

Received July 3, 2021, accepted July 19, 2021, date of publication July 26, 2021, date of current version August 3, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3099695

High-Resolution Pelvic MRI Reconstruction Using a Generative Adversarial Network With Attention and Cyclic Loss

GUANGYUAN LI¹, JUN LV¹, XIANGRONG TONG¹, CHENGYAN WANG², AND GUANG YANG^{3,4}, (Senior Member, IEEE)

¹School of Computer and Control Engineering, Yantai University, Yantai 264005, China

²Human Phenome Institute, Fudan University, Shanghai 200433, China

³Cardiovascular Research Centre, Royal Brompton Hospital, London SW3 6NP, U.K.

⁴National Heart and Lung Institute, Imperial College London, London SW7 2AZ, U.K.

Corresponding authors: Jun Lv (ljdream0710@pku.edu.cn) and Chengyan Wang(wangcy@fudan.edu.cn)

This study was funded in part by the National Natural Science Foundation of China (No. 61902338), in part by the National Natural Science Foundation of China (No. 62001120) and the Shanghai Sailing Program (No. 20YF1402400), in part by the British Heart Foundation (Project Number: TG/18/5/34111, PG/16/78/32402), in part by the Hangzhou Economic and Technological Development Area Strategical Grant (Imperial Institute of Advanced Technology), in part by the European Research Council Innovative Medicines Initiative on Development of Therapeutics and Diagnostics Combatting Coronavirus Infections Award “DRAGON: rapiD and secuRe AI imaging based diaGnosis, stratification, fOllow-up, and preparedness for coronavirus paNdemics” [H2020-JTI-IMI2 101005122], in part by the AI for Health Imaging Award “CHAIMELEON: Accelerating the Lab to Market Transition of AI Tools for Cancer Management” [H2020-SC1-FA-DTS-2019-1 952172], and in part by the U.K. Research and Innovation [MR/V023799/1].

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by Shanghai Ruijin Hospital, and performed in line with the Declaration of Helsinki.

ABSTRACT Magnetic resonance imaging (MRI) is an important medical imaging modality, but its acquisition speed is quite slow due to the physiological limitations. Recently, super-resolution methods have shown excellent performance in accelerating MRI. In some circumstances, it is difficult to obtain high-resolution images even with prolonged scan time. Therefore, we proposed a novel super-resolution method that uses a generative adversarial network with cyclic loss and attention mechanism to generate high-resolution MR images from low-resolution MR images by upsampling factors of $2\times$ and $4\times$. We implemented our model on pelvic images from healthy subjects as training and validation data, while those data from patients were used for testing. The MR dataset was obtained using different imaging sequences, including T2, T2W SPAIR, and mDIXON-W. Four methods, i.e., BICUBIC, SRCNN, SRGAN, and EDSR were used for comparison. Structural similarity, peak signal to noise ratio, root mean square error, and variance inflation factor were used as calculation indicators to evaluate the performances of the proposed method. Various experimental results showed that our method can better restore the details of the high-resolution MR image as compared to the other methods. In addition, the reconstructed high-resolution MR image can provide better lesion textures in the tumor patients, which is promising to be used in clinical diagnosis.

INDEX TERMS Super-resolution reconstruction, pelvic, generative adversarial network, cyclic loss, attention.

I. INTRODUCTION

Magnetic resonance imaging (MRI) has the characteristics of non-invasive, non-radiation, and high contrast, etc. It is an important medical imaging modality. However, its imaging speed is relatively slow, which is mainly limited by the physiological factors including magnetic field strength, slew rate, nerve stimulation, etc. [1], [2]. Therefore, the slow imaging

The associate editor coordinating the review of this manuscript and approving it for publication was Eduardo Rosa-Molinar¹.

speed has become the main factor affecting the development of MRI technology. The slower imaging speed will cause long-term data collection and cause discomfort to the patient. Severe motion artifacts can also be introduced from the patient’s movements. Therefore, improving the imaging speed of MRI has great clinical and economic significance. The current research proposes that the method of image reconstruction can be used to accelerate MRI scanning [3]–[5], and has achieved desired results. These methods used deep learning to reconstruct under-sampled images without

artifacts. With the introduction of super-resolution methods, we can also apply super-resolution methods to accelerated MRI imaging. We can get low-resolution (LR) images through fast scanning, and high-resolution (HR) images through trained neural networks, and can also achieve the purpose of accelerating MRI imaging.

The first method of applying deep learning to super-resolution reconstruction was SRCNN [6], and its network structure was composed of three convolutional layers. The proposal of FSRCNN [7] improved the previous problem of SRCNN that needed to interpolate LR images before they can be used as the input of the network. The FSR-CNN can directly input the original LR images. ESPCN [8] proposed an efficient method to extract features directly on the size of LR images to obtain HR images. VDSR [9] added ResNet to the super-resolution processing method, had a positive impact on the subsequent deep learning super-resolution method. DRCN [10] used the previous RNN structure in super-resolution processing for the first time, and deepened the network structure through residual learning. The RED [11] used an encoder-decoder framework in the super-resolution method. The network structure was composed of a symmetrical convolutional layer-deconvolutional layer, which can learn an end-to-end mapping from LR images to the corresponding HR images. DRRN [12] combined the ideas of ResNet, VDSR, and DRCN methods, and proposed a deeper network structure to the super-resolution task. Experiments proved that a deeper network can obtain better results. SRDenseNet [13] input the features of each layer in the dense block to all subsequent layers to make their features concatenate, reducing the risk of gradient disappearance. SRGAN [14] used the GAN [15] in the super-resolution method, using perceptual loss and generative adversarial loss, which effectively improved the effect of the network. EDSR [16] was an improvement to SRResNet [14], and further deepening the network structure to obtain deeper details and improved the effectiveness of the network.

Many research methods had shown that super-resolution can not only be used in natural image processing but also can achieve excellent results in MRI reconstruction. Qui *et al.* [17] proposed an effective medical image super-resolution for the reconstruction of the knee images. The network proposed by this method is composed of three hidden layers and a sub-pixel convolutional layer. Lyu *et al.* [18] applied deep learning and ensemble learning to the super-resolution of MR images. Steeden *et al.* [19] used the a novel method to super-resolution reconstruction of the LR whole heart bSSFP data quickly collected in the clinical environment, and used the 3D residual U-net [20] as the training network, and obtained the expected effect. Zheng *et al.* [21] proposed a hybrid network (Hybrid-Net) to be used in super-resolution MR image reconstruction. Du *et al.* [22] proposed an anisotropic MR image super-resolution reconstruction method based on convolutional neural network (CNN) [23] residual learning with long-skip connections and short-skip connections.

Sun *et al.* [24] applied the deep convolutional GAN [25] to the HR reconstruction of Breast MR images and obtained good reconstruction results. Chaudhari *et al.* [26] proposed a method of using CNN to perform knee MR image super-resolution. Lyu *et al.* [27] proposed a method using progressive network for multi-contrast super-resolution MRI.

In addition, researches have proven that the use of GAN can effectively reconstruct under-sampled images into images close to clinical medical standards for accelerating MRI. Yang *et al.* [28] uses deep de-aliasing GAN to perform CS reconstruction. Mardani *et al.* [29] proposed a GAN for under-sampling MRI reconstruction, and used least-squares GAN and pixel-wise L1 as the cost function during training. Quan *et al.* [30] proposed the use of cyclic loss in the GAN for end-to-end MRI reconstruction. Lv *et al.* [31] proposed the use of GAN to remove g-factor artifacts in SENSE reconstruction and proposed the application of a combination of parallel imaging and GAN to multi-channel MRI Image reconstruction [32]. Korkmaz *et al.* [33] proposed a novel unsupervised SLATER for MRI reconstruction, which was composed of a deep adversarial network with cross-attention transformer blocks. Dar *et al.* [34] used conditional GAN to perform synergistic recovery of undersampled multi-contrast acquisitions. The GAN-based method can also be applied to image synthesis. Dar *et al.* [35] proposed the use of conditional GAN for multi-contrast MR image synthesis. This method enhances the synthesis performance of the network by using a pixel-wise loss in registered multi-contrast images and a cyclic loss in unregistered images. Yurt *et al.* [36] proposed a semi-supervised depth generation model for MRI image synthesis, which can generate MR images with the same quality as the fully-supervised model. Shamsolmoali *et al.* [37] summarized some recent researches on the use of GAN for image synthesis.

It is worth noting that the resolution of MR images is an important factor in clinical medical diagnosis. HR MR images can provide radiologists with more information and facilitate future analysis. However, it takes a longer time to obtain a higher resolution MR image. In some circumstances, even if the scan time is extended, it is still difficult to obtain HR images. Therefore, the use of super-resolution methods to transform LR MR images into HR MR images through deep learning is helpful for diagnosis in clinical medicine. Based on the above research methods, we proposed a novel super-resolution reconstruction method that uses the attention mechanism and cyclic loss in a GAN. In the following sections, we will introduce our methods, experiment settings, experiment results, discussion, and conclusion in detail.

A summary of our main contributions in this study are listed below:

- (1) We proposed a novel GAN-based super-resolution reconstruction method, and used the pelvic dataset in the experiment;
- (2) We proved that the use of cyclic loss and attention in GAN can be well applied to super-resolution reconstruction;

(3) We analyzed the performance of each super-resolution method under different upsampling factors.

