International Journal of Sustainable Energy and Environmental Research
2025 Vol. 14, No. 2 pp. 65-75
ISSN(e): 2306-6253
ISSN(p): 2312-5764
DOI: 10.18488/13.v14i2.4453

Research on super-resolution reconstruction of transmission lines based on SinGAN using a single image

Xinyu Song[1]
Yifan Zhang[2]
Zhe Yin[3]
Ruijin Zhu[4+]

[1,2,3,4]*College of Clean Energy and Electrical Engineering, Xizang Agricultural and Animal Husbandry University, Nyingchi, China.*
[1]*Email: 925275903@qq.com*
[2]*Email: 1205949897@qq.com*
[3]*Email: 1211026133@qq.com*
[4]*Email: zhuruijin@xza.edu.cn*


(+ Corresponding author)

Article History
Received: 30 July 2025
Revised: 2 September 2025
Accepted: 23 September 2025
Published: 3 October 2025

Keywords
Transmission lines
SinGAN
Super-resolution
Image clarity
Detection performance
Smart grid inspection.

ABSTRACT

The purpose of this study is to address the problem of blurry transmission line images caused by foggy and cloudy weather, which severely hinders defect detection and intelligent recognition. To improve the quality and usability of inspection images, a SinGAN-based single-image super-resolution reconstruction method is proposed. The methodology exploits SinGAN's advantage of requiring no paired datasets and adopts a self-supervised framework with an image pyramid structure. By learning texture features at multiple scales, the model progressively reconstructs high-resolution images and restores structural and textural details from a single blurry input. The findings demonstrate that the proposed method significantly enhances image clarity and improves structural fidelity. Quantitative evaluation is performed using PSNR and SSIM metrics for image quality and YOLOv7 mAP for detection performance. Results show clear improvements in detail restoration and detection accuracy compared with conventional super-resolution approaches, especially in scenarios with scarce samples. The practical implications of this study highlight that the SinGAN-based approach provides a scalable and efficient solution for real-world transmission line inspections. By eliminating the need for paired training data and enabling rapid deployment, this method enhances smart grid inspection capabilities under challenging conditions and offers strong potential for practical engineering applications.

Contribution/Originality: This study contributes to the existing literature by proposing a SinGAN-based single-image super-resolution method for blurry transmission line images. Unlike traditional approaches that require paired datasets, it employs unsupervised self-learning with an image pyramid, significantly improving clarity and YOLOv7 detection accuracy. This approach offers a scalable solution for real-world inspections.

1. INTRODUCTION

Transmission lines are a core component of power systems, responsible for transmitting energy from power generation to consumption (Alcayde-García, Salmerón-Manzano, Montero, Alcayde, & Manzano-Agugliaro, 2022). Their safe and stable operation directly affects the reliability of power supply and the normal functioning of the social economy. In remote areas, transmission lines are often located in complex terrain and variable climatic conditions, exposed to adverse weather conditions such as rain, fog, clouds, and snow over long periods of time. This not only makes transmission lines susceptible to natural disasters such as corrosion, ice accumulation, and wind deflection, but also poses significant challenges for manual inspections and intelligent maintenance (Zhang, 2025).

In traditional manual inspections and image acquisition processes, images captured in remote areas are often affected by factors such as insufficient lighting, cloud cover, and image shaking, resulting in blurry images, unclear target contours, broken edges, and incomplete structural information. Such low-quality images not only reduce the detection accuracy of image recognition systems for equipment defects but also significantly limit the development of smart grids and the deployment of unmanned inspection technologies (Fang, Misra, Xue, & Yang, 2012). Therefore, improving image resolution and clarity, and restoring structural and textural details, has become a critical issue in power line inspection image processing (Faisal et al., 2025).

Traditional super-resolution (SR) reconstruction methods, such as SRCNN (Dong, Loy, & He, 2014) based on deep convolutional networks, VDSR (Kim, Lee, & Lee, 2016) and ESRGAN (Wang et al., 2018) optimized by generative adversarial networks, typically rely on large-scale paired low-resolution (LR) and high-resolution (HR) images for supervised training. While this approach performs exceptionally well on public datasets, its reliance on the quality and quantity of training data makes it difficult to apply in scenarios where samples are scarce and high-quality paired images are unavailable, posing significant limitations in power transmission image scenarios (Yang, Wang, Diao, Qi, & Han, 2018).

In recent years, some unsupervised or few-shot SR methods have been proposed to address the problem of training with unpaired images or single images. For example, ZSSR (Shocher, Cohen, & Irani, 2018) constructs training samples by downsampling the images themselves, while DPSR (Zhang, Liang, Van Gool, & Timofte, 2021) attempts to enhance the training process by using a degraded model. However, these methods still require additional image construction or prior knowledge design, resulting in high model complexity and limited generalization ability.

