



OPEN A texture guided transmission line image enhancement method

Yu Zhang^{1,2,3}, Liangliang Zhao¹, Yinke Dou^{1,2,3,4}✉, Yangyang Jiao¹ & Qiang Liu¹

With the increasing demand of transmission line inspection, image detection technology plays a crucial role in foreign object detection on transmission lines. However, due to the influence of natural environment, the image quality is often interfered with by rain, fog, blur and other factors, which affects the accuracy of detection. In order to improve image quality and accuracy of foreign object detection, a texture guided transmission line image enhancement (TGTIE) method is proposed. Firstly, the texture inference network (TINet) is used to extract texture information from input interference images. Then, the texture information is used as a priori knowledge to use the texture-based conditional generative adversarial network (TCGAN) to achieve adaptive image deraining, defogging and deblurring. In order to improve the detail quality and robustness of image generation, a neural gradient algorithm and dual path attention mechanism are proposed in the generative network of GAN. At the same time, in order to prevent local artifacts of spatial variation, an effective global-local discriminator structure is introduced into the discriminant network to perform global and local inspections on the generated images. In addition, the Charbonnier loss, SSIM loss and global-local generative adversarial loss are used in multiple stages to train the model to achieve the best performance. Both quantitative and qualitative results on different interference UAV datasets show that TGTIE can effectively remove noise interference under different conditions and improve image quality, PSNR and SSIM reached 33.218dB, 34.921 dB, 30.725 dB and 0.956, 0.962, 0.926 respectively, and show excellent performance in foreign object detection tasks. The method in this paper provides effective technical support for intelligent inspection and fault warning of transmission lines.

Keywords Image enhancement, Texture inference, Generative adversarial network, Neural gradient, Attention mechanism

Transmission lines are the main carrier of power transmission, and their safety and stability are directly related to the reliability of power supply. However, due to the wide distribution of transmission lines and the complex natural environment, foreign objects such as balloons, kites, plastics, and bird nests are easily attached to the lines. If these foreign objects are not discovered and cleaned in time, they will affect the normal operation of the transmission lines and even cause power outages or safety accidents^{1,2}. The traditional manual inspection is not only high cost and low efficiency, but also easy to be affected by the environment and climate, and the rate of missing is high. Therefore, in order to improve the operation efficiency and safety of transmission lines, intelligent inspection technology has become an important means, among which image recognition technology is widely used in foreign body detection, line damage identification and other tasks³. However, in practical applications, the foreign object detection images of transmission lines are often affected by a variety of unfavorable factors, such as rain, fog, and wind disturbances, which will reduce the image quality and thus affect the accuracy of foreign object detection. Therefore, it is of great significance to improve the accuracy and reliability of foreign body detection by enhancing the image processing to improve the image quality^{4–6}.

At present, image enhancement methods are divided into traditional image processing methods and deep learning-based methods. The traditional image deraining and defogging technology usually rely on physical model, they analyze the impact of raindrops or fog on the image, and use inversion algorithms to restore clear images^{7–9}. The deblurring technology often use motion compensation methods to reduce image blur caused by camera shake^{10,11}. Although these traditional methods can achieve good results when dealing with specific problems, they often need to rely on artificially designed features and rules, lack enough flexibility, and are difficult to meet the requirements of efficiency and accuracy. With the rapid development of deep learning technology, image enhancement methods based on deep neural networks have gradually become a hot topic

¹Shanxi Energy Internet Research Institute, Taiyuan 030032, China. ²College of Electrical and Power Engineering, Taiyuan University of Technology, Taiyuan 030024, China. ³Department of Automation, Taiyuan Institute of Technology, Taiyuan 030008, China. ⁴Key Laboratory of Cleaner Intelligent Control on Coal Electricity, Ministry of Education, Taiyuan 030024, China. ✉email: douyinke@tyut.edu.cn

of research¹². Deep learning methods can handle more complex image enhancement tasks without manually designing features by automatically learning complex patterns and features in images. Deep learning methods are highly automated and adaptable, and can simultaneously consider the influence of various interference factors on the image, such as rain, fog, wind disturbance, etc. Through the learning of multi-layer features, it fully considers the mutual influence of these factors, so as to comprehensively optimize the image and restore clearer and more detailed images^{13–15}. However, the current deep learning methods are not robust, and the generated images are prone to the problem of unclear details.

In view of the above problems, this paper analyzes the existing image enhancement methods and combines the characteristics of power line images to maximize the removal of rain, fog and blur from a single input image while retaining the image detail information, a texture guided transmission line image enhancement (TGTLE) method is proposed. The novelty of the TGTLE method lies in the design of texture inference network (TINet) and texture-based conditional generative adversarial network (TCGAN). First, TINet is used to learn the texture information guidance map of the transmission line image. On this basis, TCGAN is used to guide the removal of rain, fog and blur in the image. In addition, in the generative network of GAN, the neural gradient algorithm is designed to guide the GAN network to generate finer image details by calculating the gradient information of each pixel in the image and the dual path attention mechanism is designed to improve the accuracy of image generation by focusing on important areas and features in the image. At the same time, the discriminant network combines global and local feature discrimination to ensure that the generated image is more realistic in both overall and local details. Finally, the multi-stage joint loss function is used to optimize the training process, which can not only improve the accuracy of image generation, but also speed up the training speed and improve the training efficiency. In summary, the main contributions of this paper are as follows:

- (1) TINet is designed to mine texture information from the input transmission line images, and then a powerful TCGAN is built to achieve adaptive image deraining, defogging and deblurring.
- (2) The neural gradient algorithm is proposed to make the network more robust in feature extraction and more stable and adaptive in image detail extraction.
- (3) The dual path attention mechanism is proposed to enable the network to focus on important details and structures in the image at different spatial levels, thereby improving the sharpness of the enhanced image.
- (4) Design global and local feature discrimination network to help the generator optimize the global structure and local details at the same time, so as to improve the image enhancement effect and retain the image details, making the result more natural and realistic.
- (5) Design a multi-stage loss function to optimize the training process, including Charbonnier loss, SSIM loss and global-local generative adversarial loss, which helps to gradually improve the image detail recovery ability, thereby achieving higher quality image enhancement effect.

Related work

Image deraining

Li et al. first used a Gaussian mixture model to approximate the priors of the background image and the rain streak layer, and then removed the rain streak based on the maximum a posteriori probability (MAP)¹⁶. Kim et al. first detected the rain streak region, and then used non-local mean filtering to remove the rain streaks¹⁷. Chen et al. first used a guided image filter to decompose the input rain pattern into low and high frequency parts, then used a set of mixed handcrafted features to separate rain streaks from the high-frequency part¹⁸. Tao et al. proposed a hybrid feature fusion network (MFFDNet) for single image deraining. The network cleverly utilizes the advantages of CNN and Transformer networks to extract image features and fuse global and local features to produce more discriminative feature representations, which greatly improves the rain removal performance of the model¹⁹. Wang et al. proposed an effective pyramid feature decoupling network (PFDN), which can simultaneously model rain-related features to remove rain streaks, and model rain-independent features to restore contextual detail information from multiple angles²⁰. Yan et al. proposed a joint high-frequency channel and Fourier frequency domain guided network (FPGNet), which uses a Fourier-based multi-scale feature extraction module to learn the rain layer information in the rain map, and simultaneously learns the natural distribution of rain streaks from the high-frequency channel to provide the model with rain streak prior information²¹.

Image defogging

Li et al. divided the image into small areas based on the histogram equalization method and optimized the image by local parameters to achieve cable tunnel image defogging²². Li et al. proposed a method combining multi-scale Retinex and wavelet transform to achieve effective image defogging by processing the I component of HIS space and applying logarithmic transformation²³. Xin et al. combined the finite contrast adaptive histogram equalization method with dark channel, bright channel and other methods to remove fog in transmission line images²⁴. Zhang et al. proposed a CycleGAN image defogging method based on the residual attention mechanism. In the residual network, channel and spatial attention mechanism were used to form the attention residual fast, which was added to the two generation networks of CycleGAN, and the cycle-perceptual consistency loss is combined to achieve image dehazing²⁵. Wang et al. proposed a defogging model based on the separation of features and collaborative networks, it used neural networks to extract spatial information and detailed features of different depths, so that the restored images had natural colors and good details²⁶. Zhou et al. proposed a transmission line inspection images defogging method Diff-EaT, which combines the diffusion model structure of transformer and a hybrid scale gated feedforward network to achieve image defogging²⁷.

