IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received September 3, 2019, accepted September 24, 2019, date of publication October 1, 2019, date of current version October 17, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2944862

Medical Image Super Resolution Using Improved Generative Adversarial Networks

XINYANG BING, WENWU ZHANG, LIYING ZHENG, AND YANBO ZHANG
School of Computer Science and Technology, Harbin Engineering University, Harbin 150001, China
Corresponding author: Liying Zheng (zhengliying@hrbeu.edu.cn)

This work was supported by the National Natural Science Foundation of China under Grant 61771155.

ABSTRACT Details of small anatomical landmarks and pathologies, such as small changes of the microvasculature and soft exudates, are critical to accurate disease analysis. However, actual medical images always suffer from limited spatial resolution, due to imaging equipment and imaging parameters (e.g. scanning time of CT images). Recently, machine learning, especially deep learning techniques, have brought revolution to image super resolution reconstruction. Motivated by these achievements, in this paper, we propose a novel super resolution method for medical images based on an improved generative adversarial networks. To obtain useful image details as much as possible while avoiding the fake information in high frequency, the original squeeze and excitation block is improved by strengthening important features while weakening non-important ones. Then, by embedding the improved squeeze and excitation block in a simplified EDSR model, we build a new image super resolution network. Finally, a new fusion loss that can further strengthen the constraints on low-level features is designed for training our model. The proposed image super resolution model has been validated on the public medical images, and the results show that visual effects of the reconstructed images by our method, especially in the case of high upscaling factors, outperform state-of-the-art deep learning-based methods such as SRGAN, EDSR, VDSR and D-DBPN.

INDEX TERMS Generative adversarial network, medical image reconstruction, squeeze and excitation block, super resolution.

I. INTRODUCTION

Details of small anatomical landmarks and pathologies are critical to accurate disease analysis. For example, small changes of the microvasculature around a tumor are an important biomarker for cancer diagnosis [1], and unapparent soft exudates are important pathologies for retinal condition diagnosis [2]. However, many actual medical images suffer from the limited spatial resolution, due to imaging equipment and imaging parameters (e.g. scanning time of CT images). Such low resolution of medical images impedes the accurate detection or segmentation of small anatomical landmarks and pathologies, and impedes the accurate diagnosis of some serious diseases at its early stage.

In the past 30 years, a large amount of work has been reported for improving the resolution of actual medical images. Early resolution enhancement methods, such as basic cubic interpolations and its variants, usually suffer from the great loss of sharp-edged details and high local contrast [3] Super Resolution (SR) reconstruction techniques then came to be popular in the community of medical images resolution enhancement. Based on sparse representation, Yang et al proposed a regularized single image SR method for medical images [4]; Rueda et al reconstructed a high-resolution version of a low-resolution brain MR image [5]; Wei et al proposed a medical image SR algorithm [6] with good Peak Signal to Noise Ratio (PSNR) and visual effect. Recently, based on the random forest model selection strategy, Dou et al proposed an SR method for obtaining more information from a low resolution medical image [7]. Based on multi-kernel support vector regression, Jebadurai and Peter proposed an SR algorithm for retinal images [8]. Though these methods are more effective than traditional interpolation-based techniques, they are still unable to restore high quality images in the case of high upscaling factor.

Motived by the tremendous achievements of deep learning in computer vision, some new SR techniques have been reported, too. Based on the VGG-net, Kim et al. presented a highly accurate SR method with Very Deep CNN

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

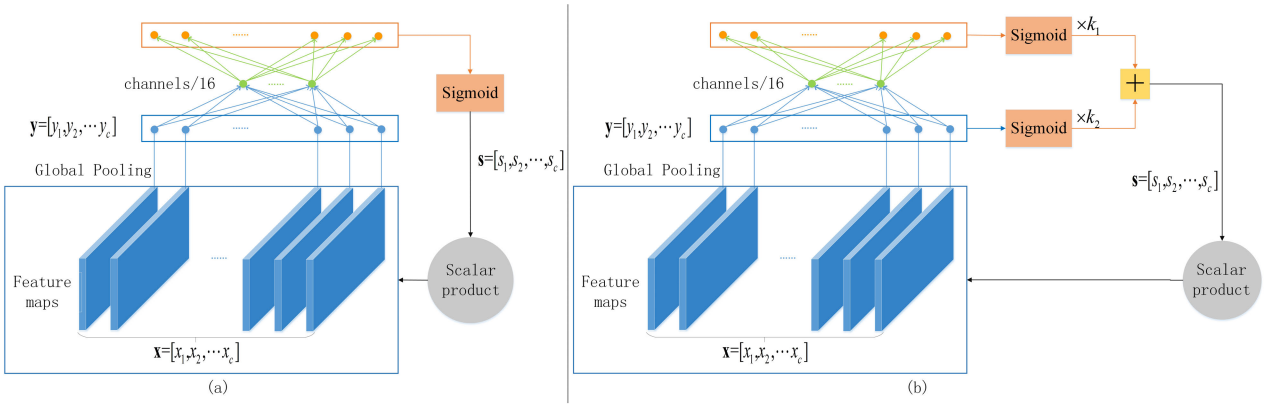

FIGURE 1. The original SE block(a) and the improved SE block(b).

(VDSR) [9]. Dong et al. first introduced Convolutional Neural Networks (CNNs) to single image SR, and proposed SRCNN model [10]. Based on the basic structure of CNNs, an SR method for grayscale medical images is proposed in [11]. He et al. proposed Residual neural Network (ResNet) to make the training procedure of SR model more easier [12]. Tai et al. proposed a 52-layer recursive network to further improve the SR performance of the ResNet [13], and Lim et al. removed unnecessary modules in the ResNet while expanding the model size [14], and achieved significant improvement. Zhang *et al.* [15] adopted effective residual dense block in SR model. They then further explored a deeper network with channel attention [16], and achieved the state-of-the-art PSNR performance.

