Better Visual Image Super-Resolution with Laplacian Pyramid of Generative Adversarial Networks

Ming Zhao[1], Xinhong Liu[1], Xin Yao[1, *] and Kun He[2]

Abstract: Although there has been a great breakthrough in the accuracy and speed of super-resolution (SR) reconstruction of a single image by using a convolutional neural network, an important problem remains unresolved: how to restore finer texture details during image super-resolution reconstruction? This paper proposes an Enhanced Laplacian Pyramid Generative Adversarial Network (ELSRGAN), based on the Laplacian pyramid to capture the high-frequency details of the image. By combining Laplacian pyramids and generative adversarial networks, progressive reconstruction of super-resolution images can be made, making model applications more flexible. In order to solve the problem of gradient disappearance, we introduce the Residual-in-Residual Dense Block (RRDB) as the basic network unit. Network capacity benefits more from dense connections, is able to capture more visual features with better reconstruction effects, and removes BN layers to increase calculation speed and reduce calculation complexity. In addition, a loss of content driven by perceived similarity is used instead of content loss driven by spatial similarity, thereby enhancing the visual effect of the super-resolution image, making it more consistent with human visual perception. Extensive qualitative and quantitative evaluation of the baseline datasets shows that the proposed algorithm has higher mean-sort-score (MSS) than any state-of-the-art method and has better visual perception.

Keywords: Single image super-resolution, generative adversarial networks, Laplacian pyramid.

1 Introduction

Image super-resolution refers to reconstructing one or more low-resolution (LR) images into high-resolution (HR) images [Harris (1964); Goodman and Gustafson (1968)]. Super-resolution is an important research problem in the field of computer vision and image processing, but there are still many shortcomings in high-resolution super-resolution reconstruction of low-resolution images. First, since the existing methods are optimized by algorithms that reduce the mean square error (MSE), the picture is too smooth [Wang, Simoncelli and Bovik (2003); Wang, Bovik, Sheikh et al. (2004); Gupta,

---

[1] School of Computer Science and Engineering, Central South University, Changsha, 410000, China.

[2] Apple Inc., California, 95014, USA.

[*] Corresponding Author: Xin Yao. Email: xinyao@csu.edu.cn.

Strivastava, Bhardwaj et al. (2011)]. Images with the highest peak signal-to-noise ratio (PSNR) and MSE do not necessarily have a better visual effect in perception. Second, the existing methods first up sample the image and then reconstruct the SR image. It will cause the model to not effectively learn the corresponding mapping function when the reconstruction multiple is too large (such as 8×). Finally, the residual block (RB) reconstruction results currently utilized are not good, and because of the use of the batch normalization (BN) layer, it will lead to greater model computational complexity.

To solve these problems, we proposed ELSRGAN for image super-resolution. We use RRDB as the basic unit of the network to make the model easier to train and capture more high-frequency details. We use the advanced mapping features of the VGG network to define the loss function [Simonyan and Zisserman (2015); Joshson, Alahi and Li (2016)] to prevent the image from being too smooth. At the same time, we combined the improved generative adversarial network (GAN) model and Laplacian pyramid to capture more high-frequency texture features. Our main work are:

1) This paper first proposes a combination of generative adversarial network and Laplacian pyramid to achieve SR of the image. By introducing the Laplacian pyramid, more high-frequency details of the image can be captured and the SR image can be reconstructed step by step. This makes our models more flexible, and our high-magnification model can perform low-magnification image reconstruction work alone, which is not possible with other methods.

2) This paper modified the GAN network, modified the basic network element from the regular residual block to the RRDB. And we removed the BN layer, which makes our model easier to train and can achieve higher reconstruction quality. A VGG-based perceptual function is used to reconstruct an image with better visual sense and better high-frequency texture features.

3) The experimental results show that the proposed ELSRGAN is better than the state-of-the-art method such as (LapSRN, DRCN and DRRN) after the MSS test.

2 Related work

2.1 Image super-resolution

Harris and Goodman proposed the concept of super-resolution. Early super-resolution problems were solved by interpolation-based methods. A general interpolation model was proposed by Rajan et al. [Rajan and Chaudhri (2001)], which divides into three steps: decomposition, interpolation and fusion. Yang et al. [Yang and Yang (2013)] implemented a statistical-based super-resolution reconstruction method to divide the feature space into multiple sub-spaces, and learn the multiple sub-spaces through the sample to create an effective mapping function. These two methods can quickly reconstruct images with super-resolution, but cannot reconstruct high-magnification images, and the generated image texture is too smooth.