II. METHOD

A. LOW-RESOLUTION (LR) DATA

We designed a simulation process to obtain LR MR images. We used a LR mask L_m , sampling from the center in the horizontal direction. The sampling rate is 50% and 25%, which represent downsampling $2\times$ and $4\times$, respectively. We used Fourier Transform (FT) to transform the original image HR_{gt} into k-space image HR_k , and then used L_m to remove part of the high-frequency information in HR_k to obtain L_k . Then through the inverse Fourier Transform (IFT) to transform L_k into the image domain space, and the LR image of the same size as HR_{gt} can be obtained as the input of the network. The process can be expressed as:

$$LR = \overline{\mathcal{F}}(L_k) = \overline{\mathcal{F}}(\mathcal{F}(HR_{gt}) \odot L_m) \quad (1)$$

where \mathcal{F} stands for FT. $\overline{\mathcal{F}}$ stands for IFT. \odot means dot multiplication operation.

B. GENERATIVE ADVERSARIAL NETWORK

Many studies demonstrated the excellent effect of GAN for accelerating MRI scanning [28], [32], [38]–[40]. The super-resolution method we proposed is also based on GAN, which has strong generation and learning capabilities. GAN is implicitly used as the optimization target through the discriminator. GAN is composed of generator G and discriminator D. The generator learns and extracts details of the input real image *Real*. Its purpose is to generate a fake image *Fake* through the learned information to let D distinguish. This is exactly the work of D, to determine whether the input image is real image *Real* or image *Fake* generated by G. G can continuously learn the effective information in the real image to continuously generate fake image. In order to correctly distinguish *Real/Fake*, D also continuously optimizes in the process. In the continuous learning and optimization of G and D, the generative adversarial loss function L_{gan} is generated. GAN continuously optimizes the network through L_{gan} , so that the image generated by G can be infinitely close to the real image. In our proposed network, L_{gan} can be expressed as:

$$\min_G \max_D L_{gan}(D, G) \quad (2)$$

where

$$L_{gan}(D, G) = E_z[1 - \log D(G(LR))] + E_x [\log D(HR_{gt})] \quad (3)$$

where $E(*)$ represents the expected value of the distribution function, $z \sim P_G(z)$ represents the distribution of data learned by the generator, and $x \sim P_{data}(x)$ represents the distribution of data in the real MR image training set. $D(HR_{gt})$ means to discriminate the real HR image, and we expect its discrimination result to be as close to 1 as possible. $G(LR)$ represents the HR image generated by the generator,

namely HR_{re} . For the generated image, we expect the discriminant result $\log D(G(LR))$ of the discriminator to be as close to 0 as possible, which is \max_D . For the optimization of G, it is only necessary to make the discrimination result $\log D(G(LR))$ close to 1, which is \min_G . In this adversarial process to find the global optimal solution and make G achieve the best generation effect.

1) GENERATOR ARCHITECTURE

In our research, the generator architecture consists of two residual u-nets with encoder and decoder [30]. As shown in Figure 1(a)(b), each residual u-net consists of four encoder blocks and four decoder blocks. There is a full-attention behind each decoder block to further extract the details of the feature maps. There is a global long-skip connection between the encoder and the corresponding decoder, which is used to deepen the network level in order to extract the details in the image. The orange part represents the encoder block, the green part represents the decoder block, and the red part represents full-attention. The number of feature maps for each encoder and decoder block is shown in Figure 1(a). In addition, the input of each module is a 4D tensor, and 2D convolution is used for convolution operation, the filter size is 3×3 , the stride is 2. The encoder block adopts the down-sampling method, and each encoder block is composed of two residual blocks with short-skip connections (SR-block). SR-block has three layers: 1) The first layer is 2D convolution with a stride of 2, as shown in the blue part of Figure 1(b). 2) The second layer is a residual block with short-skip connections, including conv0, conv1, conv2, and full-attention. Full-attention will be introduced in the following section. In the residual block, conv0 and conv2 are 2D convolutions with a filter size of 3×3 and stride of 1, and the number of feature maps is 64. Conv1 is a 2D convolution with a stride of 1, the filter size is 3×3 , and the number of feature maps is 32, as shown in the cyan part of Figure 1(b). 3) The third layer is 2D convolution with a stride of 1, which is used to adjust the feature maps that obtain attention activation. Up-sampling is used for the decoder blocks, and each decoder block is also composed of two SR-blocks. The difference between it and the encoder block is that the first layer of the SR-block is 2D convolution with a stride of 1, and the third layer is 2D convolution with a stride of 2. After the last decoder block, 2D convolution is used to obtain the global information of the feature maps, with a filter size of 3×3 and stride of 1. During the convolution operation, the same mode is used to obtain an output image of the same size as the original image. In addition, studies have confirmed [30] using two residual u-nets can extract more image details than using a single residual u-net to generate an image closer to HR_{gt} .

2) DISCRIMINATOR ARCHITECTURE

The architecture of the discriminator we used is composed of a 7-layer network sequential cascade, as shown in Figure 1(c). The first layer is 2D convolution with a filter size of 4×4 , a stride of 2, and the number of feature maps is 64. The second

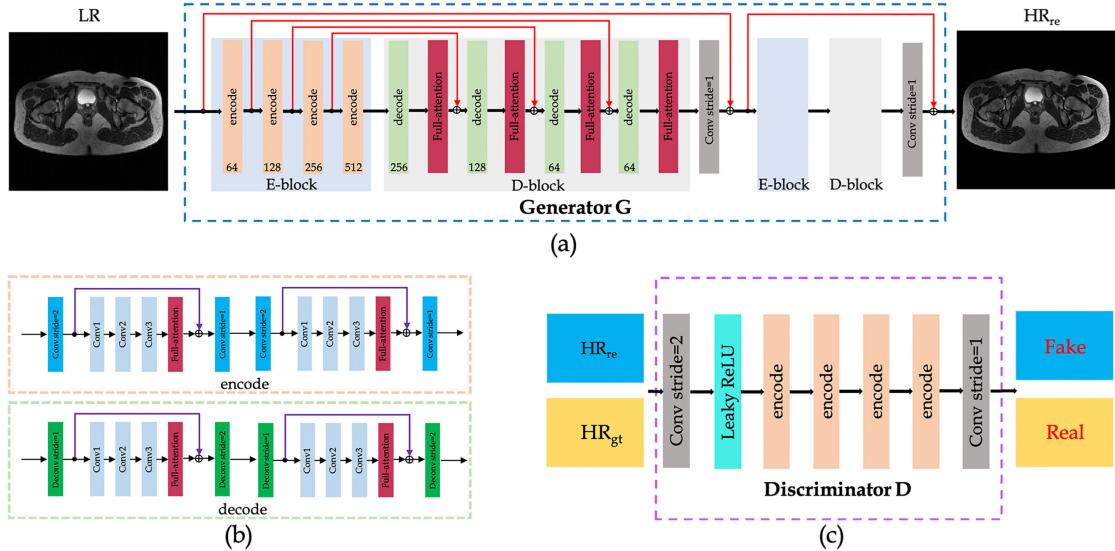


FIGURE 1. (a) The generator network architecture proposed by our research. The generator takes LR images as input to generate HR_{re} images. The red arrow is global long shortcut. The red block is full-attention. (b) The architecture of encode and decode, in which full-attention is used in each residual block. The purple arrow is local short shortcut. (c) The architecture of the discriminator, D discriminates the generated image HR_{re} and the real image HR_{gt}, and the result is Real/Fake.

layer uses the Leaky ReLU activation function to activate. The third to sixth layers use the same network structure as the encoder in G. The seventh layer uses 2D convolution with a stride of 1 to obtain global information. In addition, we use the output of the sixth-level residual block to evaluate the generative adversarial loss L_{gan} . There are two evaluation results of D: one is the real HR image HR_{gt} , and D should give the result *True*. The other is the HR image HR_{re} generated by G. The result of D should be *False*.

C. CYCLIC LOSS

In this study, we used cyclic consistency loss as an additional constraint when training the network. With enough datasets, any HR image HR_{re} generated by G has a complete mapping with the low-resolution image LR. But this is only an ideal situation. In reality, when the dataset is limited, the cyclic consistency loss can be used to correctly map $LR[m]$ and $HR_{re}[m]$, where m is the entire mapping of the data. The cyclic loss L_{cyc} in our method is divided into the cyclic loss L_{fre} in the frequency domain and the cyclic loss in the image domain L_{img} . 1) The frequency domain cycle loss is that in the i -th cycle, the $HR_{re}[i]$ generated by G is transformed into the frequency domain by the FT to obtain $KS_{re}[i]$. Then minimize the difference between $KS_{re}[i]$ and $KS_{lr}[i]$. Next, $KS_{lr}[i]$ is converted to $LR[i]$ as the input of G through the IFT. Among them, i is used to marking each cycle, and the input of the i -th G is $LR[i - 1]$. Therefore, L_{fre} can be expressed as:

$$L_{fre}(G) = \mathcal{M}(KS_{lr}[i], KS_{re}[i]) \quad (4)$$

where \mathcal{M} represents the mean absolute error.

