CT-Guided, Unsupervised Super-Resolution Reconstruction of Single 3D Magnetic Resonance Image

Jiale Wang[1], Alexander F. Heimann[2], Moritz Tannast[2], and Guoyan Zheng[1(✉)]

[1] Institute of Medical Robotics, School of Biomedical Engineering, Shanghai Jiao Tong University, No. 800, Dongchuan Road, Shanghai 200240, China
guoyan.zheng@sjtu.edu.cn
[2] Department of Orthopaedic Surgery, HFR Cantonal Hospital, University of Fribourg, Fribourg, Switzerland

Abstract. Deep learning-based algorithms for single MR image (MRI) super-resolution have shown great potential in enhancing the resolution of low-quality images. However, many of these methods rely on supervised training with paired low-resolution (LR) and high-resolution (HR) MR images, which can be difficult to obtain in clinical settings. This is because acquiring HR MR images in clinical settings requires a significant amount of time. In contrast, HR CT images are acquired in clinical routine. In this paper, we propose a CT-guided, unsupervised MRI super-resolution reconstruction method based on joint cross-modality image translation and super-resolution reconstruction, eliminating the requirement of high-resolution MRI for training. The proposed approach is validated on two datasets respectively acquired from two different clinical sites. Well-established metrics including Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index Metrics (SSIM), and Learned Perceptual Image Patch Similarity (LPIPS) are used to assess the performance of the proposed method. Our method achieved an average PSNR of 32.23, an average SSIM of 0.90 and an average LPIPS of 0.14 when evaluated on data of the first site. An average PSNR of 30.58, an average SSIM of 0.88, and an average LPIPS of 0.10 were achieved by our method when evaluated on data of the second site.

Keywords: Unsupervised image super-resolution · Cross-modality image translation · CT-Guided · Magnetic resonance imaging

1 Introduction

High-resolution magnetic resonance (MR) images (MRI) provide a wealth of structural details, which facilitate early and precise diagnosis [1]. However, images obtained in clinical practice are anisotropic due to the limitation of scan

time and signal-noise ratio [2]. In order to speed up clinical scanning procedures, only a limited number of two-dimensional (2D) slices are acquired, despite the fact that the interested anatomical structures are in three-dimensional (3D). The acquired medical images have low inter-plane resolution, i.e., large spacing between slices. Such anisotropic images will lead to misdiagnosis and can greatly impact the performance of various clinical tasks, including computer-aided diagnosis and computer-assisted interventions. Therefore, we investigate the problem of reducing the slice spacing [3] via super-resolution (SR) reconstruction. Specifically, we refer to the image with large slice spacing as a low-resolution (LR) image and the image with small slice spacing as a high-resolution (HR) image. Our goal is to reconstruct the HR image from the LR input, which is an ill-posed inverse problem and presents significant challenges.

Deep learning-based algorithms for single MR image super-resolution show great potential in restoration of HR images from LR inputs [4]. Pham et al. [5] proposed the SRCNN method, which applied convolutional neural networks (CNN) to image super-resolution of MRI and achieved a better performance than the conventional methods, such as B-spline interpolation and low-rank total variation (LRTV) [6] method. Chaudhariet al. [7] proposed a 3D residual network, which learned the residual-based transformations between paired LR and HR images for the SR reconstruction of MRI. Chen et al. [8] proposed a densely connected super-resolution network (DCSRN), which reused the block features through the dense connection in the SR reconstruction of MRI. Chen et al. [9] extended this work by using generative adversarial network (GAN) [10] in SR reconstruction of MRI in order to improve the realism of the recovered images. Feng et al. [11,12] proposed a multi-contrast MRI SR method, which aimed to learn clearer anatomical structure and edge information with the help of auxiliary contrast MRI. Despite significant progress, however, there are still spaces for further improvement. Most networks require a large amount of paired LR and HR MR images for training, which are unrealistic in clinical practice. To address the challenge of organizing paired images, methods based on unpaired images have been proposed [13,14]. However, HR MR images are still difficult to obtain, as acquiring HR MR images in clinical settings requires a significant amount of time. In contrast, CT images are acquired in clinical routine. Therefore, it is of great significance to use HR CT images as a guidance to synthesize HR MR images from LR MR images.