Image deblurring

Tan et al. incorporated non-local statistical priors and exploiting the robustness of wavelet tight frames to effectively solved the ill-posed problem in image deblurring²⁸. Liang et al. realized the deblurring of power images by generating adversarial networks²⁹. Liu et al. proposed a network based on channel and spatial attention mechanisms, and combined L2 loss and edge loss as a comprehensive loss function for insulator image deblurring³⁰. Wang et al. proposed an insulator target detection method based on WGAN image deblurring. By introducing residual network module and Wasserstein distance in the training process of WGAN, higher quality insulator images are generated and the target detection rate of fuzzy insulator images is improved³¹. Wang et al. proposed a Transformer based architecture Uformer for image deblurring, which uses non-overlapping windowing based self-attention to reduce the computational load, and uses deep convolution in the forward network to further improve its ability to capture local context, while optimizing the skip connection mechanism to effectively transfer information from the encoder to the decoder³².

Network structure

The main purpose of the transmission line image enhancement method proposed in this paper is to use the image texture information to guide the network to perform adaptive deraining, defogging and deblurring, so that the enhanced transmission line image looks more natural and real. Therefore, a new transmission line image enhancement network TGTLE is proposed, which consists of two sequential subnetworks TINet and TCGAN. Its network structure is shown in Fig. 1.

Texture inference network (TINet)

The texture inference network (TINet) plays an important role in TGTLE, which extracts rich texture features from input transmission line images to guide TCGAN for image enhancement. As shown in Fig. 1, TINet consists of an encoder and decoder. The encoder inputs the transmission line image O_s , and then extracts the feature through a series of convolutional layers, deconvolution layers and multi-level feature attention layers to learn the high-level context information in the image and efficiently extract the texture information in the image. The texture analysis network can better understand the structural information and detail changes in the image. Finally, the decoder outputs the texture feature map R_{s1} , which is used to guide processing of subsequent networks and guide the network to restore the clear texture structure in the image.

The encoder in TINet is consists of 4 convolution blocks and 4 multi-level feature attention layers. Each convolution block contains a 3×3 convolution layer with step 2 and a 3×3 convolution layer with step 1. Each convolutional layer is followed by a BatchNorm layer and a Silu activation function layer. The multi-level feature

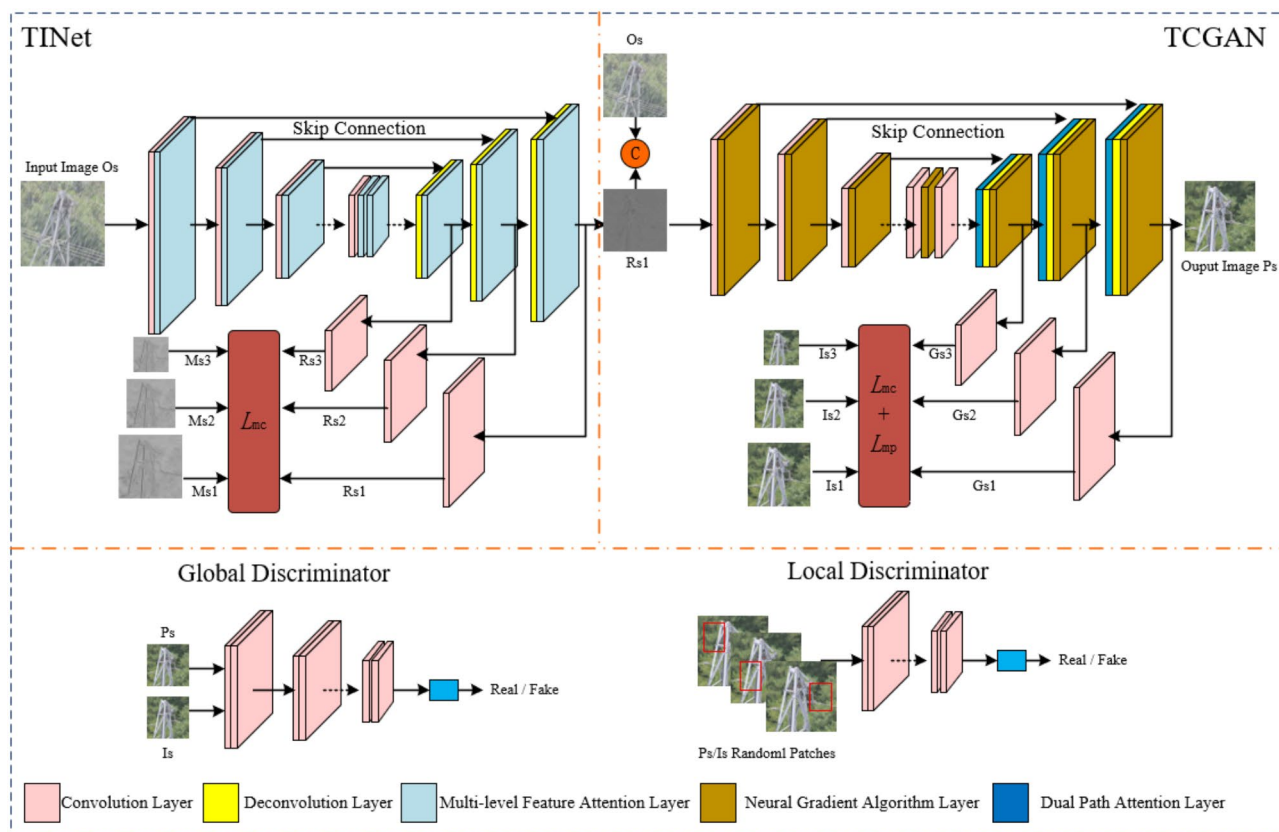


Fig. 1. Network structure diagram.

attention layer is introduced in 3.3. The decoder contains 4 deconvolution blocks and 4 multi-level feature attention layers.

In order to better mine the texture information of the input image, a multi-scale Charbonnier Loss is introduced into the decoder to measure the distance between the texture feature map and the true value of the output map at different scales. As shown below:

$$L_{mc} = \sum_{i=1}^N \sqrt{(\varepsilon^2 + (R_{si} - M_{si})^2)} \quad (1)$$

Where R_{si} represents the i -th output feature map extracted by the decoder, M_{si} represents the true feature map with the same scale as R_{si} , N represents the number of multiple scales, and ε is a small positive number, which is usually used to avoid the denominator being zero.

Texture-based conditional generative adversarial network (TCGAN)

The texture-based conditional generative adversarial network (TCGAN) uses the texture feature map generated by TINet to guide the network to perform adaptive image deraining, defogging and deblurring, accurately identify and process the details information in the image, and generate a more natural and real transmission line image. TCGAN consists of a generator and a global-local discriminator. The generator input transmission line image O_s and TINet output texture map R_{s1} , and then through a series of convolutional layers, neural gradient algorithm layers, dual path attention layers and deconvolution layers, and finally generates high-quality images P_s using multi-scale context information through encod-decoder architecture.

The encoder in the generator is consists of 5 convolution blocks and 5 neural gradient algorithm layers. Each convolution block contains a 3×3 depthwise separable convolution layer with step 2 and a 3×3 convolution layer with step 1. Each convolutional layer is followed by BatchNorm layer and Relu activation function layer. The neural gradient algorithm layer is introduced in 3.4. The decoder has a similar structure, contains 5 deconvolution blocks, 5 dual path attention layers, and 5 neural gradient algorithm layers. Each deconvolution block consists of a deconvolution layer, a Concat layer and a 1×1 convolution layer. The decoder first upsamples the image by deconvolution, then fuses it with the low-level features through the Concat layer, and finally adjusts the channels through the 1×1 convolution layer. The dual path attention layer is introduced in 3.5.

In the past, researchers used the GAN architecture for image enhancement and only used the global image discriminator to distinguish whether the images generated by the generator were real or fake. However, this is not enough to obtain a good generator, because if some areas of the generated image contain artifacts, the single global discriminator cannot effectively prevent these local artifacts. Therefore, we introduced a local discriminator based on the global discriminator to form a global-local discriminator structure. The global discriminator is used to focus on the overall visual effect of the image, and the local discriminator is used to focus on the local details of the image. Through the joint training of the local and global discriminator, the generator is prompted to improve the image generation effect.

The global image discriminator takes the generated image P_s and the real image I_s as inputs, and contains 7 convolution blocks and 1 fully connected layer. Each convolution block contains two 3×3 convolution layers, and each convolution layer contains a BatchNorm layer and a Relu activation function layer. The local image discriminator takes the random clipping blocks of the generated image P_s and the real image I_s as input. Because the input image size is small, it only contains 4 convolution blocks and 1 fully connected layer.