2) Image domain loss is to ensure that $HR_{gt}[i]$ and $HR_{re}[i]$ that generated by G in the i -th cycle are as similar as possible. In the i -th cycle, $HR_{gt}[i]$ is transformed by FT and IFT to obtain $LR[i]$, and next $HR_{re}[i]$ is generated by G. L_{img} is expressed as:

$$L_{img}(G) = \mathcal{M}(HR_{gt}[j], HR_{re}[j]) \quad (5)$$

In addition, $L_{fre}(G)$ and $L_{img}(G)$ only restrict G, and do not affect D. The loss function of our proposed network is composed of generative adversarial loss (2)(3) and cyclic loss (4)(5), so it can be expressed as:

$$L_{HPMR} = L_{adv}(G, D) + \alpha L_{fre}(G) + \beta L_{img}(G) \quad (6)$$

where α and β are hyperparameters used to constrain $L_{fre}(G)$ and $L_{img}(G)$ to stabilize the network during training. We set $\alpha = 1.0$ and $\beta = 10.0$.

D. FULL-ATTENTION

We used the attention mechanism to extract the details in the MR image. The network can learn the attention mechanism autonomously. By assigning different attention weights, the neural network can focus on different key points in the image. There have been studies using the attention mechanism in MR image reconstruction and have shown excellent results [41]–[44]. In our proposed method, full-attention refers to all-round attention, spatial and channel-wise, as shown in Figure 2. When the feature maps are input to the full-attention module, the spatial weight of the feature map is obtained through spatial attention, and then the weight of each channel is obtained through channel-wise

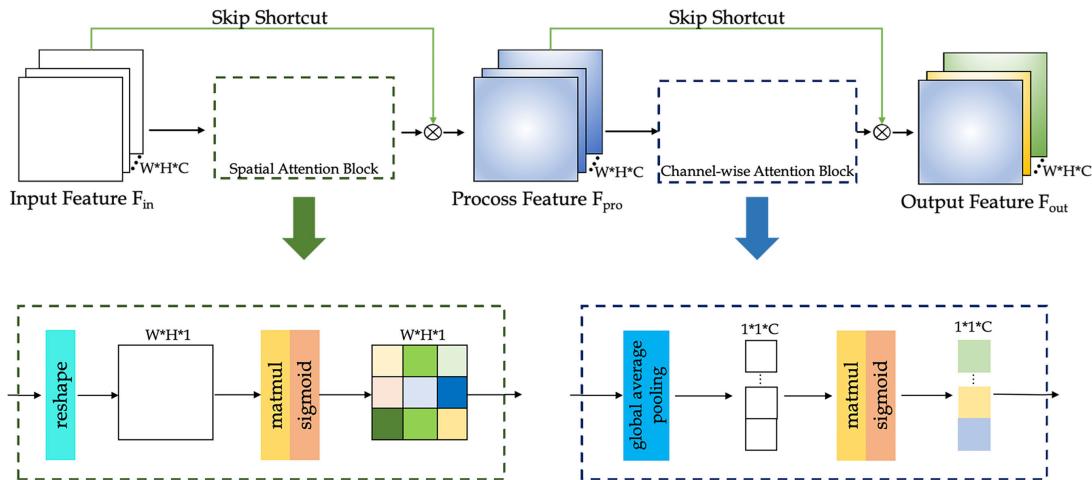


FIGURE 2. Full-attention architecture, including spatial attention module and channel-wise attention module.

attention. Full-attention can enable feature maps to obtain the distribution of attention weights in both breadth and depth, and extract deep-level image details.

1) SPATIAL ATTENTION

The feature maps F_{in} (shape $W \times H \times C$) is input into the spatial attention module (SAB), as shown in Figure 2 (green dashed box). Through the reshape operation, the input F_{in} is calculated at the same position to mean, the number of channels is compressed to 1, and the feature map of $W \times H \times 1$ is obtained, named $F_s[i]$, ($1 \leq i \leq m$), where m is the size of the entire feature map, ($m = W \times H$), which means different areas in the feature map. We define a random weight $W_s[i]$ in the matmul operation to assign spatial attention weights to $F_s[i]$. After passing the sigmoid activation function, F_s obtains the distribution of spatial attention weights. Different colors in Figure 2 represent different weights. Then according to the idea of residual learning [45], we use a short-skip connection with F_{in} for fine-tuning. Therefore, feature maps with spatial attention are obtained, which we call process feature F_{pro} , and the shape is $W \times H \times C$. The spatial attention is expressed as:

$$F_{pro} = \Phi(\Theta(W_s[i], F_s[i]) + b_s) \otimes F_{in} \quad (7)$$

where Θ represents the multiplication of matrix $W_s[i]$ and matrix $F_s[i]$, \otimes represents the multiplication of corresponding elements in the matrix, Φ represents the sigmoid activation function, and b_s is a biases term we set.

2) CHANNEL-WISE ATTENTION

As shown by the blue dashed box in Figure 2, this block assigns the channel attention weight to F_{pro} . And it uses global average pooling to extract feature maps with spatial information into feature map $F_c[i]$ on each channel, with the shape of $1 \times 1 \times C$, $1 \leq i \leq c$, where c represents the total number of channels. As same as spatial attention distribution, the distribution of channel attention is completed by random

weight $W_c[i]$ and sigmoid activation function, so that each channel gets different weights, as shown in different colors in Figure 2. Finally, through residual learning with F_{pro} for detail correction, the full-attention activated feature maps F_{out} are obtained. Full-attention can be expressed as:

$$F_{out} = \Phi(\Theta(W_c[i], F_c[i]) + b_c) \otimes F_{pro} \quad (8)$$

Using full-attention in the decoder process can ensure that the detailed information in the image is better restored during the up-sampling process.

III. EXPERIMENT

A. PELVIC DATASETS AND MODEL TRAINING

All subjects gave their informed consent for inclusion before they participated in the study with approval from the local institutional review board (in accordance with the Declaration of Helsinki). The institutional review board (at Shanghai Ruijin Hospital) has approved the MRI scanning.

As mentioned in the previous section, The pelvic dataset we used is composed of healthy subjects and patients, and they were all obtained through different MR sequences, which were T2, T2W SPAIR, and mDIXON-W. Acquisition parameters of the pelvic dataset are shown in Table 1. In the pelvic dataset of healthy subjects, a total of 80 people's pelvic MR data participated in the training and testing of the network, of which 64 people were used for network training and 16 people were used for network testing. In addition, we took the pelvic MR dataset of 7 patients for network testing. In the T2 dataset, we took the middle 100 slices of each subject as the experimental data. In addition, we resized the original MR image obtained by scanning to 512×512 as the HR label.

We used Tensorpack [46], a neural network training visualization tool based on TensorFlow, for network visualization training. When training and testing the network, the graphics processing unit (GPU) used is NVIDIA Tesla V100 (4×16 GB). The optimizer used by our network is the ADAM

TABLE 1. Acquisition parameters of the pelvic dataset.

Dataset	Contrast	MR sequences	Scanning direction	Parameters	Scan time
T2	T2W	TSE-3D	TRA	TR: 1250ms; TE: 130ms; Slices: 200; Slice gap: -1mm; FOV: 400mm; Reconstruction matrix: 1120	2min30s
T2W SPAIR	T2W	TSE-SPAIR	SAG	TR: 4000ms; TE: 70ms; Slices: 24; Slice gap: 1mm; FOV: 250mm; Reconstruction matrix: 528	2min16s
mDIXON-W	T1W	TSE-mDIXON W	TRA	TR: shortest; TE: shortest; Slices: 80; Slice gap: -2mm; FOV: 250mm; Reconstruction matrix: 512	1min04s

algorithm [47], [48], the learning rate is $1e^{-4}$, the epoch is 500, and the batch size is 100.

B. ASSESSMENT METHOD

We used Bicubic, SRCNN, SRGAN, EDSR as the comparison method of our method HPMR. We upsampled the healthy subjects pelvic LR datasets and the patients pelvic LR datasets by $2\times$ and $4\times$ respectively for testing. In order to evaluate the difference between our method and each comparison method, we used peak signal to noise ratio (PSNR), structural similarity (SSIM), root mean square error (RMSE), and variance inflation factor (VIF) to evaluate the image quality of the HR generated by each method. In addition, we also analyzed the histogram and KL divergence of the result images.

IV. RESULT

A. HEALTHY SUBJECTS

We used the pelvic dataset of healthy subjects to test with an upsampling factor of $2\times$. The results of each method are shown in Figure 3. We selected slices with similar layers in the three data to display the results. Intuitively, the LR image cannot clearly see the details in the pelvic. Although the BICUBIC method effectively improve the LR image, it still need to be improve intuitively. SRCNN and other comparison methods can effectively reconstruct HR images, but in some details, as indicated by the red arrow in the figure, it need to be improve compared to HR. The method we proposed, the HR image reconstructed by HPMR, not only have a great improvement in the senses, but also restore the details of the original HR image. In addition, the generate HR image can better distinguish the uterine and the surrounding parts.

Figure 4 shows the HR image reconstruct after upsampling $4\times$ the pelvic LR dataset of healthy subjects. It can be seen from the crop images of each method that the quality of the LR image is poor, and the details in the pelvic cannot be distinguished normally. Although the BICUBIC method can effectively recover the information lost in the LR, it still cannot get the desired effect from the sensory point of view. Compared with our proposed method, contrast methods such as EDSR have slightly worse recovery at the edge of the uterus, as shown by the red arrow in the figure. In addition, compared with the original HR, HPMR recover the detailed

information in the image as much as possible, and obtained the desired effect.

