to eliminate the reflected light, prior information is artificially extracted to distinguish the transmitting layer and the reflecting layer in the images. However, it is impossible to apply the prior information to all types of reflection scenes due to different imaging conditions. In recent years, data-driven learning through deep convolutional networks has become the mainstream method to replace artificial prior knowledge. Literature [2, 3] used the codec structure that was similar to U-Net network as the image conversion model. However, although the skip connection at the same level was realized, the connection of features at different levels was not considered, causing information loss due to lack of connection at different levels. In order to solve this problem, Literature [4] adopted the multi-scale fusion to characterize features on multiple scales. However, connection of features from a smaller scale to a larger scale was only considered during fusion, and thus there was a lack of the interactivity in different scales.

Thus, in response to the Panel Cabinet glass reflection interference problem in the pressure plate state monitoring of the extra high voltage DC converter screen, the (MA-Net) network based on multi-stage connection and adaptive region attention is proposed. First, an interactive multi-level connection structure is proposed to give different ratios to different levels of elements so as to solve the background detail problem of the multi-scale features and restore the reflection images. Secondly, after the channel-level focal layer is introduced to the multi-stage connection [5], useful network channels are focused to solve adaptive adjustment problems of channel features in multi-stage connection process. Finally, to solve the problems such as the non-natural reflected light, dispersed spot distribution and large light intensity in the monitor room of the station, a wide regional non-local (WRNL) block is proposed, in which the feature map is divided into wide meshes so that the reflected light is evenly distributed to capture long-term spatial dependencies between long-range pixels. After experimental verification is

carried out against the public data set SIR2 and the monitored image data set of extra high voltage DC converter station, the results prove that the proposed network is obviously superior to the existing methods, as it can efficiently remove reflection interference existing in the public data set and pressure plate images and greatly enhance the effect of the pressure plate state monitoring.

II. MATERIALS AND METHODS

2.1 De-reflection Demand Analysis of Surveillance Images

The light source reflected onto panel cabinet surveillance images of extra high voltage DC converter station is mainly daylight, featured by indoor and unnatural light, scattered light spots and high light intensity. As shown in Fig 1 below, the scattered light spots are likely to cover the display screen, the pressure plate and other components to be monitored. In addition, restricted by environmental factors such as waterproof measures and cable routing, small indoor surveillance camera can only be installed over the panel cabinet and look down at the cabinet. This makes it impossible to read correctly the device status information due to excessive light intensity, thus imposing a threat to image processing and disturbing subsequent machine vision tasks such as target detection and semantic segmentation.



Fig 1: Schematic diagram of panel cabinet glass reflection

As most of the existing deep neural networks for reflection removal are directed towards real natural scenes, which are usually outdoor and under natural light, the reflected light is characterized by low intensity and uniform spot distribution, which is different from the light reflected onto extra high voltage DC converter station panel cabinet. Therefore, during construction of the anti-reflective deep neural network for extra high voltage DC converter station panel cabinet, the operating environment and light characteristics of the equipment in the stations must be fully considered to efficiently and accurately eliminate the reflection interference and obtain the characteristic information of the equipment operating

state.

