IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received July 13, 2021, accepted August 28, 2021, date of publication September 1, 2021, date of current version September 9, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3109634

Positron Image Super-Resolution Using Generative Adversarial Networks

FANG XIONG, JIAN LIU, (Member, IEEE), MIN ZHAO, MIN YAO, AND RUIPENG GUO

College of Automation Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China
Nondestructive Detection and Monitoring Technology for High Speed Transportation Facilities, Key Laboratory of Ministry of Industry and Information Technology, Nanjing 211106, China

Corresponding author: Jian Liu (jliu@nuaa.edu.cn)

This work was supported in part by the Natural Science Foundation of China under Grant 62071229, Grant 51875289, and Grant 61873124; in part by the Aeronautical Science Foundation of China under Grant 2020Z060052001 and Grant 20182952029; in part by the Fundamental Research Funds for the Central Universities under Grant NJ2020014 and Grant NS2019017; and in part by the Nondestructive Detection and Monitoring Technology for High Speed Transportation Facilities, Key Laboratory of Ministry of Industry and Information Technology, Graduate Innovation Base (Laboratory) Open Fund under Grant kfjj20200303.

ABSTRACT Positron images generated by positron non-destructive testing technology under rapid detection scenes such as low concentration dose, low exposure time and short imaging time, which have some problems like low-resolution and poor definition. These issues cannot be solved for the being time. To solves these problems, this research super-resolves the low-resolution positron images to generate images with high-resolution and clear details. To make the generated super-resolution images more capable of restoring the features of low- resolution images, this research proposed a positron image super-resolution reconstruction method based on generative adversarial networks. In order to improve the input information utilization rate, long skip connections were added into the generator. In addition, the discriminant model, where composed of an image discriminator and a feature discriminator, can stimulate the generator to generate clearer super-resolution images which contain more details. In attempting to solve the problem of dataset matching, a special positron image super-resolution dataset is constructed for network application scenarios. In the adversarial training stage, perceptual similarity loss and adversarial loss are used to replace the traditional mean squared error loss to improve the images perception quality. Experimental results show that the proposed model can reconstruct low-resolution images by four times super-resolution in 0.16 seconds. The super-resolution images obtained are superior to other algorithms in visual effect, which have clearer detail structure and higher objective performance values. Hence this model can meet the requirements of rapid non-destructive testing of industrial parts.

INDEX TERMS Super-resolution reconstruction, deep learning, generative adversarial networks, positron image.

I. INTRODUCTION

Positron non-destructive testing technology (PNDTT) uses γ photons' strong penetrating and non-invasive properties, which generated by positron annihilation, to bring about the rapid imaging of the non-destructive testing inside cavity. PNDTT can conduct on-site rapid testing of a hydraulic cavity and has many advantages (i.e., low cost, wide applicability, efficient, and high security). Most hydraulic parts (i.e., hydraulic pump, hydraulic motor, hydraulic cylinder, hydraulic valve, and supercharger etc.) are important executive components of modern industrial equipment hydraulic systems. Consequently, these are used widely in machine tools, walking machinery, aerospace, and other fields. Because long-term in the state of pressure, hydraulic parts may cause some inconspicuous defects such as cracking, bulging, depression, corrosion, foreign matter, wear, and other structural damage. The defects mentioned above, which are not easily detected, will make the whole hydraulic system vulnerable to failure. Failure will directly affect system safety and even cause irreversible damage to the entire assemblage of equipment [1], [2]. At present, conventional non-destructive testing technology includes eddy current testing [3] and industrial CT testing [4], which are used in a limited cavity represented by hydraulic parts in industry. However, conventional non-destructive testing technology,

The associate editor coordinating the review of this manuscript and approving it for publication was Wuliang Yin.

which is limited by some difficulties (i.e., testing depth, reflection characteristics of cavity structure, testing technology itself, etc.), has a series of problems such as low efficiency, serious waste of resources and high cost. In consequence, conventional non-destructive testing technology cannot meet the increasing demands of lower-cost and faster testing.

Currently, PNDTT is used primarily to analyse the characteristics of semiconductors, metals, high temperature superconductors, polymers and other industrial materials, as well as to conduct non-destructive testing of internal defects of these industrial materials. The German Timo team [5] has proposed a single tracer particle testing method for tracking the density distribution of a positron at a cone angle, in order to measure the cone angle. Yao [6] proposed a method that uses 18F-labelled fuel to observe a combustion flow field in a confined space. However, when PNDTT is used in industrial detection imaging [7], the limitations of imaging sensors, scanning time, and scattering noise [8], [9] make it difficult to obtain high-quality images. This kind of image quality degradation problem is difficult to eliminate by improving the hardware design. Industrial applications of PNDTT need high-quality images that contain more details, more details can provide more accurate judgment for defect location, fault diagnosis and field condition monitoring.

The details of low-quality positron image are indispensable for accurate judgment of test results, but it is difficult to obtain the details by using conventional reconstruction algorithms. Hence, this research introduces deep learning, a new method which has rapidly developed owing to its efficient feature extraction ability. Image super-resolution (SR) methods are used in various computer vision applications, ranging from security and surveillance imaging [10], medical imaging [11], [12], object recognition [13] to image reconstruction [14]. In order to solve the problem that the positron image is not clear enough, super-resolution image reconstruction (SRIR) technology is used to process the images obtained by PNDTT. Low resolution (LR) and low-quality positron images are transformed into high resolution (HR) and high-quality ones, which being essential for PNDTT.

Single image super resolution (SISR) reconstruction technology aims to use the similarity between a LR image and a HR image to break through the inherent physical properties of the imaging system, eliminate the interference of external factors, and obtain HR images with rich details. Harris [15] proposed SISR technology in the 1960s by restoring a single image. In recent years, the concept of generative adversarial networks (GANs) [16] has been proposed and gradually emerges in supervised learning. GAN has been applied to SISR to overcome the aforementioned limitation and produce super-resolved images with synthesized high-frequency details. Ledig *et al.* [17] applied GAN to SISR for the first time and proposed SRGAN, which uses the perceptual similarity loss and adversarial loss to give the super-resolution image more authenticity. Using SRGAN significantly improves the visual effect of a reconstructed image compared with an original image.

Images generated by PNDTT has fewer features, poor definition and fewer sample images. Although the SISR method based on GAN has made significant progress in image perception quality, existing image super-resolution reconstruction models do not perform well on the positron images.

To solve above problems of positron images, this paper proposes positron image super-resolution using generative adversarial networks (PI-SRGAN), which is very suitable for high-quality reconstruction of positron images. Through adding long skip connections in generator, global feature fusion is achieved by making full use of input information, it can make up for the loss which caused by serious degradation of image detail information of the depth neural network. To provide the generated image with more realistic texture and the quality of the generated image, PI-SRGAN adds image and feature discriminators to optimize the generator. In the selection of loss function, traditional mean square error (MSE) is used in the pre-training stage, while perceptual similarity loss and adversarial loss are used in the adversarial training stage. At the same time, the parameters of PI-SRGAN are improved to make the network more suitable for positron images. PI-SRGAN network can return better to a realistic feature distribution and produce more realistic high-frequency details. The overall network architecture of PI-SRGAN is shown in Fig.1.

For this paper, the main contributions are as follows:

(1) To improve the utilization of image feature information and the quality of reconstructed images, this research proposes a new generator network with long skip connections. This network can super-resolve positron images of different sizes with achieving the generation scarce image data in PNDTT field.

(2) To produce a more realistic HR image, this research proposes two different discriminators that one based on the pixel space and the other based on the feature space. They can generate the same high-frequency features as the original HR image.

(3) To generate high quality image in visual perception and texture details, perceptual similarity loss and adversarial loss are used in the adversarial training stage. Because super-resolution images do not have corresponding high-resolution images in practical applications, this study introduces blind image quality assessment indices (i.e., the standard deviation (SD) and average gradient (AG)) to evaluate the reconstruction results.

(4) Since there is currently no publicly available super-resolution dataset of positron images, this research builds a dataset that includes 9100 pairs of positron images. We use GATE software to simulate multiple sets industrial parts to get high-resolution positron images, while use data augmentation technology to obtain enough high-resolution positron images.

Section II introduces the related work for SISR and section III describes the proposed PI-SRGAN. Section IV analyses experimental results and compares them with the