In order to generate more realistic and clear images, we use global-local generation adversarial Loss and multi-scale Charbonnier Loss and SSIM loss to train the network. The global-local generation adversarial loss is:

$$L_{D_{local}} = \sum_{i=1}^K \{ -E_{I_s \sim P_{real}} [\log D_{local}(I_{si})] - E_{P_s \sim P_{fake}} [\log(1 - D_{local}(P_{si}))] \} \quad (2)$$

$$L_{D_{global}} = -E_{I_s \sim P_{real}} [\log D_{global}(I_s)] - E_{P_s \sim P_{fake}} [\log(1 - D_{global}(P_s))] \quad (3)$$

$$L_G = \sum_{i=1}^K \{ -E_{P_s \sim P_{real}} [\log D_{local}(P_{si})] \} + \{ -E_{P_s \sim P_{real}} [\log D_{global}(P_s)] \} \quad (4)$$

Where I_s represents the real image, I_s represents the image generated by the generator, I_{si} represents the i -th random patch of the real image, P_{si} represents the i -th random patch of the generated image, and K represents the number of random blocks.

Multi-scale Charbonnier Loss and SSIM loss is:

$$L_{mc} = \sum_{i=1}^N \sqrt{(\varepsilon^2 + (G_{si} - I_{si})^2)} \quad (5)$$

$$\begin{aligned} l(G_{si}, I_{si}) &= \frac{2 \bullet \mu_{G_{si}} \bullet \mu_{I_{si}} + c_1}{\mu_{G_{si}}^2 + \mu_{I_{si}}^2 + c_1} \\ c(G_{si}, I_{si}) &= \frac{2 \bullet \sigma_{G_{si}} \bullet \sigma_{I_{si}} + c_2}{\sigma_{G_{si}}^2 + \sigma_{I_{si}}^2 + c_2} \\ s(G_{si}, I_{si}) &= \frac{\sigma_{G_{si}, I_{si}} + c_3}{\sigma_{G_{si}} \bullet \sigma_{I_{si}} + c_3} \\ SSIM(G_{si}, I_{si}) &= \sum_{i=1}^N l \bullet c \bullet s \\ L_{mc} &= 1 - SSIM(G_{si}, I_{si}) \end{aligned} \quad (6)$$

Where, G_{si} represents the output graph of the i -th scale extracted by the generator, I_{si} represents the real graph with the same scale as G_{si} , N represents the number of multi-scales, ε is a small positive number, usually

used to avoid zero denominator, μ represents the mean, σ represents the variance, and c_1, c_2, c_3 represent constants used for stable calculations.

Multi-level feature attention layer

In order to enhance the model's ability to capture image details and improve the accuracy of image feature extraction, a multi-level feature attention layer is proposed for different scales and different features extraction, as shown in Fig. 2.

In the multi-level feature attention layer, different convolution kernels are used to extract features to capture features of different scales and levels. The smaller convolution kernels are more suitable for capturing local features of images, while larger convolution kernels can be used to obtain more extensive contextual information, so enhance the model's understanding of images by using them to fuse multi-level and multi-scale feature information. In order to reduce the number of parameters, we use dilated convolution instead of traditional convolution, using 3×3 convolution with expansion rates of 2 and 3 instead of 5×5 and 7×7 convolution. After the module feature fusion, SE attention mechanism is used to obtain the global information of the entire feature map through global pooling operation, and then the full connection layer is used to calculate the weight of each channel, dynamically adjust the importance of features at each scale, and adaptive attention is paid to the most important feature information for image texture feature extraction, so as to improve the performance of the model. The multi-level feature attention layer can focus on different scale features of the image, so as to capture details information in the image more precisely and extract image features layer by layer. At the same time, the attention mechanism can focus on changes in local frequency features of the image, effectively retain and enhance details and textures, which helps TCGAN to use texture information to perform image defogging, dehazing and deblurring tasks. The calculation method is as follows:

$$F_a = C_1(\text{Concat}(C_1(F_i), C_3(F_i), C_{D2}(F_i), C_{D3}(F_i))) \quad (7)$$

$$F_M = C_3(SE(\text{Concat}(C_3(F_a), F_i))) \quad (8)$$

Where F_i represents the input feature, C_i represents the convolution operation, C_{Di} represents the dilated convolution operation, $\text{Concat}(\bullet)$ represents the concatenation operation, $SE(\bullet)$ represents the SE attention mechanism operation, and F_M represents the output feature.

Neural gradient algorithm Layer(NGAL)

Let τ represent the number of layers of the neural network, m_t represent the number of neurons in the t layer, $t \in 1, 2, \dots, \tau$. We said t layer for the output of the $x^{(t)} = [x_1^{(t)}, x_2^{(t)}, \dots, x_{m_t}^{(t)}] \in \mathbb{R}^{m_t}$, and $x^{(0)}$ is the input of the network.

Suppose the network has N inputs, denoted $x^{(0)}(N)$, $N = 1, 2, \dots, n$, for the n -th input, the i -th output of the t -th layer is as follows:

$$x_i^{(t+1)}(n) = \phi(v_i^{(t)}) \quad (9)$$

$$v_i^{(t)} = \sum_{j=0}^{m_t} \theta_{i,j}^{(t)} x_j^{(t)}(n) + z_i^{(t)}(n) \quad (10)$$

Where $x_j^{(t)}(n)$ represents the j -th input of the n -th data in the t -th layer; $\theta_{i,j}^{(t)}$ represents the weight of the i -th input in the t -th layer; $v_i^{(t)}$ represents the i -th output of the t -th layer; ϕ represents the activation function,

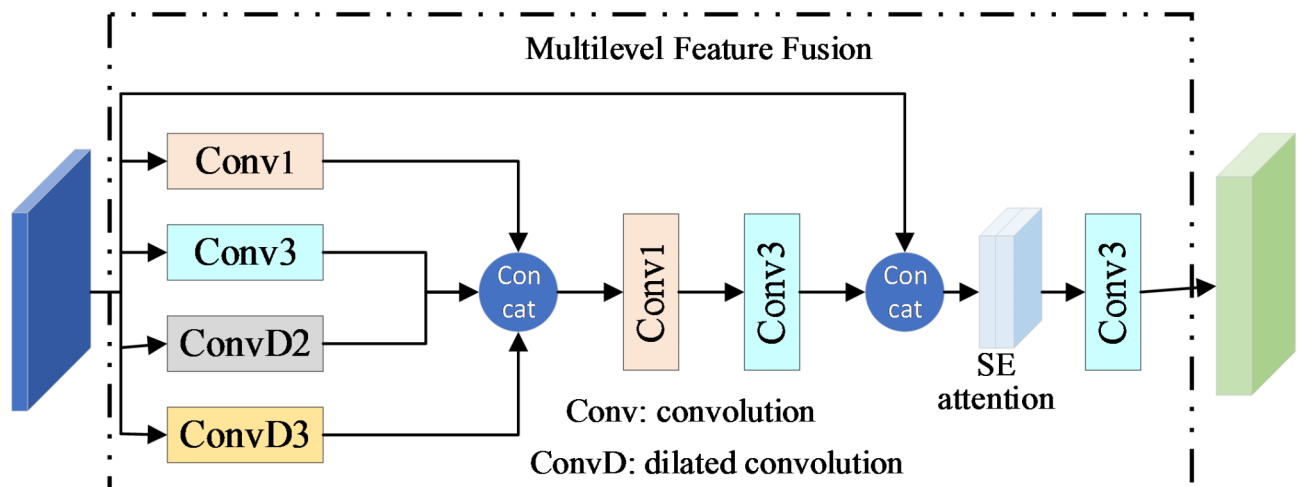


Fig. 2. Multi-level feature attention layer diagram.

and $z_i^{(t)}(n)$ represents the Gaussian disturbance with gradient added to the i -th neuron of the t -th layer by the n -th data.