Recently, due to the good performance of generative adversarial networks (GANs) in producing very realistic images, GAN-based image SR models are emerging and still growing. For example, SRGAN [17], Neural Enhance [18]and ESRGAN [19] are all GAN-based SR models. Specifically, Mahapatra et al. proposed a medical image SR algorithm using progressive generative adversarial networks (P-GANs) [2]

Though, as mentioned above, there are so many methods have been reported, medical image SR is still an open problem, and the reconstruction results are still unsatisfied for high upscaling factors. Therefore, in this paper, we propose a new medical image SR method based on the GAN framework. We first improve the original Squeeze and Excitation (SE) block [20] by strengthening important features while weakening non-important ones. Then, after simplifying the original EDSR [14], we embed the improved SE block in the simplified EDSR model. Finally, we design a new fusion loss that can further strengthen the constraints on low-level features to train the proposed image SR model. Our experimental results on two medical image datasets show that the strategies of embedding the improved SE block and using the fusion loss benefit the proposed GAN-based SR model with better visual effect than several state-of-the-art models, such as EDSR, VDSR, SRGAN, and D-DBPN, especially for high upscaling factors.

The remainder of the paper is organized as follow. Section II describes the method for improving SE block. Section III gives the details of the proposed GAN-based SR model. Section IV presents performance assessments followed by some concluding remarks in Section V.

II. IMPROVED SE BLOCK

The SE building block shown in Fig. 1(a) is proposed by Hu *et al.* [20]. The basic function of the SE block is to adaptively recalibrate channel-wise feature responses by explicitly modelling interdependencies between channels. First, by using the global pooling (1), the SE block squeezes each feature map.

$$y_c = \frac{1}{HW}\sum_{i=1}^{H}\sum_{j=1}^{W} x_c(i,j) \tag{1}$$

where y_c represents the squeezed feature corresponding to the c-th feature map x_c. H and W are the height and the width of x_c, respectively. Then, the squeezed features are fed to a fully connected 3-layer neural network, whose input layer has the same dimension as that of the output layer.

The activation function of the original SE block is the following Sigmoid function:

$$\boldsymbol{s} = E_{org}(\boldsymbol{y}) = \sigma(\boldsymbol{W_2}\delta(\boldsymbol{W_1}\boldsymbol{y})) \tag{2}$$

where $\boldsymbol{s} = [s_1, s_2 \ldots s_c]$ is the scale vector of original feature maps, and $Eorg(\cdot)$ means the original activation function. $\boldsymbol{y} = [y_1, y_2 \ldots y_c]$ is an input feature vector. σ and δ are respectively the Sigmoid function and the ReLU function $\boldsymbol{W}_1$ and $\boldsymbol{W}_2$ are weights of the input layer and the output layer, respectively.

The final output of an SE block is obtained with (3).

$$\tilde{x_c} = x_c \cdot s_c \tag{3}$$

where "$\cdot$" means elementwise product.

On one hand, the activation function in (2) doesn't thoroughly utilize the response of the hidden layer. On the other hand, E_{org} in (2) ranges from 0 to 1. In the case that multiple SE blocks are embedded in a network, such E_{org} of (0,1) will

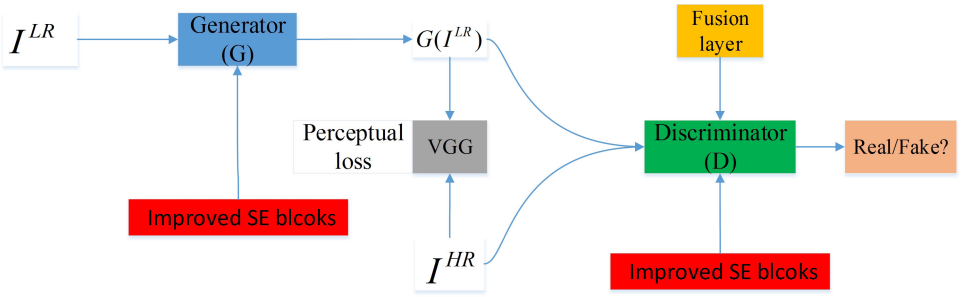

FIGURE 2. The proposed SR model. I^{LR}: Low Resolution(LR) images, I^{HR}: High Resolution(HR) images.

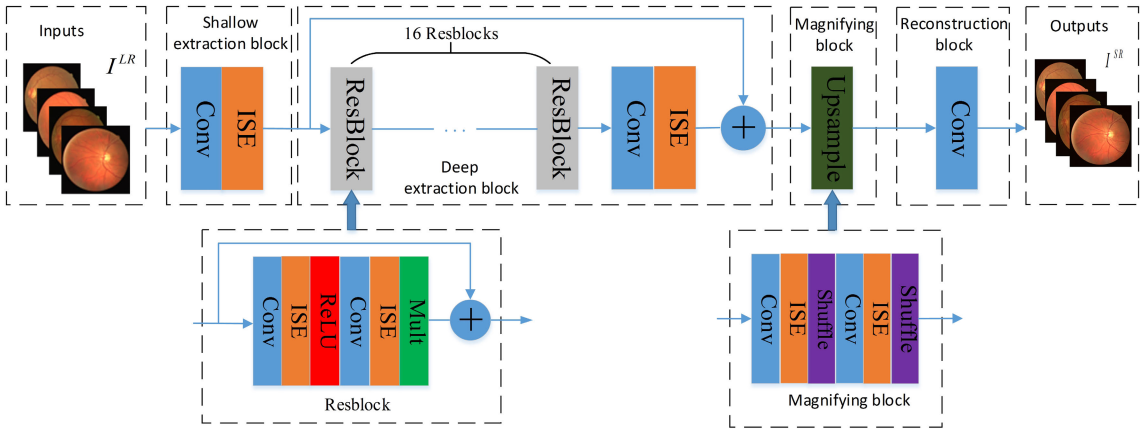

FIGURE 3. The generator of our SR model. ISE: Improved SE block.

make the responses of middle layers very small, and thus will greatly degrade the performance of the network. Therefore, as shown in Fig. 1(b), we substitute the activation function in (2) with (4) in this paper, and get an improved SE block.

$$\boldsymbol{s} = E_{imp}(\boldsymbol{y}) = \{k_1 \times [\sigma(\boldsymbol{W_2}\delta(\boldsymbol{W_1}\boldsymbol{y}))] + k_2 \times \sigma(\boldsymbol{y})\} \times 2 \quad (4)$$

where k_1 and k_2 are positive numbers and $k_1 + k_2 = 1$, and they control the contribution of the input layer and the output layer of the SE block, respectively.