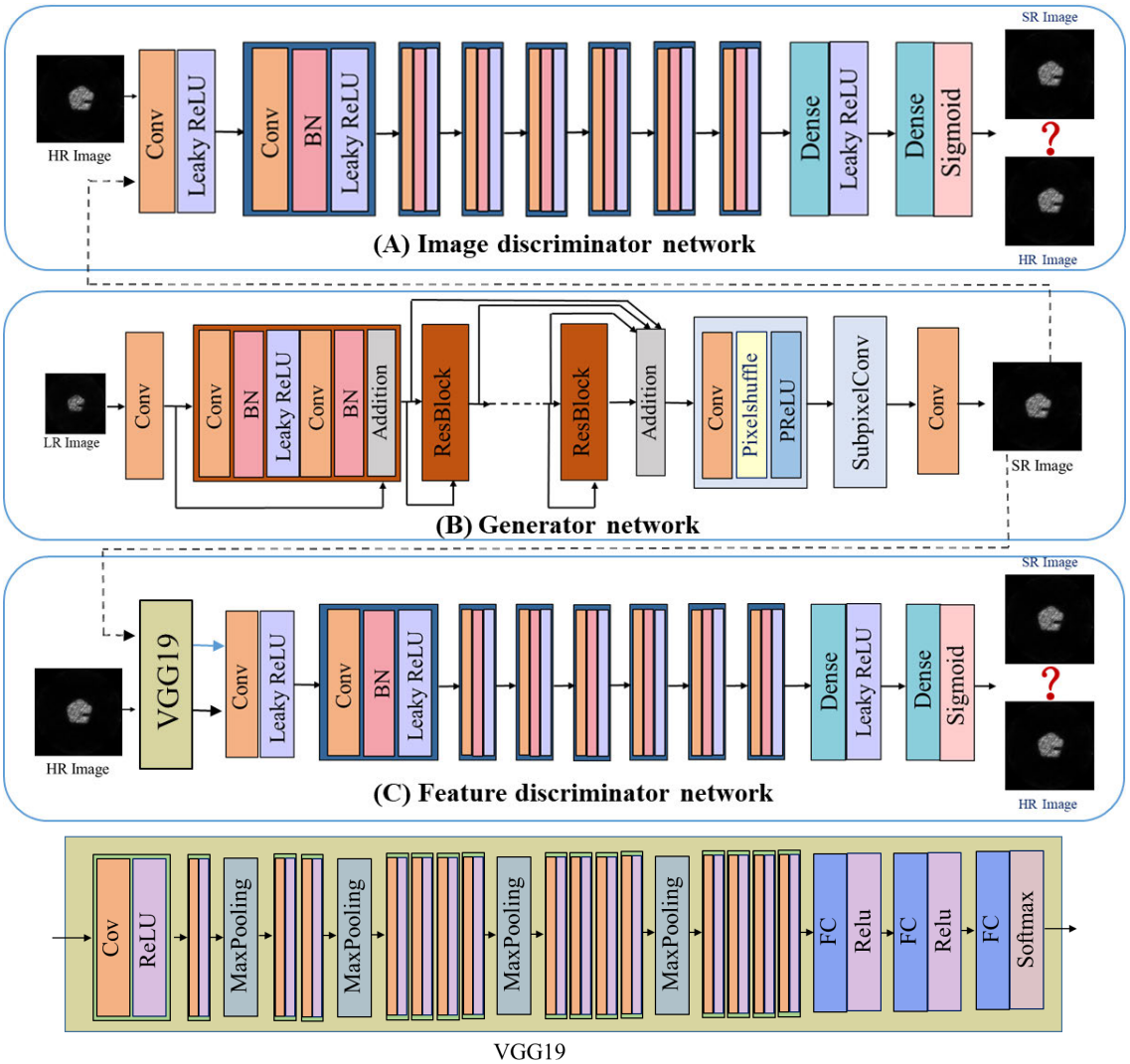


FIGURE 1. Description of PI-SRGAN network architecture and data flow direction. It contains three sub-networks: generator network, image discriminator network, and feature discriminator network. Part B is generator network, that can convert LR image to SR image. Addition residual blocks and long skip connections are added to strengthen the spread of features in generator. Then the SR image is obtained through two identical sub-pixel convolution layers. Part A is an image discriminator network that input are pixel space images, it can distinguish SR images generated by part B with real HR images in image space. Part C is a feature discriminator network uses VGG-19 to extract feature maps from SR images generated by part B and real HR images as the input images, it can distinguish input images in feature space. Feature map contains high-frequency components and structural components in part C. The bottom figure shows the network structure of VGG19 in Part C. The detailed implementation of above three structs is given in section III of this article.

leading methods in the objective assessment indicators. In Section V, this paper concludes the whole work and look forward to the future.

II. RELATED WORK

There are three principal types of existing SRIR technologies: interpolation-based, reconstruction-based and learning-based SR technologies. Among them, the SR methods based on interpolation include primarily nearest-neighbour interpolation [18], [19], bilinear interpolation [20] and bicubic interpolation [21]. These methods can improve image resolution to some extent, but the reconstructed image is too smooth, and the details are not sufficiently rich. The SR algorithms based on reconstruction include convex set projection [22], iterative inverse projection [23], and maximum posterior probability [24]. The learning-based super-resolution algorithm was first proposed by Freeman *et al.* [25] in 2002. Currently, the learning-based super-resolution algorithms include primarily example-based super-resolution and image super-resolution via sparse representation [26]–[28]. These methods solve the problem of high magnification, which could not be solved using previous algorithms, and require only a single LR image to generated the SR image. However, these methods also have their own deficiencies. For example, the example-based methods require averaging the high-frequency information of the repeated region, resulting in an excessively smooth image edge. In addition, the methods based on sparse dictionary requires many iterations and large quantities of computation, as well as lacking sufficient computational efficiency.

With the development of deep learning in image recognition, classification and other fields, researchers began to apply deep learning methods to image SR tasks [29]–[31]. Image SR has greatly progressed with the development of CNNs. Subsequently, various deep learning-based methods with different network architectures [32]–[38] and training strategies [39], [40] have been proposed to improve the SR performance. These networks have obtained an improved performance. Among them, Dong *et al.* [32]applied convolutional neural networks (CNNs) to SR for the first time and proposed the super-resolution convolutional neural network (SRCNN). The end-to-end training method was adopted in SRCNN, and the reconstruction effect was greatly improved. Kim *et al.* [33]deepened the number of convolution layers, and proposed a super-resolution model of 20 layer very deep convolutional networks (VDSR). Combining the residual idea, adaptive gradient clipping and other techniques, the quality of reconstructed images was improved effectively. At the same time, Shi *et al.* [34] proposed an efficient sub-pixel convolutional neural network (ESPCN), which achieved good reconstruction results and reduced the amount of computation effectively. In super-resolution, the traditional MSE loss function is typically used to restore the texture of LR images. MSE is used to restore the high-frequency details lost in LR images by pixel averaging, but this typically leads to image smoothing and poor image perception quality [35]. Johnson *et al.* [36] proposed using a loss function closer to perceptual similarity to restore a more convincing HR image, and achieved good results. However, although these methods are flexible in generating relatively high-quality HR images, the HR images reconstructed by them still have problems such as insufficient definition and unclear details.

GAN is a new unsupervised learning framework with wide application prospects, with its core idea being the 'two-person zero-sum game'. Ledig *et al.* [17] proposed a super-resolution method using a generative adverse network (SRGAN), the first neural network that can recover images from 4× downsampling. In recent years, Xiao *et al.* [41] proposed to model the downscaling and upscaling processes from a new perspective, that is an invertible bijective transformation, which can largely mitigate the ill-posed nature of image upscaling. However, existing SR approaches neglect to use attention mechanisms to transfer high-resolution textures from reference images, which limits these approaches to some extent. Yang *et al.* [42] propose a novel texture transformer network for image super-resolution (TTSR), in which the LR and the reference images are formulated as queries and keys in a transformer, respectively. This year, Kong *et al.* [43] proposed to combine classification and SR in a unified framework (ClassSR) to help most existing methods save up to 50% floating point operations on DIVerse 8K resolution image dataset (DIV8K), which uses a class-module to classify the sub-images into different classes according to restoration difficulties and then applies an SR-module to perform SR for different classes.

Currently, GAN is widely used in computer vision and image processing. Compared with other algorithms, the GAN-based algorithms can make full use of image information, and the model's performance is better. Image SR typically obtains the corresponding HR images by extracting the details and reconstruction steps of LR images and has a low utilization rate on feature information. Compared with CNNs, SRGAN shows substantial improvement in super-resolution image quality, but the positron image generated by SRGAN still involves fuzziness, and the algorithm typically tends to add meaningless high-frequency noise to the generated super-resolution image, leading to poor visual perception. Consequently, the PI-SRGAN method used GAN for positron image super-resolution, and improved the network structure and the loss function to generate HR images with more realistic texture details.

III. POSITRON IMAGE SUPER-RESOLUTION USING GENERATIVE ADVERSARIAL NETWORKS (PI-SRGAN)

In this section, we first give an overview of the proposed positron image super-resolution using generative adversarial networks (PI-SRGAN) and then present the modules in detail.

A. OVERVISE

The purpose of PI-SRGAN is to improve on SRGAN to enable the network to produce clearer and more realistic images when processing positron images. The structure of PI-SRGAN as shown in Fig.1. In PI-SRGAN method, a generator with long skip connections is designed, and an image discriminator network working in the pixel space and a feature discriminator working in the feature space are attached at the same time. Generator network in PI-SRGAN uses long skip connections to improve the utilization rate of intermediate feature information and enhance the generated image quality. The image discriminator network takes the image in the pixel space as the input, and the feature discriminator takes the feature map extracted from the VGG-19 network as the input.

To show this algorithm more clearly, the pseudocode of PI-SRGAN algorithm is written and shown as follow. The pseudocode contains forward and back propagation process of PI-SRGAN.

B. GENERATOR ARCHITECTURE

As a deep neural network, GAN is prone to network degradation in the process of training. The so-called network degradation problem is that with the increase of network layers, the accuracy of the training set no longer increases and might even decrease. PI-SRGAN uses residual blocks and adds long skip connections to form a generator network. Using residual networks to deepen the network depth and long skip connections to make full use of the intermediate features not only retains the forward-propagation characteristics, but also helps to alleviate the problem of gradient disappearance. The generator network structure of PI-SRGAN is shown in Fig. 2.

Algorithm 1 Forward and Back Propagation Process of PI-SRGAN

Input: pretrained model parameters *Para*;
model hyperparameters *Hyper*;
positron-image training dataset *trainSet*;
positron-image testing dataset *testSet*;

Initialization (*Para*, *Hyper*);
while not convergent **do**
  // forward time - O(n2)
  LR_{batch}, HR_{batch} = randomSelect (*trainset)*;
  SR_{batch} = generator (LR_{batch});
  $Feat_{batch}$ = VGG (SR_{batch}, HR_{batch});
  F_{result} = featureDisc ($Feat_{batch}$);
  I_{result} = imageDisc (SR_{batch}, HR_{batch});
  // back propagation
  Loss = LossFunction ($F_{result}, I_{result}, Hyper$);
  // update model parameters by minimizing
  // the object:

$$l^{SR} = \frac{1}{W_n H_n C_n} \sum_{i=1}^{W_n} \sum_{j=1}^{H_n} \sum_{k=1}^{C_n} \left(\phi^n \left(I^{HR}\right)_{i,j,k} - \phi^n \left(G_{\theta_G}\left(I^{LR}\right)\right)_{i,j,k}\right)^2 + \lambda \left(-\log\left(D_{\theta_D}^{img}\left(G_{\theta_G}\left(I^{LR}\right)\right)\right) + \left(-\log\left(D_{\theta_G}^{feat}\left(\phi^n\left(G_{\theta_G}(I^{LR})\right)\right)\right)\right)\right)$$

  updateModelParameters (*Loss*);
  // Evaluate performance
  result = forward(*testSet*);
end

Generator network in PI-SRGAN consists of one 9 × 9 convolution layer, 16 residual blocks, multiple long skip connections and two sub-pixel convolution layers. The input of the generator network is the LR image. The low-level features are extracted through the first 9 × 9 convolution layer, and then the high-level features with more nonlinearities and larger receptive fields are extracted through multiple residual blocks. Finally, the SR image is obtained by up-sampling by a factor of four through two identical sub-pixel convolution layers consisting of a 3 × 3 convolution layer, an up-sampling layer and a parametric rectified linear unit (PReLU) activation function.