To this end, this paper introduces a single-image super-resolution method based on SinGAN for detail restoration and structure enhancement of blurred images of power transmission lines. SinGAN is a multi-scale generative adversarial network that uses a single image as a training sample. By constructing an image pyramid model, it learns the structural and textural features of images at each scale under unsupervised conditions, achieving high-resolution reconstruction from coarse to fine (Shaham, Tali, & Tomer, 2019). Compared with traditional methods, SinGAN has the following significant advantages:

- Completely independent of paired images or datasets, suitable for image processing at construction sites.
- Lightweight network structure and simple training process, suitable for rapid deployment.
- Generates high-detail images while maintaining the original style of the image, improving detection friendliness.

This study contributes to the existing literature by proposing a SinGAN-based single-image super-resolution method for blurry transmission line images. Unlike traditional approaches requiring paired datasets, it employs unsupervised self-learning with an image pyramid, significantly improving clarity and YOLOv7 detection accuracy, offering a scalable solution for real-world inspections.

2. GENERATIVE ADVERSARIAL NETWORKS

Generative Adversarial Networks (GANs) are deep learning-based models consisting of two main components: a generator and a discriminator (Goodfellow et al., 2014). The generator's task is to synthesize data samples that are as close as possible to real data from random inputs, such as images, speech, or text; the discriminator's role is to determine whether the input data is a real sample or a generated sample. During training, the two components engage in a continuous adversarial process: the generator strives to "deceive" the discriminator into generating more realistic samples, while the discriminator works to enhance its ability to distinguish between real and generated samples. Through this competitive mechanism, the model as a whole gradually improves, enabling the generator to learn to produce increasingly realistic samples.

2.1. GAN Structure

Figure 1 illustrates the basic structure of a Generative Adversarial Network (GAN), which comprises a generator and a discriminator. The generator takes random noise from the latent space and produces synthetic samples, while the discriminator evaluates both real and synthetic samples to determine their authenticity. During training, the generator aims to "deceive" the discriminator by creating increasingly realistic samples; simultaneously, the discriminator enhances its ability to distinguish between real and fake samples. Through adversarial training, both components iteratively improve their performance, ultimately allowing the synthetic samples to approximate the distribution of real data more closely.

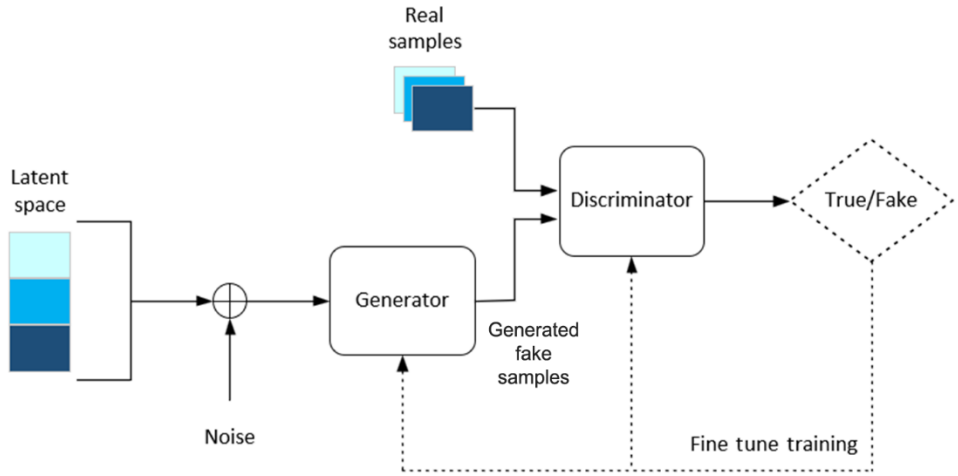

Figure 1. The structure of GAN.

2.2. GAN Structure Optimal Solution of the Objective Function

GAN consists of a generator G and a discriminator D, and its objective function is a two-person zero-sum game (Raghavan, 2012) as shown in Formula 1.

$$\min_G \max_D V(D,G) = \mathbb{E}_{x\sim p_{\text{data}}(x)}[log\, D\,(x)] + \mathbb{E}_{z\sim p_z(z)}[log(1 - D(G(z)))] \quad (1)$$

P_{data} represents the distribution of real data, p_z represents the noise distribution in the latent space, $G(z)$ represents the pseudo samples generated by the generator, and $D(x)$ represents the probability that the discriminator outputs a sample as real.

In this adversarial structure, the discriminator aims to maximize the probability of correctly classifying real samples while minimizing the probability of misclassifying synthetic samples as real; the generator, on the other hand, seeks to minimize the entire function, making the generated samples as realistic as possible.

