Lightweight Image Super-Resolution Reconstruction Algorithm Based on Spectral Norm Regularization GAN and ShuffleNet

Xinyu Zhang
School of Art & Design, Sanming University, Sanming, 365004, China
zhangxy@fjsmu.edu.cn

Keywords: improved GAN, lightweight images, super resolution

Received: October 17, 2025

In response to the low feature extraction ability of current image super-resolution models, an image reconstruction algorithm with an improved generative adv.9ersarial network is proposed. On the basis of the image super-resolution algorithm based on the generated adversarial network, the spectral norm and least square relative discriminator are introduced, and then the latest version ShuffleNetV is added to improve the accuracy of the model. A lightweight image super-resolution reconstruction algorithm based on improved generative adversarial network is tested. The test results show that the evaluation scores for WOMAN images, HEAD images, BUTTERFLY images, and BABY images using the research method were 43.6, 33.8, 27.9, and 46.3, respectively. The K values of the reconstructed samples in the Set5 dataset were mainly concentrated in the range of 0.4008.5, while in the Set14 dataset, the K values were roughly distributed in the range of 0.20 to 1.10. In the ablation experiment, the PI value of the research model is 2.11, indicating that the research model can generate high-quality images that are closest to the real high-resolution images in terms of perceptual features and texture details. From this, lightweight image super-resolution reconstruction algorithms with improved generative adversarial networks have significant performance advantages, which can promote technological progress in image super-resolution reconstruction.

Povzetek: Lahki algoritem SRLRGAN-SN z vgrajeno ShuffleNetV, spektralno normo in relativističnim diskriminatorjem izboljša teksturne podrobnosti ter ohrani barvno natančnost, ustvarjajoč bolj kvalitetno rekonstrukcijo kot metoda SRCNN ali VDSR.

1 Introduction

Currently, as an important carrier of information transmission, image resolution has become one of the key factors in measuring image quality [1-2]. However, in practical applications, due to factors such as the performance of image acquisition devices, network transmission bandwidth, and storage conditions, the obtained images often have low resolution and are difficult to meet the requirements of high-quality image processing [3]. Super Resolution Image Reconstruction (SRIR) utilizes computer processing of Low-Resolution Images (LRI) or image sequences to restore HRI, where HRI have high pixel density and can provide more details [4]. The application of Deep Learning (DL) has achieved good results in SRIR, making the results more realistic and clearer [5]. The Yu team proposed a DL based image SRIR algorithm to address the issue of low super-resolution in single images, which had good performance and could improve the super-resolution of single images [6]. Nagayama et al. Suggested SRIR with DL to measure the image quality of coronary computed tomography angiography, which improved the depiction of stent support, with better image quality [7]. The Gao team put forth an innovative approach, termed SRIR, which aims to leverage the contribution details of original characteristics and enhance the model's capacity to represent characteristics effectively. Its performance was good [8]. Generative Adversarial Networks (GANs) are a type of DL that can generate more realistic image details through adversarial training of generators and discriminators. For example, an et al. suggested a GAN with SRIR technique to improve the performance of super-resolution methods in feature extraction and expression. The HRI formed by this technique had the best subjective perception of the human eye [9]. Based on previous research, this study proposes an improved SRIR algorithm, SRLRGAN-SN, by introducing Spectral Regularized Least Squares Relative (SRLR) and Spectral Normalization (SN) into GAN. It is assumed that SRLRGAN-SN has excellent results in detail preservation, sharpness improvement and algorithm efficiency in image reconstruction tasks. By fusing spectral norm and least square method, it is expected that SRLRGAN-SN algorithm can reconstruct image details more accurately, reduce the situation of detail blur and loss, and make the reconstructed image closer to the real high-resolution image. The final improvement goal of the research is to improve the definition of the reconstructed image, reduce the noise and halo artifacts, optimize the algorithm structure, reduce the computational complexity, and improve the operation efficiency of the algorithm, so that it can complete the task of image super-resolution reconstruction more quickly in practical applications. In terms of training parameters, the

study epoch number is set to 200 epochs. The initial learning rate is set to 0.0001, and the learning rate attenuation strategy is adopted in the training process to reduce the learning rate to 0.1 times of the original every 50 epochs; Data enhancement methods such as rotation, flipping and scaling were adopted. The range of random rotation Angle was [-15°,15°], and the range of random horizontal and vertical flipping was [0.8,1.2]. In the aspect of evaluation, we use peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM) to evaluate the reconstructed images objectively. At the same time, comparative experiments were conducted with other advanced image super-resolution reconstruction algorithms GANs, SRGAN, SRCNN and SOTA to verify the superiority and effectiveness of SRLRGAN-SN algorithm through comparative analysis. A summary of the relevant methods is shown in Table 1.

Table 1: Summary of relevant methods

Method	PSNR	SSIM	Computational complexity
GANs	28.4	0.85	34.25
SRGAN	30.7	0.92	48.76
SRCNN	29.6	0.89	12.34
SOTA	32.1	0.95	65.87

As can be seen from Table 1, GANs performs well in image generation tasks because the adversarial training mechanism of its generator and discriminator makes the generated images more visually realistic. However, the images generated by GANs in super-resolution tasks lack details, are prone to artifacts, and the training process is unstable. SRGAN introduces Perceptual Loss and Adversarial Loss, significantly improving the visual quality of the generated images, but the generated images are still not realistic enough in terms of texture and structural details. SRCNN directly learns the mapping relationship from low resolution images to high resolution images through convolutional neural network (CNN). The network structure is relatively simple, but the generated images are not rich in details and textures, and the visual quality is limited. In addition, SRCNN has high computational complexity and long training and reasoning time. SOTA represents the current state of the art in the field of super resolution, performing well on PSNR and SSIM metrics. However, SOTA method still has shortcomings in perception quality, and the generated images may lack realism. SRLRGAN-SN uses the least squares Relativistic Generative Adversarial Network (LSR-GAN) to generate a more realistic image by driving the generated image closer to the high-quality perceptual manifold through the least square's relativistic discriminator. At the same time, SRLRGAN-SN introduces a perception-loss component that can better retain structural texture details, resulting in a more visually realistic image.

2 Methods and materials

2.1 GAN based image super resolution algorithm

SRIR technology increases the resolution of an image or video frame through hardware or software, by obtaining a HRI or video frame through a series of low-resolution frames. The fundamental principles of image SRIR is presented in Figure 1.