Each residual block (as shown in Fig.3) in the generator network consists of a 3 × 3 convolution layer, a batch normalization (BN) layer, a leaky rectified linear unit (LeakyRELU) activation function, a 3 × 3 convolution layer and a BN layer. The gradient of LeakyReLU can only take two values, negative with a tiny gradient when the input is less than zero, the gradient is 1 when the input is greater than 0. The equation of LeakyReLU is shown in Eq.1.

$$LeakyReLU(x) = \max(0, x) + negative\_slope * \min(0, x) \tag{1}$$

The short skip connection used in the residual block can effectively alleviate the problem of gradient disappearance in the process of deep network training. The residual block retains the signal from the previous layer of the network through the short skip connection, with the network needing to learn only the changed parts (residuals), highlighting the small changes, increasing the sensitivity of the network, and not increasing the network parameters. It will not greatly increase the complexity of generator network calculations.

Generator network in PI-SRGAN uses long skip connections to aggregate the features from different residual blocks, connecting the output feature map of each residual block directly to the end of all residual blocks. The long skip connections can further encourage back-propagation of gradients, and give potentials to re-use intermediate features to improve the final feature. At the same time, to extract the feature map from the residual block and obtain the SR image with the same size as the real HR image, the sub-pixel convolution layer is used to up sample the feature map to the target resolution. The sub-pixel convolution layers obtain feature maps through the convolution layer with r^2c filters, where c is the channel number of the original LR image and r is the image super-resolution magnification. Then, the resolution of the feature map is expanded by the shuffling layer that rearranges data from channels into different spatial locations. Two sub-pixel convolution layers are used in the generator network in a row with each sub-pixel convolution layer enlarging an input feature map by the scale factor 2 because the goal of PI-SRGAN is to conduct super-resolution on the LR positron image with quadruple resolution.

C. DISCRIMINATOR ARCHITECTURE

The discriminant network in GAN distinguishes between the generated data and the real data. Continuous training and optimization of the discriminant network to modify and optimize the generation network results in the generated image having better clarity and more image details. SRGAN consists of a

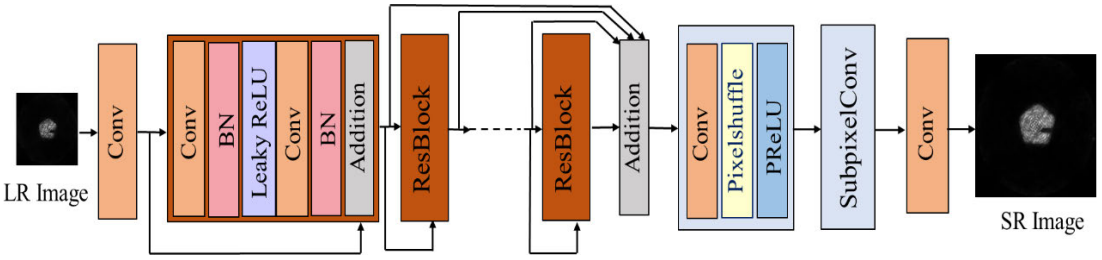


FIGURE 2. Generator network of PI-SRGAN.

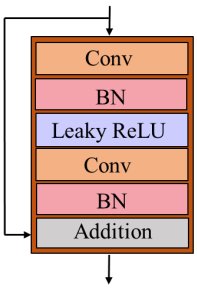


FIGURE 3. Residual block.

generator and a discriminator. Compared with convolutional neural networks, the quality of images obtained by SRGAN has been greatly improved, but the image perception quality still requires improvement owing to the lack of realistic texture details. For this reason, two discriminators are designed in PI-SRGAN. One image discriminator and one feature discriminator are used for discriminative adversarial at the same time. The inputs of the image discriminator are the SR image generated by the generator and the real HR image, which distinguishes the two kinds of images in the pixel space of the images. The SR image generated by the generator and the real HR image are extracted feature maps by a VGG-19 network as the inputs of the feature discriminator. The feature discriminator distinguishes the two kinds of images in the feature space. Through the adversarial training between the two discriminators and the generator, the generator can generate SR images with realistic structural features and texture details.

The image discriminator and feature discriminator of PI-SRGAN use the same network structure, which consist of multiple convolution layers, BN layers and LeakyReLU activation functions. The structure of the discriminator network is shown in Fig. 4.

D. LOSS FUNCTION

Since the pixel is the smallest unit of a digital image, reducing the gap between pixels can ensure the accuracy of image information more quickly and effectively. To reduce the gap between pixels and improve image generation, PI-SRGAN uses pre-training and adversarial training to train the network, with different loss functions being used in different training stages.

The MSE loss is closer to that of human visual perception and more sensitive to abnormal pixels with large numerical differences, and this improves the recovery of the low frequency content of images. Therefore, this paper selects the traditional MSE loss function in pre-training stage and. the MSE definition is as follows:

$$l_{MSE}^{SR} = \frac{1}{WHC}\sum_{i=1}^{W}\sum_{j=1}^{H}\sum_{k=1}^{C}\left(I_{i,j,k}^{HR} - G_{\theta_G}\left(I^{LR}\right)_{i,j,k}\right)^2 \quad (2)$$

where, $I_{i,j,k}^{HR}$ is the pixel value of the real HR image at its location (i, j, k), $G_{\theta_G}\left(I^{LR}\right)$ is the SR image generated by the generator network, and $G_{\theta_G}\left(I^{LR}\right)_{i,j,k}$ is the pixel value of the generated SR image at position of (i, j, k). Here W, H and C describe the dimensions of the respective feature maps within the VGG-19 network.

It is difficult to capture the difference in image perception because MSE loses part high-frequency information of the image. When takes MSE as the loss function, the image generated by the network model tends to be smooth. Although the pre-trained image can have a better peak signal to noise ratio (PSNR), but it cannot produce a pleasant effect on visual perception because of the lack of high-frequency details. To overcome the above problems, this paper trains the pre-trained generator network with discriminators, a loss function defined is minimized as:

$$l^{SR} = l_p^{SR} + \lambda l_{gen}^{SR} \quad (3)$$

where, l^{SR} is the loss function of generator, l_p^{SR} is the perceptual similarity loss that enforces SR results to look similar to the ground real HR images in the training set, l_{gen}^{SR} is the GAN loss for the generator, and λ is a weight for the GAN loss terms. The calculation process of this method is shown in Fig. 5.

The perceptual similarity loss measures the difference between the SR image and the HR image in the feature space. During training process, minimizing the perceptual similarity can achieve the visual perception consistency between the SR image and the real HR image. PI-SRGAN use the trained VGG-19 network to extract features, I^{HR} and I^{SR}(that is $G_{\theta_G}(I^{LR})$ in Eq. (4)) are put into VGG-19 network, then the feature maps of the two images at the n^{th} layer are extracted. The perceptual similarity loss between I^{HR} and I^{SR}

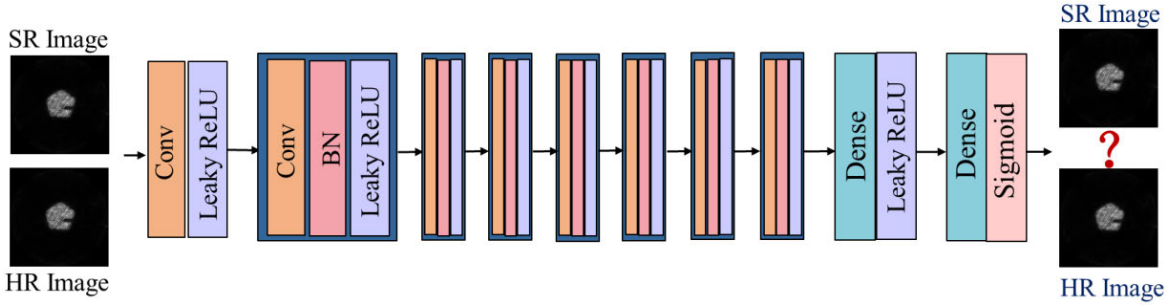


FIGURE 4. Discriminator network of PI-SRGAN.