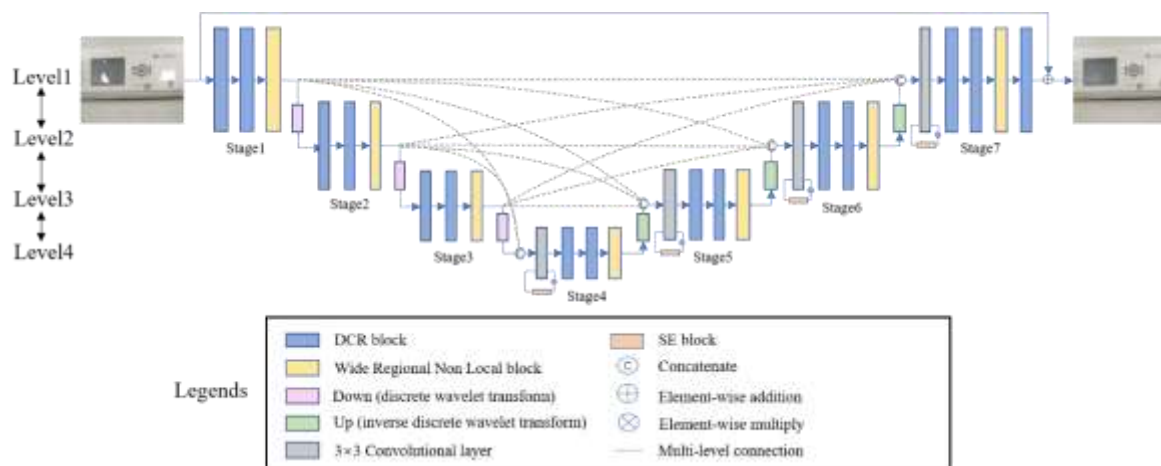


Fig 2: MA-Net network structure diagram

2.2 Architecture Design of Anti-Reflective Deep Neural Network

In response to the above-mentioned de-reflection needs of surveillance image, this paper proposes a MA-Net anti-reflective network structure as shown in Fig 2.

The MA-Net network consists of an encoder and a decoder, between which the former consists of the first three stages and the latter is composed of the remaining four stages. Classification of levels is subject to the size of the feature map, and a block is defined as a stage. Through multi-stage connection, MA-Net can connect all the outputs of the encoder to all the inputs of the decoder, so that features at different sizes can be used simultaneously in the process of image restoration.

2.2.1 Multi-stage connection mechanism.

In a network structure similar to U-Net, connection of features at the same level can compensate the defect that low-level features in the decoder cannot utilize multi-scale information. However, as image de-reflection is a low-level vision task, features on different scales are required to restore image details. Learning from Literature [6-8], this paper constructs a multi-stage connection structure, in which each stage is composed of two densely connected residual (DCR) blocks (as shown in Fig 3). Each DCR is composed of three convolutional layers and PReLU layers [9] and a WRNL block. In the up-sampling part of the network, information of down-sampled features on all scales can be aggregated through multi-stage connection. Since the features at different levels have different scales, in order to adaptively

adjust the channel characteristics after the multi-stage connection, a squeeze-and-excitation (SE) block is added at each decoder stage (as shown in Fig 4), and the number of channels is later adjusted through a 1×1 convolutional layer.

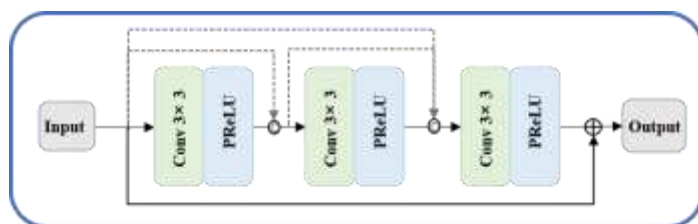


Fig 3: Densely connected residual block

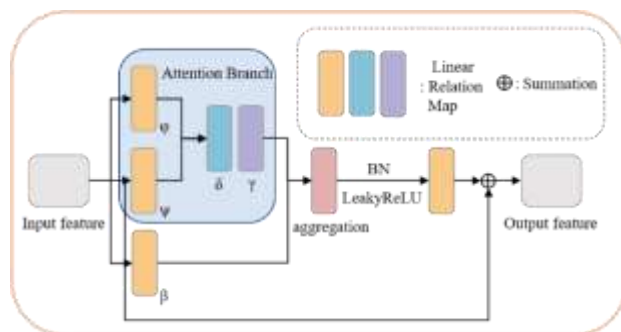


Fig 4: Squeeze-and-excitation block

Assuming that E_{out}^l is the output feature of the level l ($l=1,2,3,4$) in the encoder, the input feature D_{in}^l of each level l ($l=1,2,3,4$) in the decoder can be defined as:

$$D_{concat}^l = (\oplus_{i=1}^3 H_i^l(E_{out}^i)) \oplus H_{up}(D_{out}^{l+1}) \quad (1)$$

$$D_{in}^l = W_{1 \times 1}(f_{SE}(D_{concat}^l)) \quad (2)$$

Where, $\oplus$ represents the cascade operation, $H_{up}(\cdot)$ the up-sampling operation, D_{out}^l the output feature of the level l in the encoder, $W_{1 \times 1}$ the 1×1 convolutional layer, $f_{SE}(\cdot)$ the SE block, $H_i^l(\cdot)$ the sampling operation from level i to level l , namely the $l-i$ th down-sampling operation and the $i-l$ th up-sampling operation at $l > i$, $l=i$ and $l < i$. Here, set $D_{in}^5 = 0$.

Multi-stage connection structure allows to use features at higher levels when processing low-level features, which helps the network utilize multiple scale representations to recover large-scale objects. Here, discrete wavelet transform is used for up-sampling /down-sampling operations to find out the mapping relationship in feature shapes on different scales. In addition, to solve the problem of information loss, this paper chooses two-dimensional Haar wavelet for sampling.

