Image Deblurring Method for Safety Wire Based on GAN

Lingchao Bu, Yao Dai

***Abstract*—In order to prevent the loosening of train bolts, anti-loosening iron wires are commonly used to secure the bolts. However, the anti-loosening iron wires may experience breakage, leading to loosening or even detachment of the secured bolts, posing a safety threat to trains. In this paper, a GAN-based image deblurring network is proposed, with the generator utilizing U-Net-GAM and the discriminator based on PatchGAN. Specifically designed to address the blurred images of anti-loosening iron wires captured by the TEDS system on trains, the method achieves a PSNR of 34.63 dB and SSIM of 95.01% on the anti-loosening iron wire dataset, outperforming existing mainstream methods with good overall performance.**

***Index Terms*—Image Deblurring, Safety Wire, GAN**

I. INTRODUCTION

In recent years, machine learning technologies including deep learning [1][2] have been rapidly developing. More and more fields are actively seeking transformation by combining deep learning technologies to achieve industrial technological innovation and further improve work efficiency. Particularly in the field of transportation, which involves the safety of people's lives and property, related technological updates iterate slowly. As of the end of 2023, China's railway operating mileage is 159,000 kilometers, with high-speed railways accounting for 45,000 kilometers, ranking first in the world. Therefore, railway safety has always been one of the research focuses in the transportation field [3][4].

Anomaly detection technology [5][6] plays a crucial role in ensuring the safe operation of trains. As a key component affecting the safety of train operation, bolts have always been a focus of inspection in railway systems, as many important parts of trains are fixed by bolts. To prevent bolts from loosening, anti-loosening wires are usually used to secure the bolts. However, with the long-term operation of trains, anti-loosening wires may break due to reasons such as vibration, oxidation, and foreign object impacts, leading to loosening or even detachment of the fixed bolts. In severe cases, abnormal situations occur in key parts where loosened or detached bolts are located, posing a threat to the safety of trains and even causing serious accidents. The Trouble of Moving EMU Detection System (TEDS) is a networked application system for real-time monitoring of EMU operation states. In the TEDS system, images of train bodies collected by cameras are transmitted to the control room, where inspectors can directly conduct inspections on computers [8], reducing external environmental interference and eliminating the need for on-site inspections by inspectors, thereby saving labor. However, this method still requires inspectors to conduct inspections and does not fundamentally improve inspection efficiency.

The TEDS system of trains, as a key link in modern railway transportation safety, captures real-time image data of moving trains through advanced industrial high-speed camera technology. The main goal of this system is to continuously monitor the operation status of trains to timely discover potential problems and take corresponding measures to ensure the safety and stability of train operation. However, due to the extremely fast speed of train operation, industrial cameras face significant challenges in capturing images of dynamic trains. Frequent image blurring caused by high-speed motion makes it difficult to accurately identify some image information, undoubtedly increasing the difficulty of analyzing train status for the system.

In addition to the high-speed movement of trains, other factors resulting from long-term operation of the system also affect the quality of images. Cameras may accumulate dust due to long-term operation, leading to a decrease in image clarity during shooting. Furthermore, changes in light source intensity and camera shake during shooting further exacerbate the degree of image blurring. These factors together result in uneven image quality of captured train images. To solve this problem, it is particularly important to design effective image deblurring methods. Through deblurring algorithms, we can extract more useful information from blurred images, making the images clearer and easier for subsequent analysis and processing. In this chapter, a dedicated image deblurring network is proposed to process blurred images of anti-loosening wires captured by the TEDS system of trains.

II. RELATED WORK

In the railway industry, researchers have proposed various target detection algorithms based on convolutional neural networks (CNNs) to handle fault images, aiming to improve the efficiency and accuracy of railway equipment fault diagnosis. Wang et al. [17] proposed a deep learning-based YOLOV2 algorithm for automatic detection and positioning of railway tracks, which can detect components such as rails and bolts. These CNN-based target detection algorithms provide important technical support for fault diagnosis and equipment maintenance in the railway industry and are expected to play a significant role in practical applications, improving the safety and reliability of railway equipment.

The main objective of image deblurring is to restore the clarity and details of blurred images of anti-loosening wires.

Manuscript received March 05, 2024

Lingchao Bu, School of Computer Science and Technology, Tiangong University, Tianjin, China

Yao Dai, School of Computer Science and Technology, Tiangong University, Tianjin, China

Blurring occurs due to factors such as camera motion, lens instability, environmental vibrations, or defects in optical systems, so appropriate methods are needed to alleviate or eliminate the blurring effects.