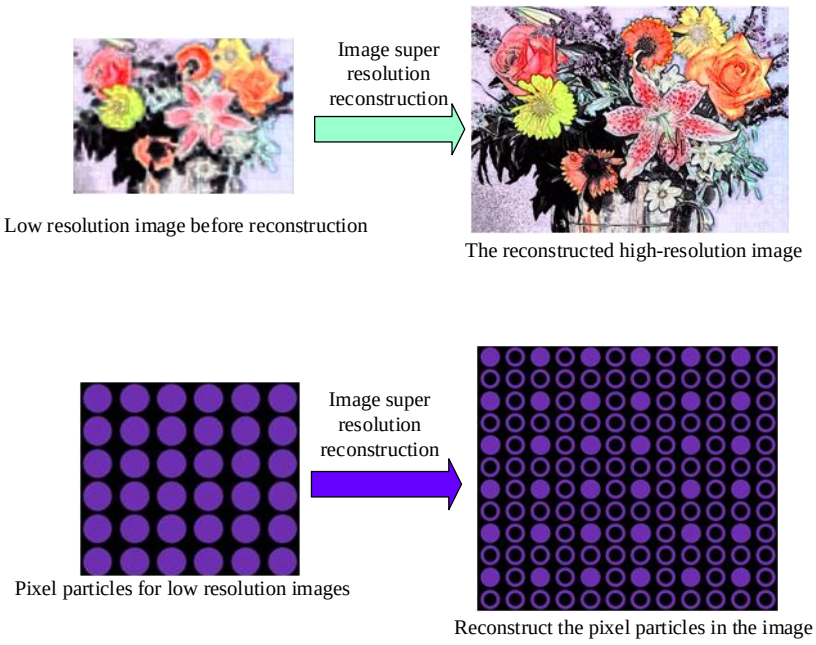


Figure 1: Schematic diagram of image super resolution technology

In Figure 1, SRIR technology utilizes details from LRI to estimate and reconstruct HRI, which can be seen as an enhancement and restoration of image information. Common image reconstruction techniques include Projection onto Convex Sets (POCS) and Maximum A Posteriori Probability (MAP) [10]. POCS is an iterative algorithm applied to handle issues in such as SRIR, signal processing, and data analysis. Its basic idea is to search for solutions in the intersection of a series of convex constraint sets, and gradually approach the true solution through an iterative process. Assuming there are m conditional constraints, equation (1) shows the calculation formula for POCS.

$$H^{(n)} = X^{m},...,X^{2}X^{1}H^{(n-1)} \quad (1)$$

In equation (1), $H^{(n)}$ is the LRI input; X^{k} is the mapping of the input image in the constraint set. MAP is a Bayesian based statistical method applied to assess the mode of a Posterior Distribution (PD). This method combines likelihood function and prior distribution to form a PD, and then finds the maximum point of this PD as the estimated parameter value [11]. Assuming the original image is X, equation (2) shows the optimal estimation expression for the reconstructed HRI Y.

$$Y = \max P(Y|X) = \max(\frac{P(Y|X)P(X)}{P(Y)}) \quad (2)$$

GAN is comprised of two models: a generative model (GM) and a Discriminative Model (DM). The GM is responsible for accurately representing the distribution of the sample data, whereas the DM is designed to ascertain whether the input data is authentic or a generated sample. In the process of optimizing a model, one side is maintained as a fixed reference point throughout the training phase, while the other side updates its parameters in an alternating sequence of iterations. Ultimately, the GM is capable of estimating the distribution of sample data. The GAN principle is shown in Figure 2.

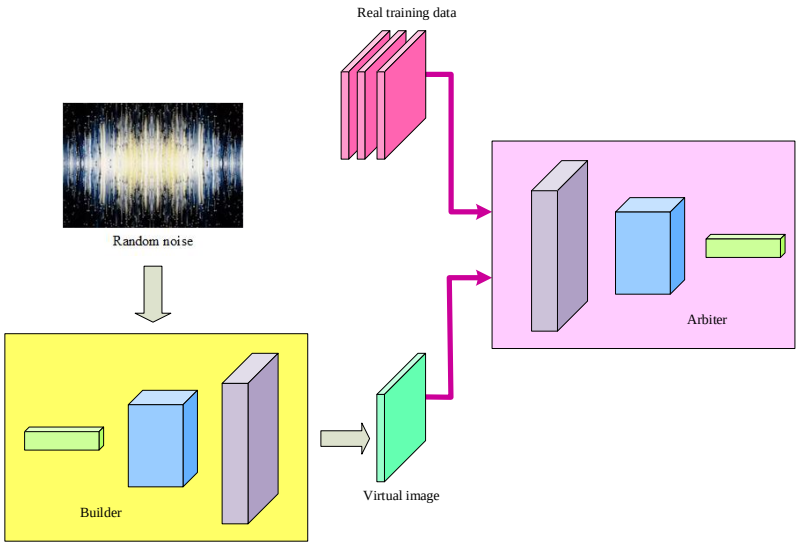


Figure 2: Generate the adversarial network schematic

In Figure 2, generator G and discriminator D are randomly initialized. A batch of real samples are sampled from the real data distribution, and a batch of noisy samples are sampled from the noisy distribution, which are input into G to obtain the generated samples. Real samples and generated samples are input into D separately, and the output of D is calculated. The parameters of D are kept unchanged, while the parameters of G are updated to provide samples that are closer to real data, thereby deceiving D [12-13]. G is repeatedly sampled and trained until it can generate high-quality data, and D cannot accurately distinguish between real data and generated data. Assuming the generative network is g and the DM is d, the image reconstructed by the generative network $G(z)$'s expression is shown in equation (3).

$$G(z) = \arg\min_{g}\max_{d} v(g,d) \quad (3)$$

In equation (3), z represents the signal input to the generative network. The default choice for determining the network's benefits v is shown in equation (4).

$$v(\theta^{(g)};\theta^{(d)}) = E_{X\sim P_{data}} \log d(x) + E_{X\sim P_{data}} \log(1-d(x)) \quad (4)$$

In equation (4), x is the real sample; the probability of the DM judging whether the image is true is represented by the symbol $d(x)$. Super Resolution GAN (SRGAN) is a method that uses convolutional neural networks to achieve Super Resolution (SR) for a single image [14]. The perceptual loss function of SRGAN l^{SR} is expressed as equation (5).

$$l^{SR} = l_X{}^{SR} + 10^{-3} l_{Gen}{}^{SR} \quad (5)$$

In equation (5), $l_X{}^{SR}$ is content loss; $l_{Gen}{}^{SR}$ is adversarial loss. SRGAN uses VGG loss, which includes VGG16Q and VGG19. Among them, VGG19 has 19 layers, including 16 convolutional layers and the last 3 fully connected layers. The middle is similar to usual, using a pooling layer and finally passing through softmax [15]. The breakdown diagram of VGG19 network structure is shown in Figure 3.