To this end, we propose a CT-guided, unsupervised MRI super-resolution reconstruction method based on joint cross-modality image translation (CIT) and super-resolution reconstruction, eliminating the requirement of HR MR images for training. Specifically, our network design features a super-resolution Network (SRNet) and a cross-modality image translation network (CITNet) based on disentanged representation learning. After pretraining, the SRNet can generate pseudo HR MR images from LR MR images. The generated pseudo HR MR images are then taken together with the HR CT images as the input to the CITNet, which can generate quality-improved pseudo HR MR images by combining disentangled content code of the input CT data with the attribute
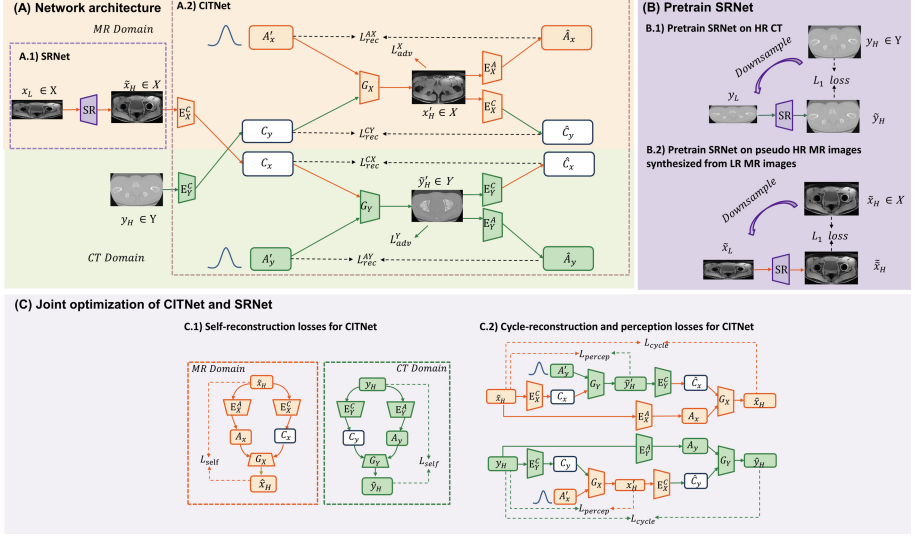
Fig. 1. A schematic illustration of our CT-guided, unsupervised MRI super-resolution reconstruction method. (A) Network architecture, including a SRNet and a CITNet; (B) Pretraining the SRNet; and (C) Joint optimization of the CITNet and the SRNet. Different colors represent different domains, i.e., orange represents the MR domain, green represents the CT domain, and white shows the shared content space. (Color figure online)

code of the input pseudo HR MR images. Joint optimization of the CITNet and the SRNet leads to better and better pseudo HR MR image generation. When converged, we can use the SRNet to generate high-quality pseudo HR MR images from given LR MR images. The contributions of our work can be summarized as follows:

– We propose a CT-guided, unsupervised MRI super-resolution reconstruction method based on joint cross-modality image translation and super-resolution reconstruction, eliminating the requirement of HR MRI for training. Our cross-modality image translation is based on disentangled representation leanring.
– Our network design features a SRNet and a CITNet. They work jointly to generate high-quality pseudo HR MR images from given LR MR images. Concretely, a better trained SRNet will help to generate a better input to the CITNet. On the other hand, the CITNet, taking the SRNet-generated pseudo HR MR images and the HR CT images as input, provides better supervision of the SRNet training. Joint optimization of the CITNet and the SRNet leads to the generation of high-quality pseudo HR MR image at the end.
– We validate the proposed method on two datasets collected from two different clinical centers.

2   Methodology

Figure 1 presents a schematic illustration of our CT-guided, unsupervised MRI super-resolution reconstruction method. It features two networks: the SRNet and the CITNet (Fig. 1-(A)). Figure 1-(B) shows how to pretrain the SRNet while Fig. 1-(C) presents how to conduct joint optimization. Below we first present the design of the SRNet and the CITNet, followed by a description of the traing strategy.

2.1   Super-Resolution Network (SRNet)

We choose to use the residual dense network (RDN) as the SRNet. The RDN utilizes cascaded residual dense blocks (RDBs), a powerful convolutional block that leverages residual and dense connections to fully aggregate hierarchical features. For further details on the structure of the RDN, please refer to the original paper [15]. Mathematically, we denote the SRNet as $\mathcal{F}_s(\cdot; \Theta_s)$ with trainable parameters Θ_s.