2.2.2 Wide area non-local block design.

The lighting characteristics of the converter station cabinet make the conventional blocks inaccessible. Therefore, first define the WRNL block, and then apply statistical knowledge to analyse the effectiveness of the WRNL block.

Assume that $X \in \mathbb{R}^{H \times W \times C}$ represents the input characteristics of WRNL. Divide X into $a \times b$ grids. $\{X^k\}, (k=1, \dots, K=ab)$, where K is the number of grids. The linear embedding process in which X^k generates the output Z^k is as follows:

$$\Phi(X^k)_i^j = \phi(X_i^k, X_j^k) = \gamma(\mu(X_i^k, X_j^k)) = \gamma(\exp\{\theta(X_i^k)\varphi(X_j^k)^T\}) \quad (3)$$

$$\begin{aligned} \theta(X_i^k) &= X_i^k W_\theta \\ \varphi(X_i^k) &= X_i^k W_\varphi \\ G(X)_i^k &= X_i^k W_g \end{aligned} \quad (4)$$

Where, X_i^k and X_j^k respectively represent the features of the position i and j , $\gamma(\cdot)$ and $\mu(\cdot)$ respectively represent the relational function and the hybrid mapping composed of linear and non-linear functions. The dimensions of the weight matrix W_θ , W_φ and W_g are $C \times L$, $C \times L$ and $C \times C$ respectively. Let $L = C / 2$, and the regional non-local operation can be expressed as:

$$Z_i^k = \frac{1}{\delta_i(X^k)} \sum_{j \in S_i} \Phi(X^k)_i^j G(X^k)_i, \forall i \quad (5)$$

Where, $\delta_i(X^k) = \sum_{j \in S_i} \phi(X_i^k, X_j^k)$ represents the correlation between X_i^k and each X_j^k in a set of regional locations S_i , Z_i^k represents the output feature of Z^k at location i . If $a > b$, the grid is wider than that at $a = b$. Therefore, when $a > b$, $a = b$ and $a < b$, they are called wide area rectangular block,

square block and high area rectangular block, respectively. In the WRNL blocks, set the size of the grid $a \times b$ to 16×4 , 8×2 , 4×1 , and 4×1 , corresponding to level 1, 2, 3, and 4 respectively.

Assuming that the non-local block recovers specific pixels according to the information of other pixels in the spot, each spot is required provide adequate background information. However, due to the uneven distribution of the reflective layer, it is difficult for regional non-local blocks to make full use of background information. Wide-area rectangular spots provide richer background information than square and high-area rectangular spots.

Divide the image into 16×4 , 8×8 , and 4×16 grids to obtain wide area rectangle, square and high area rectangle blocks. If the pixel difference between the input reflection image and the corresponding de-reflection image exceeds a certain threshold, the pixel is considered to belong to the reflective layer. The results of ablation experiments on different types of areas are shown in Table I. Compared with square and high area rectangular blocks, wide area rectangular blocks have better peak signal to noise ratio (PSNR) and structural similarity (SSIM) [10]. At this point, the reflecting layer is evenly distributed on all spots, which helps better restore the image.

TABLE I: Regional type ablation experiment of regional non-local block

Regional type	PSNR	SSIM
High area rectangle block	28.82	0.897
Square	28.92	0.898
Wide area rectangle block	29.03	0.899

2.3 Loss Function Design

The loss function of the MA-Net de-reflection network is defined as follows:

$$L_1 = ||x_{gt} - f(x_{input})||_1 + ||x_{gt} - f(x_{input})||_2 \quad (6)$$

Where, x_{input} represents the input reflection image, x_{gt} the corresponding de-reflection image, and f the output of MA-Net.

In addition, in order to better separate the reflecting layer and the transmission layer, a gradient domain-based repulsion loss is defined here. It can be seen by analysing the relationship between the edges of the two layers, that the transmission layer and the reflecting layer basically do not overlap at the edges. As the edge in the image should be caused by either the transmission layer or the reflecting layer

rather than both, the correlation between the transmission layer and the reflecting layer predicted in the gradient domain can be minimized, and the rejection loss is expressed as the product of the normalized gradient fields of the two layers at multiple spatial resolutions:

$$L_{excl}(\theta) = \sum_{I \in D} \sum_{n=1}^N \|\psi(f_T^{\downarrow n}(I; \theta), f_R^{\downarrow n}(I; \theta))\|_F \quad (7)$$

$$\psi(T, R) = \tanh(\lambda_T \|\nabla T\|) \odot \tanh(\lambda_R \|\nabla R\|) \quad (8)$$