When the generator is fixed, the optimal strategy for the discriminator is shown in Formula 2:

$$D^*(x) = \frac{p_{\text{data}}(x)}{p_{\text{data}}(x)+p_g(x)} \quad (2)$$

Substituting the optimal discriminator into the original objective function, the optimization objective of the generator is transformed into minimizing the Jensen-Shannon divergence (JSD), as shown in Formula 3:

$$\min_G V(D^*,G) = -log\,4 + 2\cdot JSD(p_{\text{data}} \parallel p_g) \quad (3)$$

Therefore, when the distribution p_g generated by the generator is completely consistent with the true distribution P_{data}, the JSD value is zero, the objective function reaches its minimum value $-log\,4$, and the GAN reaches the optimal solution. At this point, the discriminator output is constant at 0.5, unable to distinguish between real and fake samples, and the network training tends to converge.

3. MATERIALS AND METHODS

3.1. SinGAN Model Basics

SinGAN (Single-Image GAN) is an innovative unsupervised generative adversarial network that does not require a large number of paired high- and low-resolution image samples. It can perform a variety of tasks, such as image enhancement, reconstruction, style transfer, image editing, and super-resolution reconstruction, using only a single image. SinGAN is particularly suitable for real-world applications where image acquisition is challenging and sample scarcity is a concern, offering advantages such as lightweight structure, efficient training, and rich generation results. In this study, we apply SinGAN's complete multi-scale modeling mechanism to the super-resolution reconstruction task of blurry power transmission line images, aiming to achieve dual improvements in image clarity and subsequent recognizability.

3.2. Multi-scale Pyramid Structure and Process

SinGAN uses a multi-scale pyramid structure to gradually learn image features. The original image is scaled to form a sequence of images from coarse to fine, with each layer of images serving as the training input for the current scale. Each scale contains a generator and a discriminator D_n. Training starts from the lowest resolution and progresses upward layer by layer to the original resolution. The multi-scale pyramid structure is shown in Figure 2.

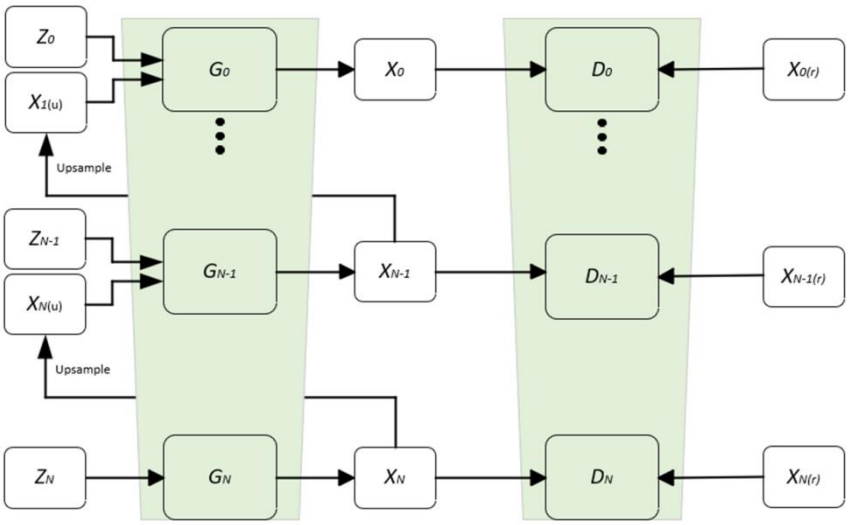

Figure 2. Multi-scale pyramid structure.

The model consists of a pyramid-shaped GAN, where both training and inference are performed in a coarse-to-fine manner. At each scale, G_N learns to generate image samples such that all overlapping image patches cannot be distinguished by the discriminator D_N from patches in the downsampled training image $X_{N\ (r)}$; the effective patch size gradually decreases from the bottom to the top of the pyramid. The inputs to the generator G_N include: a random noise image Z_N, and an image X_N generated from the previous scale, which is upsampled to match the current resolution (except for the bottom layer, which relies entirely on random noise to generate images).

3.3. Generator Network

As shown in Figure 3, the generator network structure of SinGAN consists of three parts: the head module, the body module, and the tail module. The entire network uses a 3×3 convolution kernel for shallow feature extraction and image generation (Mairal, Koniusz, Harchaoui, & Schmid, 2014). The input is the sum of the random noise generated at the current scale and the image upscaled from the previous layer. The output is the residual image after superimposition with the input image.

- Head Module: Composed of a set of convolutional layers, batch normalization, and LeakyReLU activation functions, it primarily performs preliminary linear and nonlinear mappings on the input features.
- Body module: Includes three sets of convolutional structures, each independently composed of a 3×3 convolutional layer, BatchNorm layer, and LeakyReLU activation function, stacked sequentially to progressively extract high-level semantic and local texture features, enhancing the model's ability to capture image details.
- Tail Module: Consists of a single 3×3 convolutional layer and a Tanh activation function, mapping the network output to the $[-1, 1]$ interval to restore image pixels and generate the final image output.