Such improvement on an SE block is beneficial from the following aspects:

i) The residual manner in (4) utilizes both the inputs and the outputs of the 3-layer network, and only fine-tuning on weights is required. Thus, the difficulties in the training process is alleviated.

ii) $E_{imp}(\cdot)$ in (4) ranges from 0 to 2 rather than (0,1). Therefore, the problem of feature weakening caused by performing many multiplications with a scale less than 1 can be effectively alleviated.

III. SUPER RESOLUTION METHOD WITH GAN AND IMPROVED SE

As shown in Fig. 2, our image SR model is built based on the GAN framework and the improved SE blocks. Specifically, the improved SE blocks are respectively embedded in the generator and the discriminator, and a fusion layer is appended to the discriminator.

TABLE 1. Average PSNR on test dataset with different w_{MSE}.

w_{MSE}	0	0.01	0.1	0.5	1	2	10
PSNR(dB)	40.16	Fail	44.43	**47.10**	46.67	45.60	46.61

A. THE GENERATOR AND THE DISCRIMINATOR

As shown in Fig. 3, the EDSR model proposed by Lim *et al.* [14] are simplified to serve as the generator of our GAN-based image SR model. After simplification, the new EDSR has 16 Resblocks and 64 kernels, and other parameters are the same as those in the original EDSR. We then embed the improved SE blocks in the convolutional layers of the simplified EDSR.

The discriminator of our SR model is shown in Fig. 4. It consists of 8 main convolutional layers with the increasing kernels from 64 to 512 as those in VGG [21]. We then embed the improved SE block in each convolutional layer to improve the accuracy of the discriminator. Next, a fusion layer that fuses the features of the last three convolutional layers together is added to the discriminator. By doing so, the discriminator can pay more attention to the low frequency features, and the freedom of our SR model can be reduced, too. Finally, the classification is completed by sequentially performing global pooling, convolution structure and Sigmoid activation function. Here, the convolution structure consists of two layers with 1×1 kernels.

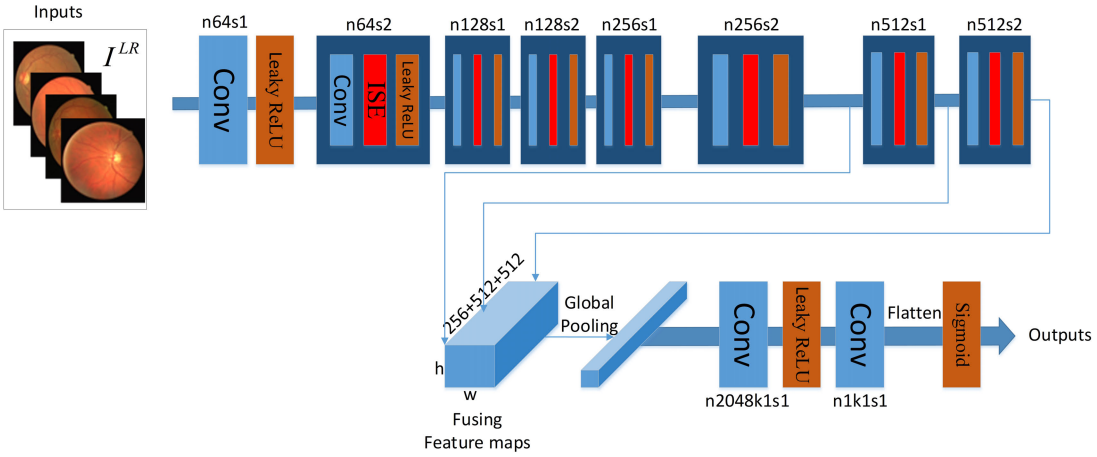

FIGURE 4. The discriminator of our SR model. ISE: Improved SE block.