Where, λ_T and λ_R is the normalization factor; $\|\cdot\|_F$ the Frobenius norm, $\odot$ the product in the order of the unit, n the image down-sampling factor, f_T and f_R down-sampling operations by bilinear interpolation using a sampling factor of $2n-1$. Here, $N=3$ $\lambda_T = \sqrt{\|\nabla R\|_F / \|\nabla T\|_F}$, $\lambda_R = \sqrt{\|\nabla T\|_F / \|\nabla R\|_F}$.

Then the total loss function is:

$$L = L_1 + L_{excl} \quad (9)$$

III. EXPERIMENTAL RESULTS AND ANALYSIS

3.1 Data Set Description

In order to verify the feasibility and effectiveness of the method proposed in this paper, the method was verified against the public data set SIR2 [11] and the state image data set of panel cabinet pressure plate in a UHV converter station protection room. In this paper, the datasets totaling 4260 images (of Pressure-plate (1400), Object (1500), Postcard (560) and Zhang et al. (800)) are used for training, and the remaining 1820 images (Pressure-plate (600), Object (640), Postcard (240) and Zhang et al. (340)) are used for quantitative evaluation. 4 data sets are randomly divided into training set and test set at a ratio of 7:3.

Among them, the pressure-plate data set is an image data set taken by Canon EOS 750D camera of the panel cabinet pressure plate state in an indoor environment. It contains 220 pairs of real images, which include images reflected and corresponding reference transmission layers. In order to simulate different imaging conditions, the following factors are considered when images are taken: 1) environment: indoor; 2) lighting conditions: incandescent lamp; 3) glass plate thickness: 3 mm and 8

mm; 4) the distance between the glass and the camera: 3-15 cm; 5) camera view: front and oblique; 6) camera exposure value: 8.0-16.0; 7) camera aperture (affecting reflection glossiness): f/4.0-f/16.

3.2 Experimental Results and Analysis

3.2.1 Ablation experiment.

The multi-stage connection structure allows aggregation of information of down-sampled features on all scales into the up-sampling part of the network. However, too many scales may reduce the weight of key information, while too little scales may produce the unobvious effect of feature information extraction. Therefore, how to determine the appropriate number of stages is very important. Fig 5 shows the multi-stage connection stage ablation experiment results of the public data set (PSNR1, SSIM1) and the panel cabinet pressure plate state image data set (PSNR2, SSIM2).

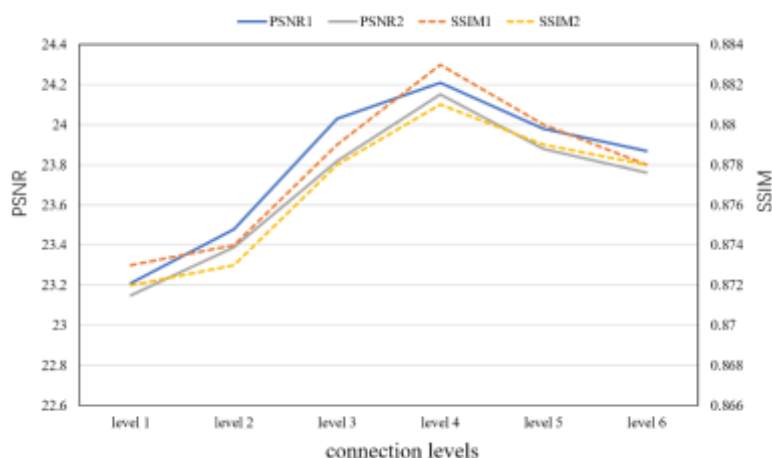


Fig 5: Effect of different connection levels on various indicators

It can be seen from Fig 5 that as there are more and more stages in multi-stage connection structure, the PSNR and SSIM indicators gradually increase, and reach their peak at stage 4. If there are more and more stages in multi-stage connection structure but PSNR and SSIM gradually decrease, it indicates that the built deep neural network has gradually reduced its ability to aggregate information at various scales. Therefore, the number of stages for subsequent multi-stage connection is selected as 4.