The final generated image is added to the image of the previous scale through a skip connection to maintain structural consistency. This network structure not only has low computational complexity and parameter count but also ensures image generation quality and style consistency, demonstrating good detail synthesis capabilities and training stability in multi-scale structures.

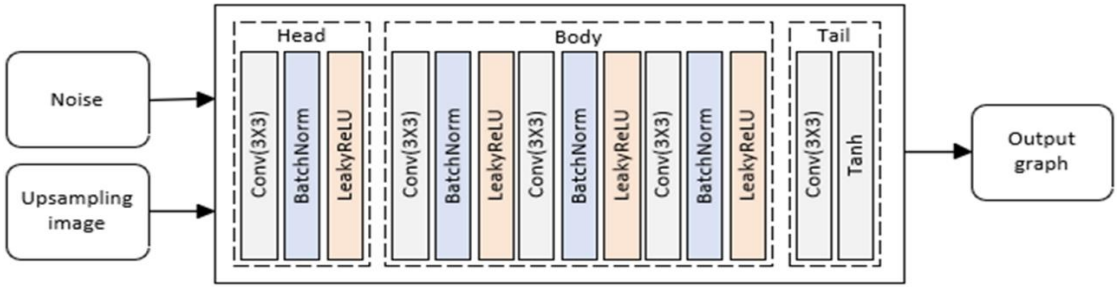

Figure 3. Generator network structure.

3.4. Discriminator Network

The discriminator network structure in the SinGAN model is a shallow CNN network, as shown in Figure 4. The discriminator uses Leaky-ReLU (α=0.2) as the activation function to suppress the gradient vanishing problem and enhance the expressive ability of negative features (Hanin, 2018; Rasamoelina, Adjailia, & Sinčák, 2020). The entire network consists of five convolutional layers, with the first layer being a basic convolutional block featuring a 3×3 convolutional kernel followed by BatchNorm and LeakyReLU directly applied to the input image (Dubey & Jain, 2019). Each convolutional layer (except the output layer) is followed by BatchNorm and LeakyReLU activation to enhance network stability and training convergence speed. As the network depth increases, the spatial dimensions of the image gradually decrease, and the number of feature map channels also decreases layer by layer, enabling the discriminator to have stronger texture recognition capabilities. The final output layer is a single-channel convolutional structure, producing a locally perceptible real-valued feature map to measure the distribution consistency between the generated image and the real image at the current scale (Yang, Yan, & Lei, 2015). The overall network structure is lightweight and efficient, providing the generator with fine-grained realism guidance.

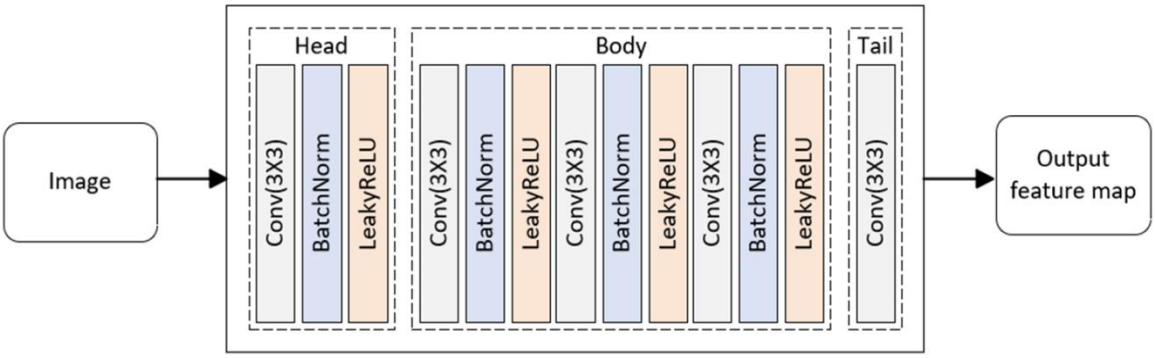

Figure 4. Discriminator network structure.

3.5. Loss Functions and SR Module Implementation Mechanism

During the training process of SinGAN, to achieve consistency in texture and structure between generated images and real images at various scales, the model introduces two key loss functions: adversarial loss and reconstruction loss. These are combined with Wasserstein distance and gradient penalty strategies to enhance training stability (Panaretos & Zemel, 2019)

Discriminator Loss: At each scale, the discriminator calculates the difference between the distributions of real and generated images using the Wasserstein distance and introduces a gradient penalty term to ensure the 1-Lipschitz condition is satisfied. The loss function expression is as follows:

$$\mathcal{L}_{D_n} = \mathbb{E}[D_n(G_n(z_n, y_n))] - \mathbb{E}[D_n(x_n)] + \lambda_{gp} \cdot \mathbb{E}[(\| \nabla_{\hat{x}} D_n(\hat{x}) \|_2 - 1)^2] \quad (4)$$

Where x_n represents the real image of the nth layer, and $G_n(z_n, y_n)$ represents the generated image. $\hat{x} = \epsilon x_n + (1-\epsilon)G_n(z_n, y_n)$,Where $\grave{O}$ follows a uniform distribution $\epsilon \sim \mathcal{U}(0,1)$, λ_{gp} represents the gradient penalty weight.