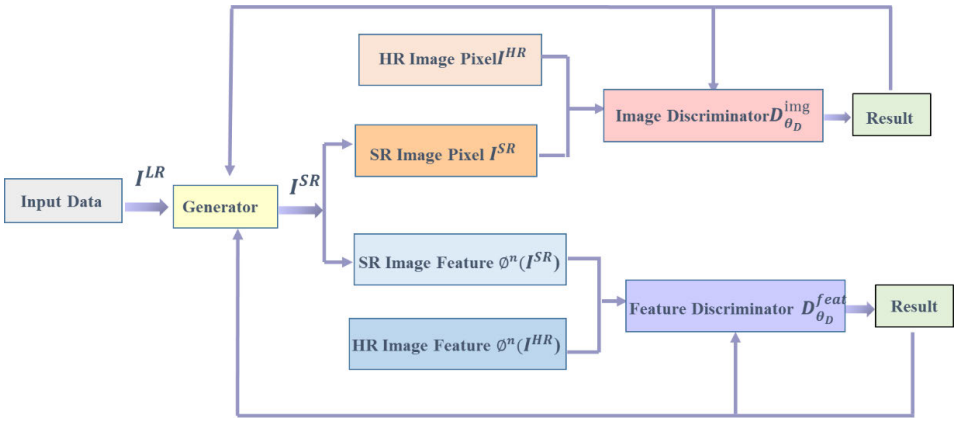


FIGURE 5. Calculation process diagram of PI-SRGAN.

is expressed as follows:

$$l_p^{SR} = \frac{1}{W_n H_n C_n} \sum_{i=1}^{W_n} \sum_{j=1}^{H_n} \sum_{k=1}^{C_n} \times \left(\phi^n \left(I^{HR} \right)_{i,j,k} - \phi^n \left(G_{\theta_G} \left(I^{LR} \right) \right)_{i,j,k} \right)^2 \quad (4)$$

where, ϕ^n is the output of the ReLU layer after the convolution before the n^{th} pooling in VGG-19 network. As well, W_n is width, H_n is height, and C_n is channel number of the n^{th} feature map.

To get SR image with more realistic structure and texture details, the adversarial loss is added, as follows:

$$l_{gen}^{SR} = l_{gen}^{img} + l_{gen}^{feat} \quad (5)$$

where, l_{gen}^{img} is the image GAN loss term for the generator, corresponding to the loss function l_{gen}^{img} for the image discriminator $D_{\theta_D}^{img}$. And l_{gen}^{feat} is the feature GAN loss of the generator, corresponding to the loss function l_{gen}^{feat} of the feature discriminator $D_{\theta_D}^{feat}$. The generator and discriminators are trained by alternatingly minimizing l^{SR}, l_{dis}^{img}, and l_{gen}^{feat}.

The adversarial loss of image l_{gen}^{img} is used to synthesize high-frequency texture details in the pixel space of SR images, as follows:

$$l_{gen}^{img} = -\log \left(D_{\theta_D}^{img} \left(G_{\theta_G} \left(I^{LR} \right) \right) \right) \quad (6)$$

The loss function of the image discriminator can be expressed as follows:

$$l_{dis}^{img} = -\log \left(D_{\theta_D}^{img} \left(I^{HR} \right) \right) - \log \left(1 - D_{\theta_D}^{img} \left(G_{\theta_G} \left(I^{LR} \right) \right) \right) \quad (7)$$

where, $G_{\theta_G}\left(I^{LR}\right)$ is the SR image generated by generator, $D_{\theta_D}^{img}\left(G_{\theta_G}\left(I^{LR}\right)\right)$ is the probability that the image $G_{\theta_G}\left(I^{LR}\right)$ is an image sampled from the distribution of real HR images, and $D_{\theta_D}^{img}\left(I^{HR}\right)$ is the probability that the image I^{HR} is an image sampled from the distribution of real HR images.

The GAN loss of feature l_{gen}^{feat} is used to synthesize structural details in the feature space of SR images, as follows:

$$l_{gen}^{feat} = -\log \left(D_{\theta_G}^{feat} \left(\phi^n \left(G_{\theta_G}(I^{LR}) \right) \right) \right) \quad (8)$$

The loss function of the feature discriminator can be expressed as follows:

$$l_{dis}^{feat} = -\log \left(D_{\theta_D}^{feat} \left(\phi^n \left(I^{HR} \right) \right) \right) - \log \left(1 - D_{\theta_D}^{feat} \left(\phi^n \left(G_{\theta_G} \left(I^{LR} \right) \right) \right) \right) \quad (9)$$

where $G_{\theta_G}\left(I^{LR}\right)$ is the SR image generated by generator, $\phi^n\left(G_{\theta_G}\left(I^{LR}\right)\right)$ represents the feature map obtained by inputting the generated SR image into the VGG19 network and passing the nth convolution layer and activation function, $D_{\theta_D}^{feat}\left(\phi^n\left(I^{HR}\right)\right)$ is the probability that the feature map $D_{\theta_D}^{feat}\left(\phi^n\left(I^{HR}\right)\right)$ is sampled from the distribution of the real HR image feature maps, $D_{\theta_D}^{feat}\left(\phi^n\left(G_{\theta_G}\left(I^{LR}\right)\right)\right)$ is the probability that the feature map $\phi^n\left(G_{\theta_G}\left(I^{LR}\right)\right)$ is sampled from the distribution of the real HR image feature maps.

E. PARAMETERS OF PI-SRGAN NETWORKS

The accuracy of the selection of neural network parameters has an effect on the complexity of network calculation, network training time and the accuracy of the target task; in particular the selection of generator network parameters plays a vital role in the quality of the generated image. Consequently, PI-SRGAN choose primarily the adjustable weight coefficient λ in Eq. (2) and the number n in Eq. (3) of residual blocks in the generator loss function, and the paper will discuss the necessity of long skip connections.

In this process, the ImageNet dataset, which contains 1,000 categories of images, with each category containing millions of images, is used as the training e dataset. Selecting randomly 10,000 images in the ImageNet dataset with width and height greater than 400 pixels and cut them to 224×224 as real HR images, and then using bicubic interpolation to downsample these HR images to obtain the 56×56 low-resolution input training images One hundred images were selected randomly from the Set5 dataset. The real HR

images were obtained by clipping and the LR images were obtained by downsampling as the test dataset.

TABLE 1. Impact of different values on network performance.

λ	0	**1**	10	100
PSNR	27.18	**32.49**	32.04	31.58
SSIM	0.6952	**0.9026**	0.8782	0.8512

1) ADJUSTABLE WEIGHT COEFFICIENT λ

The adjustable weight coefficient λ is used to balance the overall visual perception and high-frequency structural details of a reconstructed image. In this paper, the weights are set to 0, 1, 10, 100 respectively. The extracted images in the ImageNet dataset are used for training, and the images in the Set5 dataset are used for testing. Currently, the commonly used full reference objective evaluation criteria for image quality are the PSNR and structural similarity index measurement (SSIM). The PSNR reflects the error between the corresponding pixels of two images. The higher the value of PSNR is, the less distortion the output image has and the better the quality of image reconstruction is. SSIM is an evaluation index of similarity between two images. The closer the value of SSIM is to one, the closer the output image is to the original HR image, that is, the better the reconstruction effect is. Therefore, when testing images, the PSNR and SSIM of all tested images are calculated and their sum is averaged. The results are shown in Table 1.

As can be seen from Table 1, the PSNR and SSIM of an SR image are the lowest when λ was set to zero. At this time, the loss function of the generator includes only the perceptual similarity loss. The PSNR and SSIM are improved after adding the adversarial loss, but with increasing values of λ, the PSNR and SSIM decrease. It can be seen that with the increasing of the λ value, the network pays increasingly more attention to the high-frequency details of the image and ignores the overall quality of the image, thereby reducing the performance index of the image. Therefore, the adjustable weight coefficient λ was be set to one.

2) RESIDUAL BLOCK NUMBER *n* OF GENERATOR

The more residual blocks there are, the more accurate the features extracted by the residual blocks are and the closer SR images are to the real HR images. However, too many residual blocks will result in the saturation of accuracy and an increase in network computational complexity. Consequently, choosing the appropriate number of residual blocks can reduce both the computation amount and the training time of the network on the premise of ensuring the SR image quality.

The SRGAN algorithm uses 5 residual blocks in the generator network. With referring to this value, setting 5 to 20 residual blocks are set for the generator, respectively. The images from the ImageNet dataset are extracted for training and the images in the Set5 dataset are used for testing. The PSNR and SSIM of all SR images were calculated and their sum averaged. The results are shown in Table 2.

It can be seen from Table 2 that when the number of residual blocks is small, both the PSNR and SSIM improved with the increase in the number of residual blocks. However, the improvement range of the performance indicators decreases with the increase in the number of residual blocks. Until the number of residual blocks is increased to 16, both the PSNR and SSIM reach the best values. Subsequently, the continuous increase in the number of residual blocks had no obvious effect on performance. When the number of residual blocks is 16, the generator network of PI-SRGAN achieves the best performance. Therefore, the number of residuals in the generator network of PI-SRGAN is set to 16.

3) THE NECESSITY OF LONG SKIP CONNECTIONS

To verify the necessity of long skip connections, the generation network with long skip connections is shown in Fig.6(a), and the generation network without long skip connections is shown in Fig.6(b).

The images in the ImageNet dataset are extracted to train the GANs that remove the long skip connections and retain the long skip connections, respectively. The images in the Set5 dataset are used for testing. The PSNR and SSIM of all testing images are calculated and their sum is averaged. The results are shown in Table 3.

It can be seen from Table 3 that the generator network with long skip connections has higher values of PSNR and SSIM than the generator network without long skip connections. Because long skip connections can indeed improve the quality of SR image generation, the generator in PI-SRGAN adds the long skip connections.

IV. EXPERIMENT AND ANALYSIS

In Section III.B, Section III.C and Section III.D, the network structures and parameters have been determined. However, to apply PI-SRGAN method to super-resolve the positron images, it is necessary to train the network parameters with the positron images dataset. The trained network is used for super-resolving experimental images. In this section, the network training process and experimental processing will be described in detail, and the comparison results between PI-SRGAN and other image SR algorithms will be shown.