(1) Classical blind deblurring algorithms: These algorithms are based on the assumption that the shape, size, and orientation of the image and the blur kernel are unknown, hence referred to as "blind" deblurring. Common methods include algorithms based on statistical models, least squares methods, and variational Bayesian inference algorithms.

(2) Deep learning-based image deblurring methods: With the development of deep learning technology, an increasing number of studies are adopting deep neural networks to address image deblurring problems [24][25]. These methods typically use convolutional neural networks (CNNs) to learn image features and restore the clarity of blurred images through end-to-end training. Common network architectures include CNNs, generative adversarial networks (GANs), and recurrent neural networks (RNNs).

(3) Frequency domain-based image deblurring methods: Frequency domain processing methods typically involve processing images in the frequency domain and then converting the processed images back to the spatial domain. Common frequency domain processing methods include Fourier transforms, wavelet transforms, etc.

(4) Prior knowledge-based image deblurring methods: These methods typically use prior knowledge of images to assist in image deblurring, such as using gradient information, texture information, edge information, etc. This prior knowledge can be obtained through image statistical characteristics, structural characteristics, etc.

Image deblurring methods mainly include image enhancement, image restoration, and super-resolution reconstruction, among others. In the field of machine learning, pre-processing images with deblurring methods before dataset processing can effectively improve the clarity of images, thereby enhancing the efficiency and accuracy of model training.

Image enhancement mainly involves enhancing useful information in images through certain algorithms and techniques to improve the visual effects of images. This method does not directly address the causes of blurring but highlights certain features or suppresses uninteresting features in images to make them visually clearer. For example, low-pass filtering and high-pass filtering are two commonly used image enhancement techniques. Low-pass filtering removes noise from images by allowing only low-frequency signals to pass through, while high-pass filtering enhances high-frequency signals such as edges to make the blurred edges of images clearer.

Image restoration [26] is a more precise method that analyzes the causes of image degradation and attempts to restore the image to its original state. This typically requires modeling the image and using some prior knowledge or assumptions to restore the clarity of the image. Image restoration methods are usually more complex than image enhancement methods but can provide more accurate and effective deblurring effects.

Super-resolution reconstruction uses multiple low-resolution images to generate a high-resolution image. This method is typically based on deep learning and machine learning algorithms, where models are trained to learn the mapping from low resolution to high resolution. Super-resolution reconstruction not only improves the clarity of images but also to some extent restores the details in the images.

In practical applications, image deblurring technology has been widely used in various fields such as medical image analysis, security monitoring, satellite remote sensing, etc. Although image deblurring technology can improve the clarity of images, it may not completely restore all the details of all blurred images. For some severely blurred or distorted images, there may still be limitations. Therefore, when using image deblurring technology, it is necessary to assess and choose according to specific situations.

In the TEDS system, most of the images captured by industrial cameras are blurry, so it is necessary to deblur the images. As a type of image restoration technique, image deblurring has significant research significance. The mathematical model formula for image deblurring is shown as equation (1).

$$I_{\text{blur}} = k \cdot I_{sharp} + N \#(1)$$

Where I_blur represents the blurry image, I_sharp represents the sharp image, N is noise, · denotes image convolution operation, and k represents the blur kernel matrix. The formula indicates that image deblurring is an inverse convolution problem. Depending on whether the blur kernel k is known, image deblurring can be classified into two categories: non-blind deblurring and blind deblurring algorithms.

Non-blind deblurring algorithms assume that prior knowledge of the blur kernel (blur function) can be obtained during the deblurring process, i.e., processing is performed under the condition of knowing the blur kernel before image processing. Typically, non-blind deblurring algorithms first estimate the blur kernel of the image or directly use a known blur kernel, and then use inverse filtering, Wiener filtering, or other complex image restoration algorithms to restore the sharp image. Since the blur kernel is known, accurate image restoration can be performed without introducing additional noise, resulting in more accurate restoration results.

Blind deblurring algorithms assume that the blur kernel information cannot be directly obtained when processing the image and need to be estimated based on the image's own features. Typically, an attempt is made to estimate the parameters of the blur kernel while processing the image, followed by image restoration. This estimation is usually achieved by minimizing the error between the convolution result of the image and the blur kernel and the observed blurry image. Blind deblurring algorithms do not require prior knowledge of the blur kernel, making them more flexible in practical applications.