The loss function is represented by L , and for the n -th data $x^{(0)}(n)$ labeled as $Y(n)$, there is a loss value represented by $L(x^{(\tau)}(n), Y(n))$. In our work, we optimize the size of the noise level $\sigma_i^{(t)}$ of the central normal random Gaussian perturbation for each neuron, namely $z_i^{(t)}(n) = \sigma_i^{(t)} \epsilon_i^{(t)}(n)$, Where $\epsilon_i^{(t)}(n)$ is a standard normal random variable. The residual of the n -th data propagated backward through the neural network at the i -th neuron in the t -th r is defined as:

$$\delta_i^{(t)}(n) = \begin{cases} e_i^{(\tau)}(n) \phi' \left(v_i^{(\tau-1)}(n) \right) & t = \tau \\ \phi' \left(v_i^{(t-1)}(n) \right) \left(\sum_{j=0}^{m_k} \theta_{i,j}^{(t)} \delta_j^{(t+1)}(n) \right) & t < \tau \end{cases} \quad (11)$$

Where $e_i^{(\tau)}(n)$ is defined as:

$$e_i^{(\tau)}(n) = \left. \frac{\partial L(x, Y(n))}{\partial x_i} \right|_{x=x^{(\tau)}(n)} \quad (12)$$

Neural network backpropagation essentially provides information about all parameters $\theta_{i,j}^{(t)} (t = 1, 2, \dots, \tau - 1)$, estimate the path random derivative of the loss function L . As shown by formula (13), $j \in \{0, 1, \dots, m_t\}, i \in \{0, 1, \dots, m_{t+1}\}$.

$$\frac{\partial L(x^{(\tau)}(n), Y(n))}{\partial \theta_{i,j}^{(t)}} = \delta_j^{(t+1)}(n) x_j^{(t)}(n) \quad (13)$$

neural gradient algorithm calculation steps are as follows:

- First input training data $P = \{(x^{(0)}(n), Y(n))\}_{n=1}^N$, loss function L .
- Construct neural network.
- Use Eq. (9) to calculate the output $x^{(\tau)}(n)$.
- Calculate loss function $L(x^{(\tau)}(n), Y(n))$.
- Use Eqs. (10) and (11) respectively to estimate the gradient of loss to weight and Gaussian perturbation.
- Update weights and noise levels.
- Repeat steps c through f until the parameters meet the requirements of the model.

Dual path attention layer (DPAL)

In order to enable the model to adjust the attention to the features of different spatial regions and different channels adaptively, enhance the model's feature extraction capability, and achieve more efficient image enhancement. Therefore, the dual path attention layer is proposed, and the structure is shown in Fig. 3.

It can be seen from the figure that it consists of two attention paths, the first path is based on the average and max pooling of all channels in the feature graph, and then the two results are combined to enhance the feature learning ability of the model. The functions of the two operations are different, but complementary. Average pooling aims to capture the overall features of the transmission line image, while max pooling aims to capture the prominent features (edges, textures) of the transmission line image. Subsequently, the attention weights are

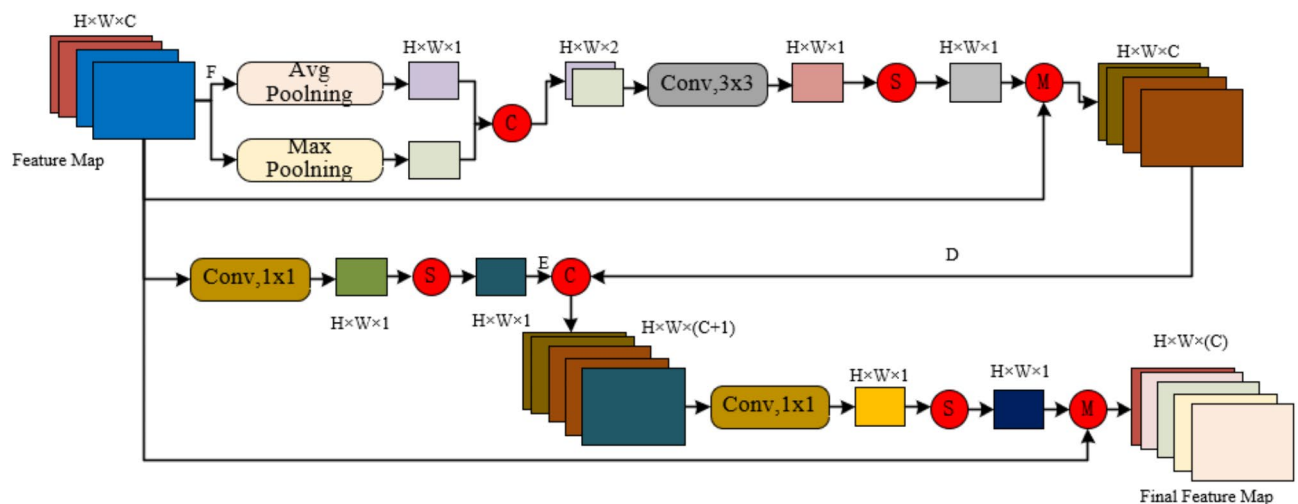


Fig. 3. Multi-level feature attention layer diagram.

generated through convolution operations and Sigmoid activation functions, and feature weighting is performed on the feature map. The calculation process is as follows.

$$\text{MeanPooling}(x) = \frac{1}{C} \sum_{c=1}^C F(i, j, c) \quad (14)$$

$$\text{MaxPooling}(x) = \text{Max}_{c=1}^C F(i, j, c) \quad (15)$$

$$F_1 = \text{Concat}(\text{MeanPooling}(x), \text{MaxPooling}(x)) \quad (16)$$

$$D = \sigma(\text{Conv}^{1 \times 1}(F_1)) \times F \quad (17)$$

Where i, j represent the index value of the feature graph space, c represents the number of channels, σ represents the activation function, and $\text{Conv}^{1 \times 1}$ represents the 1×1 convolution operation.

The second path directly compacts the number of channels in the input feature map from C to 1 through a 1×1 convolution, fuses the information of all channels together, and generates a feature map of “global channel information”. Then, through Sigmoid activation function, this feature map generates a channel attention map to highlight or suppress different channel information. The calculation process is as follows.

$$E = \sigma(\text{Conv}^{1 \times 1}(F)) \quad (18)$$

Then, the spatial attention weighting result of the upper path is concatenated with the channel attention of the lower path, and then the channel number of the input feature map is compressed from C to 1 by a 1×1 convolution. and then the attention map is generated by the Sigmoid activation function, finally, the feature weighting is performed on the original feature map. The calculation process is as follows.

$$F_{final} = \sigma(\text{Conv}^{1 \times 1}(\text{Concat}(D, E))) \times F \quad (19)$$

F_{final} represents the final generated feature map, and D, E represent the results of different path respectively. By concatenating the results of the two paths, the channel attention of the second path can be directly combined with the spatially weighted feature map to form a richer feature map, which contains the combination information of spatial attention and channel attention, without limiting the complementarity of the two kinds of attention, thus enhancing the diversity of feature expression. Finally, the model generates the overall weight map through 1×1 convolution and Sigmoid activation. Combining these two attention mechanisms, the model can adaptively adjust the features of different spatial regions and the importance of different channels, so as to achieve more efficient and accurate image enhancement effects.

Experiment

Dataset

In this work, we use datasets collected by drones and extended to existing data. Among them, there are 700 rainy images, 600 foggy images and 800 blurry images. Each type of image data is divided into train set, var set and test set according to the ratio of 7:2:1, as shown in Table 1.

Implementation details

In this paper, the TGTIE method is trained in a phased way. First, TINet is trained, and then its parameters are fixed, and the TCGAN network is trained. This experiment uses the Pytorch framework to train the model on the Ubuntu system NVidia RTX-4090. In the TINet network training process, Batch size was 2, Adam optimizer was used, the initial learning rate was 1×10^{-4} , and the training rounds were 50. In the training process of TINet network, the Batch size is 4, the Adam optimizer is used, the initial learning rate is 5×10^{-3} , and the training rounds are 200. We scaled the pixel resolution of the input image to 512×512 pixel resolution for network training. The Peak Signal-to-Noise Ratio (SSIM) and Structural Similarity (PSNR) were used as experimental evaluation indicators to evaluate the performance of network training.

Ablation experiment

Network architecture analysis

In order to verify the effect of texture inference network (TINet) on guiding image enhancement and the effect of local discriminator on preventing local artifacts in im-ages. We performe ablation experiments on the TGTIE variant. (1) TCGAN: removes TINet network from TGTIE and do not use texture information to guide; (2) TINet + TCGAN w/o LD: Remove the local discriminator from TCGAN. The quantitative ex-perimental results are shown in Table 2, and the qualitative experimental results are shown in Fig. 4.

	Rain	Fog	Blur
train	490	420	560
val	140	120	160
test	70	60	80

Table 1. Data set category and quantity.