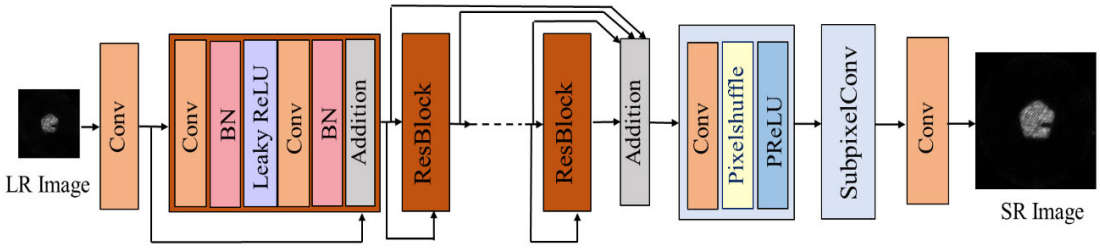
A. DATA PRE-PROCESSING

Different learning-based SISR methods use different training sets. For example, SRCNN uses ImageNet as its training set, ESPCN uses VDCL as its training set, SRGAN uses ImageNet as its training set. SolidWorks was used to design a group of industrial components to obtain positron images to train final PI-SRGAN network. As shown in Fig.7, all ten of the models can be regarded as foreign objects with different shapes placed in the inner cavity of circular pipes. The ten models are simulated by GATE to obtain the projection data. Then, the positron images of the ten groups of models are reconstructed by using the OSEM algorithm, as shown in Fig.8.

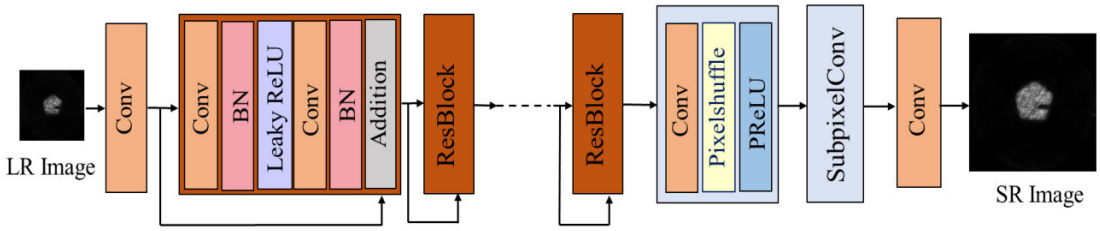
TABLE 2. Influence of the number n of residual blocks on network performance.

n	5	6	7	8	9	10	11	12
PSNR	22.23	24.66	26.60	28.12	29.31	30.22	30.91	31.45
SSIM	0.6667	0.7211	0.7649	0.7993	0.8262	0.8469	0.8624	0.8746

n	13	14	15	**16**	17	18	19	20
PSNR	31.86	32.15	32.34	**32.49**	32.48	32.50	32.49	32.49
SSIM	0.8839	0.8905	0.8948	**0.9026**	0.9023	0.9025	0.9027	0.9026



(a) Generating network with long skip connections



(b) Generating network without long skip connections

FIGURE 6. Comparison of generated networks with and without long skip connections.

TABLE 3. Impact of long skip connections on network performance.

Are there long skip connections	PSNR	SSIM
Yes	**32.49**	**0.9026**
No	32.27	0.8974

A training network requires a large number of LR and HR images. However, the number of designed models is limited. The research cut, rotated, panned and zoomed the positron images to expand the positron images dataset and to increase the generalization ability of PI-SRGAN model and avoid overfitting. Then, the study obtained the HR images dataset with 224 × 224 pixels, and down sampled (4×) the HR images to obtain the corresponding 56 × 56 LR images dataset.

For each group of models, 1,170 positron images were selected randomly and down sampled to obtain 1,170 sets of datasets consisting of HR and LR images. Eighty percent of the dataset is used as the training set, and the remaining 20% is used as the verification set. The dataset is used to train PI-SRGAN network, and the TensorBoard tool is used to visualize the change in the loss function of the training process of the generator, as shown in Fig. 9.

Fig.9(a) shows the change in the loss function of the image generated by the generator in the pixel space, and Fig.9(b) shows the change in the loss function of the image generated by the generator in the feature space. It can be seen from Fig.9 that the loss of PI-SRGAN network tends to converge in both the pixel space and the feature space, with the result that the obtained SR image gradually approaches the real HR image from the pixel space and the feature space.

B. EXPERIMENTAL RESULTS AND ANALYSIS

To verify the SR effect of the PI-SRGAN algorithm for positron images, two groups of PET simulation experiments and two groups of PET actual experiments were designed. And the research used different experimental models to verify the effects and advantages of PI-SRGAN from different aspects.

To verify the performance of PI-SRGAN, two groups of simulation experiments were designed. The first model is the classical Derenzo model, as shown in Fig.10(a). The second model is a hydraulic cylinder model. A small nut was put in the inner cavity of the hydraulic cylinder for simulation experiments, as shown in Fig.10(b).

The two groups of models are simulated by GATE to obtain the projection data. Then, the projection data is reconstructed

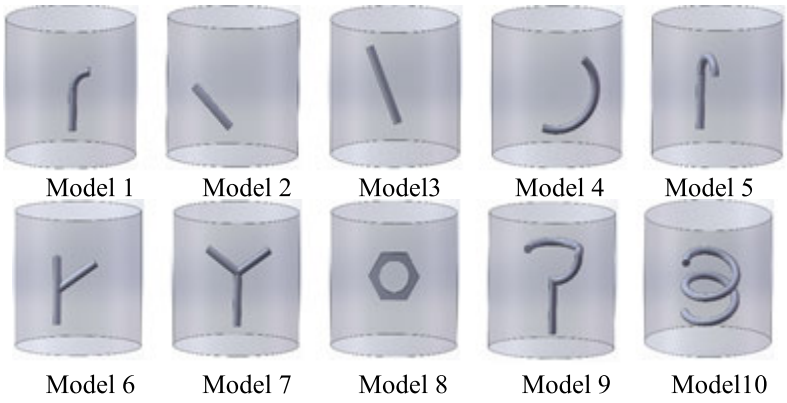


FIGURE 7. Ten groups of solidworks simulation models.

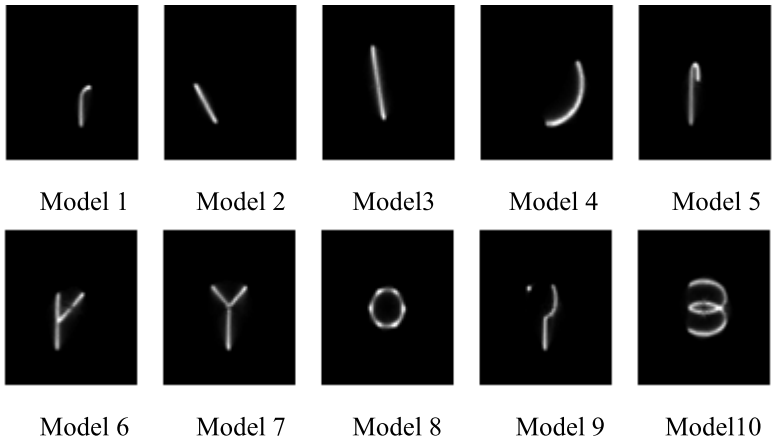
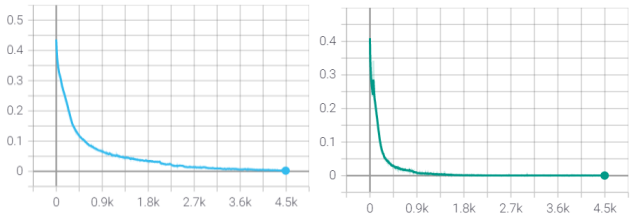


FIGURE 8. Partial positron images of the ten models.



(a) In the pixel space (b) In the feature space

FIGURE 9. Loss function changes of generator.

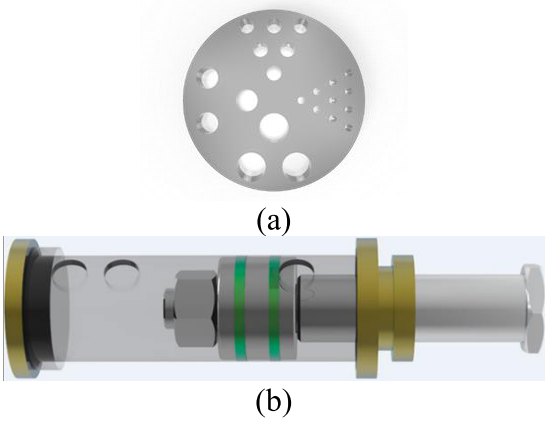


FIGURE 10. (a) Derenzo model, (b) Hydraulic cylinder model.

by OSEM to obtain the positron images of the two models. The resolution of the positron image of the Derenzo model is 128 × 128, as shown in Fig.11(a). The resolution of the positron image of the cavity of the hydraulic cylinder model is 278 × 278, as shown in Fig.11(b). In Fig.11, it can be seen that the phenomenon of the positron image being very fuzzy; it is impossible to see the details clearly. Therefore, it is necessary to SR the images in Fig.11 to improve image clarity and enhance image quality.

To test whether PI-SRGAN is superior to other SR algorithms in positron image, the paper use Bicubic algorithm, SRCNN algorithm, ESPCN algorithm, SRGAN algorithm and PI-SRGAN algorithm to super-resolve (4×) the positron images in Fig.11. The positron images of the Derenzo model and the cavity in hydraulic cylinder model are reconstructed using these methods, and the reconstructed images are shown in Fig.12 and Fig.13, respectively. From these figures, it can be seen that compared with other SR algorithms, the SR image reconstructed using PI-SRGAN method has improved visual perception after zooming in on details, the texture of the image is clearer, there are almost no fuzzy or high-frequency artefacts and the high-frequency structure details are better.

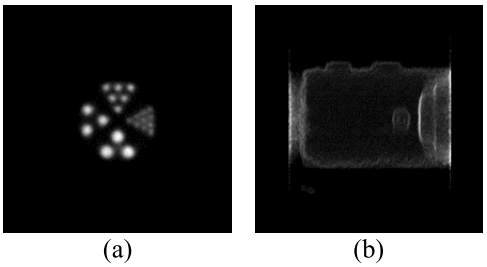


FIGURE 11. (a) Positron image of the derenzo model, (b) Positron image of the hydraulic cylinder model.