III. IMPLEMENTATION

A. GAN

GAN (Generative Adversarial Network) is a type of deep learning model, as shown in Figure 1, which generates realistic data samples through adversarial training between the generator and the discriminator. This model has wide applications in various fields such as image generation, speech synthesis, text generation, etc. The basic principles and workflow of GAN are as follows:

1. GAN samples a vector from the latent space randomly (usually a noise vector following a uniform or normal distribution). This vector is fed into the generator network. The generator network gradually generates a data sample through a series of transformations and mapping operations. This sample can be an image, audio, text, or other types of data, depending on the application field of GAN.

2. The generated samples and real training samples are fed into the discriminator network together. The goal of the discriminator network is to distinguish real samples from generated samples, i.e., to determine which samples are real and which are generated. To achieve this goal, the discriminator learns the features of real data and tries to identify the differences in generated data.

3. The generator and discriminator undergo adversarial training. The goal of the generator is to generate samples realistic enough so that the discriminator cannot accurately distinguish between real and generated samples. The goal of the discriminator is to classify samples as accurately as possible, making the differences between generated samples and real samples more pronounced. This adversarial training process drives the generator and discriminator to continuously optimize their performance.

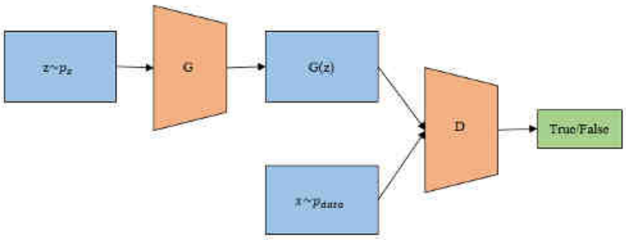

Figure 1 overall architecture of GAN

The loss function of the generator is usually inversely related to the cross-entropy loss function of the discriminator's output on fake data, to encourage the generator to generate more realistic samples. Meanwhile, the loss function of the discriminator is related to the classification accuracy of real and fake data to optimize its classification ability. Through iterative training, both the generator and discriminator gradually improve their performance. Eventually, the generator can generate new data that is difficult to distinguish from real data, while the discriminator struggles to accurately determine the source of the data.

In this study, a Generative Adversarial Network (GAN) is employed to train the image deblurring model. The network architecture is illustrated in Figure 2. During the experimental process, clear train images need to be collected in advance to train the discriminator.

The generator and discriminator mutually improve each other during the training process. The generator is typically a deep neural network, taking low-dimensional vectors (such as random noise) as input and producing high-dimensional vectors (such as images, text, or speech) as output. Similarly, the discriminator is also a deep neural network, taking high-dimensional vectors (such as real or generated images, text, or speech) as input and producing a scalar output representing the probability of the input data being real. The training process can be divided into several stages, including parameter initialization, training the discriminator, training the generator, and repeating the training steps. In each training iteration, the discriminator is trained first to distinguish between real and generated data, and then the generator is trained to generate more realistic data to deceive the discriminator. The training process stops when the gap between the fake data generated by the generator and the real data reaches a certain threshold.

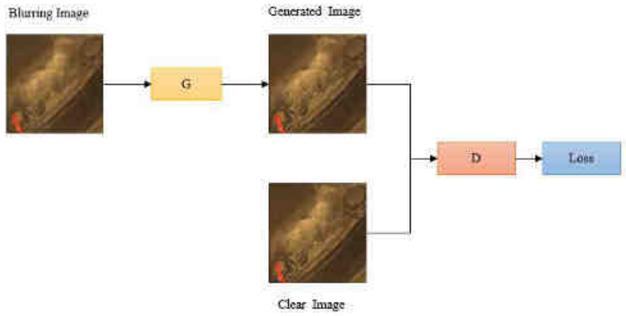

Figure 2 Architecture of Image Deblurring Network

B. Generator

The study adopts an improved U-Net network as the generator network, whose structure resembles the letter "U" as shown in Figure 3. It consists of an encoder and a decoder connected through a bottom-level convolutional layer, with skip connections between corresponding layers of the encoder and decoder. This special design allows the network to reduce spatial dimensions while capturing and reusing more contextual information, thereby better reconstructing the target structure.