2.2   Cross-Modality Image Translation Network (CITNet)

The CITNet is inspired by MUNIT [16]. As depicted in Fig. 1-(A.2), it comprises two content encoders $\{E^{\mathcal{C}}_{\mathcal{X}}, E^{\mathcal{C}}_{\mathcal{Y}}\}$, two attribute encoders $\{E^{\mathcal{A}}_{\mathcal{X}}, E^{\mathcal{A}}_{\mathcal{Y}}\}$, and two generators $\{G_{\mathcal{X}}, G_{\mathcal{Y}}\}$. The encoder in each domain disentangles an input image separately into a domain-invariant content space $\mathcal{C}$ and a domain-specific attribute space $\mathcal{A}$. And the generator networks combine a content code with an attribute code to generate translated images in the target domain. For instance, when translating CT image $y_H \in \mathcal{Y}$ to MR image $x^{'}_H \in \mathcal{X}$, we first randomly sample from the prior distribution $p(\mathcal{A}^{'}_x) \sim \mathcal{N}(0, \mathbf{I})$ to obtain an MRI attribute code $\mathcal{A}^{'}_x$, which is empirically set as a 8-bit vector. We then combine $\mathcal{A}^{'}_x$ with the disentangled content code of the CT image $\mathcal{C}_y = E^{\mathcal{C}}_{\mathcal{Y}}(y_H)$ to generate the translated MRI image $x^{'}_H \in \mathcal{X}$ through the generator $G_{\mathcal{X}}$. Similarly, we can get the the translated CT image $\tilde{y}^{'}_H \in \mathcal{Y}$ through the generator $G_{\mathcal{Y}}(\mathcal{C}_x, \mathcal{A}^{'}_y)$, where $\mathcal{C}_x = E^{\mathcal{C}}_{\mathcal{X}}(\mathcal{F}_s(x_L; \Theta_s))$ and $\mathcal{A}^{'}_y$ is also sampled from the prior distribution $p(\mathcal{A}^{'}_y) \sim \mathcal{N}(0, \mathbf{I})$.

   Disentangled Representation Learning. Cross-modality image translation is based on disentangled representation learning, trained with self- and cross-cycle reconstruction losses. As shown in Fig. 1-(C.1, C.2), the self-reconstruction loss L_{self} is utilized to regularize the training when the content and attribute code originate from the same domain, whereas the cross-cycle consistency loss L_{cycle} is used when the content and attribute code come from different domains. The self-reconstruction and cross-cycle reconstruction losses are defined as follows:

$$L_{\text{self}} = \left\| G_{\mathcal{X}}\left(E^{\mathcal{C}}_{\mathcal{X}}(\tilde{x}_H), E^{\mathcal{A}}_{\mathcal{X}}(\tilde{x}_H)\right) - \tilde{x}_H \right\|_1 + \left\| G_{\mathcal{Y}}\left(E^{\mathcal{C}}_{\mathcal{Y}}(y_H), E^{\mathcal{A}}_{\mathcal{Y}}(y_H)\right) - y_H \right\|_1 \quad (1)$$

$$L_{\text{cycle}} = \| G_{\mathcal{X}}(E^{\mathcal{C}}_{\mathcal{Y}}(\tilde{y}^{'}_H), E^{\mathcal{A}}_{\mathcal{X}}(\tilde{x}_H)) - \tilde{x}_H \|_1 + \| G_{\mathcal{Y}}(E^{\mathcal{C}}_{\mathcal{X}}(x^{'}_H), E^{\mathcal{A}}_{\mathcal{Y}}(y_H)) - y_H \|_1 \quad (2)$$

where $\tilde{x}_H = \mathcal{F}_s(x_L; \Theta_s)$, $x'_H = G_{\mathcal{X}}(E^{\mathcal{C}}_{\mathcal{Y}}(y_H), \mathcal{A}'_x)$, $\tilde{y}'_H = G_{\mathcal{Y}}(E^{\mathcal{C}}_{\mathcal{X}}(\tilde{x}_H), \mathcal{A}'_y)$. Specially, in the cross-cycle translation processes, we employ a latent reconstruction loss to maintain the invertible mapping between the image and the latent space. In details, we have:

$$L_{\text{latent}} = \|\hat{\mathcal{C}}_x - \mathcal{C}_x\|_1 + \|\hat{\mathcal{C}}_y - \mathcal{C}_y\|_1 + \|\hat{\mathcal{A}}_x - \mathcal{A}'_x\|_1 + \|\hat{\mathcal{A}}_y - \mathcal{A}'_y\|_1 \qquad (3)$$