We also calculated the KL divergence of each result in Figure 3 and Figure 4 to show the similarity of the distribution of HR images reconstructed by each method, as shown in Table 2. It can be seen that the LR and BICUBIC methods are different from the original HR. Deep learning-based methods such as SRCNN, SRGAN, and EDSR have a smaller KL divergence value, which is similar to the original HR. Among them, our proposed method HPMR is the most similar to the original HR. In addition, we have counted the average values of the indicators of each method when the upsampling factors are $2\times$ and $4\times$ in the three datasets of healthy subjects pelvic, including PSNR, SSIM, RMSE, and VIF, as shown in Figure 5. It can be seen from the figure that the various index values of our method HPMR are better than other comparison methods.

B. PATIENTS

In order to verify the performance of the trained network in the patients dataset, we tested the reconstruction effect of each method on three patients pelvic data with upsampling factors of $2\times$ and $4\times$. The reconstruct HR image is shown in Figure 6. In Figure 6(a), the LR image can only see the contours of the uterus and uterine fibroids, but cannot clearly see the details. The BICUBIC method can only recover part of the detailed information in the original HR. SRCNN, SRGAN and EDSR can effectively restore the sharpness of the image, but the recovery of the details of the uterine fibroids still needs to be improved. Our proposed method HPMR can better restore the details in the original HR, as shown by the red arrow in Figure 6(a). In addition, we can clearly see the junction of the uterus and uterine fibroids and the internal information of uterine fibroids, which is helpful for radiologists to diagnose.

Figure 6(b) shows the results of each method under the patients pelvic dataset when the upsampling factor is $4\times$. We can see that the LR image has a poor quality and cannot effectively distinguish the uterine part and the uterine fibroids. Although BICUBIC and SRCNN have improved the quality of part of the image, they still cannot effectively distinguish the tumor part in the image from the sensory point of view. Although the SRAGN and EDSR methods are better than the previous comparison methods, the image quality

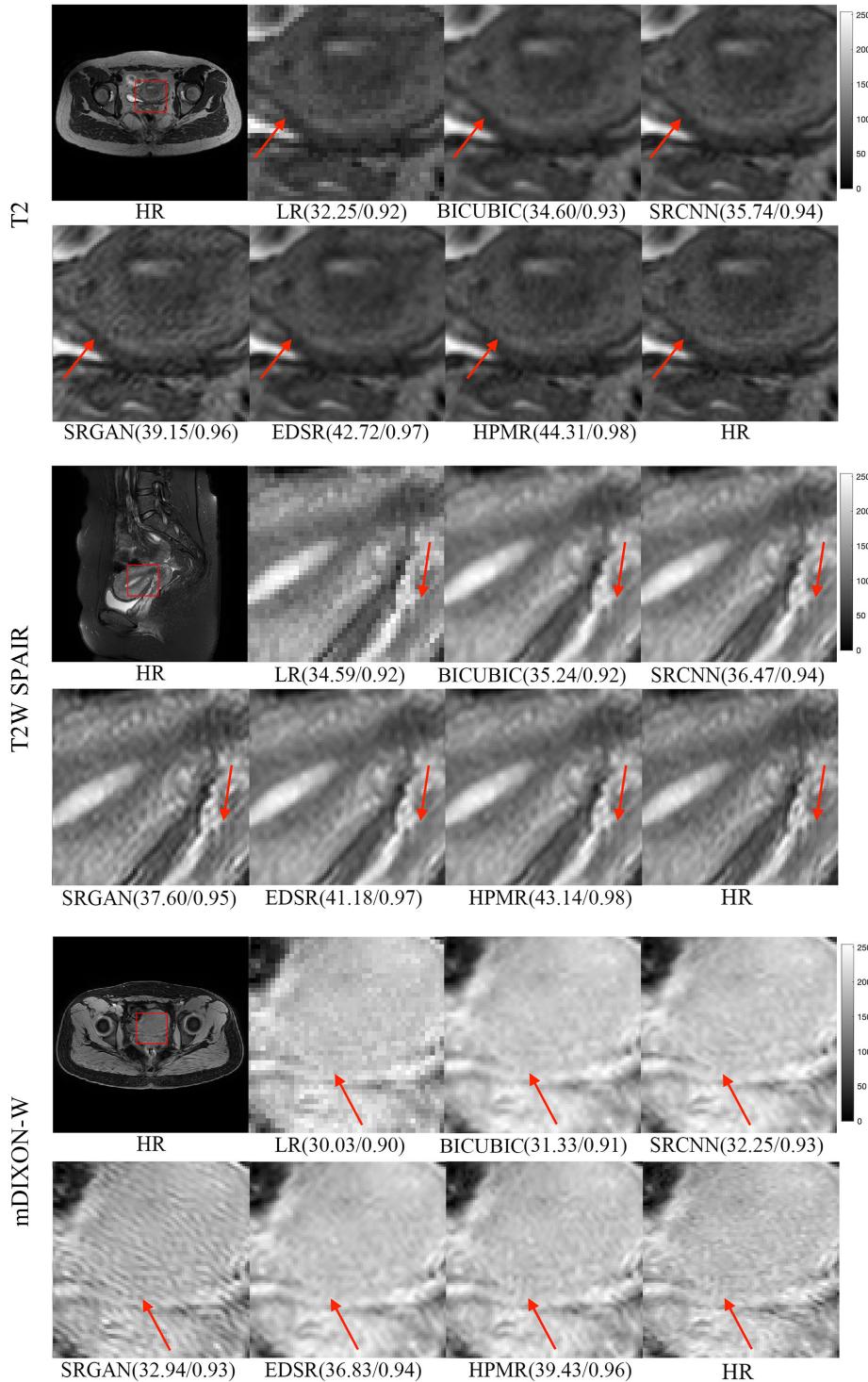


FIGURE 3. The results of each method when the upsampling factor is 2 \times . The figure shows the three datasets under healthy subjects pelvic, and the names of the corresponding datasets have been marked in the figure. Among them, the index value (PSNR/SSIM) of each method is marked in the figures.

still does not reach the desired effect, and the restoration of the details of the tumor needs to be improved. The HPMR method we proposed, from the junction of the uterus and uterine fibroids, such as the part pointed by the red arrow in Figure 6(b), can effectively distinguish the junction, which

is helpful for the doctor to diagnose the size of the patient's uterine fibroids and period. In addition, it can be seen from the image index values (PSNR/SSIM) marked in Figure 6(b) that the quality of the HR image reconstructed by HPMR is higher than other comparison methods.

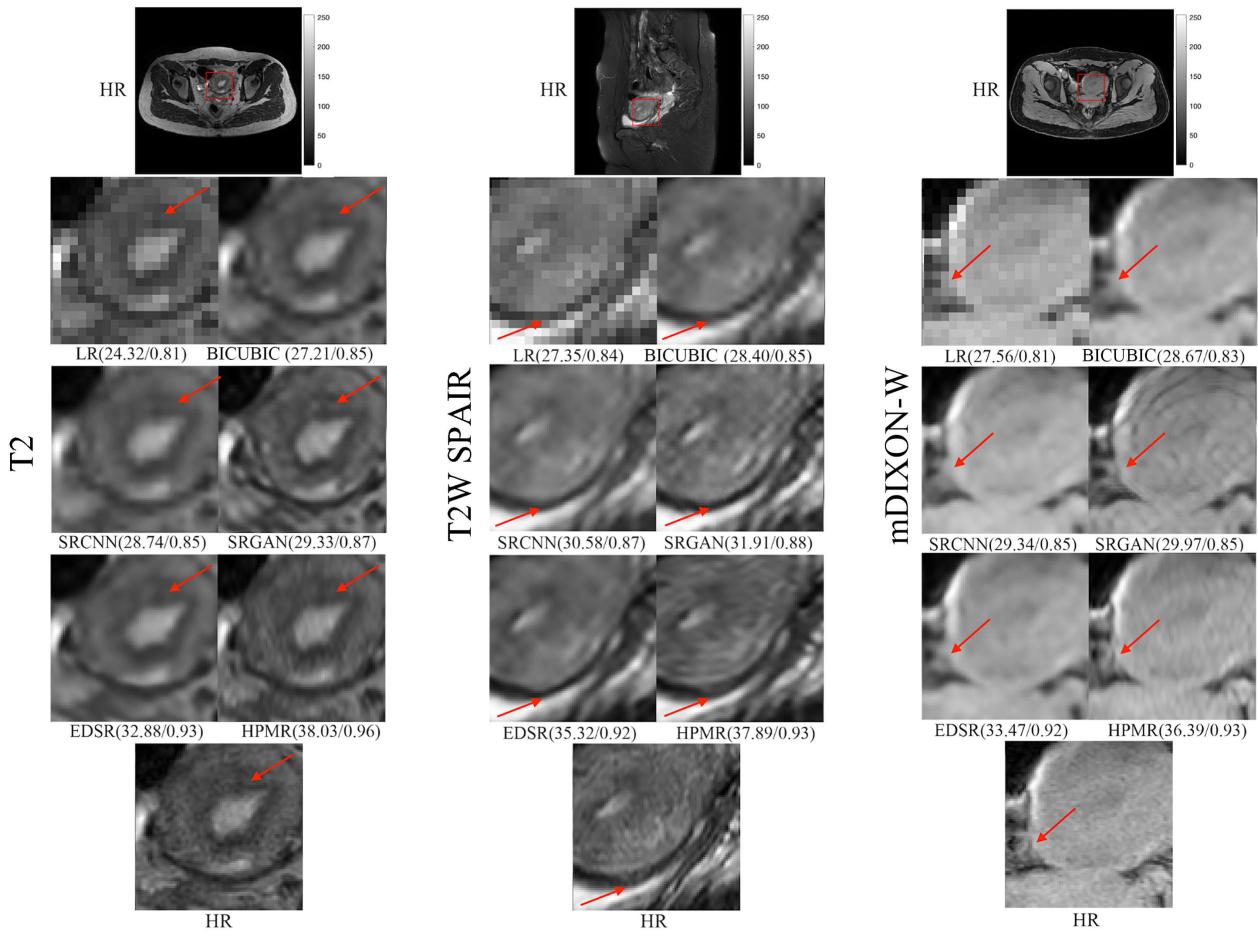


FIGURE 4. The results of each method when the upsampling factor is $4\times$. The first column is the T2 data; the second column is the T2W SPAIR data; the third column is the mDIXON-W data. The cropped part in HR_{re} reconstructed by each method intuitively shows the reconstruction effect of each method. It can be seen from the index values of PSNR and SSIM that our proposed method is better than other methods.