Subsequently, the super-resolution reconstruction method based on learning was used extensively. By using a large number of LR images and HR images for comparison, inferring the missing high frequency information, and finally completing the super-resolution reconstruction. Ni et al. [Ni and Nguyen (2017)] utilized the support vector

regression method to learn the mapping relationship between LR and HR images, and learned the regression amount through the structural characteristics of the DCT domain to obtain high-resolution images. Chang et al. [Chang, Yeung and Xiong (2004)] proposed an image SR method based on local linear embedding. By finding the K nearest neighbors in the training library for the input low-resolution image blocks, the high-resolution reconstruction of the nearest neighbor feature blocks is finally performed using the linear combination. Resolution image block. Singh et al. [Singh and Ahuja (2014)] decomposed the mapping patch into directional frequency subbands and independently matched each subband pyramid to obtain better results. In order to expand the size of the training set, Huang et al. [Huang, Sigh and Ahuja (2015)] made small changes and shape changes to the image. Kim et al. [Kim and Kwon (2010)] used kernel ridge regression to estimate the high frequency information of high-resolution images. In order to reduce the practical complexity of the method, a sparse kernel ridge regression model was proposed based on matching tracking and gradient descent methods. Compared to, it is more general. Yue et al. [Yue, Sun, Yang et al. (2013)] proposed to retrieve similar HR images from the network and propose perceptual matching criteria for alignment structures. However, this method is limited by the training set size, and the number of mappings between LR and HR images may not be sufficient to cover texture changes in the image, and the counterparts requiring HR are known.

Dong et al. [Dong and Loy (2015)] first used deep learning [Hinton and Salakhutdinov (2006)] to solve the super-resolution problem. Wang et al. [Wang, He, Sun et al. (2019)] used convolutional neural networks to remove image noise. Shocher et al. [Schocher, Cohen and Irani (2018)] proposed an unsupervised SR method based on CNN by using the repetitiveness of image internal information. In order to solve the problem that LR degradation is not downsampled from HR bicubic, Zhang et al. [Zhang, Zuo and Zhang (2018)] proposes a general framework with dimensionality stretching strategy. A discriminator network attached to the feature domain is used by Park et al. [Perk, Son, Cho et al. (2018)] to generate high frequency features. Kim et al. [Kim, Lee and Lee (2016)] proposed a network with multiple recursive layers, proposing a high performance architecture that allows long distance pixels Dependencies also maintain a small number of parameter models. These super-resolution methods all learn the mapping function through the MSE loss. Although it can effectively obtain high PSNR and SSIM, it will cause the output image to be too smooth.

In order to prevent smooth output, a method for extracting the high-level feature-aware loss of the trainer is proposed by Bruna et al. [Bruna, Sprechmann and Lecun (2015)] and Johnson et al. [Johnson, Alahi and Li (2016)]. Huang et al. [Huang, He, Sun et al. (2019)] proposed a method for generating wavelet domain, which will be low, the resolution of the face is enlarged to a higher multiple. To create more realistic texture details, Sajjad et al. [Sajjadi, Scholkopf and Hirsch (2017)] proposed a texture-based matching loss function. Lai et al. [Lai, Huang, Ahuja et al. (2018)] proposed a new loss function Charbonnier to correct the difference between the SR image and the HR image. Ledig et al. [Ledig, Theis and Huszar (2016)] utilized the generated confrontation network to perform super-resolution reconstruction by optimizing the damage and content loss. We refer to this form of network architecture and modify its basic network elements. And Ledig the difference is that SRGAN focuses on 4×, our method has higher magnification,

such as 8× to 16×. And Ledig focuses on learning the mapping function between low frequency image and high frequency image, and our network architecture is focused on capturing the exact high frequency details.

2.2 Laplacian pyramid

Burt et al. [Burt and Adelson (1983)] proposed the concept of Laplacian pyramid. The Laplacian pyramid interprets the image with multiple resolutions and restores the image to the greatest extent by residual prediction. Denton et al. [Denton, Chintala and Fergus (2015)] proposed a LAPGAN (Laplacian Pyramid GAN) is used to produce real images. Recently, Lai et al. [Lai, Huang, Ahuja et al. (2018)] proposed LapSRN for image super resolution. Our method has the following differences compared to LapSRN:

1) LapSRN used a deep convolution network to combine with the Laplacian pyramid. We use the GAN network to combine with the Laplacian pyramid. By combining the GAN network, we can enhance super-resolution image quality, have better results, and recover more high-frequency details.