Method	Rain		Fog		Blur	
	PSNR(dB)	SSIM	PSNR(dB)	SSIM	PSNR(dB)	SSIM
TCGAN	31.924	0.940	33.354	0.949	27.515	0.887
TINet+ TCGAN w/o LD	32.547	0.945	34.372	0.954	28.547	0.905
TGTLIE	33.218	0.956	34.921	0.962	30.725	0.926

Table 2. Results of network architecture ablation experiment.

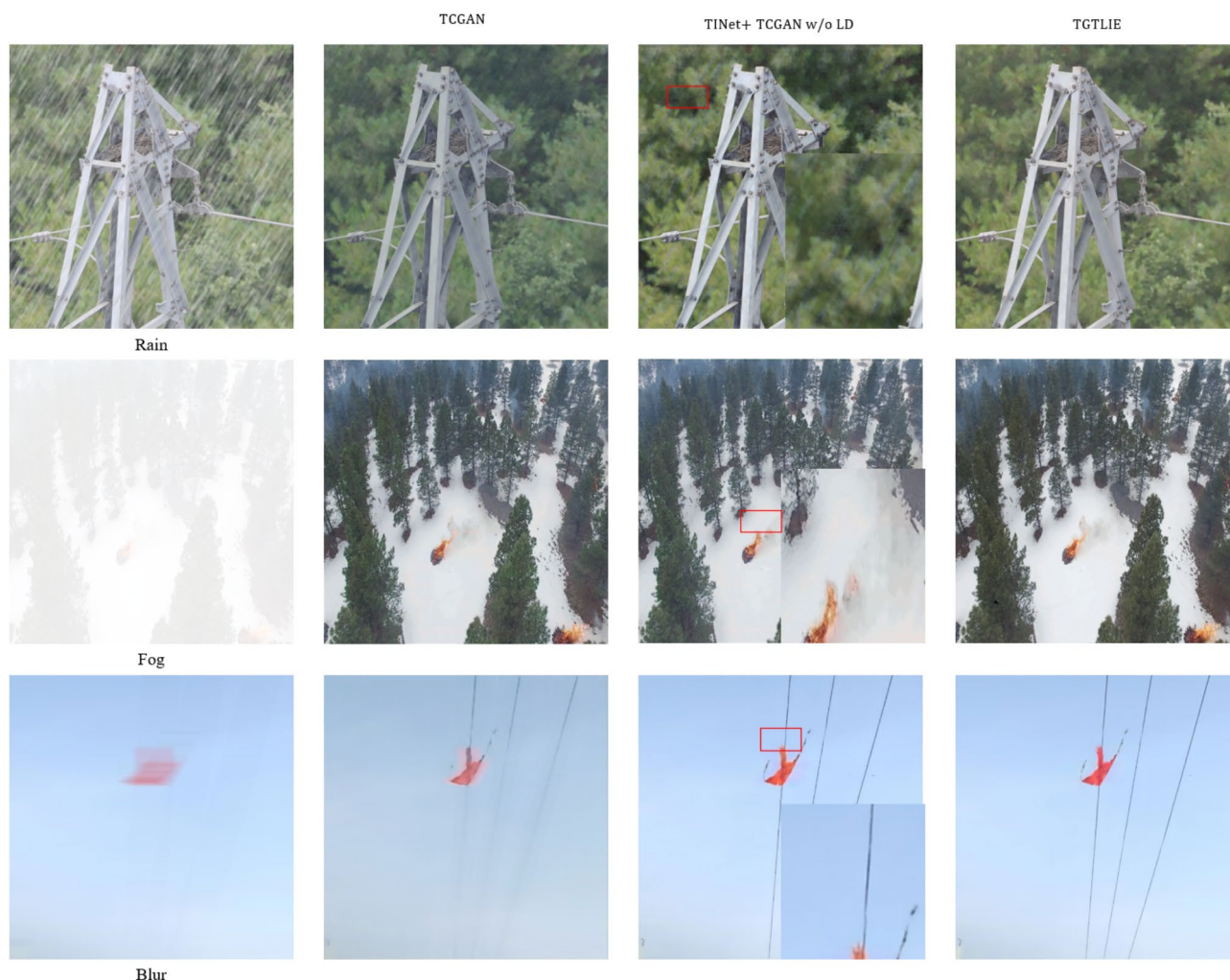


Fig. 4. Visualization results of network architecture ablation experiment.

As can be seen from Table 2, the image enhancement performance of TGTLIE and TINet+TCGAN w/o LD is always better than that of TCGAN. The PSNR and SSIM of TGTLIE are improved by 1.294dB, 1.567dB, 3.210dB and 0.016, 0.013, 0.039 respectively compared with TCGAN. The PSNR and SSIM of TINet+TCGAN w/o LD are improved by 0.623dB, 1.018dB, 1.032dB and 0.005, 0.005, 0.018 respectively compared with TCGAN. It is shown that the guidance of texture information can effectively improve the performance of model image enhancement. The PSNR and SSIM of TGTLIE are improved by 0.671dB, 0.549dB, 2.178 dB and 0.011, 0.008, 0.021 respectively compared with TINet+TCGAN w/o LD, respectively, indicating that the local discriminator can focus on the local details of the image and improve the image enhancement effect.

As can be seen from Fig. 4, TCGAN lacks the guidance of texture information, and the object details in the generated images are relatively fuzzy. For example, the texture structure of the tower in the generated image of rain image is unclear, surface artifacts exist in the generated image of fog image, and kites and cables are not clearly displayed in the generated image of blur image. TINet+TCGAN w/o LD lacks local discriminator, and local details in the generated map are fuzzy, such as local artifacts in the generated map of rain map, color shadows around mountain fires in the generated map of fog map, and partial missing kites and cables in the generated map of blur map. The texture structure and detail information of all generated images by TGTLIE are

Method	Rain		Fog		Blur	
	PSNR(dB)	SSIM	PSNR(dB)	SSIM	PSNR(dB)	SSIM
TGTLIE w/o NGAL, DPAL	31.622	0.940	33.065	0.947	27.684	0.881
TGTLIE w/o NGAL	32.354	0.947	33.821	0.952	28.786	0.908
TGTLIE w/o DPAL	32.725	0.951	34.254	0.955	29.345	0.911
TGTLIE	33.218	0.956	34.921	0.962	30.725	0.926

Table 3. Results of network module ablation experiment.

Method	Params /M	FLOPs /G	Time /S	Rain		Fog		Blur	
				PSNR(dB)	SSIM	PSNR(dB)	SSIM	PSNR(dB)	SSIM
SE	21.59	58.26	0.095	31.635	0.940	33.086	0.948	27.736	0.891
CA	21.82	58.28	0.112	31.923	0.943	33.351	0.950	28.204	0.897
CBAM	21.46	58.25	0.105	31.934	0.943	33.462	0.951	28.342	0.901
DPAL	21.42	58.22	0.100	32.354	0.947	33.821	0.952	28.786	0.908

Table 4. Results of attention module ablation experiment.

clearer, rain lines, fog and fuzzy areas are effectively removed, and local details of the image are more complete and clear. This proves the effectiveness of the texture inference network and the local discriminator, making the generated images clearer in overall visual effect.

Network module analysis

In order to verify the effectiveness of the designed components in TGTLIE, different ablation experiments are performed on the dataset by using different components, including the neural gradient algorithm layer (NGAL) and the dual path attention layer (DPAL). (1) TGTLIE w/o NGAL, DPAL: neural gradient algorithm layer (NGAL) and dual path attention layer (DPAL) are not used in TGTLIE, only convolution operation is used. (2) TGTLIE w/o NGAL: No neural gradient algorithm layer (NGAL) is used in TGTLIE; (3) TGTLIE w/o DPAL: dual path attention layer (DPAL) is not used in TGTLIE. The results are shown in Table 3.

It can be seen from the table, when TGTLIE does not use the NGAL and DPAL, the PSNR and SSIM of the model are 31.622dB, 33.065dB, 27.684dB and 0.940, 0.947, 0.881, respectively. After adding DPAL, the PSNR and SSIM of TGTLIE are increased by 0.732dB, 0.756dB, 1.102 dB and 0.007, 0.005, 0.027 respectively, indicating that the DPAL can make the model pay more attention to important areas in the image and improve the image enhancement effect. After adding NGAL, the PSNR and SSIM of TGTLIE increased by 1.103dB, 1.189dB, 1.661dB and 0.011, 0.008, 0.030, respectively, indicating that the NGAL enhances the robustness of the model and enables the model to better retain the image structure details, and improves the model's feature extraction ability. After using the NGAL and the DPAL at the same time, the image enhancement effect of TGTLIE is optimal. This shows the effectiveness and necessity of the various components proposed in our method for transmission line image enhancement.