We further use pretrained vgg16 network, denoted as $\phi(\cdot)$, to extract high-level features for computing the perceptual loss [17]:

$$L_{\text{percep}} = \frac{1}{CHW}\left\|\phi(\tilde{y}'_H) - \phi(\tilde{x}_H)\right\|^2_2 + \frac{1}{CHW}\left\|\phi(x'_H) - \phi(y_H)\right\|^2_2 \qquad (4)$$

where C, H, W indicate the channel number and the image size, respectively.

Adversarial Learning. As shown in Fig. 1-(A.2), we use GAN [10] to learn the translation between MR and CT image domains better. A GAN typically contains a generation network and a discrimination network. We use the discriminator $D_{\mathcal{X}}$ to judge whether the image is from MR image domain, and the discriminator $D_{\mathcal{Y}}$ to judge whether the image is from CT image domain. The auto-encoders try to generate the image of the target domain to fool the discriminators so that the distribution of the translated images can match that of the target images. The minmax game is trained by:

$$L^{\mathcal{X}}_{adv} = \mathbb{E}_{\tilde{x}_H \sim P_{\mathcal{X}}(\tilde{x}_H)}\left[\log D_{\mathcal{X}}(\tilde{x}_H)\right] + \mathbb{E}_{y_H \sim P_{\mathcal{Y}}(y_H)}\left[\log(1 - D_{\mathcal{X}}(x'_H))\right] \qquad (5)$$

$$L^{\mathcal{Y}}_{adv} = \mathbb{E}_{y_H \sim P_{\mathcal{Y}}(y_H)}\left[\log D_{\mathcal{Y}}(y_H)\right] + \mathbb{E}_{\tilde{x}_H \sim P_{\mathcal{X}}(\tilde{x}_H)}\left[\log(1 - D_{\mathcal{Y}}(\tilde{y}'_H))\right] \qquad (6)$$

Joint Optimization. The SRNet and the CITNet are jointly optimized by minimizing following loss function:

$$L_{disentangle} = \left(L^{\mathcal{X}}_{adv} + L^{\mathcal{Y}}_{adv}\right) + \lambda_1(L_{self} + L_{cycle}) + \lambda_2 L_{latent} + \lambda_3 L_{percep} \qquad (7)$$

where λ_1, λ_2, and λ_3 are parameters controlling the relative weights of different losses.

2.3   Training Strategy

Empirically, we found that training the network shown in Fig. 1-(A) end to end did not converge. We thus design the following three-stage training strategy.

Stage 1. Let's denote the downsampling function as $\mathcal{D}(\cdot)$. In this stage, we pretrain the SRNet using the HR CT images, as shown in Fig. 1-(B.1), for T iterations. At each iteration, we sample a batch of HR CT images. We then downsample the sampled HR CT images y_H to get the paired LR CT images $y_L = \mathcal{D}(y_H)$. The SRNet is trained with the paired LR-HR CT images by minimizing $L1$ loss $\|y_H - \mathcal{F}_s(\mathcal{D}(y_H); \Theta_s)\|_1$. In this stage, we are aiming to train the SRNet to learn the upsampling kernels.