Subjective assessment might have visual differences as a result of individual differences. To evaluate the results more objectively, two common testing indices were used in image processing: PSNR and SSIM, which show the superiority of the PI-SRGAN algorithm more accurately in comparison with other SR algorithms. When the SR factor is four, the PSNR and SSIM values of SR images reconstructed using these two models are calculated. Table 4 shows the PSNR and SSIM values of Derenzo model positron images and hydraulic cylinder model positron images. The results of PI-SRGAN method are better than the comparison algorithms in PSNR and SSIM, and the SR images generated by PI-SRGAN have the best performance index. Consequently, these results show that PI-SRGAN is superior to other SR algorithms.

To verify the effect of PI-SRGAN on the positron image in actual experiments, two groups of intracavity defect detection experiments were designed. In the first group, the defect detection experiment of the inner cavity of the U-shaped pipe was designed. The object of the first experiment is a model with grooves in the U-shaped pipe; the figure in Fig.14 (a) is the solid model drawn by SolidWorks. To further verify the effectiveness of PI-SRGAN in practical applications, the second group of experimental models adds expansion and irregular crack defects that the narrowest crack defect is 1 mm on the annular pipe to verify the imaging effect of PI-SRGAN for small defects. The second model is shown in Fig.14(b).

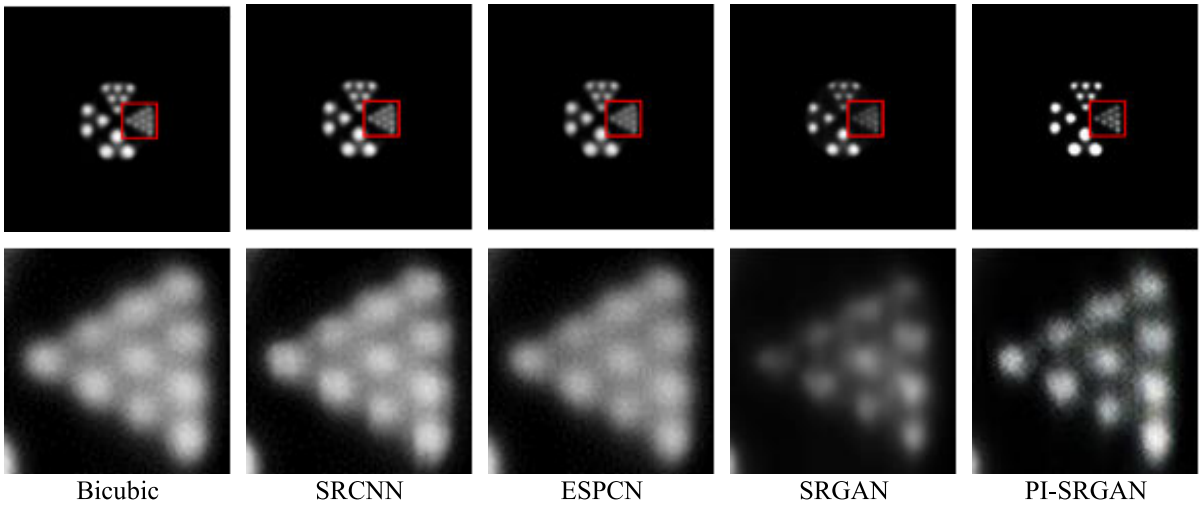


FIGURE 12. SRIR results of the positron image of the derenzo model.

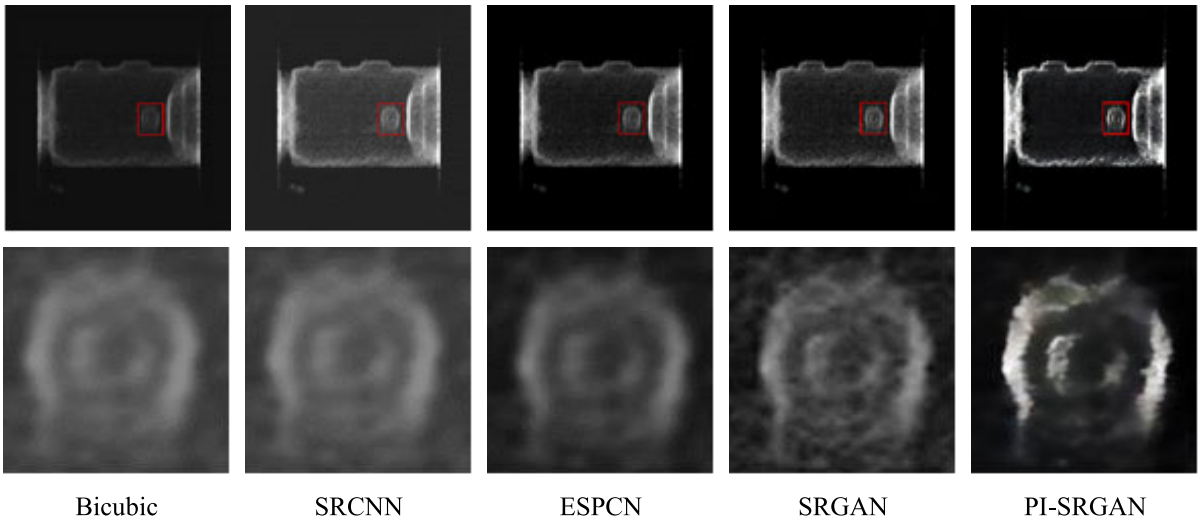


FIGURE 13. SRIR results of the positron image of the inner cavity in hydraulic cylinder model.

TABLE 4. Quantitative comparison of different SR algorithms for positron images in simulation.

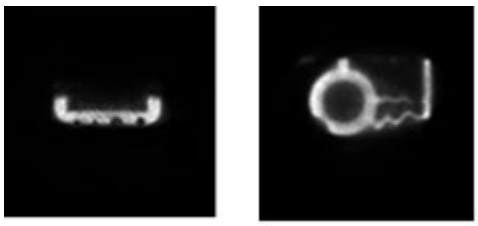
Sets	Bicubic PSNR/SSIM	SRCNN PSNR/SSIM	ESPCN PSNR/SSIM	SRGAN PSNR/SSIM	**PI-SRGAN PSNR/SSIM**
Positron images of Derenzo model	30.72/0.8837	30.86/0.8865	32.47/0.8969	36.35/0.9188	**38.79/0.9296**
Positron images of inner in hydraulic cylinder model	24.31/0.8158	24.92/0.8247	26.26/0.8571	27.13/0.8960	**29.53/0.9150**

TABLE 5. Quantitative comparison of different SR algorithms for positron images in actual experiment.

Sets	Bicubic SD/AG	SRCNN SD/AG	ESPCN SD/AG	SRGAN SD/AG	**PI-SRGAN SD/AG**
Positron images of pipeline detection experiment	47.7544/0.7517	51.8064/0.7943	55.7392/0.8389	64.0038/1.0176	**74.0308/2.3774**
Positron images of crack experiment	60.9962/0.8282	63.5891/0.8519	69.9287/0.9767	78.2899/1.5969	**84.3000/3.7665**

(a) (b)

FIGURE 14. (a) Experimental model of inner cavity in a U-shaped pipe, (b) Experimental model of a crack.



(a) (b)

FIGURE 15. (a) Positron image of the U-shaped pipe cavity defect experiment, (b) Positron image of the crack experiment.

The two groups of experiments used Trans-Pet Explorer-180 to detect and sample the gamma photons in a short scan time and low count state. The experiments used a deoxyglucose solution labelled by 18F with a total activity of 1mCi as radionuclides, and the scan time was set for 2 minutes. The study used OSEM to reconstruct the experimental projection data quickly and obtained the corresponding positron image, as shown in Fig.15. Fig.15(a) is the positron image of the defect detection experiment in the inner cavity of the pipeline, and the pixel size of the image is 128×128. Fig.15(b) shows the positron image of the crack experiment. The pixel size of the image is 398×398.

As can be seen from Fig.15, the positron image reconstructed by fast sampling has problems such as low resolution and fuzziness. It is difficult to see the specific texture details inside the image and to identify the specific shapes of the parts, which cannot meet industrial testing standards. Consequently, SRIR processing is required for Fig.15 to enrich image details, to improve image clarity and to enhance image quality.

The paper used the Bicubic, SRCNN, ESPCN, SRGAN and PI-SRGAN algorithms to SRIR the images in Fig.15, respectively. These SR images are compared from the subjective effect and objective indexes. To ensure objective and accurate comparison, all network models are set to the amplification factor of four, and the resolution of the reconstructed image was four times the original image. The results of the SR images of Fig.15 (a) are shown in Fig.16, and the SR images of Fig.15 (b) are shown in Fig.17.

As shown in Fig.16, the image obtained by PI-SRGAN shows the shape marked by the red box clearly, and the shapes are consistent with the model. As shown in Fig.17, the crack details of the image reconstructed using PI-SRGAN can be seen clearly. In addition, it can be seen from Fig.16 and Fig.17 that the SR image obtained by PI-SRGAN is improved significantly in visual perception, the texture details of the image are clearer, and there are almost no fuzzy and high-frequency artefacts. This further shows that PI-SRGAN can enrich the details of the image, can improve the quality of the image and is superior to other SR algorithms.

Generally, the original reference images cannot be obtained in many practical applications. PSNR and SSIM are not suitable for evaluating actual experimental images. The study used the blind image quality assessment indices [44], [45] such as the standard deviation (SD) and average gradient (AG) to evaluate the quality of these positron images. The SD is the degree of dispersion of grey values of image pixels relative to the average value. The larger the SD is, the more scattered the grey level is in the image and the better the image quality is. The AG is the expressive ability of the image for the contrast of small details and the texture features. The greater the average gradient value is, the clearer the image will be. The SD and AG of the SR images reconstructed by the five algorithms are calculated respectively. The results are shown in Table 5.