Attention module analysis

In order to verify the effectiveness of the dual path attention layer (DPAL) designed in TGTLIE in terms of computational efficiency and accuracy. Different ablation experiments are being conducted by using DPAL, SE, CoordAttention (CA), and CBAM, respectively. The results are shown in Table 4.

It can be seen from the table, PSNR and SSIM are improved after DPAL, SE, CA and CBAM are introduced into the model. It shows that the introduction of attention mechanism can improve the image enhancement performance of the model. The detection speed of DPAL is slightly lower than that of SE, but compared with SE, CA and CBAM, DPAL has fewer parameters and floating point computation, and better detection performance. Taking these factors into account, DPAL has better overall performance and can improve the enhancement of the model without adding too much complexity to the model.

Contrast experiment

In order to verify the advanced performance of TGTLIE in image enhancement of transmission lines and accurately evaluate the performance of TGTLIE, quantitative experiments were conducted with CBDNet, Uformer, MSPFN and DGNL-Net, and the results are shown in Table 5.

It can be seen from Table 5, the PSNR and SSIM values of TGTLIE achieve the best performance in image enhancement under different environmental interferences, the PSNR and SSIM values are 33.218dB, 34.921dB, 30.725dB and 0.956, 0.962, 0.926, respectively. Compared with DGNL-Net, the PSNR and SSIM of TGTLIE are improved by 0.777dB, 0.667dB, 3.410 dB and 0.006, 0.007, 0.049, respectively. Compared with Uformer, the PSNR and SSIM of TGTLIE are improved by 1.594dB, 1.109dB, 1.343 dB and 0.011, 0.011, 0.009, respectively. It is proved that the proposed method TGTLIE has good image enhancement effect and can recover clear images, verifying the effectiveness and advancement of TGTLIE in transmission line image enhancement.

Method	Rain		Fog		Blur	
	PSNR(dB)	SSIM	PSNR(dB)	SSIM	PSNR(dB)	SSIM
CBDNet	29.561	0.935	31.843	0.938	25.068	0.846
MSPFN	29.732	0.937	32.563	0.943	25.531	0.853
Uformer	31.624	0.945	33.812	0.951	29.382	0.915
DGNL-Net	32.441	0.95	34.254	0.955	27.315	0.877
TGTLIE	33.218	0.956	34.921	0.962	30.725	0.926

Table 5. Experimental results of different methods.

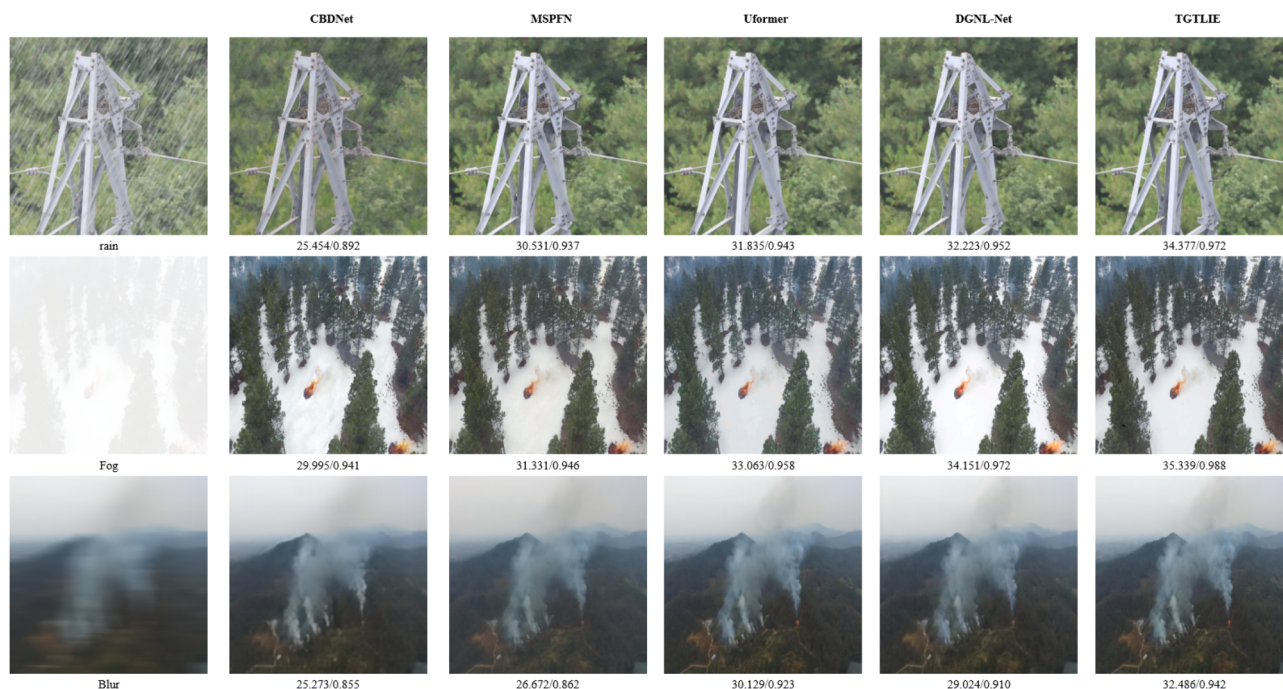


Fig. 5. Visualization results of different methods.

Qualitative experiment

In order to comprehensively evaluate the performance of the proposed method, in addition to relying on objective quantitative indicators, qualitative visualization experiments are also indispensable. Through intuitive visual observation, we can reveal the feeling of the image in terms of detail preservation, color restoration, noise suppression, etc. We conducted qualitative experiments with CBDNet, Uformer, MSPFN and DGNL-Net to demonstrate the performance of the TGTLIE in different image types. The experimental results are shown in Fig. 5.

As can be seen from Fig. 5, there are artifacts on the surface of the image enhanced by CBDNet method, and the image details are fuzzy and the texture is not clear. Although the enhanced image by MSPFN method reduces the artifacts on the image surface, it is easy to cause the image color deviation. Although the image enhanced by Uformer and DGNL-Net methods has no artifacts, the local details of the image are missing, especially in the processing of blurred images. The TGTLIE can better remove rain streaks, fog and deformation, and the enhanced image has no artifact interference, accurately restore the true color of the image, and retain the details in the image. The enhanced image is clearer, true and more pleasant in visual experience. The excellent performance in quantitative and qualitative evaluation consistently proves the great effectiveness of the TGTLIE in transmission line image enhancement, which lays a good foundation for the subsequent transmission line monitoring.

Object detection application experiment

In order to verify the effectiveness of TGTLIE in the actual application of transmission line monitoring after image enhancement, this paper conducted object detection experiments before and after image enhancement. Through the transmission line foreign body detection algorithm ETSD-YOLO developed by us before³³, the detection results in different environments are shown in Fig. 6.

As can be seen in Fig. 6, after the transmission line image is disturbed by rain, fog and blur, the false detection rate and missing rate of the image also increase, and the detection algorithm mistakenly identifies non-existent