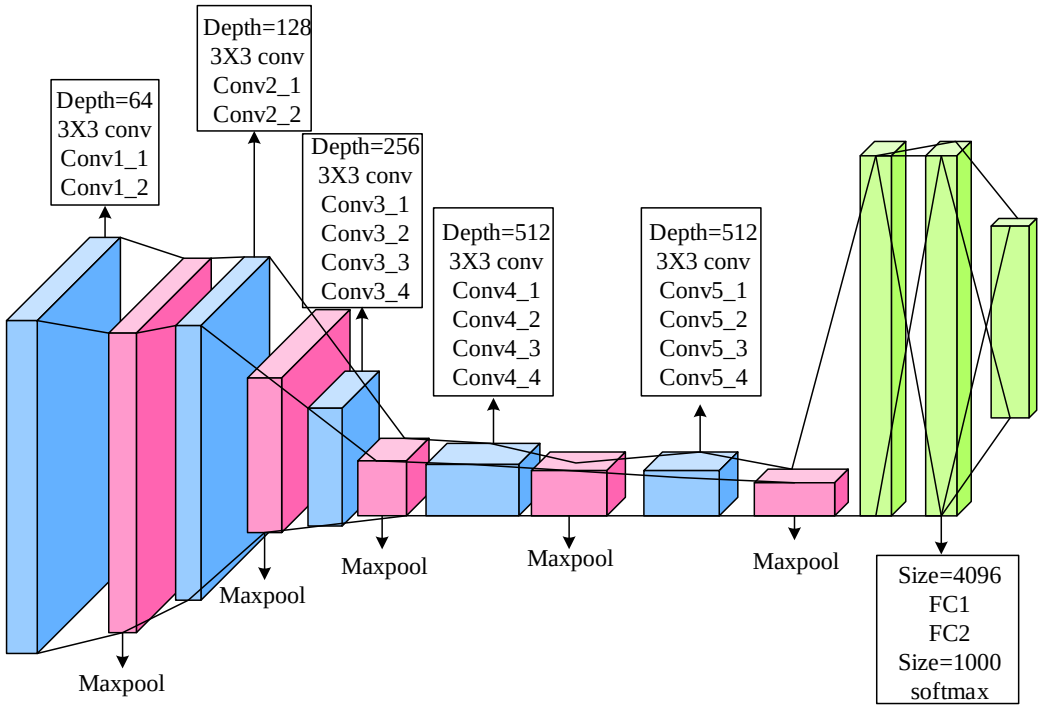


Figure 3: VGG19 Network structure decomposition diagram

In Figure 3, VGG19 consists of multiple convolutional blocks and fully connected layers. The former contains multiple convolutional layers and ReLU activation layers, and after each convolutional block is completed, a max pooling layer is used to decrease the size of the feature map [16]. VGG19 replaces large convolution kernels by stacking multiple 3×3 small convolution kernels, which can increase the depth of the network while maintaining the same receptive field, thereby improving the network's feature extraction ability. Equation (6) shows the expression for the VGG19 loss function.

$$l_{VGG/i,j}^{SR} = \frac{1}{W_{i,j}H_{i,j}}\sum_{x=1}^{W_{i,j}}\sum_{y=1}^{H_{i,j}}(\phi_{i,j}(I^{HR})_{x,y} - \phi_{i,j}(G_{\theta G}(I^{HR}))_{x,y})^2 \quad (6)$$

In equation (6), $W_{i,j}$ is the dimension of one feature map in the model; $H_{i,j}$ is the dimension of another feature map in the model; $\phi_{i,j}$ is the feature map of the j-th convolutional layer in VGG19. For adversarial loss, SRGAN mainly indirectly improves the quality of the network generated images by adding discriminators. The expression for adversarial loss l_{Gen}^{SR} is shown in equation (7).

$$l_{Gen}^{SR} = \sum_{n=1}^{N} -\log D_{\theta D}(D_{\theta G}(I^{LR})) \quad (7)$$

2.2 Lightweight SRIR Method with SRLRGAN-SN

To achieve precise control over the parameters of the generative and DM while ensuring training stability, this study introduces the SN and SRLR relative discriminators based on SRGAN and proposes the SRLRGAN-SN image reconstruction method. Figure 4 illustrates the flowchart.

In Figure 4, SRLRGAN-SN first receives LRI as input and performs necessary preprocessing operations. Then, the LRI is mapped to a high-resolution space through a generative network to generate the corresponding SR image. Then, the SRLR relative discriminator is applied to discriminate the relative difference, and the evaluation of the relative realism is output. Then, content loss, feature loss, texture loss, and SRLR relative discriminative adversarial loss are calculated. Finally, the above process is repeated until the preset training epochs are reached or convergence conditions are met. To simplify the complexity of the model problem, this study uses the SRLR loss function as the criterion for discrimination, transforming the absolute discrimination problem between generated images and real samples into relative difference discrimination. The discriminator loss function based on SRLR is shown in equation (8).

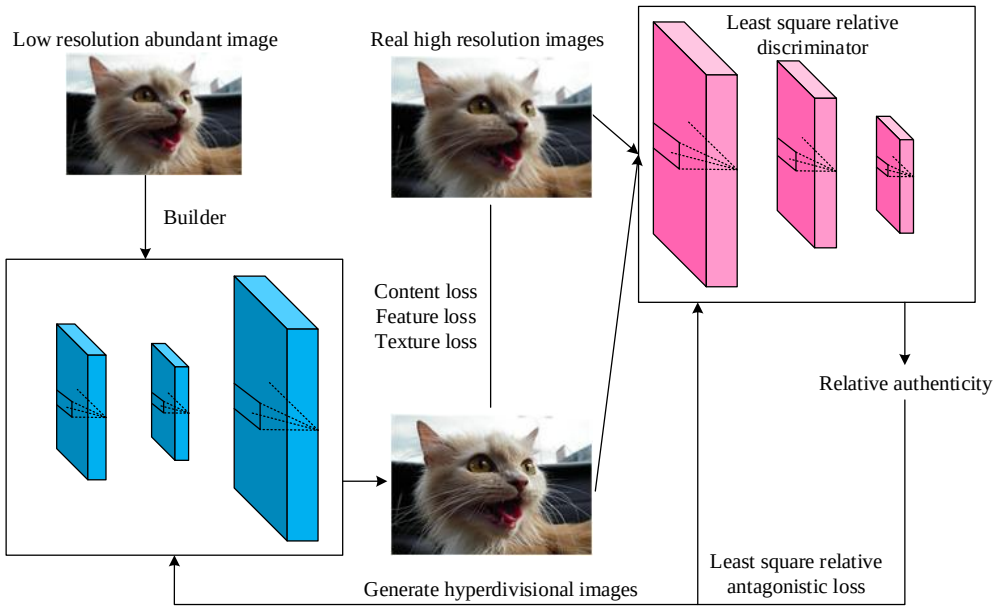


Figure 4: SRLRGAN-SN process

$$\min_D V_{LSGAN}(D) = \frac{1}{2} E_{x\sim Pdata(x)}\left[(D(x)-b)^2\right] + E_{z\sim Pg(z)}\left[(D(G(z))-a)^2\right] \quad (8)$$