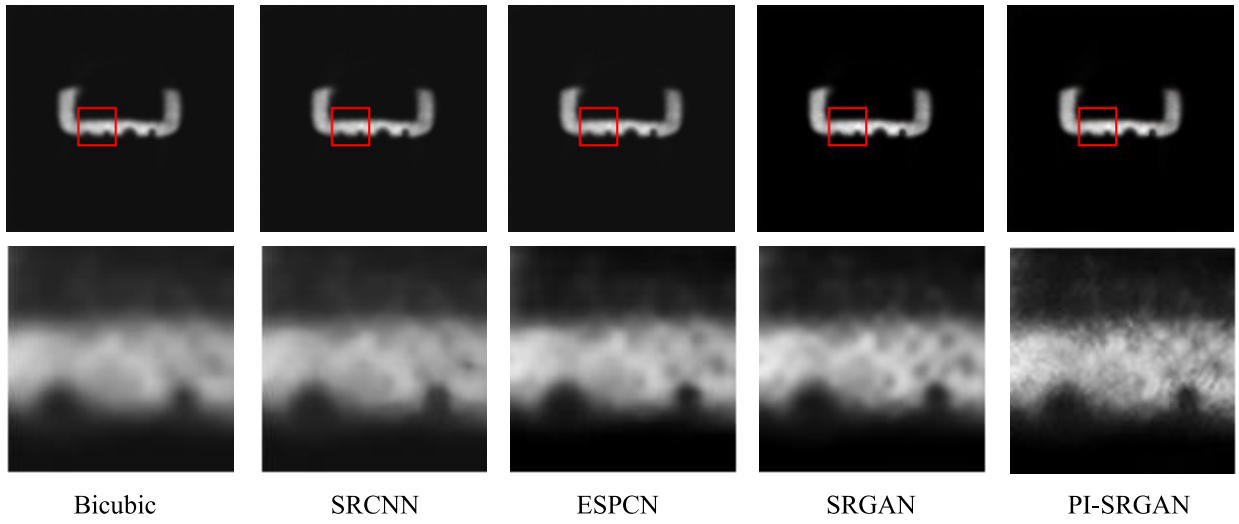


FIGURE 16. SRIR results of positron image of the U-shaped pipe cavity defect experiment.

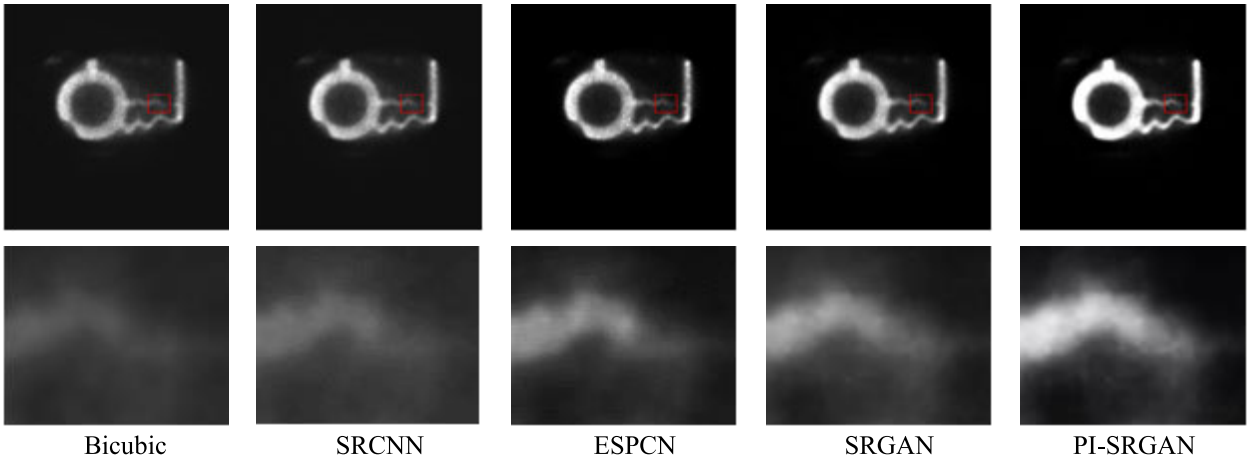


FIGURE 17. SRIR results of positron image of the crack experiment.

From the calculation results in Table 5, it can be seen that the SD and AG of the SR image reconstructed by PI-SRGAN are higher than the corresponding values of the images obtained by the comparison algorithms. It can be seen that the SR images generated by PI-SRGAN have better performance in the sense quality. The experiment shows that PI-SRGAN algorithm is helpful in restoring high-frequency information such as texture details to obtain better visual effect and enhance image quality effectively, and it is also better than the other algorithms in practical application.

V. CONCLUSION

This research proposes a deep neural network (PI-SRGAN) for super-resolution of low-resolution positron images based on GANs, which can perform 4× super-resolution reconstruction of positron images of industrial parts. For positron images in simulation experiments, the evaluation values (i.e. PSNR and SSIM evaluation indexes) of SR positron images generated by PI-SRGAN are higher than evaluation values of other methods. Meanwhile, the SR images have better subjective perception which was judged by professionals. In actual experiments, directly generating HR positron images will incur high costs. Existing quantitative assessment measures, such as PSNR and SSIM, are not appropriate to measure the SR images when the input of PI-SRGAN is LR positron image obtained in actual experiments. Thereby, this research evaluates the SR images about actual experiments by the standard deviation (SD) and average gradient (AG) of the blind image quality assessment indexes. The subjective effect and objective data of this work are superior to other comparison methods. PI-SRGAN method introduces additional residual blocks and adds long skip connections, which strengthen feature propagation and accelerate feature reuse, thereby improving feature utilization. Furthermore, an image discriminator and a feature discriminator simultaneously distinguish between generated SR images and real HR images from pixel space and feature space. In consequence, the discriminators can render the generator to pay more attention to high-frequency texture and structural details to reconstruct high quality SR positron images. Compared with other super-resolution algorithms such as SRGAN, PI-SRGAN can extract the features of LR positron images

more comprehensively, improve the utilization of feature information, and restore the texture information of images more fully. PI-SRGAN further improves the quality of the reconstructed image while accelerating the convergence of the model. This research fully proves the effectiveness and practicability of the PI-SRGAN. Thus, it can improve the ability of positron rapid non-destructive testing in industrial applications.

In the future, this research can be expanded in the following parts. First of all, due to the limitation of the size of dataset, the performance of PI-SRGAN in image super-resolution of unknown industrial parts may not up to the part that has already appeared in dataset. For increasing the generalization ability of the model, researchers could collect more high-resolution positron images and they can conduct transfer learning based on the dataset and weights provided by this research. What's more, by the reason of the positron images dataset is relatively pure dataset, noise can be added into the training dataset to improve robustness of the model. Similarly, the proposed model can be applied in various scenarios to improve flexibility and robustness of the model. Second, as self-attention [46], [47], transformer [46] and brand new MLP [48] come to the stage of CV field, the original pure CNN structure can be replaced by a new neural network module, or extra neural network layer can be added in the original network. This practice may improve the capacity and performance of the model, such as improving the quality of the super-resolution images, reducing the amount of computation required in model training, and increasing the authenticity of the super-resolution images inferred by the model. Thirdly, aiming to making the model super-resolve video stream data, model pruning algorithm can be applied on PI-SRGAN to remove redundant parameter branches, reduce the computational amount required by the model forward reasoning, and then speed the super-resolution of positron image up.

REFERENCES

[1] Y. Jin and Y. Dong, "Research on maintenance and fault diagnosis technology of hydraulic system of construction machinery," *Manuf. Automat.*, vol. 33, no. 9, pp. 144–146, 2011.

[2] M. Rezaei-Malek, M. Mohammadi, J.-Y. Dantan, A. Siadat, and R. Tavakkoli-Moghaddam, "A review on optimisation of part quality inspection planning in a multi-stage manufacturing system," *Int. J. Prod. Res.*, vol. 57, nos. 15–16, pp. 4880–4897, 2019.

[3] X. Chen, J. Li, and Z. Wang, "Inversion method in pulsed eddy current testing for wall thickness of ferromagnetic pipes," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 12, pp. 9766–9773, Dec. 2020.

[4] Z. Zhu, H.-H. Huang, and S. Pang, "Photon allocation strategy in region-of-interest tomographic imaging," *IEEE Trans. Comput. Imag.*, vol. 6, pp. 125–137, 2020.

[5] T. Hensler, M. Tupy, T. Strer, T. Pöschel, and K.-E. Wirth, "Positron emission particle tracking in fluidized beds with secondary gas injection," *Powder Technol.*, vol. 279, pp. 113–122, Jul. 2015.

[6] M. Yao, Y. Zhang, M. Zhao, R. Guo, and J. Xu, "Research on combustion flow field imaging method based on ray casting algorithm," *AIP Adv.*, vol. 9, no. 5, May 2019, Art. no. 055022.

[7] L. Jiantang, Z. Min, and X. Hui, "A positron injection-based 3D imaging test method for inner cavities in complex parts," *Insight, Non-Destructive Test. Condition Monit.*, vol. 58, no. 11, pp. 617–622, Nov. 2016.

[8] H. Xiao, M. Zhao, J. Liu, and H. Chen, "A study on scattering correction for γ-photon 3D imaging test method," *AIP Adv.*, vol. 8, no. 3, Mar. 2018, Art. no. 035315.

[9] H. Xiao, M. Zhao, J. Liu, J. Liu, and H. Chen, "A new method for spatial structure detection of complex inner cavities based on 3D γ-photon imaging," *AIP Adv.*, vol. 8, no. 5, May 2018, Art. no. 055205.

[10] W. W. W. Zou and P. C. Yuen, "Very low resolution face recognition problem," *IEEE Trans. Image Process.*, vol. 21, no. 1, pp. 327–340, Jan. 2012.