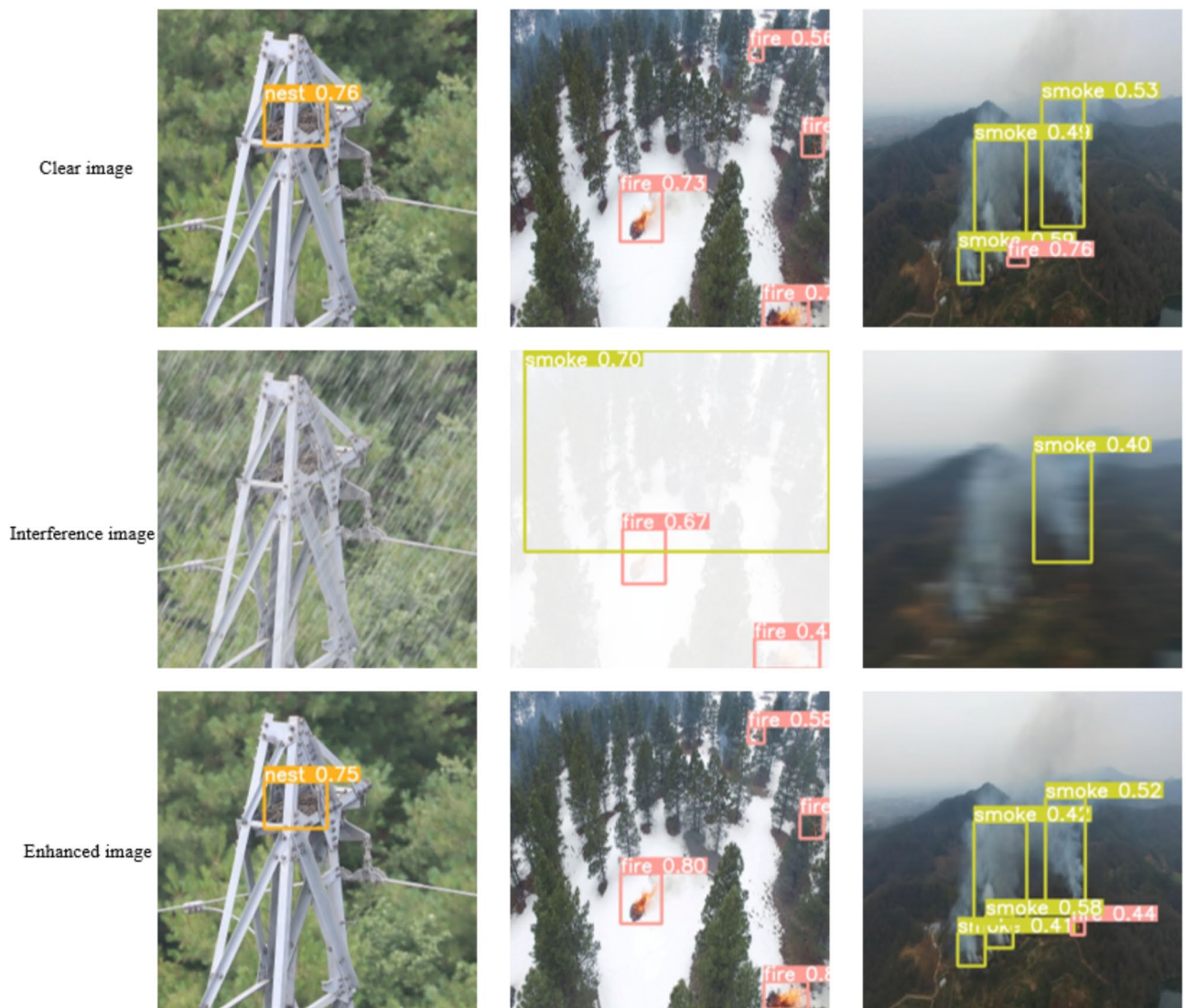


Fig. 6. Object detection results before and after image enhancement.

targets in the image. For example, the nest object cannot be detected in the rain map, the smoke object in the fog map is mistakenly detected, and the mountain fire and smoke object in the blur image are missed. After enhanced by the TGTLE, all the objects in the image are detected, and the detection accuracy does not decrease. This shows that TGTLE can effectively remove the interference of rain, fog and blur in the image, improve the quality of the image, make the image clearer and more realistic, and contribute to the data analysis and application of the transmission line monitoring.

Adaptability experiment

In order to verify the robustness and generalization of the proposed method for transmission line image enhancement, experimental validations were conducted in various transmission line infrastructure environments, snowy weather, and with low-resolution images. The experimental results are presented in Fig. 7.

As shown in Fig. 7, Figure (A) shows the inspection image of transmission line in a rural area during rainy weather. After enhanced by the method in this paper, the image is clearly displayed, and the Trash foreign object is accurately detected. Similarly, Figure (B) shows the inspection image of transmission line in urban areas, where the Kite foreign object is also accurately detected. Figure (C) shows the inspection image of transmission line under extreme snowy weather in winter. After enhancement by the method in this paper, the Nest foreign object is accurately detected. Figure (D) shows the inspection image of transmission line captured in low-resolution foggy weather. Although there are some artifacts and slight deformation of foreign objects in the image after enhancement using the proposed method, the Kite foreign object is still accurately detected. These results indicate that the method in this paper achieves good image enhancement effects in various practical scenarios, improves image quality, and enhances the accuracy of power line monitoring, thereby proving the strong robustness and generalization ability of the proposed method.

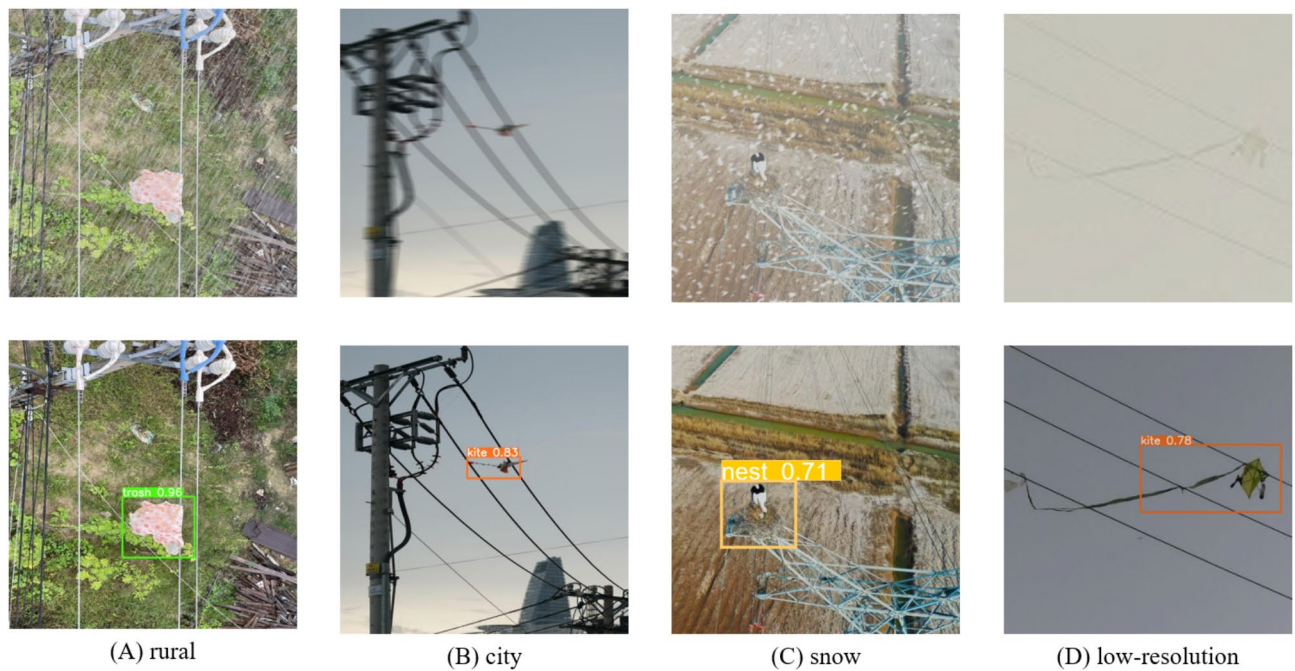


Fig. 7. Experimental results of adaptability in different environments.

Conclusion

In this paper, a texture guided transmission line image enhancement method (TGTLE) is proposed to improve image quality and the accuracy of foreign object detection, aiming at the interference of natural environmental factors (such as rain, fog, blur, etc.) in transmission line inspection. Firstly, texture inference network (TINet) is designed, which can effectively extract texture information from the input disturbed images, then the texture information is used as prior knowledge to guide the subsequent image enhancement process. Then, we use texture-based conditional generation adversarial network (TCGAN) to realize adaptive image de-rain, de-fog and de-blur processing. In order to further improve the detail quality and robustness of image generation, we introduce neural gradient algorithm and the dual path attention in TCGAN. These techniques help the model pay more attention to important details in the image and reduce noise interference. At the same time, in order to prevent local artifacts, a global-local discriminator structure is designed in the discriminant network, and the generated images are inspected globally and locally. During the training process, the multi-stage joint loss function is used, including Charbonnier loss, SSIM loss and global-local generative counter loss, which helps the model to improve image quality and detail performance while maintaining image authenticity. Through quantitative and qualitative experimental, TGTLE can effectively remove noise interference under different meteorological conditions, significantly improve image quality, and provide effective technical support for intelligent inspection and fault warning of transmission lines. In the future, we will continue to optimize TGTLE to improve its real-time, robustness and generalization to meet the challenges of a more complex and changing natural environment.

Data availability

The data that support the findings of this study are available on request from the reader upon reasonable request.