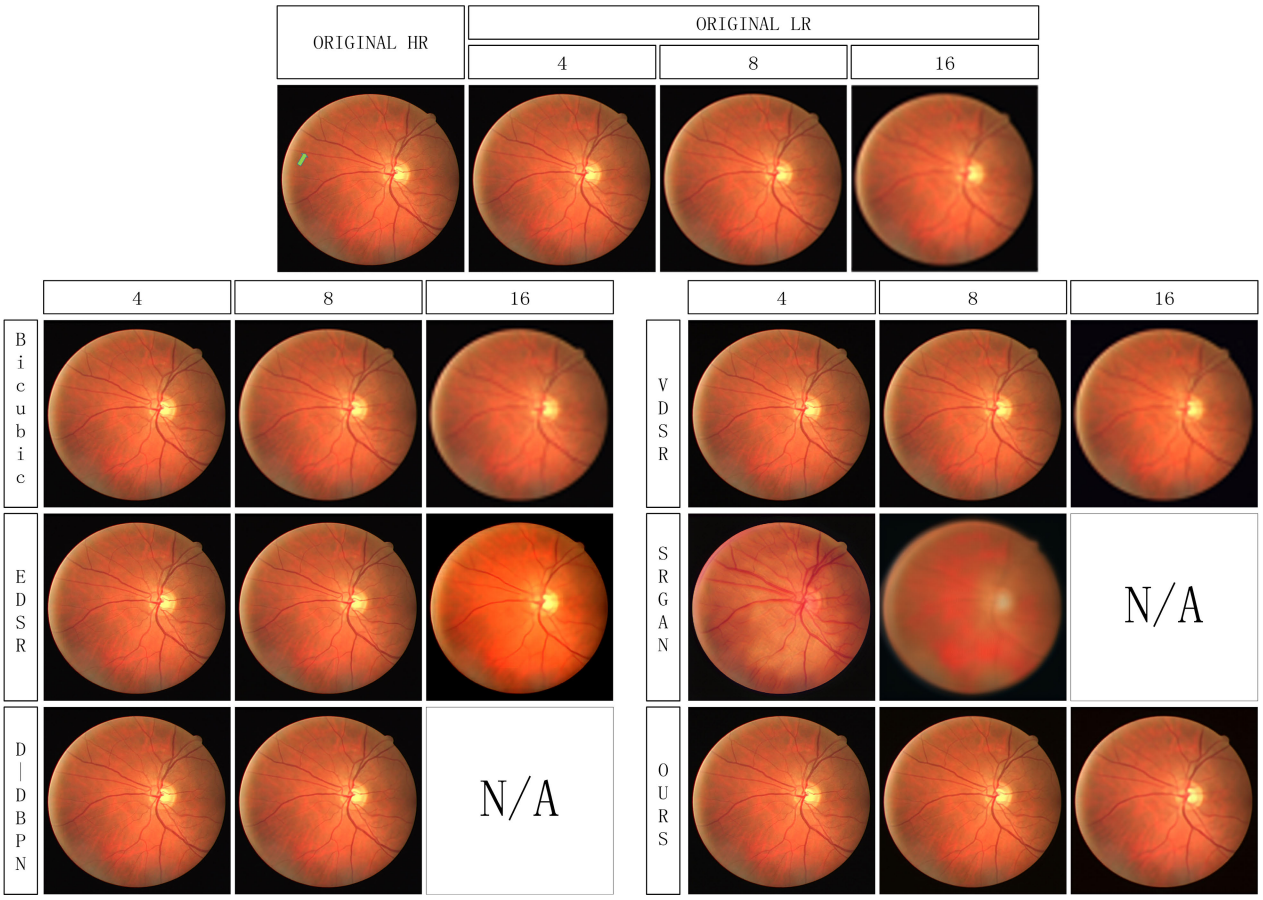

FIGURE 5. Example SR results for a test image in DRIVE database.

B. LOSS FUNCTION

In this paper, we propose a new loss function for training our GAN-based SR model shown in Figs. 2–4. As given in (5), the new proposed loss function combines L_1 loss (L_1), relativistic adversarial loss(L_{RG}) [22], perceptual loss(L_{VGG}) [19], and Mean Square Error loss (L_{MSE}) [10], [23] together.

$$L_{Fusion} = L_{VGG} + w_1 L_1 + w_{RG} L_{RG} + w_{MSE} L_{MSE} \quad (5)$$

where w_1, w_{RG} and w_{MSE} are positive real numbers. They are hyper-parameters that control the contribution of each individual loss.

In (5), L_{VGG} contributes to higher-level semantic contents rather than pixel-level structures in the feature space and it is closely related to the perceptual similarity. The second term L_1 encourages the network to get information from ground truth images. Although both L_{VGG} and L_1 lead high PSNR of the reconstructed image, a lot of high-frequency details

TABLE 2. Training configuration for training the proposed model.

Parameter	Adam	Batch size	Epoch	Initial leaning rate	Patch size
Value	$\beta_1 = 0.9, \beta_2 = 0.999, \varepsilon = 10^{-8}$	16	300	0.0001	48×48

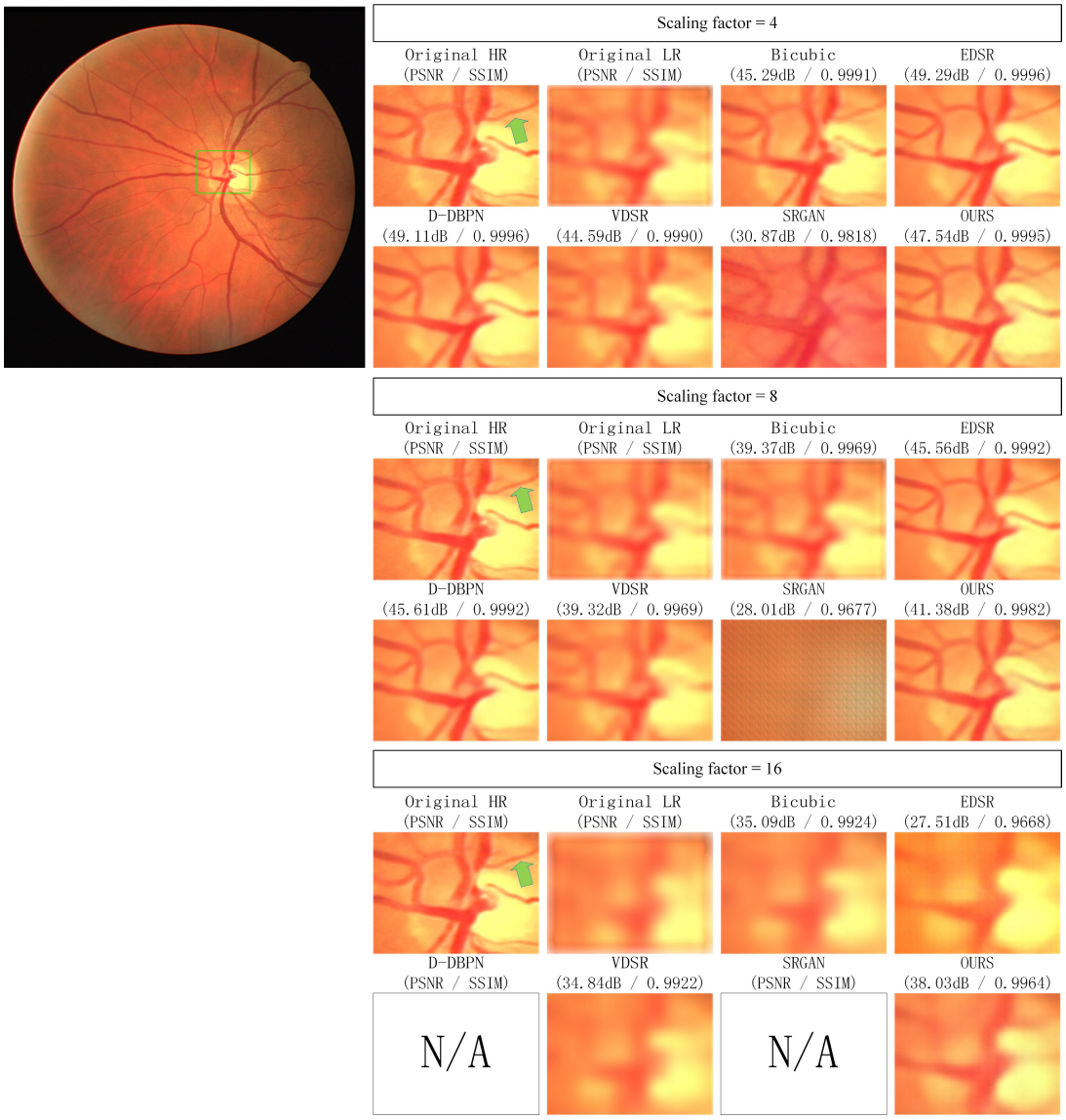

FIGURE 6. Some details of Fig. 5.

are probably lost by adopting these two losses. Therefore, the third term L_{RG} is adopted to enforce the network to produce sharp and clear images. The last term L_{MSE} is used to minimize the MSE between the generated images and the corresponding ground truth.