2) LapSRN implemented feature sharing to ensure higher inference speed, but does not focus on recovering high-frequency texture details of the image. This article focuses on recovering the high-frequency contour information of the image to generate a picture more consistent with human perception. We used LR images to train step by step to capture the missing high-frequency features at different magnifications.

3) LapSRN utilized the charbonnier loss function to better capture the difference between the SR image and the HR image. While we focus on restoring images that match the human perception and capture higher high frequency details, we used the VGG loss function.

2.3 Generative adversarial network

In 2014, Goodfellow et al. [Goodfellow, Pouget-Abadie and Mirza (2014)] proposed Generative Adversarial Network that utilized the generator network G to generate as realistic a sample as possible, using the discriminator network D to determine whether the image is a generated image or a real HR image, both iterative confrontation complete training. Ledig et al. [Ledig, Theis and Huszar (2016)] proposed to solve the problem that the discriminator in the network min-max is easy to discriminate by optimizing the loss function. A method for super-resolution using wavelets was proposed by Huang et al. [Huang, Li, He et al. (2018)].

3 Laplacian pyramid generative adversarial network for super-resolution

3.1 Network architecture

Our network works as shown in Fig. 1. By inputting the LR image to the GAN network, a residual image is generated. The upsampled image of the LR and the residual image are added pixel by pixel to recover the desired SR image. For 4× and 8× images, we need to continue to call the ELSRGAN network for super-resolution on the resulting SR results until the magnification requirement is met.
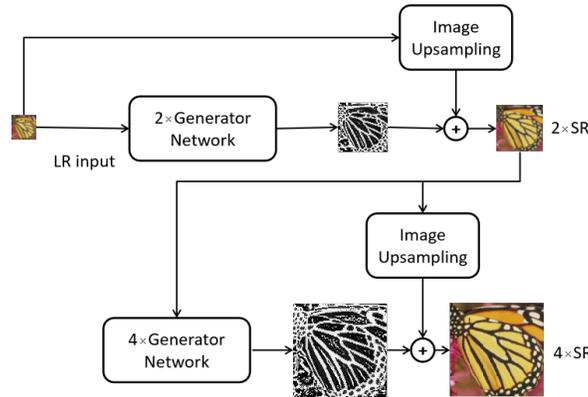
Figure 1: Super-resolution reconstruction process. This figure is a schematic diagram of 4× resolution amplification. For the case of 4× amplification, we call the GAN network twice to generate the final SR image

In the process of generator network training, we used the combination of Laplacian pyramid and GAN. This paper is different from the ordinary GAN network. The ordinary GAN network directly inputs the LR image to generate the corresponding HR image. But we in order to capture more high frequency details, the modified LR image is generated to generate a different residual image (DR) in the Laplacian pyramid. Finally, the residual image is sampled on the original image and added pixel by pixel. ELSRGAN training schematic is shown in Fig. 2.
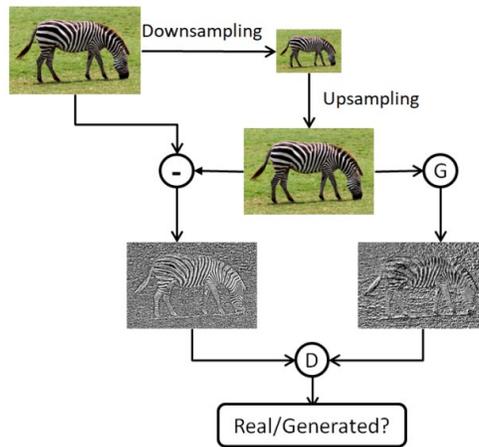

Figure 2: ELSRGAN training diagram. The difference between the upsampled HR image and the LR image is captured for training. The generator network generates a predicted residual image, determine whether the image is generated or a true HR image is performed by the discriminator network

3.2 Generative adversarial network

The generator network D and the discriminator network G form the generative adversarial network model. The generator network is responsible for generating as

realistic a picture as possible, the discriminator network discriminating whether the image is a real HR image or an image generated by the generator network. Both the generator network and the discriminator network perform iterative training on the resistance, and finally complete the training, so that the generator generates a highly similar result to the image. However, in the GAN network, the min-max problem is easy to occur. That is, the discriminator can easily identify the result and cause the training model to crash. To solve this problem, we used the function of Eq. (1).