Stage 2. As the SRNet is only pretrained with CT images in stage 1, we need to generalize the learned upsampling kernels to the MR image domain. We thus

---

Algorithm 1. Training procedure

(Stage1) Pretrain SRNet with CT based self-supervision:
  GET HR CT images y_H
  FOR $t = 1$ to T
    Train SRNet by minimizing $\|y_H - \mathcal{F}_s(\mathcal{D}(y_H); \Theta_s)\|_1$
  END FOR
(Stage2) Pretrain SRNet with pseudo MR based self-supervision:
  GET LR MR images x_L
  FOR $t = T$ to $2T$
    Train SRNet by minimizing $\|\mathcal{F}_s(x_L; \Theta_s) - \mathcal{F}_s(\mathcal{D}(\mathcal{F}_s(x_L; \Theta_s)); \Theta_s)\|_1$
  END FOR
(Stage3) Joint optimization of CITNet and SRNet:
  GET unpaired LR MR images x_L and HR CT images y_H
  FOR $t = 2T$ to $10T$
    Train $D_\mathcal{X}$, $D_\mathcal{Y}$ by maximizing $\left(L_{adv}^\mathcal{X} + L_{adv}^\mathcal{Y}\right)$
    Train $E_\mathcal{X}^\mathcal{C}$, $E_\mathcal{Y}^\mathcal{C}$, $E_\mathcal{X}^\mathcal{A}$, $E_\mathcal{Y}^\mathcal{A}$, $G_\mathcal{X}$, $G_\mathcal{Y}$ and SRNet by minimizing $L_{disentangle}$
  END FOR

---

further pretrain the SRNet with pseudo MR images, as shown in Fig. 1-(B.2), for another T iterations. At each iteration, we first sample a batch of LR MR images x_L and input them into the SRNet to get the pseudo HR MR images $\tilde{x}_H = \mathcal{F}_s(x_L; \Theta_s)$. We then downsample $\tilde{x}_H$ to get corresponding pseudo LR MR images $\tilde{x}_L = \mathcal{D}(\tilde{x}_H)$. The SRNet is trained with the paired pseudo LR-HR MR images by minimizing $L1$ loss $\|\mathcal{F}_s(x_L; \Theta_s) - \mathcal{F}_s(\mathcal{D}(\mathcal{F}_s(x_L; \Theta_s)); \Theta_s)\|_1$. The idea behind such a pretraining stategy is that since both CT and MR images share the common structural information, the model pretrained with CT images in stage 1 facilitates the super-resolution reconstruction of pseudo HR MR images in stage 2. On the other hand, the training done in stage 2 can help the SRNet to learn MRI-specific domain information.

Stage 3. The MR images generated by the model pretrained at the first two stages can be further improved. In stage 3, we conduct joint optimization of the SRNet and the CITNet as shown in Fig. 1-(C), for another $8 \times T$ iterations. At each iteration, we first train $D_\mathcal{X}$, $D_\mathcal{Y}$ by maximizing $\left(L_{adv}^\mathcal{X} + L_{adv}^\mathcal{Y}\right)$. We then train $E_\mathcal{X}^\mathcal{C}$, $E_\mathcal{Y}^\mathcal{C}$, $E_\mathcal{X}^\mathcal{A}$, $E_\mathcal{Y}^\mathcal{A}$, $G_\mathcal{X}$, $G_\mathcal{Y}$ and the SRNet by minimizing $L_{disentangle}$ as defined in Eq. (7).

The training procedure of our method is illustrated by Algorithm 1.

Implementation Details. To train the proposed network, each training sample is unpaired LR MRI and HR CT images. All images are normalized to the range between -1.0 and 1.0. Optimization is performed using Adam with a batch size of 1. The initial learning rate is set to 0.0001 and decreased by a factor of 5 every 2 epochs. We empirically set $\lambda_1 = 10$, $\lambda_2 = \lambda_3 = 1$ and $T = 100,000$.

Table 1. The mean and the standard deviation when the proposed method was compared with the state-of-the-art (SOTA) unsupervised [18–20] and supervised [15,21] methods on both datasets. Paired T-Tests of all evaluation metrics achieved by ours and other methods are all smaller than 0.0001.

Dataset	Site1			Site2		
Method	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓
Bicubic	30.67±1.96	0.86±0.03	0.32±0.05	28.69±1.92	0.83±0.03	0.26±0.02
TSCN [18]	29.00±2.03	0.83±0.05	0.24±0.03	28.00±1.71	0.85±0.02	0.14±0.01
ZSSR [19]	30.95±2.12	0.88±0.03	0.16±0.02	28.94±1.69	0.83±0.03	0.16±0.01
SMORE [20]	31.78±1.98	0.89±0.03	0.21±0.03	29.93±2.00	0.86±0.03	0.14±0.02
Ours	**32.23±1.98**	**0.90±0.02**	**0.14±0.02**	**30.58±1.97**	**0.88±0.02**	**0.10±0.01**
Supervised [15]	32.99±2.07	0.91±0.02	0.10±0.02	31.66±1.72	0.90±0.02	0.08±0.01
ReconResNet [21]	32.93±3.12	0.88±0.05	0.09±0.02	29.97±1.50	0.84±0.03	0.07±0.01

Table 2. Results of ablation study on dataset from Site1.