In equation (8), a represents the generated sample label; b represents the true sample label; $x \sim p_{data}(x)$ represents the probability distribution of the true sample; $z \sim p_g(z)$ represents the probability distribution that the generated sample follows; $E_{x\sim Pdata(x)}$ represents the expected value of the true sample distribution; $E_{z\sim Pg(z)}$ represents the expected value of the generated sample distribution. The loss function of the generator is shown in equation (9).

$$\min_G V_{LSGAN}(G) = \frac{1}{2} E_{z\sim Pg(z)}\left[(D(G(z))-c)^2\right] \quad (9)$$

In equation (9), c represents the value that the GM wishes to discriminate as true. The calculation formula for the JS divergence optimized by the original GAN is shown in equation (10).

$$JS(P_r \| P_f) = \log 2 + \max\left\{\frac{1}{2} E_{x_r\sim P_r}\left[\log(D((x_r)) + \frac{1}{2} E_{x_f\sim P_f}\left[\log(1-D(x_f))\right]\right\} \quad (10)$$

In equation (10), P_r represents the true distribution; P_f represents the generated distribution; x_r is the true sample; x_f is the generated sample. The expression for the discriminative mechanism of the original GAN is shown in equation (11).

$$D_{GAN}(x) = \sigma(C(x)) \quad (11)$$

In equation (10), C is the output of the DM before activation; σ is the sigmoid activation function. The discriminant mechanism of the original GAN is realized through the discriminator constantly learning and distinguishing between the real data and the generated data, as well as the adversarial training between the discriminator and the generator. This mechanism enables the GAN to generate realistic data samples and show strong ability in various generation tasks. The expression for the discriminative mechanism of the original RGAN is shown in equation (11).

$$D_{RGAN}(x_r, x_f) = \sigma(C(x_r) - C(x_f)) \quad (12)$$

The discriminant mechanism of the original RGAN reduces the randomness of the gradient and does not increase the time complexity of the algorithm by averaging the original output of the discriminator of a set of reference samples and taking the mean value as the reference. The loss function of the RGAN discriminator is shown in equation (13).

$$L_D^{RGAN} = -E_{(x_r, x_f)\sim(P,Q)}\left[\log(\sigma(C(x_r) - (C(x_f)))\right] \quad (13)$$

The loss function of the RGAN generator is shown in

equation (14).

$$L_G^{RGAN} = -E_{(x_r, x_f)\sim(P,Q)}\left[\log(\sigma(C(x_f) - (C(x_r)))\right] \quad (14)$$

The relative discriminator considers both real and generated samples when evaluating samples, and makes decisions by comparing their relative authenticity [17]. The relative discriminator not only compensates for the shortcomings of the original GAN in training stability and generating sample diversity, but also promotes a more balanced and dynamic game between G and D, which may achieve higher quality generation results. The DM structure of SRLRGAN-SN is shown in Figure 5.

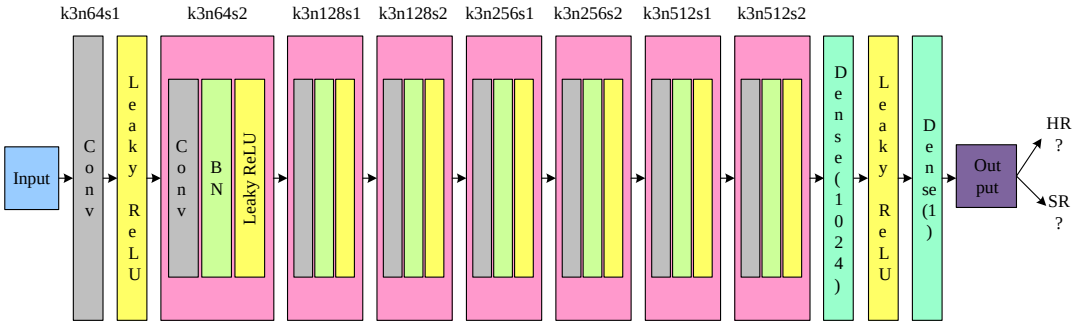


Figure 5: The discriminant network structure of SRLRGAN-SN

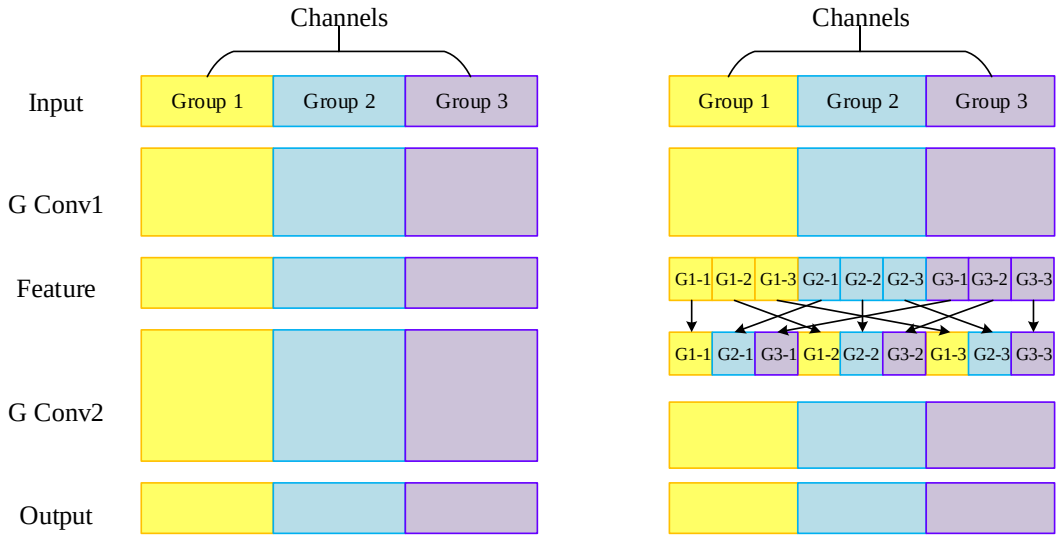
In Figure 5, the front-end of the DM consists of 8 convolutional layers, each of which extracts features at different levels from the input image. The features are gradually abstracted from low to high levels. After each convolutional layer, an SN layer is introduced to control the SN of the network weights, thereby satisfying the Lipschitz constraint. The SN layer forms a soft constraint by limiting the maximum singular value of the weight matrix to ensure the Lipschitz continuity of the network. Lipschitz constraint can ensure that the output change of the network does not exceed a certain multiple of the input change, which helps to stabilize the training process and prevent gradient explosion and gradient disappearance.