[11] X. Bing, W. Zhang, L. Zheng, and Y. Zhang, "Medical image super resolution using improved generative adversarial networks," *IEEE Access*, vol. 7, pp. 145030–145038, 2019.

[12] C. K. Chong and E. T. W. Ho, "Synthesis of 3D MRI brain images with shape and texture generative adversarial deep neural networks," *IEEE Access*, vol. 9, pp. 64747–64760, 2021, doi: 10.1109/ACCESS.2021.3075608.

[13] M. Adil, S. Mamoon, A. Zakir, M. A. Manzoor, and Z. Lian, "Multi scale-adaptive super-resolution person re-identification using GAN," *IEEE Access*, vol. 8, pp. 177351–177362, 2020.

[14] Y. Li, K. Zhao, F. Ren, B. Wang, and J. Zhao, "Research on super-resolution image reconstruction based on low-resolution infrared sensor," *IEEE Access*, vol. 8, pp. 69186–69199, 2020, doi: 10.1109/ACCESS.2020.2984945.

[15] J. L. Harris, "Diffraction and resolving power," *J. Opt. Soc. Amer.*, vol. 54, no. 7, pp. 931–936, 1964.

[16] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Proc. Adv. Neural Inf. Process. Syst.*, 2014, pp. 2672–2680.

[17] C. Ledig, L. Theis, F. Huszar, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, and W. Shi, "Photo-realistic single image super-resolution using a generative adversarial network," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jul. 2017, pp. 4681–4690.

[18] R. R. Schultz and R. L. Stevenson, "A Bayesian approach to image expansion for improved definition," *IEEE Trans. Image Process.*, vol. 3, no. 3, pp. 233–242, May 1994.

[19] R. Timofte, V. De Smet, and L. Van Gool, "Anchored neighborhood regression for fast example-based super-resolution," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2013, pp. 1920–1927.

[20] H. S. Hou and H. Andrews, "Cubic splines for image interpolation and digital filtering," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-26, no. 6, pp. 508–517, Dec. 1978.

[21] R. Keys, "Cubic convolution interpolation for digital image processing," *IEEE Trans. Acoust., Speech, Signal Process.* vol. ASSP-29, no. 6, pp. 1153–1160, Dec. 1981.

[22] H. Stark and P. Oskoui, "High-resolution image recovery from image-plane arrays, using convex projections," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 6, no. 11, pp. 1715–1726, 1989.

[23] M. Irani and S. Peleg, "Super resolution from image sequences," in *Proc. 10th Int. Conf. Pattern Recognit.*, 1990, pp. 115–120.

[24] R. R. Schultz and R. L. Stevenson, "Extraction of high-resolution frames from video sequences," *IEEE Trans. Image Process.*, vol. 5, no. 6, pp. 996–1011, Jun. 1996.

[25] W. T. Freeman, T. R. Jones, and E. C. Pasztor, "Example-based super-resolution," *IEEE Comput. Graph. Appl.*, vol. 22, no. 2, pp. 56–65, Mar./Apr. 2002.

[26] J. Yang, J. Wright, T. S. Huang, and Y. Ma, "Image super-resolution via sparse representation," *IEEE Trans. Image Process.*, vol. 19, no. 11, pp. 2861–2873, Nov. 2010.

[27] R. Zeyde, M. Elad, and M. Protter, "On single image scale-up using sparse-representations," in *Proc. Int. Conf. Curves Surf.*, Jun. 2010, pp. 711–730.

[28] J. Yang, Z. Wang, Z. Lin, S. Cohen, and T. Huang, "Coupled dictionary training for image super-resolution," *IEEE Trans. Image Process.*, vol. 21, no. 8, pp. 3467–3478, Aug. 2012.

[29] Y. Wang, L. Wang, H. Wang, and P. Li, "End-to-end image super-resolution via deep and shallow convolutional networks," *IEEE Access*, vol. 7, pp. 31959–31970, 2019.

[30] Y. Sun, W. Zhang, H. Gu, C. Liu, S. Hong, W. Xu, J. Yang, and G. Gui, "Convolutional neural network based models for improving super-resolution imaging," *IEEE Access*, vol. 7, pp. 43042–43051, 2019.

[31] B. Zhao, R. Hu, X. Jia, and Y. Guo, "Multi-scale residual fusion network for super-resolution reconstruction of single image," *IEEE Access*, vol. 8, pp. 155285–155295, 2020.

[32] C. Dong, C. C. Loy, K. He, and X. Tang, "Learning a deep convolutional network for image super-resolution," in *Proc. Eur. Conf. Comput. Vis.* Cham, Switzerland: Springer, 2014, pp. 184–199.

[33] J. Kim, J. K. Lee, and K. M. Lee, "Accurate image super-resolution using very deep convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 1646–1654.
[34] W. Shi, J. Caballero, F. Huszár, J. Totz, A. P. Aitken, R. Bishop, D. Rueckert, and Z. Wang, "Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 1874–1883.
[35] M. Mathieu, C. Couprie, and Y. LeCun, "Deep multi-scale video prediction beyond mean square error," 2015, *arXiv:1511.05440*. [Online]. Available: http://arxiv.org/abs/1511.05440
[36] J. Johnson, A. Alahi, and F. Li, "Perceptual losses for real-time style transfer and super-resolution," in *Proc. Eur. Conf. Comput. Vis.* Cham, Switzerland: Springer, 2016, pp. 694–711.
[37] Y. Wu, L. Lan, H. Long, G. Kong, X. Duan, and C. Xu, "Image super-resolution reconstruction based on a generative adversarial network," *IEEE Access*, vol. 8, pp. 215133–215144, 2020.
[38] J. Gu, H. Lu, W. Zuo, and C. Dong, "Blind super-resolution with iterative kernel correction," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 1604–1613.
[39] A. Bulat, J. Yang, and G. Tzimiropoulos, "To learn image super-resolution, use a GAN to learn how to do image degradation first," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2018, pp. 185–200.
[40] A. Shocher, N. Cohen, and M. Irani, "Zero-shot super-resolution using deep internal learning," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2018, pp. 3118–3126.
[41] M. Xiao, S. Zheng, C. Liu, Y. Wang, D. He, G. Ke, J. Bian, Z. Lin, and T.-Y. Liu, "Invertible image rescaling," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2020, pp. 126–144.
[42] F. Yang, H. Yang, J. Fu, H. Lu, and B. Guo, "Learning texture transformer network for image super-resolution," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 5790–5799.
[43] X. T. Kong, H. Y. Zhao, Y. Qiao, and C. Dong, "ClassSR: A general framework to accelerate super-resolution networks by data characteristic," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Mar. 2021, pp. 12016–12025.
[44] S. Xu, S. Jiang, and W. Min, "No-reference/blind image quality assessment: A survey," *IETE Tech. Rev.*, vol. 34, no. 3, pp. 223–245, May 2017.
[45] G. Zhai and X. Min, "Perceptual image quality assessment: A survey," *Sci. China Inf. Sci.*, vol. 63, no. 11, Nov. 2020, Art. no. 211301.
[46] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, "Attention is all you need," 2017, *arXiv:1706.03762*. [Online]. Available: https://arxiv.org/abs/1706.03762
[47] F. Wang, M. Jiang, C. Qian, S. Yang, C. Li, H. Zhang, X. Wang, and X. Tang, "Residual attention network for image classification," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2017, pp. 3156–3164. [Online]. Available: https://arxiv.org/abs/1704.06904
[48] I. Tolstikhin, N. Houlsby, A. Kolesnikov, L. Beyer, X. Zhai, T. Unterthiner, J. Yung, A. Steiner, D. Keysers, J. Uszkoreit, M. Lucic, and A. Dosovitskiy, "MLP-mixer: An all-MLP architecture for vision," *CoRR*, vol. abs/2105.01601, pp. 1–16, May 2020.

FANG XIONG was born in Ganzhou, Jiangxi, China, in 1995. She received the B.Sc. degree from Shandong University, China, in 2018. She is currently a Graduate Student with Nanjing University of Aeronautics and Astronautics. Her research interests include positron nondestructive testing, algorithm optimization, and image processing.

JIAN LIU (Member, IEEE) received the B.S., M.S., and Ph.D. degrees in instrumentation science and technology from Harbin Institute of Technology (HIT), Harbin, China, in 2005, 2007, and 2011, respectively. He was a Joint Cultivated Ph.D. Candidate at the Prefectural University of Hiroshima, Hiroshima, Japan, from October 2008 to September 2009. He is currently a Lecturer with Nanjing University of Aeronautics and Astronautics, Nanjing, China. His main areas of research interests include active noise control, adaptive signal processing, and sensor signal processing. He is a member of the Acoustical Society of China (ASC) and the Institute of Electronics, Information, and Communications Engineers (IEICE), Japan.

MIN ZHAO is currently a Professor and the Doctoral Supervisor with Nanjing University of Aeronautics and Astronautics. His current research interests include UAV trajectory planning and control, positron annihilation technology for nondestructive testing, computer measurement, and control system for high-speed railway nondestructive testing. He is a member of the Metrology Technology Professional Committee, Chinese Society of Aeronautics and Astronautics; the Vice Chairman of the Thermal Engineering Professional Committee, Jiangsu Institute of Metrology and Testing; and the Chairman of the Supervisory Board, Jiangsu Institute of Instrumentation.

MIN YAO received the B.Sc., M.Sc., and Ph.D. degrees from Nanjing University of Aeronautics and Astronautics, China, in 1997, 2002, and 2008, respectively. She is currently an Associate Professor with Nanjing University of Aeronautics and Astronautics. Her research interests include computer measurement, control and UAVs task assignment, data and signal processing, and algorithm optimization.

RUIPENG GUO received the Ph.D. degree in instrument science and technology from Shanghai Jiao Tong University, Shanghai, China, in 2011. She is currently an Associate Professor with Nanjing University of Aeronautics and Astronautics. Her research interests include non-destructive testing, computer measurement, and control and signal processing.