TABLE 2. KL divergence value of each image in figure 3 and figure 4.

Upsampling factor	Dataset	LR	BICUBIC	SRCNN	SRGAN	EDSR	HPMR
2×	T2	0.0333	0.0282	0.0142	0.0126	0.0098	0.0024
	T2W SPAIR	0.0245	0.0233	0.0119	0.0084	0.0063	0.0028
	mDIXON-W	0.0948	0.0879	0.0365	0.0251	0.0127	0.0067
4×	T2	0.1764	0.1620	0.0699	0.0585	0.0189	0.0066
	T2W SPAIR	0.0521	0.0473	0.0256	0.0217	0.0109	0.0058
	mDIXON-W	0.2157	0.2043	0.0901	0.0797	0.0350	0.0129

TABLE 3. KL divergence value of each image in figure 6.

Upsampling factor	Dataset	LR	BICUBIC	SRCNN	SRGAN	EDSR	HPMR
2×	T2	0.0274	0.0227	0.0201	0.0193	0.0151	0.0030
	T2W SPAIR	0.0172	0.0167	0.0141	0.0093	0.0073	0.0047
	mDIXON-W	0.0645	0.0579	0.0404	0.0244	0.0141	0.0089
4×	T2	0.1358	0.1246	0.0648	0.0613	0.0541	0.0068
	T2W SPAIR	0.0747	0.0676	0.0303	0.0344	0.0156	0.0076
	mDIXON-W	0.2341	0.2226	0.0834	0.0811	0.0315	0.0153

We calculated the KL divergence value of each result image in Figure 6, and the results are shown in Table 3. We can see that the HR image generated by HPMR is most similar to the original HR image. In addition, we also

analyzed the histogram of each result in Figure 6(a) when the upsampling factor is $2\times$. Figure 7 shows the histogram of each image in the T2 dataset, the histogram and quantitative indicators show that HPMR (kurtosis:5.38,

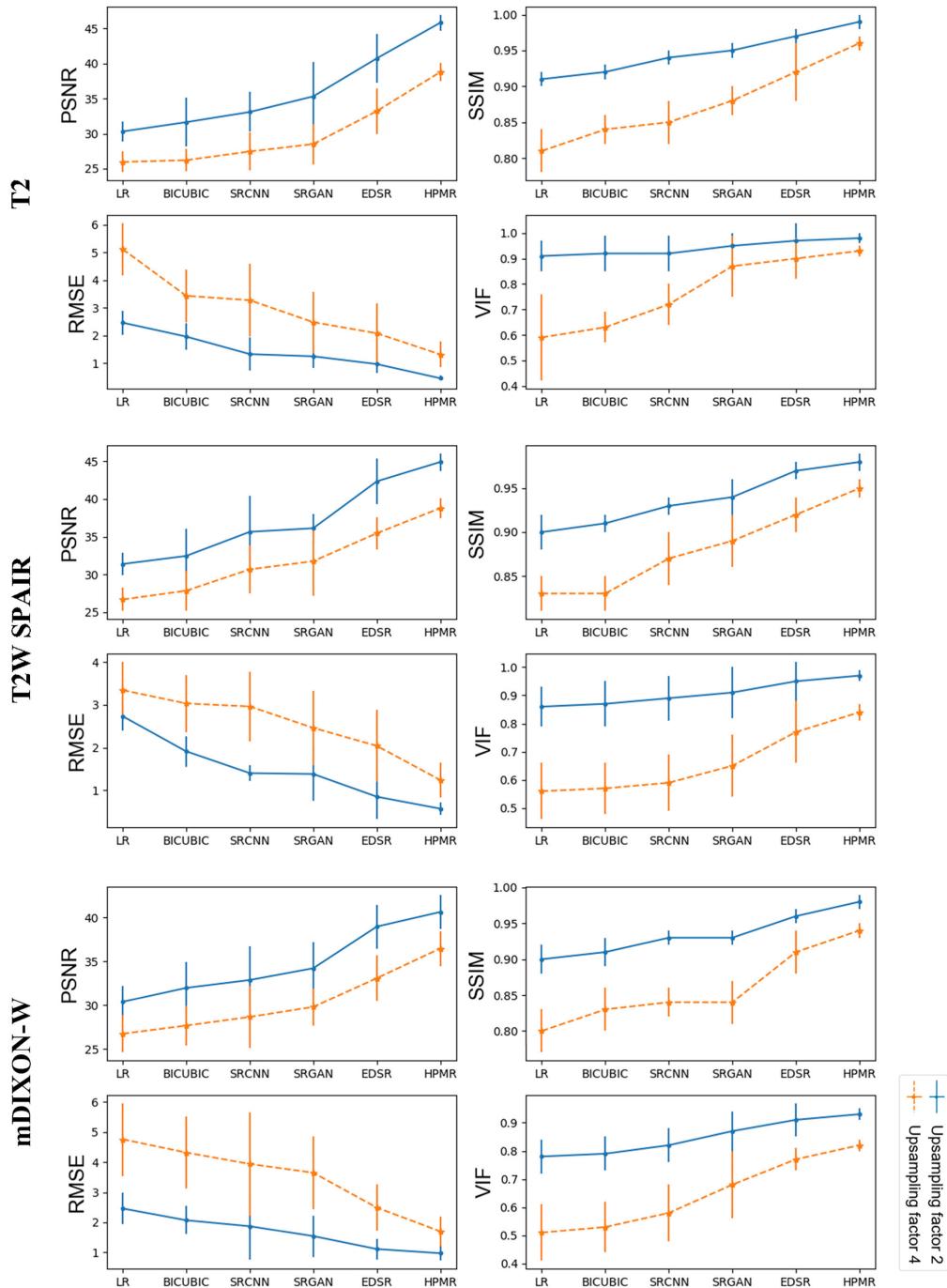


FIGURE 5. The mean value (std) of each index of the three healthy subjects pelvic dataset. The blue solid line indicates that the LR image is upsampled by 2x, and the orange dashed line indicates that the LR image is upsampled by 4x. PSNR: peak signal to noise ratio; SSIM: structural similarity; RMSE: root mean square error; VIF: variance inflation factor. The unit of RMSE is 10^{-2} .

skewness:1.35, mean:0.20, Standard:0.09) achieved the closest quantifications to the original HR(kurtosis:5.38, skewness:1.35, mean:0.20, Standard:0.10) as compared to other methods. Other histogram results are shown in the supplementary materials. In addition, we calculated the average value of the index on the three pelvic dataset of 7 patients. Table 4 shows the mean value of indicators for different

dataset under different upsampling factors. We can see that when using the patients dataset for testing, the HPMR method has a better performance than other comparison methods.

C. ABLATION EXPERIMENTS

In addition, we examined the relative contributions of the various components (cyclic loss, spatial attention, and channel