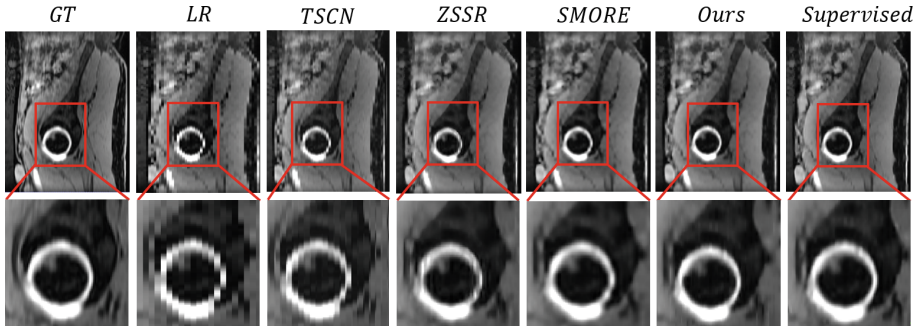| Stage 1	Stage 2	Stage 3	PSNR↑	SSIM↑	LPIPS↓
√	−	−	31.01±2.18	0.87±0.03	0.20±0.03
√	√	−	31.62±2.08	0.89±0.03	0.16±0.02
√	√	√	**32.23±1.98**	**0.90±0.02**	**0.14±0.02**

Fig. 2. Visual comparison of different methods when evaluated on dataset from Site1.
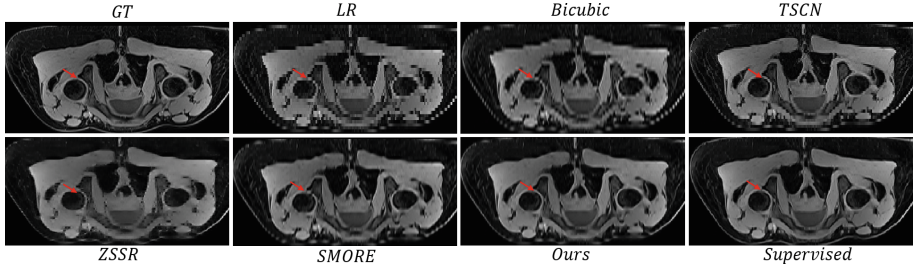
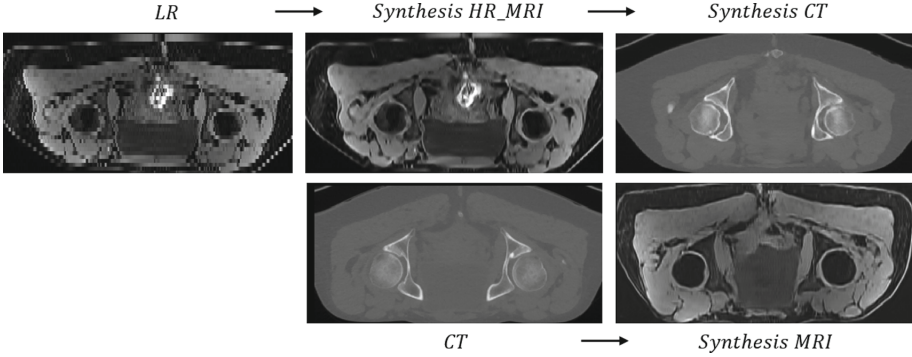**Fig. 3.** Visual comparison of different methods when evaluated on dataset from Site2.

LR ⟶ Synthesis HR_MRI ⟶ Synthesis CT

CT ⟶ Synthesis MRI

Fig. 4. Examples of cross-modality image translation between MRI and CT using data from Site2.