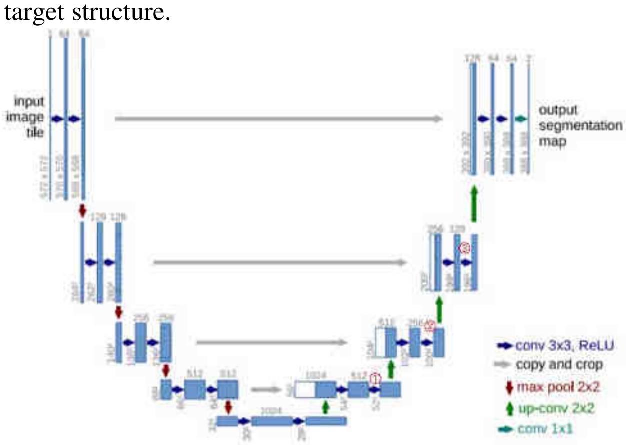

Figure 3 U-Net

The encoder consists of multiple convolutional layers, activation layers, and pooling layers, which are responsible for capturing the contextual information of the input image. The bottom convolutional layer serves as the core part of the network, deepening the feature extraction process and providing rich feature information for the subsequent upsampling process.

The decoder increases spatial dimensions gradually through upsampling and convolutional layers, while utilizing the feature information from the encoder for better feature reconstruction. U-Net utilizes skip connections to directly concatenate feature maps from the encoder to the corresponding layers of the decoder, preserving more high-frequency detail information during the reconstruction process.

U-Net has demonstrated strong performance in image segmentation tasks due to its U-shaped structure and skip connections design, especially in scenarios with limited data. Its advantages can be summarized as follows:

1. High-resolution feature preservation: It can better preserve high-resolution image features, which is crucial for image reconstruction and generation tasks.

2. Fast and efficient: It exhibits significant speed and performance advantages in real-time applications and resource-constrained environments.

3. Data efficiency: It can achieve efficient training even with small datasets, which is advantageous for training GAN networks on limited data.

Utilizing U-Net's U-shaped network structure as the main structure of the generator in GAN networks can bring significant advantages, including generating higher quality samples and faster training times. Its structural design also makes it a flexible and versatile tool suitable for various image generation and transformation tasks.

In this study, to enhance the feature extraction capability of U-Net, a Global Attention Mechanism (GAM) module is added at locations ①, ②, and ③ as shown in figure 3, GAM as shown in figure 4.

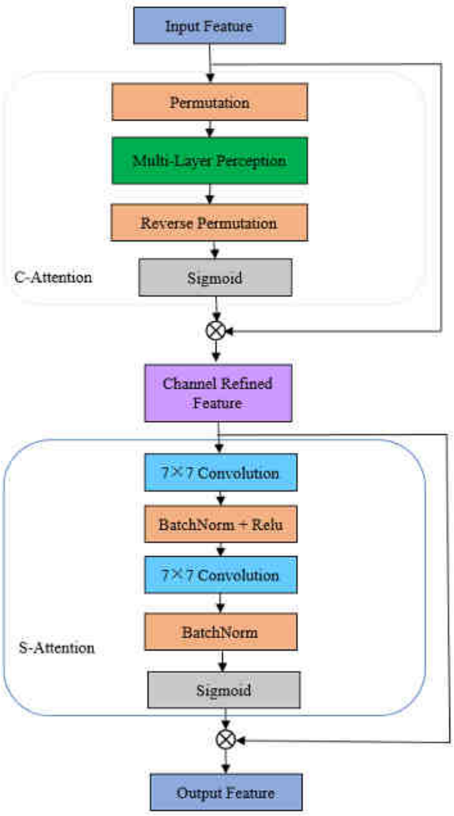

Figure 4 GAM

C. *Discriminator*

The discriminator primarily evaluates the difference between the images generated by the generator and the clear images, which then enters the loss function for calculation, thereby narrowing the gap and generating clearer images. In this study, the PatchGAN is used as the model for the discriminator, and the network structure is as shown in Figure 5. The detailed structure is as follows:

1. Input and preprocessing: The input image undergoes a random cropping step to increase the robustness and diversity of the network.

2. First convolutional layer:

- Convolution: The cropped image is input to a convolutional layer with a 4x4 kernel size.
- Activation: After convolution, the GELU activation function is used to introduce non-linearity.