Generator loss: The goal of the generator is to deceive the discriminator and generate images consistent with real images. Therefore, its loss consists of both adversarial and reconstruction terms.

$$\mathcal{L}_{G_n} = -\mathbb{E}[D_n(G_n(z_n, y_n))] + \alpha \cdot \| G_n(z_n, y_n) + y_n - x_n \|_2^2 \quad (5)$$

The first term represents the adversarial loss, and the second term represents the MSE reconstruction loss, which is used to constrain the similarity between the generated image and the real image. α represents the reconstruction loss weight, y_n represents the input after upsampling on the previous scale image, and z_n represents the current scale noise input.

Overall training objective: SinGAN uses a multi-scale, layer-wise training strategy during training, and the final loss is obtained by combining all scale losses:

$$\mathrm{L}_{total} = \sum_{n=0}^{N} (\mathrm{L}_{D_n} + \mathrm{L}_{G_n}) \quad (6)$$

This loss function design can effectively guide the generator to learn the distribution patterns of structure and texture under single-image conditions and achieve image reconstruction from coarse to fine (Wang, Ma, Zhao, & Tian, 2022).

In SinGAN, the super-resolution module gradually optimizes the reconstruction loss by enlarging the entire pixel structure of the image layer by layer. This module uses multiple trained generators $G_0, G_1, \ldots, G_N$ to reconstruct low-resolution images into higher-resolution images layer by layer in a bottom-up manner.

The core process is as follows:

- Start with the lowest resolution image (such as x_N) and perform upsampling processing.
- Input the current scale noise z_n and the upsampled feature map $x_{N(u)}$ from the previous layer, and obtain the current layer output through the generator G_n.
- Optimize the generator and discriminator through the loss function to obtain the final high-resolution image.

4. RESULTS

4.1. Data and Training Parameters

To validate the practicality and effectiveness of the proposed SinGAN-based super-resolution reconstruction method for power transmission line images in power inspection scenarios, this paper selected typical image scenarios commonly encountered in actual power operations as input samples. The SinGAN model was used to perform fourfold super-resolution reconstruction on these images to improve the image quality foundation for subsequent object detection and fault identification.

This experiment was implemented using the PyTorch framework on a high-performance computer equipped with an NVIDIA GeForce RTX 4060 Ti 8GB graphics card, running Windows 11 with Python version 3.10. During

training, the learning rate was set to 0.0005, with Adam optimizer momentum parameters β_1 and β_2 set to 0.5 and 0.999, respectively. Each scale underwent 2,000 iterations, and upon meeting convergence conditions, the model automatically progressed to the next scale for training. The experimental steps are as follows:

- Image preparation: Select a typical blurred power line image as the sole input sample for SinGAN and construct a multi-scale image pyramid.
- Model training: Following the hierarchical mechanism of SinGAN, train the generator and discriminator layer by layer, starting from the lowest resolution image to extract and learn structural and texture distribution features at different scales.
- SR Reconstruction Inference: Utilize the trained multi-scale generator to perform the reconstruction process layer by layer, starting from the lowest resolution image, ultimately obtaining a clear image with double the resolution.
- Metric Evaluation: Compare the input image with the generated image, and use PSNR and SSIM to assess the improvement in image quality. Additionally, introduce the YOLOv7 model to calculate the mAP value, verifying the practical auxiliary value of high-resolution images for object detection tasks.
- Discuss the results, analyze the model's strengths and weaknesses, and clarify the objectives for future research.

4.2. Evaluation Indicators

To verify the advanced nature of the technology used in this paper, PSNR, SSIM, and the mAP value of the YOLOv7 model are used as evaluation indicators. The formula for PSNR is:

$$PSNR = 10 \cdot \log_{10}\left(\frac{MAX_I^2}{MSE}\right) \tag{7}$$

MAX_I represents the maximum possible value of image pixels, $MSE = \frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}[I(i,j)-K(i,j)]^2$, where I and K are the input image and the generated image, respectively, and m and n are the height and width of the image, respectively. PSNR measures the pixel-level difference between the generated image and the original image, with units in decibels (dB) (Hore & Ziou, 2010). Higher values indicate more significant improvements in image quality and lower noise. SSIM is a metric for measuring the similarity between two images, with the formula as follows:

$$SSIM(x,y) = \frac{(2\mu_x\mu_y+c_1)(2\sigma_{xy}+c_2)}{(\mu_x^2+\mu_y^2+c_1)(\sigma_x^2+\sigma_y^2+c_2)} \tag{8}$$

μ_x, μ_y represent the average pixel values of images x and y, respectively; σ_x^2, σ_y^2 represent the variances of images x and y, respectively; σ_{xy} represents the covariance of images x and y; c_1, c_2 are constants used to avoid a zero denominator. SSIM evaluates the structural similarity of images, considering brightness, contrast, and structural information (Nilsson & Akenine-Möller, 2020). The value ranges from [0, 1], with values closer to 1 indicating that the generated image is more similar to the original image in terms of perceived quality and better structural preservation.