IV. EXPERIMENTAL RESULTS AND ANALYSIS

The proposed GAN-based medical image SR model has been implemented in PyTorch 0.4.1 on Ubuntu 16.04 with CUDA 8, CUDNN 5.1, and NVIDIA 1080Ti. All experiments were performed on two retina image datasets, DRIVE [24] and STARE [25]. DRIVE consists of 20 training images and 20 test images, while STARE consists of 397 images. The images in STARE are randomly divided into two parts, part A and part B. Part A includes 20 images, and part B includes the other 377 images.

The training dataset consists of 397 retina images, 20 of which come from the training images in DRIVE and others come from part B images in STARE. The test dataset consists of 40 retina images that are independent from the training images. 20 of them come from the test images in DRIVE and others come from part A in Stare. All images are first resized to 1024 × 1024 pixels that serve as reference High Resolution (HR) images.

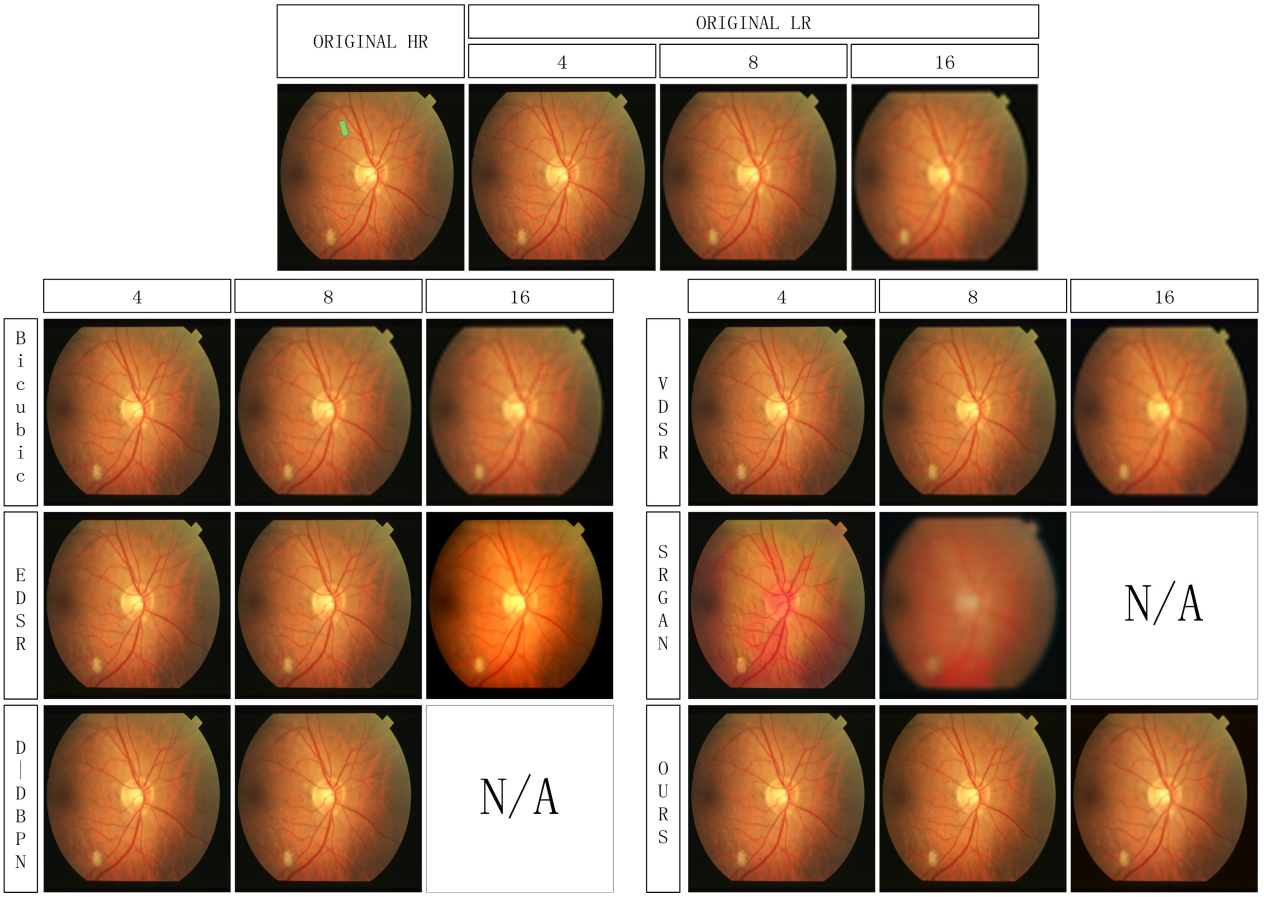

FIGURE 7. Example super resolution results for a test image in STARE database.

TABLE 3. PSNR and SSIM of different models (the best results are in bold).

Dataset		DRIVE+STARE		
Scaling factor		4	8	16
Bicubic	PSNR(dB)	46.41	41.16	37.11
	SSIM	0.9992	0.9975	0.9943
EDSR	PSNR(dB)	**49.06**	**44.97**	29.91
	SSIM	**0.9997**	**0.9993**	0.9667
D-DBPN	PSNR(dB)	48.90	44.95	N/A
	SSIM	0.9996	0.9993	N/A
VDSR	PSNR(dB)	43.04	37.92	33.43
	SSIM	0.9991	0.9975	0.9940
SRGAN	PSNR(dB)	30.13	26.43	Fail
	SSIM	0.9741	0.9501	Fail
OURS	PSNR(dB)	47.78	41.64	**38.00**
	SSIM	0.9995	0.9982	**0.9968**

TABLE 4. The PSNR of original SE and improved SE.