3. Intermediate convolutional blocks:

- Convolutional layers: After the GELU activation layer, three sets of identical convolutional blocks are set up, each beginning with a convolutional layer with a 4x4 kernel size to further extract features.
- Layer normalization: Following the convolutional layers is the layer normalization layer, which helps maintain the stability and convergence speed of network training.
- GELU activation layer: Each convolutional block ends with a GELU activation layer to introduce non-linearity and further extract important features.

4. Final convolutional layer: After three sets of convolutional blocks, the feature maps enter the final 4x4 convolutional layer for deeper feature extraction.

5. Output:

- Wasserstein distance: The Wasserstein distance is utilized to measure the image restoration effect, providing a more stable and effective optimization objective.

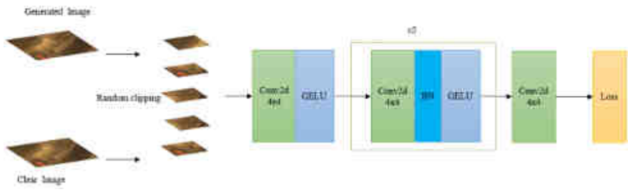

Figure 5 Discriminator

IV. EXPERIMENT AND RESULT

From Table 1, it can be seen that the method proposed in this paper achieves a PSNR of 34.63 dB and SSIM of 95.01% on the anti-loosening wire dataset, which are both superior to existing mainstream methods such as DeblurGAN, DeblurGAN-v2, DBGAN, and DMPHN. Overall, the research method based on U-Net adopted in this paper demonstrates good generalization ability and stability in image deblurring of anti-loosening wire images, with good comprehensive performance..

Table 1 The Experiment of Image Deblurring

Methods	PSNR/dB	SSIM(%)
DeblurGAN	30.26	90.07
DeblurGAN-v2	31.22	91.81
DBGAN	32.96	94.02
DMPHN	32.88	93.80
ours	**34.63**	**95.01**

V. CONCLUSION

In this paper, we propose a GAN-based image deblurring method aimed at enhancing the clarity of train anti-loosening iron wire images. The generator adopts U-Net-GAM architecture, while the discriminator is based on PatchGAN. By removing image blur, we can more accurately identify and analyze anti-loosening iron wires, thereby improving detection accuracy and reliability. The method achieves a PSNR of 34.63 dB and SSIM of 95.01% on the anti-loosening iron wire dataset, outperforming existing mainstream methods. Overall, our GAN-based approach demonstrates good generalization capability and stability in deblurring anti-loosening iron wire images, showing promising overall performance.

REFERENCES

[1] Qu Z, Huang Y, Zheng M. A novel coherence-based quantum steganalysis protocol[J]. Quantum Information Processing, 2020, 19: 1-19.
[2] Sun S, Zhou J, Wen J, et al. A DQN-based cache strategy for mobile edge networks[J]. Computers, Materials & Continua, 2022, 71(2): 3277-3291.
[3] Sansan D, Dawei C, Jiali L. Research, development and prospect of China high-speed train[J]. Chinese Journal of Theoretical and Applied Mechanics, 2021, 53(1): 35-50.
[4] Liu D, Wang T, Liang X, et al. High-speed train overturning safety under varying wind speed conditions[J]. Journal of Wind Engineering and Industrial Aerodynamics, 2020, 198: 104111.
[5] De Vitis G A, Foglia P, Prete C A. Row-level algorithm to improve real-time performance of glass tube defect detection in the production phase[J]. IET Image Processing, 2020, 14(12): 2911-2921.
[6] Rasheed A, Zafar B, Rasheed A, et al. Fabric defect detection using computer vision techniques: a comprehensive review[J]. Mathematical Problems in Engineering, 2020, 2020: 1-24.
[7] GENG Q H, LIU W M, DAI Y. Automated visual inspection of multiple missing fastening bolts on the sides of high-speed train body in railway maintenance[J]. IEEE Access, 2020,8:187060-187071.
[8] Wang T, Yang F, Tsui K L. Real-time detection of railway track component via one-stage deep learning networks[J]. Sensors, 2020, 20(15): 4325.
[9] Sun J, Cao W, Xu Z, et al. Learning a convolutional neural network for non-uniform motion blur removal[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2015: 769-777.
[10] Zhang K, Luo W, Zhong Y, et al. Deblurring by realistic blurring[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2020: 2737-2746.
[11] Jiang Y, Gong X, Liu D, et al. Enlightengan: Deep light enhancement without paired supervision[J]. IEEE transactions on image processing, 2021, 30: 2340-2349.