$$\min_{G} \max_{D} E_{I^{DR} \sim p_{train}(I^{DR})}[\log D(I^{DR})] + E_{I^{LR} \sim p_G(I^{LR})}[\log(1 - D(G(I^{LR})))] \tag{1}$$

For the generator network architecture, SRGAN implemented the residual block structure as shown in Fig. 3. We introduced RRDB as the new generator network architecture, as shown in Fig. 4. We removed the BN layer in SRGAN to ensure the stability of the training. And reduce the complexity of the calculation, and combine the multi-layer residual network with dense links.
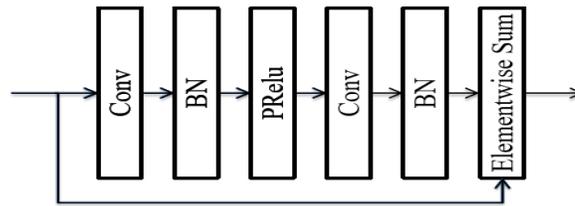

Figure 3: Residual Block. The picture shows the Residual Block in SRGAN. The network structure adopts the method of jumping connection layer. The BN layer is used after each convolution layer
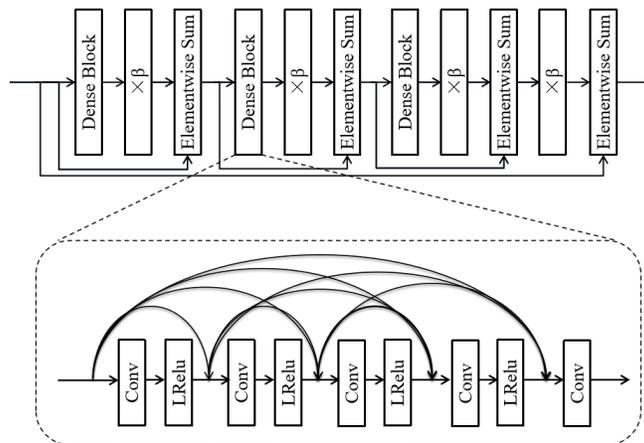

Figure 4: Residual-in-Residual Dense Block. The figure shows the RRDB used by our model, and the BN layer is removed, where β is the residual scaling parameter. The specific structure of the Dense block is given

It has been confirmed that removing the BN layer in super-resolution can improve the operation speed, reduce the computational complexity and GPU memory consumption. And we propose a more complex and deep basic network unit structure, and the network

capacity is obtained from the dense connection [Luo, Qin, Xiang et al. (2019)] which have higher benefit.

To solve the max-pooling problem, LeakyReLU is used as the activation function of the discriminator network. Our discriminator can effectively solve the min-max problem by Eq. (1). Fig. 5 is our discriminator network model. Whenever the number of features doubles, we use a convolutional layer to reduce the resolution. We implemented the Sigmoid activation function to get the final classification probability.
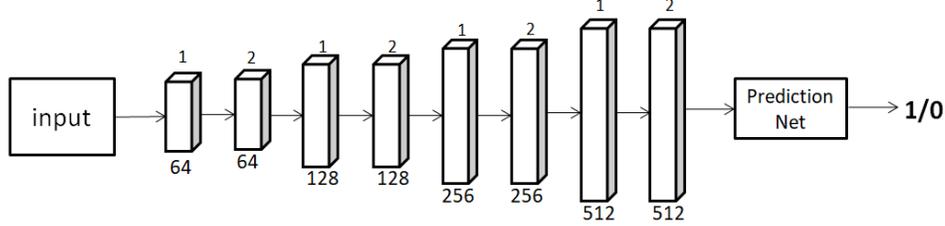

Figure 5: Discriminator network. It consists of eight 3×3 kernel convolution layers. The number above the convolution layer is the step distance. The number of feature maps is indicated by the number below

3.3 Perceptual loss function

The perceptual function used in this paper refers to the perceptual function proposed by Ledig [Ledig, Theis and Huszar (2016)], which is weighted by content loss (l_{VGG}^{SR}) and adversarial loss (l_{Gen}^{SR}):

$$l^{SR} = l_{VGG}^{SR} + 10^{-3} l_{Gen}^{SR} \tag{2}$$

Common content loss relies on MSE-optimized loss function, which can result in higher PSNR, but results in an overly smooth texture that is not satisfactory in visual perception. We utilized a VGG loss function that is closer to the perceived similarity, which is defined as the Euclidean distance between the feature representation and the reference image:

$$l_{VGG/i,j}^{SR} = \frac{1}{W_{i,j} H_{i,j}} \sum_{x=1}^{W_{i,j}} \sum_{y=1}^{H_{i,j}} (\phi_{i,j}(I^{HR})_{x,y} - \phi_{i,j}(G_{\theta_G}(I^{LR}))_{x,y})^2 \tag{3}$$