Scaling factors	original SE(PSNR/SSIM)	improved SE(PSNR/SSIM)
4	47.66 / 0.9994	**47.78 / 0.9995**
8	39.99 / 0.9978	**41.64 / 0.9982**
16	30.98 / 0.9809	**38.00 / 0.9968**

A. TRAINING DETAILS

To further augment the training dataset, we randomly choose one of the following operations on the HR images during the training: rotated by 90°, 180°, or 270°, flip horizontally, or zero-mean. The corresponding Low Resolution (LR) images are obtained by down-sampling each high resolution image with scaling factors 4, 8 and 16.

Under the constraint that the output of improved SE block should contribute more than the input layer, parameters k_1 and k_2 in (4) are experimentally determined as 0.8 and 0.2, respectively.The proposed SR model is trained with the fusion loss in (5). Here, according to previous work [19], we fix the value of w_1 and w_{RG} in (5) while vary w_{MSE} from 0 to 10. The average corresponding PSNR of the proposed model on all test images with scaling factor of 4 are listed in Table 1.

From Table 1, one can notice that the proposed SR model cannot be successfully trained with $w_{MSE} = 0.01$. In our experiments, we find that the model is unstable with small w_{MSE} (e.g. 0–0.01). From this point of view, w_{MSE} should not be too small. In this paper, according to Table 1, we choose $w_{MSE} = 0.5$.

In terms of dimension reduction ratio in the improved SE block, we set it to 16 that is the same with the original SE block [20]. The ADAM optimizer [26] is adopted for training our SR model, and the training configuration is listed in Table 2. Our models has been trained with 10^6 updates and batch size of 16.

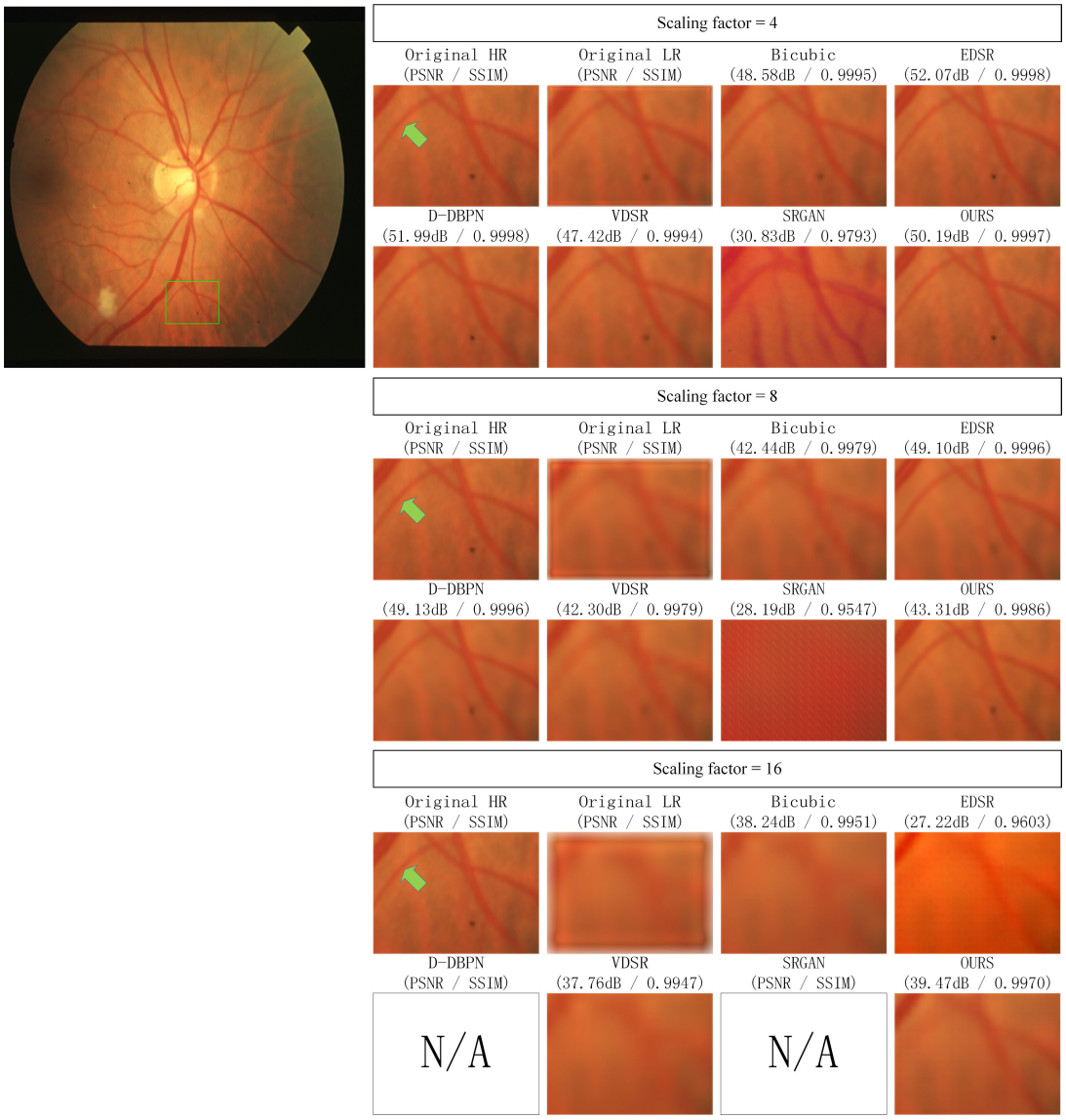

FIGURE 8. Some details of Fig. 7.

B. EVALUATION ON MEDICAL IMAGES

In this section, we evaluate our image SR model on 40 test images. The traditional model Bicubic and the state-of-the-art SR models including EDSR [14], D-DBPN [27], VDSR [9], and SRGAN [17] have been chosen for comparison. The parameter settings accompanied to each compared model are the same as those in their original paper.

Similar to [14], the last 10 images of the training dataset have been selected as training validation set on which the evaluation is conducted. The objective metrics PSNR and structural similarity index (SSIM) for above mentioned models are listed in Table 3, and some visual results are shown in Figs. 5–9. Here the sample images in Fig. 5 and Fig. 7 are from DRIVE database and STARE database, respectively. To show more details, the zoomed in small areas in the reconstructed images in Fig. 5 and Fig. 7 are respectively shown in Fig. 6 and Fig. 8. Since our model is more competitive on high upscaling factors, Fig. 9 is presented to show more visual results of scaling factor of 16.