All activation layers use the Leaky ReLU (Leaky Rectified Linear Unit) activation function. After the convolutional layer and SN layer, the DM connects two linear dimensionality reduction layers. The final output layer of the discrimination network adopts the SRLR relative discrimination mechanism, which makes discrimination by calculating the relative degree of realism. The SRLR relative discrimination output layer calculates the distance in the feature space, and uses it as the discrimination basis to generate high-quality images.

Assuming the output of layer l in the network is x^l, its expression is shown in equation (15).

$$x^l = f^l(W^l - x^{l-1} + b^l) \quad (15)$$

In equation (15), x^{l-1} represents the input of layer l in the network; W^l represents the SN of layer l in the network; b^l represents the bias of layer l in the network. ShuffleNet is a lightweight network, and its latest version ShuffleNetV uses point by point group convolution and channel shuffling to reduce computational complexity while improving model accuracy. Therefore, this study introduces ShuffleNetV into SRLRGAN-SN. The schematic diagram of ShuffleNetV's grouped convolution and channel shuffling is shown in Figure 6.



(a) Block convolution diagram (b) Channel mixing diagram

Figure 6: Schematic diagram of packet convolution and channel mixing

In Figure 6 , in grouped convolution, the input channel is divided into several groups, and the feature maps of each group are only convolved with the corresponding convolution kernels, significantly reducing computational complexity and model parameter count. In Figure 6 (b), channel shuffling shuffles and reallocates feature map channels from different groups, allowing each group to receive information from all original input channels in subsequent operations. ShuffleNetV achieves a dual improvement in model efficiency and performance by introducing grouped convolution and channel shuffling techniques.

3 Results

3.1 Performance Analysis of SRLRGAN-SN

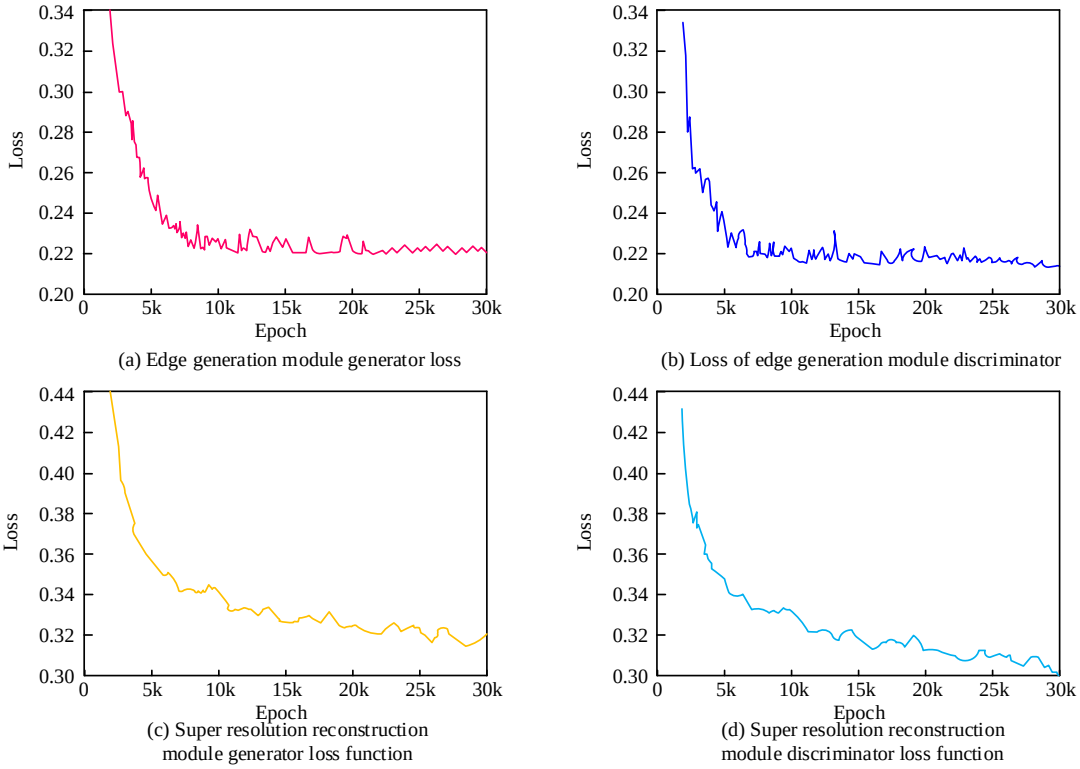
To test the effectiveness of the reconstruction model, this study conducted simulation experiments. The first stage was to train the edge generator. In the second stage, the trained edge generator was used as a guiding signal to drive the training process of the SR reconstruction module. At each training stage, the Adam optimization algorithm was used to efficiently and stably adjust the model parameters. The variation curves of the edge generation module and super-resolution reconstruction loss function are shown in Figure 7.

In Figure 7 (a), as the iterations increased, a steady trend was observed when the loss value approached 0.22. In Figure 7 (b), as the iterations were 25000, the discriminator loss curve of the edge generation module tended to stabilize. In Figure 7 (c), the loss function curve of the SR reconstruction module generator tended to stabilize at 0.32. In Figure 7 (d), as the iterations increased, the loss curve gradually decreased and stabilized when the loss value approached 0.30, indicating that the discriminator's ability to distinguish between real HRI and generated HRI had stabilized. To verify the effectiveness of SRLRGAN-SN in the reconstruction process, BRISQUE was used as an evaluation metric, and VDSR and SRCNN were introduced as comparative algorithms for testing on the Set5 dataset. Figure 8 shows the experimental results.

In Table 2, this study implemented the SRLRGAN-SN algorithm on the computer Python programming language.