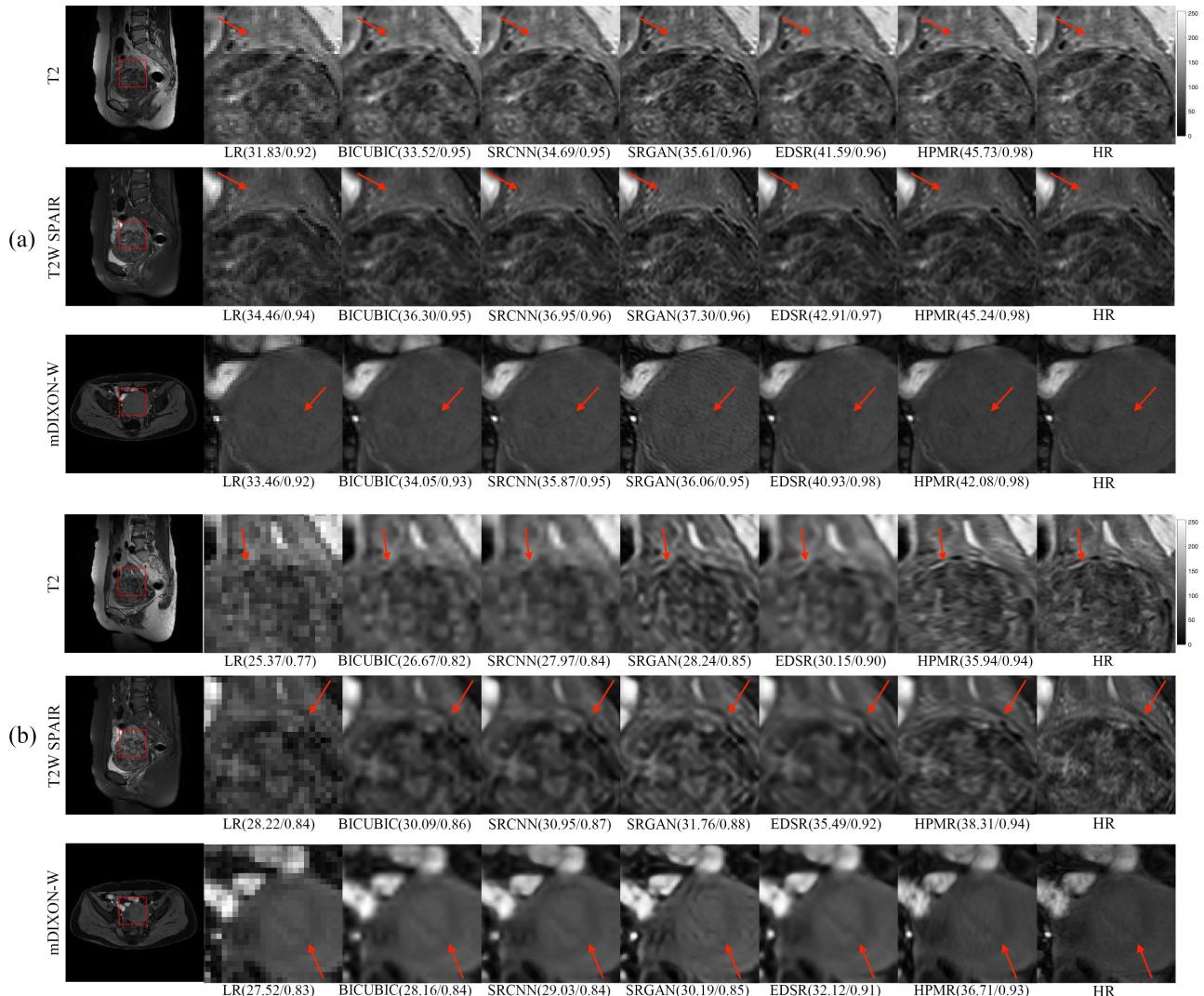


FIGURE 6. The results of HR_{re} reconstructed by each method when the upsampling factors are 2 \times and 4 \times of the patients pelvic data. Among them, (a) is the result when the upsampling factor is 2 \times , and (b) is the result when the upsampling factor is 4 \times . The first column of each row is the original HR image, and the rest are the crop images of HR_{re} reconstructed by each method. The red arrow in (a) indicates the details of HR_{re} ; the red arrow in (b) indicates the edge of the uterus and uterine fibroids in HR_{re} .

attention) of the HPMR method to network performance. Therefore, we set up three variant models, 1) using spatial attention and channel attention $HPMR_a$; 2) using cyclic loss and channel attention $HPMR_b$; 3) using cyclic loss and spatial attention $HPMR_c$. Experiments are performed on the T2 pelvic dataset of healthy subjects with an upsampling factor of 4 \times . Table 5 lists the mean (standard deviation) of each indicator. We can see that the contribution provided by the cyclic loss is the largest, and the spatial attention and channel attention have close contributions.

V. DISCUSSION

The purpose of this research is to reconstruct LR images into HR images and restore the details in the original HR images so that the generated HR images can be used in

clinical examinations. The significance of this research is that LR images can be obtained through fast MRI scans, which can shorten the scan time, and then use the trained network to reconstruct the LR images into HR images to achieve the purpose of accelerating the MRI scan. In addition, it is difficult to obtain original HR images for some parts, such as the abdomen, organs, and LR images caused by the patient's breathing movement. Therefore, the LR images obtained can be refined by the method we propose to facilitate doctors' diagnosis. The data used in our proposed method is female pelvic data, including normal and abnormal. Uterine fibroids are a common female disease. We use the network trained from normal data to test the abnormal data, and the reconstructed HR image has the expected quality. From Figure 6, it can be seen that the HR image reconstructed

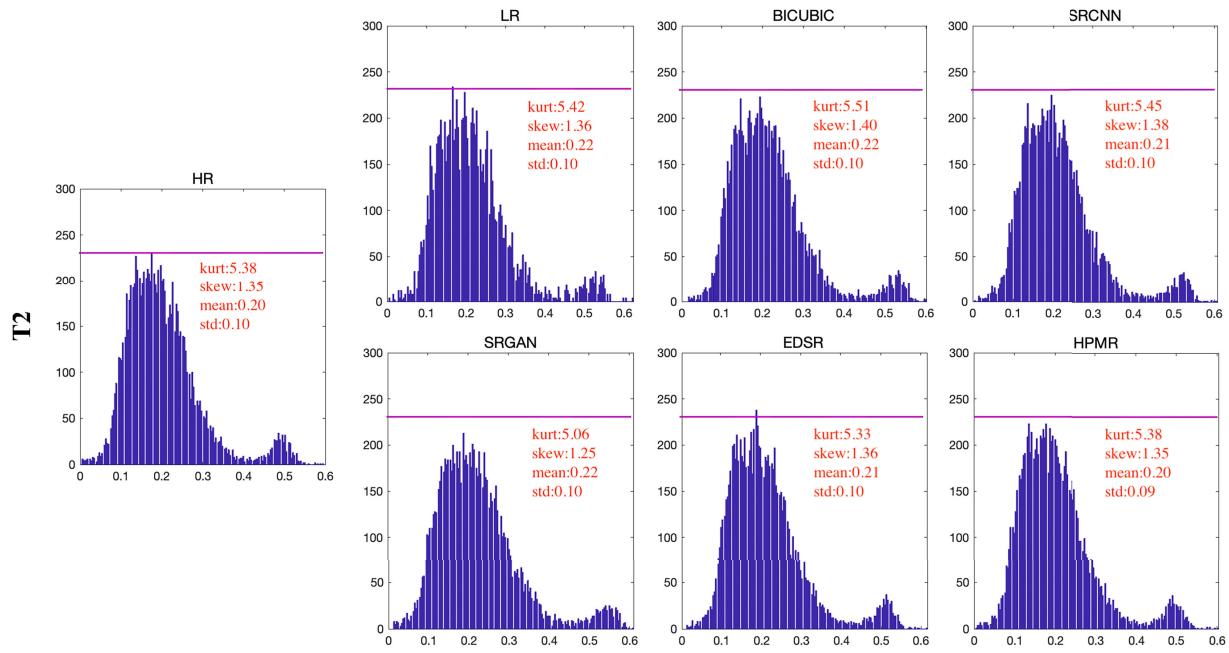


FIGURE 7. When the upsampling factor is 2×, the histogram of the HR image generated by each method under the T2 patients dataset in figure 6(a). Kurt: kurtosis; skew: skewness.

TABLE 4. Means (standard deviations) of each index in the patients pelvic dataset. UF: Upsampling factor; PSNR: peak signal-to-noise ratio; SSIM: structural similarity; RMSE: root mean square error; VIF: variance expansion coefficient. The unit of RMSE is 10^{-2} .

Dataset	UF	Method	PSNR	SSIM	RMSE	VIF
T2	2×	LR	32.14±0.94	0.92±0.01	2.48±0.28	0.90±0.04
		BICUBIC	33.48±1.04	0.95±0.01	1.93±0.26	0.91±0.04
		SRCCN	34.64±1.01	0.96±0.01	1.91±0.23	0.92±0.04
		SRGAN	35.36±1.00	0.96±0.01	1.83±0.35	0.93±0.05
		EDSR	40.22±3.41	0.97±0.01	1.68±0.57	0.99±0.04
		HPMR	46.15±0.71	0.99±0.02	0.42±0.15	0.99±0.01
	4×	LR	25.88±1.11	0.79±0.02	5.12±0.68	0.57±0.06
		BICUBIC	26.16±1.21	0.83±0.02	4.43±0.66	0.63±0.07
		SRCCN	27.45±1.15	0.85±0.02	3.65±0.83	0.65±0.14
		SRGAN	27.88±1.13	0.85±0.02	3.56±0.80	0.66±0.09
		EDSR	30.38±3.23	0.88±0.04	3.32±1.43	0.80±0.06
		HPMR	36.74±1.12	0.95±0.01	1.44±0.31	0.92±0.02
T2W SPAIR	2×	LR	34.12±1.34	0.93±0.01	1.77±0.28	0.86±0.05
		BICUBIC	37.20±1.45	0.95±0.01	1.41±0.25	0.87±0.04
		SRCCN	37.30±1.35	0.96±0.01	1.39±0.23	0.88±0.04
		SRGAN	38.94±1.38	0.96±0.01	1.13±0.38	0.94±0.05
		EDSR	41.84±2.17	0.97±0.01	0.85±0.24	0.96±0.04
		HPMR	44.53±1.12	0.98±0.01	0.59±0.10	0.96±0.01
	4×	LR	28.42±1.42	0.84±0.02	3.43±0.59	0.55±0.07
		BICUBIC	30.58±1.54	0.86±0.02	3.01±0.57	0.57±0.08
		SRCCN	30.57±1.46	0.88±0.02	3.16±0.65	0.57±0.77
		SRGAN	31.68±1.28	0.88±0.02	2.51±0.59	0.65±0.09
		EDSR	34.99±2.07	0.91±0.02	1.92±0.63	0.75±0.10
		HPMR	38.41±1.32	0.94±0.01	1.27±0.42	0.84±0.05
mDIXON-W	2×	LR	32.38±1.88	0.91±0.02	2.46±0.55	0.77±0.06
		BICUBIC	33.99±1.93	0.93±0.01	2.07±0.49	0.79±0.06
		SRCCN	34.14±1.86	0.94±0.01	2.03±0.46	0.81±0.07
		SRGAN	35.98±1.78	0.95±0.02	1.92±0.68	0.91±0.06
		EDSR	39.40±2.19	0.96±0.01	1.13±0.31	0.93±0.06
		HPMR	40.57±1.91	0.97±0.01	0.96±0.23	0.96±0.02
	4×	LR	27.19±2.01	0.81±0.03	4.49±1.11	0.48±0.09
		BICUBIC	28.17±2.11	0.83±0.03	4.04±1.06	0.49±0.10
		SRCCN	28.96±2.02	0.85±0.02	4.03±1.10	0.50±0.10
		SRGAN	29.94±1.97	0.86±0.02	3.85±1.21	0.63±0.13
		EDSR	32.76±2.58	0.92±0.02	2.46±0.73	0.75±0.16
		HPMR	36.23±1.97	0.94±0.01	1.70±0.53	0.81±0.03