From Table 3, one can see that in terms of PSNR and SSIM, our model outperforms the traditional SR method Bicubic, the state-of-the-art models VDSR and SRGAN. Moreover, though our model is a light weight network and has much less layers than EDSR and D-DBPN (e.g. the number of layers of EDSR is almost twice as large as ours), it performs just lightly worse than EDSR and D-DBPN for scaling factors 4 and 8, and is superior to them for high scaling factors (e.g. 16). Specifically, our model significantly superiors to the state-of-the-art EDSR on scaling factor 16 with the improvement margin of 8.09dB(PSNR) and 0.0301(SSIM). The major reason is that improved SE blocks embedded in our model can effectively strengthen important features while weaken non-important ones.

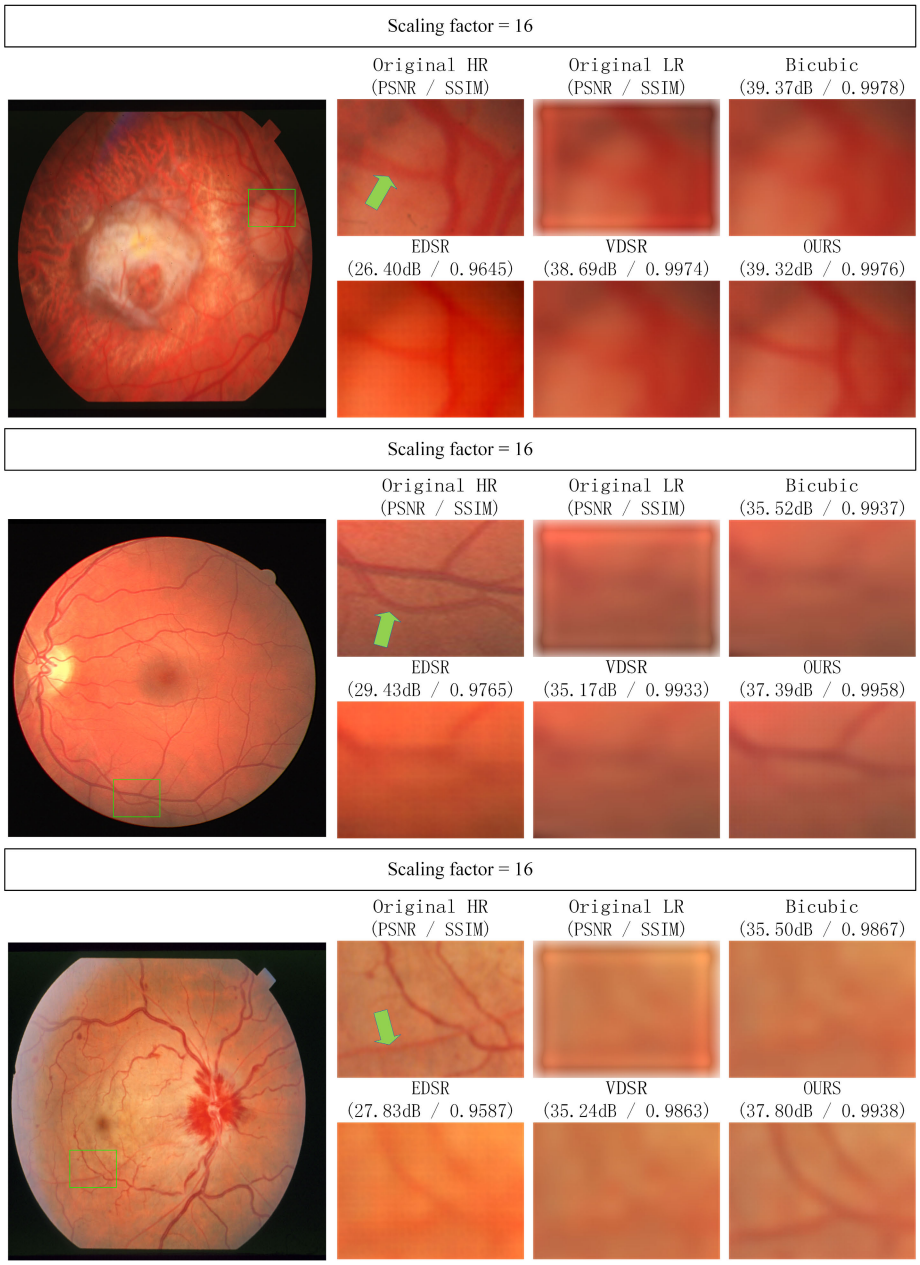

FIGURE 9. More results on scaling factor 16.

Table 4 listed the PSNR of the model with original SE blocks and improved SE blocks. Here, the same network structure as that shown in Figs. 2–4 is adopted, except that the improved SE blocks in the generator and the discriminator are substituted to the original SE blocks. We can see that our improvement strategy on SE blocks benefits the model with higher PSNR and higher SSIM, especially in the case of medium and high upscaling factors (e.g. 8 and 16). From the results in Table 4, we can see that it is the improved SE blocks make our model have higher PSNR and SSIM for high upscaling factors.

Figs. 5–9 illustrate that our model can reconstruct SR images with more visual details than other methods especially for high upscaling factors (e.g. 16). For example, all compared SR models except ours cannot clearly reconstructed the thin blood vessels pointed out with a green arrow in Fig. 5 and Fig. 7 in the case of upscaling factor of 16. Similar results can be seen in Fig. 6 and Fig. 8. Specifically, Fig. 8 and Fig. 9 show that for scaling factor of 16 the small blood vessel is lost from the image reconstructed by Bicubic, EDSR, VDSR, D-DBPN, or SRGAN. Our model can still reconstruct such small blood vessel, though very blurry. Figs. 5–9 further illustrate that though the PSNR and SSIM of our model are lower than the models without adding adversarial loss, such reduction of PSNR or SSIM doesn't degrade the visual effects of reconstructed images. The major reason is that the new

fusion loss in (5) can effectively drive the model to produce images more similar to the ground truth ones.

V. CONCLUSION

In this paper, by embedding improved SE blocks in the generator and the discriminator of the GAN, and by using new fusion loss, we have presented an effective light weight medical image SR model. The experimental results on two retina image datasets have shown that our model outperforms state-of-the-art SR methods including EDSR, SRGAN, VDSR and D-DBPN in terms of visual effects and is comparative to existing image SR models in terms of PSNR and SSIM. Moreover, our method can reconstruct images with more detail structures for higher scaling factors.

