Enhancing the MR Neuroimaging by Using the Deep Super-Resolution Reconstruction

Yu Cao[1,3], Hongzhi Kuai[2,3(✉)], and Guanqiao Peng[4]

[1] Information Department, Beijing University of Technology, Beijing, China
[2] Graduate School of Engineering, Maebashi Institute of Technology, Maebashi, Gunma, Japan
hongzhi.kuai@gmail.com
[3] International WIC Institute, Beijing University of Technology, Beijing, China
[4] Beijing 101 Middle School, Beijing, China

Abstract. As brain investigation progresses, the need has become urgent from acquiring the higher resolution neuroimaging data to give a more detailed interpretation. In particular, the technological development and innovation of the Magnetic Resonance Imaging (MRI) machine, through increasing the magnetic field from low (such as 3T) to high (such as 7T), has revealed significant advantages regarding the image quality enhancement, etc. Currently, due to the limitations of hardware, physics and physiology, the large-scale acquisition of the high-resolution MRI neuroimages is still running on the road. Hence, enhancing the quality of the low-field MRI data is critical by using the advanced artificial intelligence technology. In this study, we propose a novel image enhancement framework, namely SR-MRI, trying to improve the quality of the low-resolution neuroimage: (1) combining with the Real-ESRGAN deep learning model; (2) bridging the 3T-MRI and the 7T-MRI within the same analysis scale; and (3) systematically comparing multiple evaluation indicators, including Brenner, SMD, SMD2, Variance, Vollath, Entropy, and NIQE. The experimental results suggest that the edge, fineness and texture features of the low-resolution neuroimages are restored to a great extent from the SR-MRI framework. In addition, the evaluation results of multiple indicators show that the processed 3T-MRI can achieve the similar effect from the 7T-MRI machine.

Keywords: Magnetic Resonance Imaging (MRI) · Super-resolution · Brain informatics · Deep learning · Real-ESRGAN model

1 Introduction

As the most important information processing and control center in the human body, the brain is closely related to cognitive, emotional, psychological and behavioral functions [1]. With the rapid development of medical informatization and the popularization of medical neuroimages in this digital age, Magnetic Resonance Imaging (MRI) plays an increasingly significant role in the detection and diagnosis of various diseases for smart health [2]. The quality of MR neuroimages, as the carrier of patients' pathological

M. Mahmud et al. (Eds.): BI 2022, LNAI 13406, pp. 184–194, 2022
https://doi.org/10.1007/978-3-031-15037-1_16

information, can influence doctors' perception, reception, comprehension and diagnosis of patients' pathological information.

The magnetic resonance phenomenon is produced by using a high magnetic field and a radio signal to excite hydrogen protons in the human body. By changing the intensity of external gradient magnetic field "T", different tissues of the body can resonate at different frequencies and draw structural images in human body [3]. The letter "T" indicates Tesla in the magnetic resonance, which is the magnitude of the field strength. Theoretically, the stronger the field intensity is used, the higher the signal to noise ratio (SNR) is given, implying that more image resolution can be offered. From the hardware perspective, the 7T-MRI machine, even the machine with the higher Tesla, is developed to obtain the higher-resolution MRI data. However, due to the limitations of hardware and physics, the acquisition system cost of the high-resolution (HR) 7T MRI is high [4]. Therefore, the current mainstream still depends on the 3T-MRI technology, which has generated massive amounts of data. Furthermore, from the method perspective, the super-resolution (SR) reconstruction of low-resolution (LR) MR neuroimages is attracting greater attention, which can reduce the requirements of hardware equipment without increasing the cost of imaging technology. The reconstructed high-resolution MR neuroimages can help doctors make accurate diagnosis of patients' condition.

So far, the natural images have been the focus of academic super-resolution network processing. However, as the super-resolution advances and the medical industry's demand of high-resolution images grow, more network structures and associated approaches for medical images are being presented [5]. For instance, Liu [6] proposed a multi-scale fusion convolution network to conduct super-resolution for MRI reconstruction in order to investigate the various edge responses utilizing various convolution kernel sizes. Shi [7] put forward a new residual learning-based SR technique for MRI by combining local residual block with global residual network. Furthermore, numerous researchers have focused on different improved methods, such as the modified SRCNN (Super-Resolution Convolutional Neural Network) based global residual learning strategy [8] and the GAN based common algorithms [9]. As a result, in this paper, we introduce a novel super-resolution reconstruction framework for enhancing the MR neuroimaging quality from 3T to 7T, combining with different deep learning algorithms. Different from the common pipeline of the super-resolution reconstruction, the current work builds flexible components to process different scale neuroimaging data within a unified framework towards the goal of greater practice. In addition, considering the necessity of the systematic analysis from the Brain Informatics methodology [10], multiple evaluation indicators are calculated to verify the effectiveness of the enhanced 3T-MRI.