3.2.2 Qualitative analysis

Under the condition of choosing 4-layer connection level, the MA-Net method proposed in this paper is compared with other methods including CEILNet [12], BDN [13] and ERRNet[14]. For peer-to-peer comparison, the same public data set training samples and panel cabinet pressure plate state image data set training samples are adopted to fine-tune the parameters of each model, with best results of the fine-tuned version expressed by the suffix'-F'.



Fig 6: Visual comparison between existing approaches and the proposed method on real image dataset, where input (Line 1), CEILNET (Line 2), BDN (Line 3), ERRNet (Line 4), the proposed method (Line 5)



Fig 7: Visual comparison between existing approaches and the proposed method on pressure plate image dataset, where input (Line 1), CEILNET (Line 2), BDN (Line 3), ERRNet (Line 4), the proposed method (Line 5)

Fig 6 and Fig 7 give the anti-reflective visual processing results of the real natural landscape images and the panel cabinet platen images. It can be found that compared with other methods, the method proposed in this paper has more accurate visual effect, removes most of unnecessary reflections and has obvious advantages in processing the images characterized by indoor environment, unnatural light, scattered light spots and high light intensity. There are common problems with other methods such as insignificant reflection removal and louder noise.

3.2.3 Quantitative analysis

Table II summarizes the experimental results of different methods on four real data sets, namely Pressure-plate, Object, Postcard and Zhang et al. The number of test images in each data set is placed after the name, and the PSNR and SSIM indicators are used. The larger the value of PSNR and SSIM is, the better the performance will be.

TABLE II. Quantitative comparison of different methods on four real dataset

DATASET(SIZE)	Index	METHODS			
		CEILNet-F	BDN-F	ERRNet-F	MA-NET Ours
OBJECT(640)	PSNR	22.82	23.03	24.85	24.88
	SSIM	0.801	0.853	0.889	0.894
POSTCARD(240)	PSNR	20.07	20.72	21.99	23.39
	SSIM	0.810	0.857	0.873	0.876
ZHANG ET AL. (340)	PSNR	18.79	19.47	23.35	21.87
	SSIM	0.747	0.721	0.812	0.761
PRESSURE-PLATE (600)	PSNR	19.34	18.95	22.18	23.57
	SSIM	0.745	0.737	0.757	0.783
AVERAGE(1820)	PSNR	21.32	21.61	23.44	24.08
	SSIM	0.806	0.841	0.870	0.875

It can be seen from the table that, except for the Zhang et al. data set, the MA-Net works well on all data sets, because ERRNet[14] that is created based on the Zhang et al. [15] model has better ability to generalize this data set. Thus, this algorithm works well on Zhang et al. data set. MA-Net is superior to other methods in terms of average performance of all test data sets.

Table III aims at the existing pressure plate state recognition methods (cluster matching method [16], improved BOF method [17], OpenVINO method [18], transfer learning method [19] and improved SSD method [20]) and taking panel cabinet pressure plate state data set Pressure-plate as the object, this paper compares and analyzes the effect of reflecting network on pressure plate state recognition.

TABLE III. Effect of reflecting network on pressure plate state recognition

STATE RECOGNITION METHOD	RECOGNITION ACCURACY(BEFORE DE- REFLECTION)	RECOGNITION ACCURACY(AFTER DE- REFLECTION)
CLUSTER MATCHING METHOD	78.22%	84.50%
IMPROVED BOF METHOD	83.55%	87.21%
OPENVINO METHOD	92.90%	93.35%
TRANSFER LEARNING METHOD	89.63%	91.20%
IMPROVED SSD METHOD	88.97%	90.75%

It can be seen from the table that under the condition of presence of reflection, the recognition accuracy of these five-pressure plate state recognition methods is 78.22%, 83.55%, 92.90%, 89.63% and 84.55%, respectively. After the reflection interference is removed using the de-reflection network, the recognition accuracy of the five methods has been improved to varying degrees. Among them, the recognition accuracy of the cluster matching method and the improved BOF method is increased by 6.28% and 3.66% respectively, which were significantly higher than that of the OpenVINO method, the migration learning method and the improved SSD method, namely 0.45%, 1.57%, and 1.87% respectively. Because the traditional image recognition method is more dependent on the original images with poor anti-interference ability, the de-reflection network works better.

IV. CONCLUSIONS

To solve the reflection problem in the pressure plate state monitoring of UHV DC converter station, an anti-reflection deep neural network based on multi-level connection and self-adaptive regional attention is proposed to remove the reflection interference in the images. The MA-Net network can adaptively aggregate features through multi-stage connections and SE blocks, and make full use of abundant remote non-reflective background information based on wide-area non-local blocks. Experiments show that MA-Net can not only restore the details of the input image, but almost completely eliminate the reflection interference on the real image data set and the image data set of the panel cabinet pressure plate state, effectively improving the recognition of the pressure plate state.