The mAP of the YOLOv7 model is a commonly used performance evaluation metric in object detection tasks and is widely applied in the evaluation of modern object detection models such as YOLOv7 (Olorunshola, Irhebhude, & Evwiekpaefe, 2023). YOLOv7 significantly improves detection accuracy and efficiency through multi-level prediction and self-supervised learning techniques. In this study, the mAP value is used to quantify the auxiliary effect of super-resolution reconstructed images on the YOLOv7 object detection task. The formula is as follows:

$$Precision = \frac{TP}{TP+FP} \tag{9}$$

The above equation is the accuracy formula, where TP is the number of correctly detected targets and FP is the number of false detections.

$$Recall = \frac{TP}{TP+FN} \quad (10)$$

The above equation is the recall rate formula, where FN is the number of undetected targets.

$$AP = \int_0^1 p(r)\,dr \quad (11)$$

The above equation is the average precision formula, where $p(r)$ is the highest precision under a given recall rate r.

$$mAP = \frac{1}{N}\sum_{i=1}^{N} AP_i \quad (12)$$

The above equation is the mAP formula, which is the average value for all categories of AP, where N is the number of categories and AP_i is the average accuracy of the i category.

4.3. Image Comparison and Various Metric Analysis

This study first utilized the Power Transmission Line Dataset (IEEE DataPort) (Alcayde-García et al., 2022) for training, which contains both real and synthetic power transmission line images. Experimental results demonstrate that this dataset performs well in deep neural network training and transmission line super-resolution reconstruction tasks, with significant improvements in the quality metrics of the generated images. To further validate the model's applicability and robustness in real-world scenarios, this study additionally used a transmission line image captured by the researchers as a test sample. The intuitive comparison results between the input blurred transmission line image and the clear transmission line image generated through super-resolution reconstruction are shown in Figure 5.

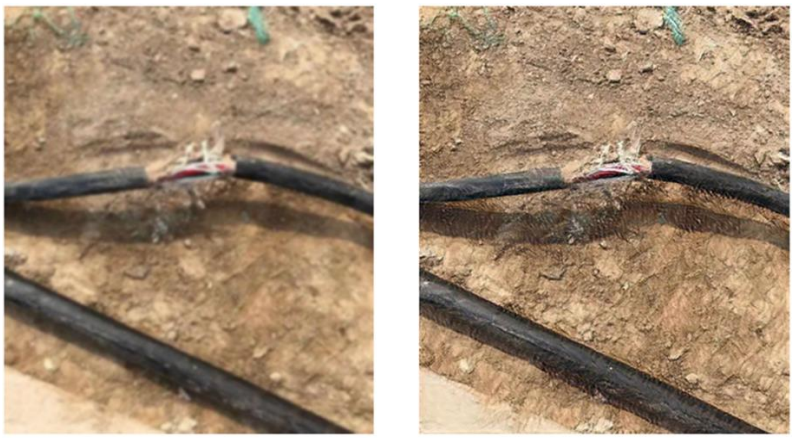

Figure 5. Comparison of input and output graphs.

Figure 5 illustrates the visual differences between the input blurred image and the reconstructed high-definition image. The input image (left) exhibits noticeable blurring, particularly in the damaged areas of the power lines, where the cable texture and background details (such as soil patterns) are difficult to discern, resulting in the loss of target features. In contrast, the reconstructed image (right) significantly enhances the clarity of the damaged cable areas, with the boundaries between the red insulating layer inside the cable and the black outer sheath becoming more distinct. The subtle textures in the soil background have also been effectively restored. The three metrics for these two images are compared in Table 1:

Table 1. Comparison of indicators for two images.

Image type	*PSNR (dB)*	*SSIM*	*YOLOv7 mAP*
Enter blurry image	14.85	0.092	0.051
Output high-definition images	23.72	0.412	0.186

From the PSNR metric, the output image improved by 8.87 dB compared to the input image, indicating that SinGAN has good error compression capabilities in pixel-level reconstruction, effectively improving image clarity. SSIM improved from 0.092 to 0.412, indicating that the model not only reconstructed high-frequency detail information but also reproduced edge and texture features at the structural level, further verifying SinGAN's superior performance in structural restoration.

More importantly, the YOLOv7 detection model achieved an mAP of 0.186 in super-resolution results, which is more than three times higher than the 0.051 achieved with blurred input, reflecting the practical application value of SinGAN-reconstructed images in downstream tasks. This result indicates that the proposed SinGAN reconstruction method not only improves perceptual quality but also demonstrates good adaptability and robustness in practical inspection and identification scenarios.