Received: 24 December 2024; Accepted: 27 February 2025

Published online: 29 March 2025

References

1. Zheng, J. et al. GEB-YOLO: a novel algorithm for enhanced and efficient detection of foreign objects in power transmission lines. *Sci. Rep.* **14**, 15769. <https://doi.org/10.1038/s41598-024-64991-9> (2024).
2. Wang, B., Li, C., Zou, W. & Zheng, Q. Foreign object detection network for transmission lines from unmanned aerial vehicle images. *Drones* **8**, 361. <https://doi.org/10.3390/drones8080361> (2024).
3. Yan, S. et al. An algorithm for power transmission line fault detection based on improved YOLOv4 model. *Sci. Rep.* **14**, 5046. <https://doi.org/10.1038/s41598-024-55768-1> (2024).
4. Zhang, M. et al. Study on the enhancement method of online monitoring image of dense fog environment with power lines in smart City. *Front. Neurobotics*. **16**, 1104559 (2023).
5. Zou, K., Zhao, S. & Jiang, Z. Power line scene recognition based on convolutional capsule network with image enhancement. *Electronics* **11**, 2834. <https://doi.org/10.3390/electronics11182834> (2022).
6. Zai, W. & Yan, L. Multi-Patch hierarchical transmission channel image dehazing network based on dual attention level feature fusion. *Sensors* **23**, 7026. <https://doi.org/10.3390/s23167026> (2023).

7. Cheng, Y. et al. A single image Derain method based on residue channel decomposition in edge computing. *Intell. Autom. Soft Comput.* **37** (2), 1469–1482. <https://doi.org/10.32604/iasc.2023.038251> (2023).
8. Li, C. et al. Attention-generative adversarial networks for simulating rain field. *IET Image Proc.* **18** (6), 1540–1549 (2024).
9. Shen, M., Lv, T., Liu, Y., Zhang, J. & Ju, M. A. Comprehensive review of traditional and Deep-Learning-Based defogging algorithms. *Electronics* **13**, 3392. <https://doi.org/10.3390/electronics13173392> (2024).
10. Liu, Y., Sheng, Z. & Shen, H. L. Guided image deblurring by deep Multi-Modal image fusion. *IEEE Access.* **10**, 130708–130718. <https://doi.org/10.1109/ACCESS.2022.3229056> (2022).
11. Xiang, Y. et al. Deep learning in motion deblurring: current status, benchmarks and future prospects. *Visual Comput.*, 1–27, (2024).
12. Zhai, L., Wang, Y., Cui, S. & Zhou, Y. A comprehensive review of deep Learning-Based Real-World image restoration. *IEEE Access.* **11**, 21049–21067. <https://doi.org/10.1109/ACCESS.2023.3250616> (2023).
13. Wang, Y., Xie, W. & Liu, H. Low-light image enhancement based on deep learning: a survey. *Opt. Eng.* **61** (4), 040901–040901. <https://doi.org/10.1117/1.OE.61.4.040901> (2022).
14. Zhang, P. Image enhancement method based on deep learning. *Math. Probl. Eng.* **2022**(1), 6797367. <https://doi.org/10.1155/2022/6797367> (2022).
15. Zhu, M. L., Zhao, L. L. & Xiao, L. Image denoising based on GAN with optimization algorithm. *Electronics* **11**, 2445. <https://doi.org/10.3390/electronics11152445> (2022).
16. Li, Y. et al. Rain streak removal using layer priors. Proceedings of the IEEE conference on computer vision and pattern recognition. 2736–2744. (2016).
17. Kim, J. H., Lee, C., Sim, J. Y., Kim, C. S. & Australia Single-image deraining using an adaptive nonlocal means filter. 2013 IEEE International Conference on Image Processing, Melbourne, VIC, 914–917, (2013). <https://doi.org/10.1109/ICIP.2013.6738189>
18. Chen, D. Y. & Chen, C. C. Kang. Visual depth guided color image rain streaks removal using sparse coding. *IEEE Trans. Circuits Syst. Video Technol.* **24** (8), 1430–1455. <https://doi.org/10.1109/TCSVT.2014.2308627> (2014).
19. Tao, W., Yan, X., Wang, Y. & Wei, M. MFFDNet: single image deraining via Dual-Channel mixed feature fusion. *IEEE Trans. Instrum. Meas.* **73**, 1–13. <https://doi.org/10.1109/TIM.2023.3346498> (2024).
20. Wang, Q., Sun, G., Dong, J. & Zhang, Y. PFDN: pyramid feature decoupling network for single image deraining. *IEEE Trans. Image Process.* **31**, 7091–7101. <https://doi.org/10.1109/TIP.2022.3219227> (2022).
21. Yan, Z., Li, G., Si, L. & Dong, H. FPGNet: Single Image Deraining with High-Frequency Channel and Frequency Domain Prior Guidance. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Seoul, Korea, Republic of, 4410–4414, (2024). <https://doi.org/10.1109/ICASSP48485.2024.10446408>
22. Li, Q. et al. Histogram Equalization Based Cable Tunnel Image Defogging Method. 2022 IEEE 6th Advanced Information Technology, Electronic and Automation Control Conference, 420–424, (2022). <https://doi.org/10.1109/IAEAC54830.2022.9929507>
23. Li, C. Enhancement Method for Fog-degraded Images based on Multi-scale Retinex and Wavelet Transform. IEEE 7th Information Technology and Mechatronics Engineering Conference (ITOEC), Chongqing, China, 652–655, (2023). (2023). <https://doi.org/10.1109/ITOEC57671.2023.10292051>
24. Xin, R. et al. Insulator umbrella disc shedding detection in foggy weather. *Sensors* **22**, 4871 (2023).
25. Zhang, Y., Hu, Q., Wei, Y., Wang, J. & CycleGAN Image Defogging Method Based on Residual Dual Attention Mechanism. 3rd International Conference on Electronic Information Engineering and Computer Science, Changchun, China, 234–237, (2023). <https://doi.org/10.1109/EIECS59936.2023.10435552>
26. Wang, W. et al. Adaptive single image defogging based on Sky segmentation. *Multimed Tools Appl.* **82**, 46521–46545. <https://doi.org/10.1007/s11042-023-15381-2> (2023).
27. Zhou Jing, T. & Zhaoxing, W. Image defogging with diffusion model integrating external attention [J/OL]. *Electron. Meas. Technol.* **11** (11), 1–10 (2024).
28. Tan, Y., Zhang, L. & Chen, Y. Joint Non-Local statistical and wavelet tight frame Information-Based ℓ_0 regularization model for image deblurring. *IEEE Access.* **12**, 117285–117297. <https://doi.org/10.1109/ACCESS.2024.3448450> (2024).
29. Liang, C., Yin, Y., Qin, Y. & He, X. A Deblurring Method of Electric Power Inspection Images Based on GAN. 2021 7th International Conference on Systems and Informatics (ICSAI), Chongqing, China, 1–6, (2021). <https://doi.org/10.1109/ICSAI53574.2021.9664028>
30. Liu, Z. & Research on Deblurring Method for Insulator Images Based on Channel and Spatial Attention Mechanisms. IEEE International Conference on Power Science and Technology (ICPST), Kunming, China, 317–321, (2023). (2023). <https://doi.org/10.1109/ICPST56889.2023.10165447>
31. Wang, D. W. & Li, Y. D. Insulator object detection based on image deblurring by WGAN. *Electr. Power Autom. Equip.* **40** (5), 188–194 (2020).
32. Wang, Z. et al. Uformer: A General U-Shaped Transformer for Image Restoration. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), New Orleans, LA, USA, 17662–17672, (2022). (2022). <https://doi.org/10.1109/CVPR52688.2022.01716>
33. Zhao, L., Zhang, Y., Dou, Y., Jiao, Y. & Liu, Q. ETLSh-YOLO: an Edge-Real-Time transmission line safety hazard detection method. *Symmetry* **16**, 1378. <https://doi.org/10.3390/sym16101378> (2024).

Acknowledgements

The authors would like to thank the Development of Key Technologies for Resonance-Based Ice Removal Robots(24CX0923) for their support.

Author contributions

Y.Z. was mainly responsible for the conceptual design of the project and the development of the research framework. He also participated in data analysis and wrote the main part of the research report. L.Z. was responsible for the design and execution of the experiments and the collection of experimental data. He was also involved in the preliminary analysis of the experimental results. Y.D. was responsible for project management and co-ordination to ensure that the project schedule met expectations. Y.J. contributed significantly in data processing and statistical analysis. He also contributed to the literature review and the development of the theoretical framework. Q.L. was responsible for the preparation of experimental materials and maintenance of experimental equipment.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to Y.D.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025