The adversarial loss in this paper is to encourage our generator network to spoof the discriminator network to achieve a variety of natural images. The generator loss is defined as follows:

$$l_{Gen}^{SR} = \sum_{n=1}^{N} \log[1 - D_{\theta_D}(G_{\theta_G}(I^{LR}))] \tag{4}$$

4 Experiment

4.1 Training data and test data

The training data in this paper is trained by the common reference data DIV2K, and the HR image domain LR upsampled image is subtracted pixel by pixel to obtain the DR

image for training. The test data was tested using the widely used benchmark datasets Set5 [Bevilacqua, Roumy, Guillemot et al. (2012)], Set14 [Zeyde, Elad and Protter (2012)], BSD100 [Yang, Wright, Huang et al. (2008)] and URBAN100 [Yang, Wright, Huang et al. (2010)], where the first three methods consist of natural scenes and the URBAN contains challenging Scene image.

4.2 ELSRGAN model experimental analysis

We conducted an experimental analysis of the ELSRGAN model structure and compared the gaps between the two residual blocks of RRDB and RB. It can be found that the used of RRDB can capture more image details.
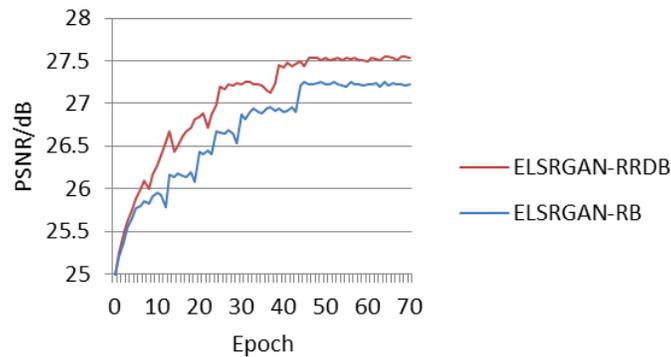

Figure 6: Comparison of results of different network unit. We will use the ELSRGAN model of RRDB as ELSRGAN-RRDB and the ELSRGAN-RB using RB ELSRGAN model. It can be found that using RRDB can obtain higher reconstruction quality, the figure is the result under 4× super resolution

We compared the number of RRDBs utilized and finally decided to use 24 RRDB units for our generator network, which guarantees that the network will not be too deep without sacrificing the reconstruction results.
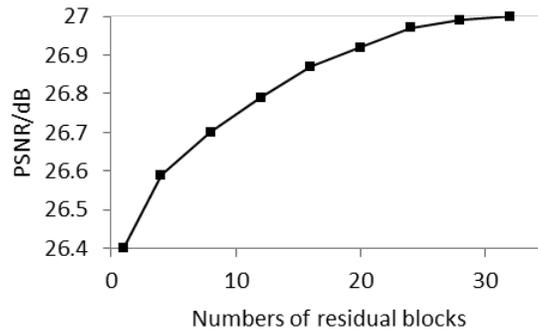

Figure 7: Comparison of the number of RRDBs. We compared the impact of the number of RRDBs on the results and finally decided to use 24 RRDBs as our final generator network

4.3 Perceptual mean-sort-score

We propose a mean-sort-score (MSS) to reflect the visual effect of the image. Ledig et al. proposed a MOS score using 26 evaluators to evaluate images from 1-5, and we found that images with two sensory differences still scored the same during the assessment. And the evaluator is not good for nearby scoring (such as the choice between 2 and 3). In order to distinguish the perceptions of different methods, we propose the concept of MSS. The images of all methods and the HR images are simultaneously provided to the judges, and the scorer is sorted for the senses of all the images, and the lowest score is 1 point, and the highest score is N points (N is the number of images of a single score). For the convenience of comparison, the score is processed in Eq. (5):

$$MSS = \frac{S}{n \times N} \tag{5}$$

where S is the total score of the picture and n is the number of test sets. Using 50 evaluators to sort a total of 219 images, we found that the HR scores were closer to full marks than the MOS method (in 4× and 8×, all graders ranked HR first in the senses). The MSS score of the bicubic method is the last one, and we can find that our scoring method can better evaluate the senses. The comparison results are shown in Section 4.4.