REFERENCES

[1] F. Lin, J. D. Rojas, and P. A. Dayton, "Super resolution contrast ultrasound imaging: Analysis of imaging resolution and application to imaging tumor angiogenesis," in *Proc. IEEE Int. Ultrason. Symp. (IUS)*, Sep. 2016, pp. 1–4.
[2] D. Mahapatra, B. Bozorgtabar, and R. Garnavi, "Image super-resolution using progressive generative adversarial networks for medical image analysis," *Comput. Med. Imag. Graph.*, vol. 71, pp. 30–39, Jan. 2019.
[3] T. M. Lehmann, C. Spitzer, and K. Gonner, "Survey: Interpolation methods in medical image processing," *IEEE Trans. Med. Imag.*, vol. 18, no. 11, pp. 1049–1075, 1999.
[4] S. Yang, Y. Sun, Y. Chen, and L. Jiao, "Structural similarity regularized and sparse coding based super-resolution for medical images," *Biomed. Signal Process. Control*, vol. 7, no. 6, pp. 579–590, Nov. 2012.
[5] A. Rueda, N. Malpica, and E. Romero, "Single-image super-resolution of brain MR images using overcomplete dictionaries," *Med. Image Anal.*, vol. 17, no. 1, pp. 113–132, Jan. 2013.
[6] S. Wei, X. Zhou, W. Wu, Q. Pu, Q. Wang, and X. Yang, "Medical image super-resolution by using multi-dictionary and random forest," *Sustain. Cities Soc.*, vol. 37, pp. 358–370, Feb. 2018.
[7] Q. Dou, S. Wei, X. Yang, W. Wu, and K. Liu, "Medical image super-resolution via minimum error regression model selection using random forest," *Sustain. Cities Soc.*, vol. 42, pp. 1–12, Oct. 2018.
[8] J. Jebadurai and J. D. Peter, "Super-resolution of retinal images using multi-kernel SVR for IoT healthcare applications," *Future Gener. Comput. Syst.*, vol. 83, pp. 338–346, Jun. 2018.
[9] 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. (CVPR)*, Jun. 2016, pp. 1646–1654.
[10] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 2, pp. 295–307, Feb. 2016.
[11] H. Liu, J. Xu, Y. Wu, Q. Guo, B. Ibragimov, and L. Xing, "Learning deconvolutional deep neural network for high resolution medical image reconstruction," *Inf. Sci.*, vol. 468, pp. 142–154, Nov. 2018.
[12] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Las Vegas, NV, USA, Jun. 2016, pp. 770–778.
[13] Y. Tai, J. Yang, and X. Liu, "Image super-resolution via deep recursive residual network," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Honolulu, HI, USA , Jul. 2017, pp. 2790–2798.
[14] B. Lim, S. Son, H. Kim, S. Nah, and K. M. Lee, "Enhanced deep residual networks for single image super-resolution," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops*, Jul. 2017, pp. 136–144.
[15] Y. Zhang, Y. Tian, Y. Kong, B. Zhong, and Y. Fu, "Residual dense network for image super-resolution," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 2472–2481.
[16] Y. Zhang, K. Li, K. Li, L. Wang, B. Zhong, and Y. Fu, "Image super-resolution using very deep residual channel attention networks," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, Sep. 2018, pp. 286–301.
[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] M. S. Sajjadi, B. Scholkopf, and M. Hirsch, "EnhanceNet: Single image super-resolution through automated texture synthesis," in *Proc. IEEE Int. Conf. Comput. Vis.*, Oct. 2017, pp. 4491–4500.
[19] X. Wang, K. Yu, S. Wu, J. Gu, Y. Liu, C. Dong, Y. Qiao, and C. C. Loy, "Esrgan: Enhanced super-resolution generative adversarial networks," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, Sep. 2018, pp. 1–16.
[20] J. Hu and L. G. Shen Sun, "Squeeze-and-excitation networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 7132–7141.
[21] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*. [Online]. Available: https://arxiv.org/abs/1409.1556
[22] A. Jolicoeur-Martineau, "The relativistic discriminator: A key element missing from standard GAN," 2018, *arXiv:1807.00734*. [Online]. Available: https://arxiv.org/abs/1807.00734
[23] W. Shi, J. Caballero, F. Huszar, 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.*, 2016, pp. 1874–1883.
[24] *DRIVE: Digital Retinal Images for Vessel Extraction*. Accessed: 2004. [Online]. Available: http://www.isi.uu.nl/Research/Databases/DRIVE/
[25] *STructured Analysis of the Retina*. Accessed: 2000. [Online]. Available: http://cecas.clemson.edu/~ahoover/stare/
[26] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2017, *arXiv:1412.6980*. [Online]. Available: https://arxiv.org/abs/1412.6980
[27] M. Haris, G. Shakhnarovich, and N. Ukita, "Deep back-projection networks for super-resolution," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 1664–1673.

XINYANG BING was born in 1995. She received the B.S. degree in computer science and technology from Harbin Normal University, in 2017. She is currently pursuing the Ph.D. degree in computer science and technology with Harbin Engineering University, under the supervision of Prof. L. Zheng.

Her research interests include image analysis and deep learning.

WENWU ZHANG was born in 1992. He received the B.S. degree in computer science and technology from the Northeastern University at Qinhuangdao and the M.S. degree in computer science and technology from Harbin Engineering University.

He is currently with The 716 Institute, China Shipping Heavy Industry Group. His research interests include image analysis and machine learning.

LIYING ZHENG was born in 1976. She received the B.S., M.S., and Ph.D. degrees in control theory and control engineering from Harbin Engineering University, in 1999, 2001, and 2003, respectively, where she has been a Professor with the College of Computer Science and Technology, since 2013.

She has authored one book, eight inventions, and more than 30 articles. Her research interests include machine learning and pattern recognition.

YANBO ZHANG was born in 1996. She received the B.S. degree in computer science and technology from the Changchun University of Technology, in 2018. She is currently pursuing the M.S. degree in computer science and technology with Harbin Engineering University, under the supervision of Prof. L. Zheng.

Her research interests include video prediction and deep learning.