Table 2 presents the comparison of training times among SinGAN, SRGAN, and DDPM, indicating that SinGAN achieves much higher training efficiency by completing single-image training within 2 hours, whereas SRGAN and DDPM require 6.3 and 24.6 hours, respectively.

Table 2. Comparison of model running time.

Types of technology	Time required for operation(h)
SinGAN	2.0
SRGAN	6.3
DDPM	24.6

5. DISCUSSION

Although this study achieved super-resolution reconstruction of blurry images of power transmission lines based on SinGAN and demonstrated significant advantages in terms of operational efficiency, adaptability, and engineering deployment, there are still some shortcomings in the quality of the generated images that need to be further optimized and expanded.

First, SinGAN has the advantages of fast training speed, unsupervised learning, and the ability to train using a single image, making it particularly suitable for power inspection scenarios with scarce data and strong task timeliness. However, in terms of image quality, the output images still exhibit defects in edge details, texture consistency, and overall clarity. For example, as seen in the image results, while the cable boundaries have been partially reconstructed, certain areas exhibit texture stretching, local distortion, and overlapping artifacts. Additionally, non-natural interference textures are present in background details (such as soil texture), which impair the image's realism and the accuracy of target recognition.

Based on the above analysis, future research can be conducted in the following directions:

- Integrating structural priors with deep supervision mechanisms: Consider incorporating edge detection, object segmentation, or structure-guided modules into the SinGAN training framework to enhance the clarity of power facility edges and structural fidelity.
- Improving the generator architecture: Introduce structural enhancement modules such as residual blocks and attention mechanisms to enhance the network's ability to model complex textures and spatial distributions.
- Introducing perceptual loss and multi-scale fusion: Adding perceptual loss based on feature maps and multi-scale constraint mechanisms to optimize the global consistency and detail restoration of the output image.

- Building a high-quality training dataset with a small sample size: Combining existing power industry inspection images to explore SinGAN expansion strategies based on a small number of images, thereby improving the model's generalization ability and universality.
- Integrating YOLOv7 object detection feedback mechanisms: Establishing a feedback mechanism between image quality and downstream tasks, utilizing object detection results to provide "reverse guidance" for the super-resolution process, and achieving end-to-end optimization.

In summary, future research will continue to focus on optimizing SinGAN's generation quality, structural fidelity, and compatibility with target tasks, driving its practical application in smart grid inspection and image enhancement fields.

6. CONCLUSIONS

This paper proposes a single-image super-resolution reconstruction method for power transmission lines based on SinGAN. Addressing the limitations of traditional super-resolution algorithms, which rely on paired data and are not suitable for scenarios with scarce samples, we designed an unsupervised image enhancement framework that uses single images as input. Without the need for paired training images, the model constructs a multi-scale pyramid structure, progressively learning the texture and structural information within the image, ultimately achieving high-quality reconstruction of blurry images. Experiments use actual transmission line images as test subjects, combining PSNR, SSIM, and YOLOv7 mAP metrics to evaluate model performance across three dimensions: pixel quality, structural integrity, and task performance. The results show that the PSNR of the reconstructed images improved by 8.87 dB, SSIM improved by 0.32, and mAP improved by more than three times, significantly enhancing image quality and detectability. At the same time, SinGAN training takes approximately two hours, demonstrating good efficiency and deployment advantages. In summary, the SinGAN image reconstruction model constructed in this paper not only achieves effective application of single-image super-resolution tasks in power inspection scenarios but also provides new ideas for improving image quality under small sample and weak supervision conditions. The research results have good practical value and promotion potential in improving the image processing capabilities of smart grids and enhancing the accuracy of target detection, laying a foundation for the development of intelligent power line inspection technology in the future.

Funding: This research was supported by the Tibet Agriculture and Animal Husbandry College Talent Team Construction Project (Grant number: NMXYRCDWJS-2025-09).
Institutional Review Board Statement: Not applicable.
Transparency: The authors state that the manuscript is honest, truthful, and transparent, that no key aspects of the investigation have been omitted, and that any differences from the study as planned have been clarified. This study followed all writing ethics.
Competing Interests: The authors declare that they have no competing interests.
Authors' Contributions: All authors contributed equally to the conception and design of the study. All authors have read and agreed to the published version of the manuscript.

REFERENCES

Alcayde-García, F., Salmerón-Manzano, E., Montero, M. A., Alcayde, A., & Manzano-Agugliaro, F. (2022). Power transmission lines: Worldwide research trends. *Energies, 15*(16), 5777. https://doi.org/10.3390/en15165777

Dong, C., Loy, C. C., & He, K. (2014). Learning a deep convolutional network for image super-resolution. In D. Fleet, T. Pajdla, B. Schiele, & T. Tuytelaars (Eds.), Computer Vision – ECCV 2014. Lecture Notes in Computer Science. In (Vol. 8692, pp. 184–199). Cham, Switzerland: Springer.