4.4 Comparison with state-of-the-arts

We compare ELSRGAN with the most advanced methods: Bicubic, ScSR, Kim, SRCNN, SelfExSR and LapSRN. Among them SRGAN is only used for comparison of 4× images. As shown in Figs. 8-10, we enumerate the detailed features of the picture after super-resolution at different magnifications. It can be seen from the comparison that our proposed ELSRGAN contains more high-frequency details and accurately reconstructs the texture lines.

Figure 8: Visual comparison at 2× magnification. Our method reconstructs the contour features and texture details of the jewelry, which is not available in other methods. The texture details on the walls of the building are also the most abundant
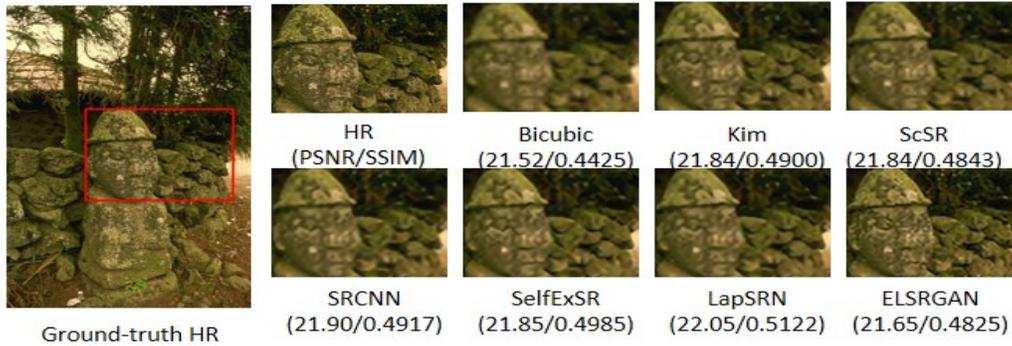

Figure 9: Visual comparisons at 4× magnification. It can be seen that at 4× magnification, ELSRGAN can reconstruct more high-frequency texture features and the angle of the statue is more obvious

Figure 10: Visual comparison at 8× magnification. At 8× magnification, our method complements the detail of the image in an illusion detail to make it look more realistic

We detail the comparison of PSNR, structural similarity index (SSIM) and MSS. It can be found that although the proposed ELSRGAN method has a small PSNR and SSIM after super-resolution, the MSS is higher than other methods and is more in line with human sensory visual effects.

Table 1: Comparison of different method results

Algorithm	Scale	Set5			Set14			BSD100			URBAN100		
		PSNR	SSIM	MSS	PSNR	SSIM	MSS	PSNR	SSIM	MSS	PSNR	SSIM	MSS
Bicubic	2×	33.69	0.931	0.160	30.25	0.870	0.145	29.57	0.844	0.150	26.89	0.841	0.113
ScSR		35.78	0.949	0.215	31.64	0.894	0.235	30.77	0.874	0.225	28.26	0.883	0.253
Kim		36.24	0.958	0.385	32.15	0.903	0.383	31.11	0.884	0.387	28.74	0.894	0.378
SRCNN		36.72	0.955	0.528	32.51	0.908	0.521	31.38	0.889	0.523	39.53	0.896	0.593
SelfExSR		36.60	0.955	0.595	32.24	0.904	0.612	31.20	0.887	0.600	29.55	0.898	0.629
LapSRN		**37.52**	**0.959**	0.737	**33.08**	**0.913**	0.733	**31.80**	**0.895**	0.732	**30.41**	**0.910**	0.792
ELSRGAN		33.94	0.941	**0.889**	30.87	0.884	**0.902**	29.56	0.847	**0.903**	26.81	0.838	**0.905**
HR				0.975			0.972			0.970			0.970
Bicubic	4×	28.43	0.811	0.158	26.01	0.704	0.149	25.97	0.670	0.162	23.15	0.660	0.164
ScSR		29.07	0.826	0.217	26.40	0.722	0.223	26.61	0.703	0.214	24.02	0.702	0.212
Kim		30.07	0.855	0.375	27.18	0.743	0.346	26.71	0.698	0.384	24.20	0.710	0.382
SRCNN		30.50	0.863	0.552	27.52	0.753	0.525	26.91	0.712	0.534	24.53	0.725	0.543
SelfExSR		30.34	0.862	0.574	27.41	0.753	0.584	26.84	0.713	0.582	24.83	0.740	0.558
LapSRN		**31.54**	**0.885**	0.784	**28.1**	**0.772**	0.762	**27.32**	**0.727**	0.779	**25.21**	**0.756**	0.779
ELSRGAN		28.81	0.820	**0.841**	26.17	0.712	**0.853**	25.88	0.681	**0.849**	23.27	0.655	**0.864**
HR				1.000			1.000						1.000
Bicubic	8×	24.40	0.658	0.173	23.10	0.566	0.169	23.67	0.548	0.182	20.74	0.516	0.191
ScSR		24.61	0.671	0.202	23.36	0.577	0.211	23.85	0.561	0.205	20.91	0.531	0.197
Kim		24.98	0.685	0.377	23.58	0.585	0.386	24.02	0.558	0.364	21.03	0.535	0.372
SRCNN		25.33	0.690	0.562	23.76	0.591	0.554	24.13	0.566	0.545	21.29	0.544	0.513
SelfExSR		25.49	0.703	0.564	23.92	0.601	0.571	24.19	0.568	0.568	**21.81**	0.577	0.582
LapSRN		**26.15**	**0.738**	0.770	**24.35**	**0.620**	0.694	**24.54**	**0.586**	0.741	**21.81**	**0.581**	0.671
ELSRGAN		24.33	0.673	**0.856**	23.03	0.558	**0.860**	23.31	0.544	**0.846**	20.45	0.511	**0.870**
HR				1.000			1.000			1.000			1.000