3 Experiments

Dataset. We conduct experiments to evaluate the proposed method on two datasets acquired from two different clinical centers. The dataset from HFR Cantonal Hospital, University of Fribourg (Site1) consists of 50 paired MR-CT volumes, which are divided into training (35 volumes), validation (5 volumes), and testing sets (10 volumes). The HR MRI are acquired by coronal plane and the voxel spacing of both HR CT and MRI are 1.0*1.0*1.0 mm^3. We downsample along the coronal axis with a scale factor $K = 4$ to generate the LR MRI with a voxel spacing 1.0*1.0*(1.0*K) mm^3. We shuffle the paired MR-CT volumes and only use the unpaired LR MRI and HR CT for training. Then we use the HR MRI to evaluate the reconstruction metrics. The dataset from the University Hospital of Bern (Site2) consists of 19 unpaired MR-CT volumes, which are divided into training (13 volumes) and testing sets (6 volumes). The HR MRI are acquired by coronal plane and the voxel spacing of both HR CT and MRI are 1.0*1.3*1.3 mm^3. We downsample along the coronal axis by a scale factor $K = 4$ to generate the LR MRI with a voxel spacing 1.0*1.3*(1.3*K) mm^3.

Experimental Results. We compare our method with the conventional algorithm bicubic interpolation, and the state-of-the-art (SOTA) unsupervised SR methods including TSCN [18], ZSSR [19], SMORE [20] as well as the SOTA supervised methods including RDN [15] and ReconResNet [21]. Well-established metrics including Peak Signal-to-Noise Ratio (PSNR) [22,23], Structural Similarity Index Metrics (SSIM) [24], and Learned Perceptual Image Patch Similarity (LPIPS) [25] are used to assess the performance of different methods.

Table 1 shows the mean and the standard deviation of the evaluation results of each method on both datasets. Figure 2 and Fig. 3 respectively show the super-resolution results on data from Site1 and Site2, when the scale factor is set as $K = 4$, as well as the corresponding LR and ground truth (GT) images. Both qualitative and quantitative results demonstrated that our method achieved better results than other SOTA unsupervised SR methods. It achieved comparable performance when compared with the supervised SR methods.

Our method is trained in two pretrain stages and one joint optimization stage. We thus conduct ablation study on dataset from Site1 to analyze the quality of the generated pseudo HR MR images at each stage. As shown in Table 2, quantitatively, the quality of the generated pseudo HR MR images is become better and better, demonstrating the effectiveness of the training strategy.

4 Conclusion

In this paper, we proposed a CT-guided, unsupervised MRI super-resolution reconstruction method based on joint cross-modality image translation and super-resolution reconstruction, eliminating the requirement of HR MRI for training. We conducted experiments on two datasets respectively acquired from two different clinical centers to validate the effectiveness of the proposed method. Quantitatively and qualitatively, the proposed method achieved superior performance over the SOTA unsupervised SR methods.

References

1. Jia, Y., Gholipour, A., He, Z., Warfield, S.K.: A new sparse representation framework for reconstruction of an isotropic high spatial resolution MR volume from orthogonal anisotropic resolution scans. IEEE Trans. Med. Imaging **36**(5), 1182–1193 (2017)
2. Plenge, E., et al.: Super-resolution methods in MRI: can they improve the trade-off between resolution, signal-to-noise ratio, and acquisition time? Magn. Reson. Med. **68**(6), 1983–1993 (2012)
3. Xuan, K., et al.: Reducing magnetic resonance image spacing by learning without ground-truth. Pattern Recogn. **120**, 108103 (2021)
4. Kim, J., Lee, J.K., Lee, K.M.: Accurate image super-resolution using very deep convolutional networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1646–1654 (2016)
5. Pham, C.H., Ducournau, A., Fablet, R., Rousseau, F.: Brain MRI super-resolution using deep 3D convolutional networks. In: 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017), pp. 197–200. IEEE (2017)
6. Shi, F., Cheng, J., Wang, L., Yap, P.T., Shen, D.: LRTV: MR image super-resolution with low-rank and total variation regularizations. IEEE Trans. Med. Imaging **34**(12), 2459–2466 (2015)
7. Chaudhari, A.S., et al.: Super-resolution musculoskeletal MRI using deep learning. Magn. Reson. Med. **80**(5), 2139–2154 (2018)
8. Chen, Y., Xie, Y., Zhou, Z., Shi, F., Christodoulou, A.G., Li, D.: Brain MRI super resolution using 3d deep densely connected neural networks. In: 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), pp. 739–742. IEEE (2018)
9. Chen, Y., Shi, F., Christodoulou, A.G., Xie, Y., Zhou, Z., Li, D.: Efficient and accurate MRI super-resolution using a generative adversarial network and 3D multi-level densely connected network. In: Frangi, A.F., Schnabel, J.A., Davatzikos, C., Alberola-López, C., Fichtinger, G. (eds.) MICCAI 2018. LNCS, vol. 11070, pp. 91–99. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-00928-1_11