by HPMR can clearly distinguish the uterine part and the uterine fibroids, and restore the detailed information, which is

beneficial to the diagnosis of doctors. Our method is expected to be used in clinical medicine. SRGAN uses the GAN for the

TABLE 5. Reconstruction performance in HPMR ablation experiments. The mean value (standard deviation) of each indicator in the T2 pelvic dataset of healthy subjects with an upsampling factor of 4 \times . The unit of RMSE is 10 $^{-2}$; + means used component; – means unused component.

	Method			Metric			
	Cyclic loss	Spatial attention	Channel attention	PSNR	SSIM	RMSE	VIF
HPMR _a	-	+	+	35.96±1.70	0.94±0.02	1.80±0.71	0.87±0.03
HPMR _b	+	-	+	37.71±1.31	0.95±0.01	1.52±0.66	0.91±0.01
HPMR _c	+	+	-	37.66±1.29	0.95±0.01	1.47±0.47	0.90±0.02
HPMR _f	+	+	+	38.76±1.30	0.96±0.01	1.32±0.46	0.93±0.02

super-resolution method, and EDSR uses a deeper network in the super-resolution method. Both methods can excellently reconstruct LR images into HR images. Based on these two points, we chose to use GAN, and use two cascaded residual u-net as the generator. Each encoder-decoder block in each residual u-net is composed of small residual blocks connected by short hops to increase the depth of the network. In addition to these two points, our proposed method also uses cyclic loss and attention mechanisms. The cyclic loss can make the network use $KS_{lr}[i]$ and $KS_{re}[i]$ cycle in the frequency domain in the case of limited datasets so that the difference between them is smaller so that the network learns more image details. Through the cycle of $HR_{re}[i]$ and $HR_{gt}[i]$ in the image domain, the HR_{re} generated by G is more and more similar to the real HR_{gt} . Cyclic loss can play a role in expanding the datasets. We use cyclic loss in HPMR to establish the mapping between LR and reconstructed HR_{re} as completely as possible. Through this mapping, the network can learn more details. The attention mechanism is to simulate the human senses, so that the system can focus on effective information, and can extract the characteristics of the image. What we use in HPMR is soft attention, which focuses on the characteristics of the spatial or channel, and can calculate the gradient through a neural network. In the previous research [41], the attention mechanism was used in residual u-net for MR image reconstruction, and higher quality pictures were obtained. In our proposed method, full-attention is to assign spatial attention weights and channel-wise attention weights to feature maps so that the network can learn the details in the image more effectively, and then generate high-quality HR_{re} . The feasibility of using spatial attention and channel-wise attention simultaneously in neural networks has been confirmed in the [49] method. According to the tests in the normal pelvic datasets and the abnormal pelvic datasets, the quality of the HR_{re} reconstructed by HPMR is the best, and when the upsampling factor is 4 \times , the obtained HR_{re} can restore the details of the original HR to the greatest extent. Therefore, our method is expected to be used in clinical medicine in the future.

Although our proposed method can achieve the expected results, there is still room for improvement. When we train the network, we need the original HR image as a label. We hope that in future work, LR images can be used as labels in between, that is, completely separate from the original HR images when training the network. This problem can use unsupervised learning methods, when training the network, let the network learn effective information in the LR label.

VI. CONCLUSION

The generative adversarial network with cyclic loss and attention is proposed by us for HR reconstruction of the pelvic datasets, including the healthy subjects pelvic datasets and the patients pelvic datasets. Each dataset is obtained through different MR sequences. This method can upsampling LR images by 2 \times and 4 \times , and the reconstructed HR_{re} can reach the quality required by radiologists for diagnosis.

REFERENCES

- [1] M. Lustig, D. L. Donoho, J. M. Santos, and J. M. Pauly, “Compressed sensing MRI,” *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 72–82, Mar. 2008.
- [2] D. L. Donoho, “Compressed sensing,” *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.
- [3] T. Eo, Y. Jun, T. Kim, J. Jang, H.-J. Lee, and D. Hwang, “KIKI-Net: Cross-domain convolutional neural networks for reconstructing undersampled magnetic resonance images,” *Magn. Reson. Med.*, vol. 80, no. 5, pp. 2188–2201, 2018.
- [4] F. Knoll, K. Hammernik, C. Zhang, S. Moeller, T. Pock, D. K. Sodickson, and M. Akcakaya, “Deep-learning methods for parallel magnetic resonance imaging reconstruction: A survey of the current approaches, trends, and issues,” *IEEE Signal Process. Mag.*, vol. 37, no. 1, pp. 128–140, Jan. 2020.
- [5] K. G. Hollingsworth, “Reducing acquisition time in clinical MRI by data undersampling and compressed sensing reconstruction,” *Phys. Med. Biol.*, vol. 60, no. 21, pp. R297–R322, Nov. 2015.
- [6] C. Dong, C. C. Loy, K. He, and X. Tang, “Learning a deep convolutional network for image super-resolution,” in *Proc. Eur. Conf. Comput. Vis.* Cham, Switzerland: Springer, 2014, pp. 184–199.
- [7] C. Dong, C. C. Loy, and X. Tang, “Accelerating the super-resolution convolutional neural network,” in *Proc. Eur. Conf. Comput. Vis.* Cham, Switzerland: Springer, 2016, pp. 391–407.
- [8] W. Shi, J. Caballero, F. Huszár, J. Totz, A. P. Aitken, R. Bishop, D. Rueckert, and Z. Wang, “Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 1874–1883.
- [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] J. Kim, J. K. Lee, and K. M. Lee, “Deeply-recursive convolutional network for image super-resolution,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 1637–1645.
- [11] X.-J. Mao, C. Shen, and Y.-B. Yang, “Image restoration using convolutional auto-encoders with symmetric skip connections,” 2016, *arXiv:1606.08921*. [Online]. Available: <http://arxiv.org/abs/1606.08921>
- [12] Y. Tai, J. Yang, and X. Liu, “Image super-resolution via deep recursive residual network,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 3147–3155.
- [13] T. Tong, G. Li, X. Liu, and Q. Gao, “Image super-resolution using dense skip connections,” in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 4799–4807.
- [14] 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. (CVPR)*, Jul. 2017, pp. 4681–4690.