4.5 Discussion

By comparing MSS, we can find that our proposed ELSRGAN method is more in line with human visual sensory effects, and can restore more texture features in image high-frequency detail comparison. Although we have improved the scoring criteria for image sensory effects, we still use the criteria for human scoring. In the future work, we will further study the sensory scoring criteria and propose corresponding formulas to evaluate visual sensory effects and high-frequency texture features. At the same time, although the paper can restore the fine structure, the more elaborate methods of illusion produced by self-similarity may not be suitable for medical applications and surveillance.

5 Conclusion

We use generative adversarial networks combined with Laplacian pyramids to achieve single-resolution super-resolution reconstruction. We used RRDB as the basic network unit to capture more content details while using the perceptual loss function to obtain finer texture features. Extensive evaluation of the baseline data set shows that using the MSS score can confirm that our proposed ELSRGAN method is more visually authentic.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

Bevilacqua, M.; Roumy, A.; Guillemot, C.; Alberi-Morel, M. (2012): Low-complexity single-image super-resolution based on nonnegative neighbor embedding. *HAL CCSD.*

Bruna, J; Sprechmann, P; LeCun, Y. (2015): Super-resolution with deep convolutional sufficient statistics. *Computer Vision and Pattern Recognition.*

Burt, P.; Adelson, E. (1983): The aplacian pyramid as a compact image code. *IEEE Transactions on Communications*, vol. 31, no. 4, pp. 532-540.

Chang, H.; Yeung, D.; Xiong, Y. (2004): Super-resolution through neighbor embedding. *IEEE Conferenceon Computer Vision and Pattern Recognition*, pp. 275-282.

Denton, E.; Chintala, S.; Fergus, R. (2015): Deep generative image models using a aplacian pyramid of adversarial networks. *In Neural Information Processing Systems.*

Dong, C.; Loy, C. (2015): Image super-resolution using deep convolutional networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 2, pp. 295-306.

Goodfellow, I.; Pouget-Abadie, J.; Mirza, M. (2014): Generative adversarial nets. *International Conference on Neural Information Processing Systems*. MIT Press.

Goodman, J.; Gustafson, S. (1968): Introduction to ourier optics. *Roberts and CompanyPublishers*, New York.

Gupta, P.; Srivastava, P.; Bhardwaj, S.; Bhateja, V. (2011): A modified metric based on for quality assessment of color images. *IEEE International Conference on*

Communication and Industrial Application, pp. 1-4.

Harris, J. L. (1964): Diffraction and resolving power. *Journal of the Optical Society of America*, vol. 54, no. 7, pp. 931-936.

Hinton, G.; Salakhutdinov, R. (2006): Reducing the dimensionality of data with neural networks. *Science*, vol. 313, no. 5786, pp. 504-507.

Huang, H.; He, R.; Sun, Z.; Tan, T. (2019): Wavelet domain generative adversarial network for multi-scale face hallucination. *International Journal of Computer Vision*, pp. 1-22.