Dubey, A. K., & Jain, V. (2019). *Comparative study of convolution neural network's relu and leaky-relu activation functions applications of Computing*. Paper presented at the Automation and Wireless Systems in Electrical Engineering: Proceedings of MARC 2018, Springer Singapore.

Faisal, M. A. A., Mecheter, I., Qiblawey, Y., Fernandez, J. H., Chowdhury, M. E. H., & Kiranyaz, S. (2025). Deep learning in automated power line inspection: A review. *Applied Energy, 385*, 125507. https://doi.org/10.1016/j.apenergy.2025.125507

Fang, X., Misra, S., Xue, G., & Yang, D. (2012). Smart grid—The new and improved power grid: A survey. *IEEE Communications Surveys & Tutorials, 14*(4), 944-980. https://doi.org/10.1109/SURV.2011.101911.00087

Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., . . . Bengio, Y. (2014). *Generative adversarial nets. In Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence, & K. Q. Weinberger (Eds.), Advances in Neural Information Processing Systems* (Vol. 27). Red Hook, NY: Curran Associates.

Hanin, B. (2018). *Which neural net architectures give rise to exploding and vanishing gradients? Advances in Neural Information Processing Systems* (Vol. 31). USA: Curran Associates.

Hore, A., & Ziou, D. (2010). *Image quality metrics: PSNR vs. SSIM*. Paper presented at the 2010 20th International Conference on Pattern Recognition, IEEE.

Kim, J., Lee, J. K., & Lee, K. M. (2016). *Accurate image super-resolution using very deep convolutional networks*. Paper presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.

Mairal, J., Koniusz, P., Harchaoui, Z., & Schmid, C. (2014). *Convolutional kernel networks. In Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence, & K. Q. Weinberger (Eds.), Advances in Neural Information Processing Systems*. USA: Curran Associates.

Nilsson, J., & Akenine-Möller, T. (2020). Understanding ssim. *arXiv preprint arXiv:2006.13846*. https://doi.org/10.48550/arXiv.2006.13846

Olorunshola, O. E., Irhebhude, M. E., & Evwiekpaefe, A. E. (2023). A comparative study of YOLOv5 and YOLOv7 object detection algorithms. *Journal of Computing and Social Informatics, 2*(1), 1-12. https://doi.org/10.33736/jcsi.5070.2023

Panaretos, V. M., & Zemel, Y. (2019). Statistical aspects of Wasserstein distances. *Annual Review of Statistics and its Application, 6*(1), 405-431. https://doi.org/10.1146/annurev-statistics-030718-104938

Raghavan, T. E. S. (2012). Zero-sum two person games. In Computational Complexity. In (pp. 3372-3395). New York: Springer

Rasamoelina, A. D., Adjailia, F., & Sinčák, P. (2020). *A review of activation function for artificial neural network[C]//2020*. Paper presented at the IEEE 18th World Symposium on Applied Machine Intelligence and Informatics (SAMI), IEEE.

Shaham, T. R., Tali, D., & Tomer, M. (2019). *Singan: Learning a generative model from a single natural image*. Paper presented at the Proceedings of the IEEE/CVF International Conference on Computer Vision.

Shocher, A., Cohen, N., & Irani, M. (2018). *"Zero-shot" super-resolution using deep internal learning*. Paper presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.

Wang, Q., Ma, Y., Zhao, K., & Tian, Y. (2022). A comprehensive survey of loss functions in machine learning. *Annals of Data Science, 9*(2), 187-212. https://doi.org/10.1007/s40745-020-00253-5

Wang, X., Yu, K., Wu, S., Gu, J., Liu, Y., Dong, C., & Change Loy, C. (2018). *Esrgan: Enhanced super-resolution generative adversarial networks*. Paper presented at the Proceedings of the European Conference on Computer vision (ECCV) Workshops.

Yang, B., Yan, J., & Lei, Z. (2015). *Convolutional channel features*. Paper presented at the Proceedings of the IEEE International Conference on Computer Vision.

Yang, M., Wang, J., Diao, H., Qi, J., & Han, X. (2018). Interval estimation for conditional failure rates of transmission lines with limited samples. *IEEE Transactions on Smart Grid, 9*(4), 2752-2763. https://doi.org/10.1109/TSG.2016.2618623

Zhang, G. (2025). Transmission line trip faults under extreme snow and ice conditions: A case study. *Energy Informatics, 8*(1), 7. https://doi.org/10.1186/s42162-024-00463-8

Zhang, K., Liang, J., Van Gool, L., & Timofte, R. (2021). *Designing a practical degradation model for deep blind image super-resolution*. Paper presented at the Proceedings of the IEEE/CVF International Conference on Computer Vision.