Table 2: Experimental environment setting table

Simulation environment	Parameter value
Operating system	Ubuntu 20.04 LTS
Python	Python 3.8
CPU	Intel(R)Core(TM) i7-9700k@3.60GHz
GPU	NVIDIA GeForce RTX 3090
DL framework	TensorFlow 2.x
Internal memory	32G



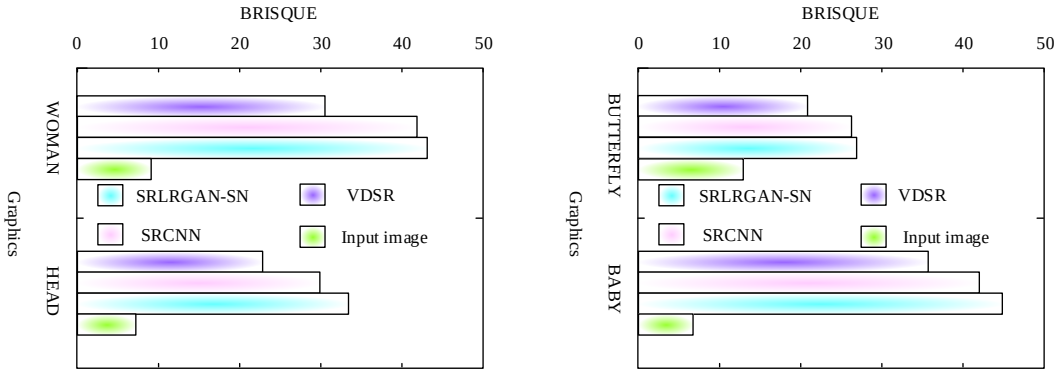
(a) Edge generation module generator loss

(b) Loss of edge generation module discriminator

(c) Super resolution reconstruction module generator loss function

(d) Super resolution reconstruction module discriminator loss function

Figure 7: Edge Generation module and superresolution reconstruction loss function change curve



(a) Economic cost objective function diagram

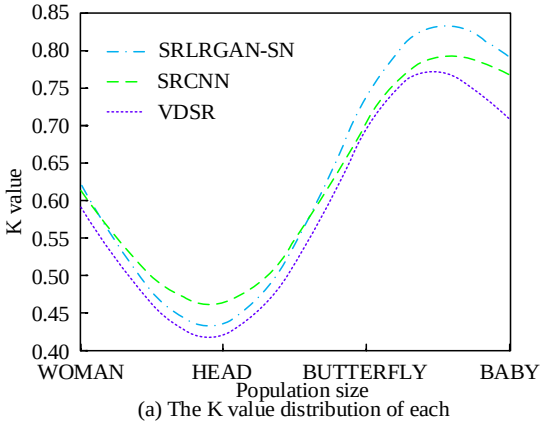
(b) Economic cost objective function diagram

Figure 8: BRISQUE evaluation index of three algorithms on Set5

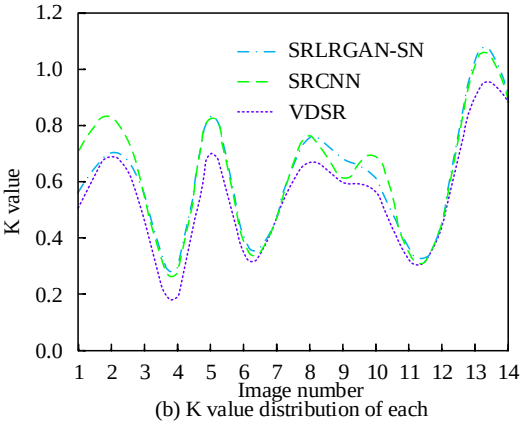
In Figure 8 (a), SRLRGAN-SN scored 43.6 for WOMAN images and 33.8 for HEAD images. In Figure 8 (b), on the BUTTERFLY image, the evaluation score of VDSR was the lowest at 21.2, followed by SRCNN algorithm at 26.5, and SRLRGAN-SN with a score of 27.9 was slightly higher than SRCNN, indicating its advantage in processing detail rich images. For BABY images, the performance of each algorithm had been improved, with VDSR score of 35.7, SRCNN score of 43.2, and SRLRGAN-SN score of 46.3. From this, SRLRGAN-SN performed excellently in the evaluation of various image types, with performance far exceeding traditional VDSR and SRCNN. This study conducted color cast factor analysis using three algorithms, VDSR, SRCNN, and SRLRGAN-SN, with Set5 and Set14 as reference datasets at a magnification of 2. For the images in the Set14 dataset, this study numbered them alphabetically one by one. Figure 9 shows the distribution of K values for each reconstructed image in different datasets.

In Figure 9 (a), the K values of the reconstructed samples within Set5 were mainly concentrated in the interval of 0.4008.5, indicating the similarity level between the reconstructed image and the original image in terms of color reproduction. In Figure 9 (b), the K values in Set14 were roughly distributed in the range of 0.20 to 1.10. Among them, SRLRGAN-SN had a lower overall color distortion level on both Set5 and Set14, indicating that it had better performance in color preservation and reconstruction quality compared to other algorithms. In summary, the K-value distribution of the reconstructed samples in Set5 and Set14

was within an acceptable color shift threshold range. In order to ensure the robustness of SRLRGAN-SN algorithm, the above three algorithms are compared with the current popular SOTA algorithm on Set5 and Set14, DIV2K and Urban100 data sets. Quantitative indicators PSNR and SSIM were used to evaluate image quality, and FLOPs was used to analyze computational efficiency. The experimental results are shown in Table 3.



(a) The K value distribution of each reconstructed image in the Set5 dataset



(b) K value distribution of each reconstructed image in Set14 dataset

Figure 9: K Value distribution of reconstructed images in different data sets

Table 3: Comparison of results of different algorithms

Algorithm	Set5			Set14			DIV2K			Urban100		
	PSNR	SSIM	FLOPs	PSNR	SSIM	FLOPs	PSNR	SSIM	FLOPs	PSNR	SSIM	FLOPs
VDSR	30.14	0.92	66543.25	28.32	0.86	132456.78	26.78	0.81	264913.50	24.56	0.76	529827.00
SRCNN	29.56	0.89	5321.34	27.98	0.85	10542.12	26.34	0.79	10684.28	24.12	0.74	21084.56
SOTA	31.23	0.91	123456.89	29.45	0.87	246910.56	27.12	0.82	493827.56	25.03	0.78	987654.32
SRLRGAN-SN	31.50	0.92	234567.11	29.72	0.88	469134.23	27.30	0.83	938271.43	25.28	0.79	1938271.65

As can be seen from Table 3, the PSNR of VDSR algorithm on Set5 test set is 30.14dB, SSIM is 0.901 and FLOPs is 66543.25 times, while the performance on Set14 test set is slightly lower. The performance of SRCNN algorithm is lower than that of VDSR on both test sets. As the current optimal algorithm, SOTA algorithm outperforms the other three algorithms on the two test sets, has the highest PSNR and SSIM values, but also has the largest computational load. Although SRLRGAN-SN is computatively more expensive than SOTA on Set5 and Set14 test sets, PSNR and SSIM are higher than SOTA, so they can produce more natural images.