Huang, H.; Li, Z.; He, R.; Sun, Z.; Tan, T. (2018): IntroVAE: introspective variational autoencoders for photographic image synthesis. *Conference and Workshop on Neural Information Processing Systems*, pp. 52-63.

Huang, J.; Singh, A.; Ahuja, N. (2015): Single image super-resolution from transformed self-exemplars. *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 5197-5206.

Johnson, J.; Alahi, A.; Li, F. (2016): Perceptual losses for real-time style transfer and super- resolution. *In European Conference on Computer Vision*, pp. 694-711.

Kim, J.; Lee, J.; Lee, K. (2016): Deeply-recursive convolutional network for image super-resolution. *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1637-1645.

Kim, K. I.; Kwon, Y. (2010): Single-image super-resolution using sparse regression and natural image prior. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 6, pp. 1127-1133.

Lai, W.; Huang, J.; Ahuja, N.; Yang, M. (2018): Fast and accurate image super-resolution with deep laplacian pyramid Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 1.

Ledig, C.; Theis, L.; Huszar, F. (2016): Photo-realistic single image super-resolution using a generative adversarial network. *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 105-114.

Liu, Q; Xiang, X; Qin, J; Tan, Y; Tan, J. et al. (2019): Coverless steganography based on image retrieval of DenseNet features and DWT sequence mapping. *Knowledge-Based Systems*.

Luo, Y; Qin, J; Xiang, X; Tan, Y; Liu, Q. et al. (2019): Coverless real-time image information hiding based on image block matching and dense convolutional network. *Journal of Real-Time Image Processing*, pp. 1-11.

Ni, K.; Nguyen, T. (2017): Image super resolution using support vector regression. *IEEE Transactionson Image processing*, vol. 16, no. 6, pp. 1596-1610.

Park, S.; Son, H.; Cho, S.; Hong, K.; Lee, S. (2018): SRFeat: single image super-resolution with feature discrimination. *Lecture Notes in Computer Science*, pp. 455-471.

Rajan, D.; Chaudhri, S. (2001): Generalized interpolation and its application in super resolution imaging. *Image and Vision Computing*, vol. 19, no. 13, pp. 957-969.

Sajjadi, M.; Scholkopf, B.; Hirsch, M. (2017): Enhancenet: single image super-resolution through automated texture synthesis. *IEEE International Conference on*

Computer Vision, pp. 4491-4500.

Shocher, A.; Cohen, N.; Irani, M. (2018): Zero-shot super-resolution using deep internal learning. *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3118-3126.

Simonyan, K.; Zisserman, A. (2015): Very deep convolutional networks for large-scale image recognition. *In International Conference on Learning Representations*, pp. 2-5.

Singh, A.; Ahuja, N. (2014): Super-resolution using sub-band selfsimilarity. *Asia Conference on Computer Vision*, pp. 552-568.

Wang, N; He, M; Sun, J; Wang, H; Zhou, L. et al. (2019): PNCC: noise processing method for underwater target recognition convolutional neural network. *Computers, Materials & Continua*, vol. 58, no. 1, pp. 169-181.

Wang, Z.; Bovik, A.; Sheikh, H.; Simoncelli, E. (2004): Image quality assessment: rom error visibility to structural similarity. *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600-612.

Wang, Z.; Simoncelli, E.; Bovik, A. (2003): Multi-scale structural similarity for image quality assessment. *IEEE Asilomar Conference on Signals, Systems and Computers*, vol. 1, no. 2, pp. 9-13.

Yang, C.; Yang, M. (2013): Fast direct super-resolution by simple functions. *IEEE International Conference on Computer Vision*, pp. 561-568.

Yang, J.; Wright, J.; Huang, T.; Ma, Y. (2008): Image super-resolution as sparse representation of raw image patches. *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1-8.

Yang, J.; Wright, J.; Huang, T.; Ma, Y. (2010): Image super-resolution via sparse representation. *IEEE Transactions on Image Processing*, vol. 19, no. 11, pp. 2861-2873.

Yue, H.; Sun, X.; Yang, J.; Wu, F. (2013): Landmark image super-resolution by retrieving web images. *IEEE Transactions on Image Processing*, vol. 22, no. 12, pp. 4865-4878.

Zeyde, R.; Elad, M.; Protter, M. (2012): On single image scale-up using sparse-representations. *In Curves and Surfaces*, pp. 711-730.

Zhang, K.; Zuo, W.; Zhang, L. (2018): Learning a single convolutional super-resolution network for multiple degradations. *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3262-3271.