ACKNOWLEDGE

This work is supported by the Science and Technology Project of State Grid Anhui Electric Power Co., Ltd. "Research and Application of Key Technologies for UHV DC Protection Fault Warning and Intelligent Decision Based on Panoramic State Intelligent Perception" (No. 52120019007Z)

REFERENCES

- [1] Zhao ZB, Zhang W, Zhai YJ, et al. (2020) Concept, Research Status and Prospect of Electric Power Vision Technology. *Electric Power Science and Engineering* 36(01): 1-8
- [2] Wang GQ, Sun CM, Arcot S (2019) Erl-net: Entangled representation learning for single image de-raining, in *International Conference on Computer Vision* 5644-5652
- [3] Yu WJ, Huang Z, Zhang W, et al. (2019) Gradual network for single image de-raining, in *ACM International Conference on Multimedia* 1795-1804
- [4] Jiang K, Wang ZY, Yi P, et al. (2020) Multi-scale progressive fusion network for single im-age deraining. In *Computer Vision and Pattern Recognition*: 8346-8355
- [5] Hu J, Shen L, Sun G (2018) Squeeze-and-excitation networks. In *Computer Vision and Pattern Recognition*: 7132-7141
- [6] Sun K, Xiao B, Liu D, et al. (2019) Deep high-resolution representation learning for human pose estimation. In *Computer Vision and Pattern Recognition*: 5693-5703
- [7] Tan MX, Pang RM, Quo VL (2019) Efficientdet: Scalable and efficient object detection. *arXiv preprint arXiv: 1911.09070*
- [8] Wang JD, Sun K, Cheng TH, et al. (2019) Deep high-resolution representation learning for visual recognition. *arXiv preprint arXiv: 1908.07919*
- [9] Ludovic T, Philippe G, Brahim CD, et al. (2017) Parametric exponential linear unit for deep convolutional neural networks. In *International Conference on Machine Learning and Applications IEEE*
- [10] Zhou Wang, Alan C Bovik, Hamid R Sheikh, et al. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 2004, 13(4): 600–612.
- [11] Renjie Wan, Boxin Shi, Lingyu Duan, et al. Benchmarking single-image reflection removal algorithms. *IEEE International Conference on Computer Vision*, 3922-3930, 2017.
- [12] Qingnan Fan, Jiaolong Yang, Gang Hua, et al. A generic deep architecture for single image reflection removal and image smoothing. *IEEE International Conference on Computer Vision*, 3238-3247, 2017.
- [13] Jie Yang, Dong Gong, Lingqiao Liu, et al. Seeing deeply and bidirectionally: A deep learning approach for single image reflection removal. *European Conference on Computer Vision*, 654-669, 2018.
- [14] Kaixuan Wei, Jiaolong Yang, Ying Fu, et al. Single image reflection removal exploiting misaligned training data and network enhancements. *IEEE Conference on Computer Vision and Pattern Recognition*, 8178–8187, 2019.
- [15] Xuaner Zhang, Ren Ng, Qifeng Chen. Single image reflection separation with perceptual losses. *IEEE Conference on Computer Vision and Pattern Recognition*, 4786-4794, 2018.

- [16] Tiecheng Li, Jiangbo Ren, Qingquan Liu, et al. Image Recognition and Model Cluster Matching of Relaying Plate. Journal of Harbin University of Science and Technology, 2021(04): 70-77[2021-09-03].
- [17] Yueqing Chen, Lin Hu, Yixiang Wu, et al. Condition Recognition System for Relay Protection Plate Based on Improved Bag of Feature Algorithm. Smart Power, 2021, 49(02): 99-106.
- [18] Xiaolei Sun, Hao Chen, Zichang Sun, et al. Status Recognition Technology of Substation Protection Platen Based on OpenVINO Model Optimization. Hubei Electric Power, 2021, 45(01): 28-34.
- [19] Xiang Chen, Qingnian Zou, Zhaoyu Xie, et al. Identification of Platen Switch State Based on Transfer Learning Strategy. Computer and Modernization, 2021(05): 120-126.
- [20] Ke Zhou, Qianwen Yang, Yaoyi Wang, et al. An improved SSD algorithm based on pressure plate status recognition. Electrical Measurement & Instrumentation, 2021, 58(01): 69-76