3.2 The ablation experiment of SRLRGAN-SN

To scientifically evaluate the advantages of each combination module in the SRLRGAN-SN method in SRIR performance, this study compared the model performance under different module combinations to quantify the independent and collaborative effects of each module in improving image reconstruction quality and optimizing visual perception characteristics. Table 4 shows the ablation experiment setup.

In Table 4, the model with Setting 1 performed the worst in all indicators, indicating the limitations of basic training configuration on model performance. When feature loss was introduced in Setting 2, the model demonstrated a notable enhancement in performance, with the PI value dropping to 3.25. This improvement was significant, indicating the effectiveness of feature loss in optimizing the output quality of the model. Setting 3 reduced the PI index to 3.28 by incorporating texture loss, indicating that the introduction of texture loss effectively improved the visual perception quality of the generated image. In Setting 4 and Setting 5, the strategy of SRLR relative adversarial loss demonstrated superior performance in improving visual perception quality compared to traditional relative adversarial loss. Setting 6 had the highest performance metrics among all Setting methods, indicating that the research approach could generate high-quality images. Figure 10 shows the texture detail comparison image of the ablation experiment reconstruction.

Table 4: Analysis and comparison of ablation experiments

Settings	Content loss	Feature loss	Texture loss	Relative counteraction loss	Least square relative antagonistic loss	SN	PI	P	LPIPS
(a)	○	×	×	×	×	×	5.36	>0.05	0.25
(b)	○	○	×	×	×	×	3.25	<0.01	0.18
(c)	○	○	○	×	×	×	3.18	<0.05	0.15
(d)	○	○	○	○	×	×	2.46	<0.05	0.12
(e)	○	○	○	×	○	×	2.24	<0.05	0.10
(f)	○	○	○	×	○	○	2.11	<0.05	0.08

Figure 10: Comparison of texture details reconstructed through ablation experiments

In Figure (10), the texture visual effect of the image generated by Setting 1 was the blurriest, indicating that only basic training had not fully optimized the texture features. Setting 2 introduced a feature loss term, which improved the visual effect of the image to a certain extent, but the texture details of the image were still not clear enough. Setting 3 introduces texture loss, which enhanced the texture detail processing effect of the image. However, the perceptual visual effect of the image was poor, indicating that although the introduction of a single texture loss helped with detail restoration, it was not sufficient to comprehensively optimize the perceptual performance of the image. Setting 4 adopted GAN adversarial loss, which significantly improved the perceptual quality of the image. The overall perception of the image was more natural and detailed, indicating that GAN adversarial loss had significant advantages in improving the realism and detail expression of image generation. Setting 5 adopted the SRLR relative adversarial loss proposed in the study to replace GAN adversarial loss, which further improved the perceptual visual effect of the image, indicating that the proposed SRLR relative adversarial loss was more effective in optimizing the quality of image generation. Based on Setting 5, Setting 6 combined SN regularization to generate images with extremely fine texture and highly similar perceptual features to real HRI. This further validated the positive role of each combination module in improving the quality of SR reconstruction image generation in the research method. For example, lightweight image super-resolution reconstruction methods based on multi-frequency features and texture enhancement, or lightweight image super-resolution reconstruction methods based on interleaved group convolution and sparse global attention, can be explored to solve the performance bottleneck of current algorithms when processing large-scale data sets.

4 Discussion and conclusion

4.1 Discussion

SRIR technology has always been a highly valued and focused hot topic in computer vision, with significant research value in healthcare, remote sensing, and surveillance. In recent years, with the rise of DL, significant

progress has been made in GAN based SRIR. To this end, research has improved GAN and proposed an image reconstruction algorithm with SRLRGAN-SN algorithm.

Firstly, in the SR reconstruction module, the super resolution reconstruction module, the generator loss curve tends to be stable at 0.32, and the discriminator loss curve tends to be stable at about 0.30, indicating that the discriminator has a strong ability in distinguishing real and generated high-resolution images. This result was consistent with the conclusion of scholar Almansour in the study of SR reconstruction of upper abdominal magnetic resonance imaging based on DL [18]. Secondly, the experimental results of VDSR, SRCNN, and SRLRGAN-SN algorithms on the Set5 dataset were analyzed using the BRISQUE evaluation index. It was found that SRLRGAN-SN had better evaluation scores than the other two algorithms on various image types. This result was consistent with the conclusion of the Orii team in the study of coronary artery computed tomography angiography SRDL reconstruction to evaluate the lumen of coronary arteries and stents [19]. Then, the above algorithm was reconstructed on Set5 and Set14 datasets, where SRLRGAN-SN's K value fell between 0.43 and 0.85 on Set5 and between 0.28 and 1.10 on Set14. Compared with the current best technology SOTA, SRLRGAN-SN is not only more accurate in color preservation, but also exceeds the SOTA method in image detail reconstruction and overall visual effect. This result further proves the leading position of SRLRGAN-SN algorithm in the field of image reconstruction, and provides strong support for future related research and practice. Finally, the study employed ablation experiments to further reveal the impact of different settings on model performance. The basic training configuration set Group a only adds content loss, and the resulting images may not be rich enough in detail and texture and visually realistic enough. Set group b to introduce feature loss on top of a, while feature loss can improve detail, it cannot generate complex textures and structures. Set group c introduces texture loss, which produces more realistic texture effects, but artifacts still exist when dealing with complex textures. Set group d introduced RaGAN counterloss, which further improved the realism of the generated image and reduced artifacts, but it was unstable in the training process. Setting group e introduces the least squares relative adversarial loss, which combines the least squares loss and relativistic adversarial loss to train the generated adversarial network more stably. The least squares relative adversarial loss function does not apply to all types of data sets, and may not perform as well as other loss functions on some specific data sets. Setting group f combined with spectral norm regularization to further stabilize the training process, reduce the generation of artifacts, and generate images that are highly close to real high-resolution images in texture details and perceptual features. Elshafai et al. also proposed similar conclusions in their review of single image super-resolution models for medical images with DL [20].

In summary, SRIR with SRLRGAN-SN has shown excellent performance in image quality improvement, color preservation, and computational efficiency, and has broad application prospects and enormous research value.