- [15] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," *Adv. Neural Inf. Process. Syst.*, vol. 27, pp. 2672–2680, Dec. 2014.
- [16] 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 (CVPRW)*, Jul. 2017, pp. 136–144.
- [17] D. Qiu, S. Zhang, Y. Liu, J. Zhu, and L. Zheng, "Super-resolution reconstruction of knee magnetic resonance imaging based on deep learning," *Comput. Methods Programs Biomed.*, vol. 187, Apr. 2020, Art. no. 105059.
- [18] Q. Lyu, H. Shan, and G. Wang, "MRI super-resolution with ensemble learning and complementary priors," *IEEE Trans. Comput. Imag.*, vol. 6, pp. 615–624, 2020.
- [19] J. A. Steeden, M. Quail, A. Gotschy, K. H. Mortensen, A. Hauptmann, S. Arridge, R. Jones, and V. Muthurangu, "Rapid whole-heart CMR with single volume super-resolution," *J. Cardiovascular Magn. Reson.*, vol. 22, no. 1, pp. 1–13, Dec. 2020.
- [20] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* Cham, Switzerland: Springer, 2015, pp. 234–241.
- [21] Y. Zheng, B. Zhen, A. Chen, F. Qi, X. Hao, and B. Qiu, "A hybrid convolutional neural network for super-resolution reconstruction of MR images," *Med. Phys.*, vol. 47, no. 7, pp. 3013–3022, Jul. 2020.
- [22] J. Du, Z. He, L. Wang, A. Gholipour, Z. Zhou, D. Chen, and Y. Jia, "Super-resolution reconstruction of single anisotropic 3D MR images using residual convolutional neural network," *Neurocomputing*, vol. 392, pp. 209–220, Jun. 2020.
- [23] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel, "Backpropagation applied to handwritten zip code recognition," *Neural Comput.*, vol. 1, no. 4, pp. 541–551, Dec. 1989.
- [24] K. Sun, L. Qu, C. Lian, Y. Pan, D. Hu, B. Xia, X. Li, W. Chai, F. Yan, and D. Shen, "High-resolution breast MRI reconstruction using a deep convolutional generative adversarial network," *J. Magn. Reson. Imag.*, vol. 52, no. 6, pp. 1852–1858, Dec. 2020.
- [25] A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," 2015, *arXiv:1511.06434*. [Online]. Available: <http://arxiv.org/abs/1511.06434>
- [26] A. S. Chaudhari, Z. Fang, F. Kogan, J. Wood, K. J. Stevens, E. K. Gibbons, J. H. Lee, G. E. Gold, and B. A. Hargreaves, "Super-resolution musculoskeletal MRI using deep learning," *Magn. Reson. Med.*, vol. 80, no. 5, pp. 2139–2154, Nov. 2018.
- [27] Q. Lyu, H. Shan, C. Steber, C. Helis, C. Whitlow, M. Chan, and G. Wang, "Multi-contrast super-resolution MRI through a progressive network," *IEEE Trans. Med. Imag.*, vol. 39, no. 9, pp. 2738–2749, Sep. 2020.
- [28] G. Yang, S. Yu, H. Dong, G. Slabaugh, P. L. Dragotti, X. Ye, F. Liu, S. Arridge, J. Keegan, Y. Guo, and D. Firmin, "DAGAN: Deep de-aliasing generative adversarial networks for fast compressed sensing MRI reconstruction," *IEEE Trans. Med. Imag.*, vol. 37, no. 6, pp. 1310–1321, Jun. 2018.
- [29] M. Mardani, E. Gong, J. Y. Cheng, S. S. Vasanaawala, G. Zaharchuk, L. Xing, and J. M. Pauly, "Deep generative adversarial neural networks for compressive sensing MRI," *IEEE Trans. Med. Imaging*, vol. 38, no. 1, pp. 167–179, Jan. 2019.
- [30] T. M. Quan, T. Nguyen-Duc, and W.-K. Jeong, "Compressed sensing MRI reconstruction using a generative adversarial network with a cyclic loss," *IEEE Trans. Med. Imag.*, vol. 37, no. 6, pp. 1488–1497, Jun. 2018.
- [31] J. Lv, P. Wang, X. Tong, and C. Wang, "Parallel imaging with a combination of sensitivity encoding and generative adversarial networks," *Quant. Imag. Med. Surgery*, vol. 10, no. 12, p. 2260, 2020.
- [32] J. Lv, C. Wang, and G. Yang, "PIC-GAN: A parallel imaging coupled generative adversarial network for accelerated multi-channel MRI reconstruction," *Diagnostics*, vol. 11, no. 1, p. 61, Jan. 2021.
- [33] Y. Korkmaz, S. U. Dar, M. Yurt, M. Özbeý, and T. Çukur, "Unsupervised MRI reconstruction via zero-shot learned adversarial transformers," 2021, *arXiv:2105.08059*. [Online]. Available: <http://arxiv.org/abs/2105.08059>
- [34] S. U. H. Dar, M. Yurt, M. Shahdloo, M. E. Ildiz, B. Tinaz, and T. Cukur, "Prior-guided image reconstruction for accelerated multi-contrast MRI via generative adversarial networks," *IEEE J. Sel. Topics Signal Process.*, vol. 14, no. 6, pp. 1072–1087, Oct. 2020.
- [35] S. U. H. Dar, M. Yurt, L. Karacan, A. Erdem, E. Erdem, and T. Çukur, "Image synthesis in multi-contrast MRI with conditional generative adversarial networks," *IEEE Trans. Med. Imag.*, vol. 38, no. 10, pp. 2375–2388, Oct. 2019.
- [36] M. Yurt, S. U. H. Dar, M. Özbeý, B. Tinaz, K. K. Oğuz, and T. Çukur, "Semi-supervised learning of mutually accelerated MRI synthesis without fully-sampled ground truths," 2020, *arXiv:2011.14347*. [Online]. Available: <http://arxiv.org/abs/2011.14347>
- [37] P. Shamsolmoali, M. Zareapoor, E. Granger, H. Zhou, R. Wang, M. E. Celebi, and J. Yang, "Image synthesis with adversarial networks: A comprehensive survey and case studies," *Inf. Fusion*, vol. 72, pp. 126–146, Aug. 2021.
- [38] M. Arjovsky, S. Chintala, and L. Bottou, "Wasserstein GAN," 2017, *arXiv:1701.07875*. [Online]. Available: <http://arxiv.org/abs/1701.07875>
- [39] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 2223–2232.
- [40] Y. Guo, C. Wang, H. Zhang, and G. Yang, "Deep attentive wasserstein generative adversarial networks for mri reconstruction with recurrent context-awareness," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* Cham, Switzerland: Springer, 2020, pp. 167–177.
- [41] Q. Huang, D. Yang, P. Wu, H. Qu, J. Yi, and D. Metaxas, "MRI reconstruction via cascaded channel-wise attention network," in *Proc. IEEE 16th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2019, pp. 1622–1626.
- [42] Y. Wu, Y. Ma, J. Liu, J. Du, and L. Xing, "Self-attention convolutional neural network for improved MR image reconstruction," *Inf. Sci.*, vol. 490, pp. 317–328, Jul. 2019.
- [43] Z. Yuan, M. Jiang, Y. Wang, B. Wei, Y. Li, P. Wang, W. Menpes-Smith, Z. Niu, and G. Yang, "SARA-GAN: Self-attention and relative average discriminator based generative adversarial networks for fast compressed sensing MRI reconstruction," *Frontiers Neuroinform.*, vol. 14, Nov. 2020, Art. no. 611666.
- [44] Z. Huang, Z. Chen, Q. Zhang, G. Quan, M. Ji, C. Zhang, Y. Yang, X. Liu, D. Liang, H. Zheng, and Z. Hu, "CaGAN: A cycle-consistent generative adversarial network with attention for low-dose CT imaging," *IEEE Trans. Comput. Imag.*, vol. 6, pp. 1203–1218, 2020.
- [45] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.
- [46] (2016). *Tensorpack*. [Online]. Available: <https://github.com/tensorpack/>
- [47] Y. Yang, J. Sun, H. Li, and Z. Xu, "Deep ADMM-Net for compressive sensing MRI," *Adv. Neural Inf. Process. Syst.*, vol. 29, pp. 10–18, 2016.
- [48] Y. Yang, J. Sun, H. Li, and Z. Xu, "ADMM-Net: A deep learning approach for compressive sensing MRI," 2017, *arXiv:1705.06869*. [Online]. Available: <http://arxiv.org/abs/1705.06869>
- [49] L. Chen, H. Zhang, J. Xiao, L. Nie, J. Shao, W. Liu, and T.-S. Chua, "SCA-CNN: Spatial and channel-wise attention in convolutional networks for image captioning," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 5659–5667.



GUANGYUAN LI was born in Jining, Shandong, China, in 1997. He is currently pursuing the M.S. degree with the School of Computer and Control Engineering, Yantai University. His research interests include deep learning and medical image reconstruction.



JUN LV was born in Yantai, Shandong, China, in 1990. She received the B.S. degree in intelligence science and technology from Xidian University, in 2013, and the Ph.D. degree in biomechanics and medical engineering from Peking University, in 2018. She is currently a Lecturer with the School of Computer and Control Engineering, Yantai University. Her research interests include medical image reconstruction, image registration, and image segmentation.



XIANGRONG TONG received the Ph.D. degree in computer science and technology from Beijing Jiaotong University. He is currently a Full Professor with Yantai University. He has published more than 50 articles in well-known journals and conferences. He has also presided and joined three national projects and three provincial projects. His research interests include computer science, intelligent information processing, and social networks.



GUANG YANG (Senior Member, IEEE) received the M.Sc. degree in vision imaging and virtual environments from the Department of Computer Science, in 2006, and the Ph.D. degree in medical image analysis jointly from the CMIC, Department of Computer Science and Medical Physics, in 2012, both from University College London. He is currently a Tenured Senior Research Fellow with NHLI, Imperial College London and a Honorary Senior Lecturer with the School of Biomedical Engineering and Imaging Sciences, King's College London. He is the Head of the Smart Imaging Laboratory funded by UKRI, BHF and ERC. His research interests include pattern recognition, machine learning, and medical image processing and analysis.



CHENGYAN WANG received the B.S. degree in biomedical engineering from the Beijing Institute of Technology, Beijing, China, in 2012, and the Ph.D. degree in biomechanics and medical engineering from Peking University, Beijing, in 2017.

From 2017 to 2018, he was a Postdoctoral Fellow with the Institute of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai, China. From 2018 to 2019, he was a Postdoctoral Fellow with the Institute of Electronic Engineering, University of Illinois at Urbana-Champaign, IL, USA. Since 2019, he has been an Assistant Professor with the Human Phenome Institute, Fudan University, Shanghai. His research interests include image reconstruction, medical image analysis, and deep learning.