10. Goodfellow, I., et al.: Generative adversarial nets. In: Advances in Neural Information Processing Systems 27 (2014)
11. Feng, C.M., Yan, Y., Yu, K., Xu, Y., Shao, L., Fu, H.: Exploring separable attention for multi-contrast MR image super-resolution. arXiv preprint arXiv:2109.01664 (2021)
12. Feng, C.-M., Fu, H., Yuan, S., Xu, Y.: Multi-contrast MRI super-resolution via a multi-stage integration network. In: de Bruijne, M., et al. (eds.) MICCAI 2021. LNCS, vol. 12906, pp. 140–149. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-87231-1_14
13. You, C., et al.: CT super-resolution GAN constrained by the identical, residual, and cycle learning ensemble (GAN-circle). IEEE Trans. Med. Imaging **39**(1), 188–203 (2019)
14. Wang, J., Wang, R., Tao, R., Zheng, G.: UASSR: unsupervised arbitrary scale super-resolution reconstruction of single anisotropic 3D images via disentangled representation learning. In: Medical Image Computing and Computer Assisted Intervention-MICCAI 2022: 25th International Conference, Singapore, September 18–22, 2022, Proceedings, Part VI, pp. 453–462. Springer (2022). https://doi.org/10.1007/978-3-031-16446-0_43
15. Zhang, Y., Tian, Y., Kong, Y., Zhong, B., Fu, Y.: Residual dense network for image super-resolution. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2472–2481 (2018)
16. Huang, X., Liu, M.Y., Belongie, S., Kautz, J.: Multimodal unsupervised image-to-image translation. In: Proceedings of the European Conference on Computer Vision (ECCV), pp. 172–189 (2018)
17. Johnson, J., Alahi, A., Fei-Fei, L.: Perceptual losses for real-time style transfer and super-resolution. In: Leibe, B., Matas, J., Sebe, N., Welling, M. (eds.) ECCV 2016. LNCS, vol. 9906, pp. 694–711. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46475-6_43
18. Lu, Z., Li, Z., Wang, J., Shi, J., Shen, D.: Two-stage self-supervised cycle-consistency network for reconstruction of thin-slice MR images. In: de Bruijne, M., et al. (eds.) MICCAI 2021. LNCS, vol. 12906, pp. 3–12. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-87231-1_1
19. Shocher, A., Cohen, N., Irani, M.: zero-shot super-resolution using deep internal learning. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3118–3126 (2018)
20. Zhao, C., Dewey, B.E., Pham, D.L., Calabresi, P.A., Reich, D.S., Prince, J.L.: Smore: a self-supervised anti-aliasing and super-resolution algorithm for MRI using deep learning. IEEE Trans. Med. Imaging **40**(3), 805–817 (2020)
21. Chatterjee, S., et al.: ReconResNet: regularised residual learning for MR image reconstruction of undersampled cartesian and radial data. Comput. Biol. Med. **143**, 105321 (2022)
22. Wu, Q., et al.: An arbitrary scale super-resolution approach for 3-dimensional magnetic resonance image using implicit neural representation. arXiv preprint arXiv:2110.14476 (2021)
23. Du, J., et al.: Super-resolution reconstruction of single anisotropic 3D MR images using residual convolutional neural network. Neurocomputing **392**, 209–220 (2020)
24. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: from error visibility to structural similarity. IEEE Trans. Image Process. **13**(4), 600–612 (2004)

25. Zhang, R., Isola, P., Efros, A.A., Shechtman, E., Wang, O.: The unreasonable effectiveness of deep features as a perceptual metric. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 586–595 (2018)