4.2 Conclusion

This study proposes an SRLRGAN-SN based on GAN, which effectively improves the quality of SRIR by introducing multiple loss functions and optimization strategies. The collaborative work of its edge generation module and SR reconstruction module enables the algorithm to restore rich texture details while maintaining image edge clarity. In terms of reconstruction quality and color preservation, SRLRGAN-SN had shown significant advantages compared to traditional VDSR and SRCNN algorithms. By comparing the BRISQUE evaluation indicators, SRLRGAN-SN could maintain high evaluation scores when processing different types of images, especially when processing images with rich details. In addition, by analyzing the K-value distribution of the reconstructed images in Set5 and Set14, the superiority of SRLRGAN-SN in color preservation and reconstruction quality was further verified. Its lower overall color distortion level indicated that the algorithm had high accuracy in color reproduction. Finally, the positive effect of SRLRGAN-SN algorithm on improving image generation quality was verified through ablation experiments. The limitation of this study is that the running time of the algorithm may become one of the limiting factors for its practical application when processing large-scale datasets or HRI. Therefore, future research can focus on lightweight design and acceleration techniques for algorithms to improve their computational efficiency and practicality.

References

[1] Y. Mengbei, W. Hongjuan, L. Mengyang (2021) Overview of Research on Image Super-Resolution Reconstruction, Proceedings of the IEEE International Conference on Information Communication and Software Engineering, IEEE, Location, pp. 131–135. https://doi.org/10.1109/ICICSE52190.2021.9404113

[2] X. Li, Z. Chen (2022) Single image super-resolution reconstruction based on fusion of internal and external features, Multimedia Tools and Applications, Springer, Location, pp. 1589–1605. https://doi.org/10.1007/s11042-021-11400-2

[3] D. Qiu, L. Zheng, J. Zhu (2021) Multiple improved residual networks for medical image super-resolution, Future Generation Computer Systems, Elsevier, Location, pp. 200–208. https://doi.org/10.1016/j.future.2020.11.001

[4] S. Wang, C. Qiao, A. Jiang (2021) Instant multicolor super-resolution microscopy, Biophysics Reports, Springer, Location, pp. 304–312.

[5] Sunkari, Venkateswarlu, Srinagesh, A. (2024) DeepExplain: Enhancing DeepFake Detection Through Transparent and Explainable AI model, Informatica (Slovenia), Slovenian Society Informatika,

Ljubljana, pp. 103–110. https://doi.org/10.31449/inf.v48i8.5792

[6] L. An, F. Dai, L. An (2021) Image Super-Resolution Reconstruction Technology Based on Deep Learning, Journal of Advances in Artificial Life Robotics, Springer, Location, pp. 163–167. https://doi.org/10.57417/jaalr.1.4_163

[7] M. Gao, W. Song, G. Zhang (2023) CT image super-resolution reconstruction via Pixel-Attention Feedback Network, International Journal of Biomedical Engineering and Technology, Inderscience, Location, pp. 21–33. https://doi.org/10.1504/IJBET.2023.131697

[8] Y. Nagayama, T. Emoto, H. Hayashi (2023) Coronary Stent Evaluation by CTA: Image Quality Comparison Between Super-Resolution Deep-Learning Reconstruction and Other Reconstruction Algorithms, American Journal of Roentgenology, American Roentgen Ray Society, Location, pp. 599–610. https://doi.org/10.2214/AJR.23.29506

[9] Li, Min (2024) Application of GAN-Based Data Encryption Technology in Computer Communication System, Informatica (Slovenia), Slovenian Society Informatika, Ljubljana, pp. 17–34. https://doi.org/10.31449/inf.v48i15.6390

[10] S. Pal, A. Roy, P. Shivakumara, U. Pal (2023) Adapting a Swin Transformer for License Plate Number and Text Detection in Drone Images, Artificial Intelligence and Applications, Springer, Location, pp. 145–154. https://doi.org/10.47852/bonviewaia3202549

[11] Y. Wang, Y. Luo, J. Ren (2022) Research on application of spatial attention mechanism in the super-resolution reconstruction of single-channel greyscale image of mouse brain, International Journal of Wireless and Mobile Computing, Inderscience, Location, pp. 259–264. https://doi.org/10.1504/ijwmc.2022.10049459

[12] L. Ma (2021) Face Recognition of Intelligent Building based on Super-Resolution Reconstruction of Visual Imag, Proceedings of the International Conference on Intelligent Computing and Control Systems, IEEE, Location, pp. 895–900. https://doi.org/10.1109/iciccs51141.2021.9432282

[13] W. Xin, Z. Shufen, L. Yuanyuan (2023) Pixel attention convolutional network for image super-resolution, Neural Computing & Applications, Springer, Location, pp. 8589–8599. https://doi.org/10.1007/s00521-022-08132-1

[14] C. Wilpert, C. Neubauer, A. Rau, T. W. HannahBenkert (2023) Accelerated Diffusion-Weighted Imaging in 3 T Breast MRI Using a Deep Learning Reconstruction Algorithm With Superresolution Processing: A Prospective Comparative Study, Investigative Radiology, Lippincott Williams & Wilkins, Location, pp. 842–852. https://doi.org/10.1097/rli.0000000000000997

[15] Yan, Xu (2024) A Face Recognition Method for Sports Video Based on Feature Fusion and Residual Recurrent Neural Network, Informatica (Slovenia), Slovenian Society Informatika, Ljubljana, pp. 137–152. https://doi.org/10.31449/inf.v48i12.5968

[16] B. Xie, F. Niu (2021) Super-resolution reconstruction algorithm for aerial image data management based on deep learning, Distributed and Parallel Databases, Springer, Location, pp. 77–94. https://doi.org/10.1007/s10619-021-07356-9

[17] X. Yang, D. Zhang, Z. Wang (2022) Super-resolution reconstruction of terahertz images based on a deep-learning network with a residual channel attention mechanism, Applied Optics, Optical Society of America, Location, pp. 3363–3370. https://doi.org/10.1364/AO.452511

[18] H. Almansour (2021) Deep Learning-Based Superresolution Reconstruction for Upper Abdominal Magnetic Resonance Imaging An Analysis of Image Quality, Diagnostic Confidence, and Lesion Conspicuity, Investigative Radiology, Lippincott Williams & Wilkins, Location, pp. 509–516. https://doi.org/10.1097/rli.0000000000000769

[19] M. Orii, M. Sone, T. Osaki (2023) Super-resolution deep learning reconstruction at coronary computed tomography angiography to evaluate the coronary arteries and in-stent lumen: an initial experience, BMC Medical Imaging, BioMed Central, Location, pp. 1–10. https://doi.org/10.1186/s12880-023-01139-7

[20] W. Elshafai, A. M. Ali, S. A. Elnabi (2024) Single image super-resolution approaches in medical images based deep learning: a survey, Multimedia Tools and Applications, Springer, Location, pp. 30467–30503. https://doi.org/10.1007/s11042-023-16197-w