2 Method

In this section, we introduce a novel super-resolution framework, namely SR-MRI, for enhancing MR neuroimaging from 3T to 7T. Figure 1 depicts the image enhancement framework, which is made up of three components: the preprocessing component, the super-resolution component and the evaluation component.

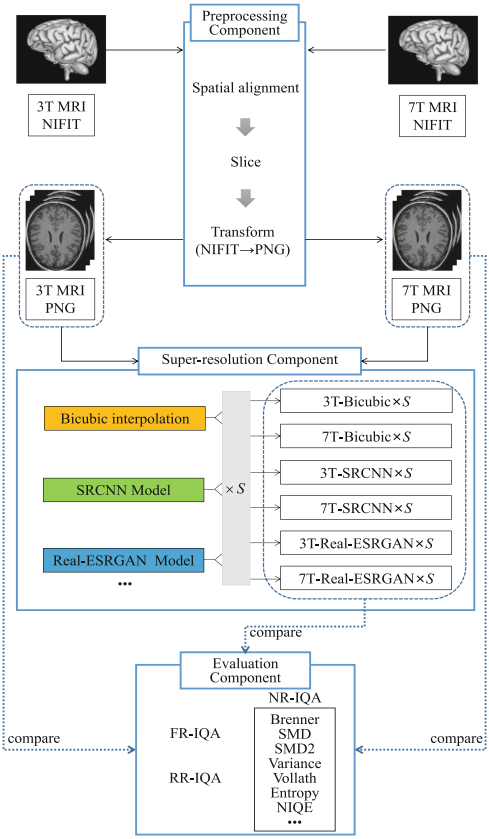

Fig. 1. The detailed components and their workflows in the SR-MRI framework. (*S* is the magnification of neuroimaging reconstruction. FR: Full Reference; RR: Reduced Reference; NR: No Reference.)

2.1 The Preprocessing Component

In the preprocessing component, through the operation of spatial alignment, the images have been aligned into the template derived from 555 healthy subjects of the IXI database (http://www.brain-development.org) [11]. To execute the nonlinear registration properly, we also need to initialize it by a linear registration to acquire the image's orientation and size close enough for the nonlinear registration. The significance of spatial alignment is to align images to the same template space for comparison, as well as to eliminate noise that may arise during PNG conversion. Subsequently, we slice both 3T and 7T MR neuroimages along the Z axis, and convert them from 3D NIfTI format to PNG format.

2.2 The Super-Resolution Component

In the component of super-resolution, three super-resolution methods are integrated into this framework, including the bicubic interpolation algorithm, and the SRCNN

and Real-ESRGAN models based on deep learning. These methods are performed to reconstruct low-resolution images, respectively. Meanwhile, images can be magnified with different scales in these three various methods. Obviously, more effective super-resolution algorithms can be added to the framework to improve the image quality.

SRCNN technology is the first time to apply the deep learning model of convolutional neural network in the field of super-resolution reconstruction [12]. Its network structure is shown in Fig. 2. The supplied low-resolution image is first enlarged to the intended size using bicubic interpolation. Secondly, to match the non-linear mapping between LR and HR images, a three-layer convolutional neural network is utilized. Finally, the output of the network is the reconstructed HR image.

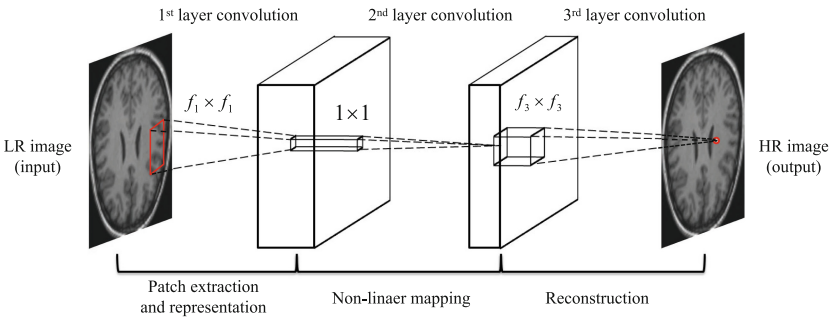

Fig. 2. The network structure of SRCNN.

Real-ESRGAN (Real-Enhanced Super-Resolution Generative Adversarial Networks) is one of the top SR technologies in recent years including the Generative Network and the Discriminator Network [13]. The model manages to achieve a decent mix between image improvement and artifact reduction. The study proposes a high-order degradation procedure, and uses the 'sinc' filters to the mimic typical ringing and overshoot problems in order to synthesize more realistic degradations. The Generative Network (as shown in Fig. 3) is a deep network constructed by several residual-in-residual dense blocks (RRDB) without batch normalization, making it easier to train deeper and more complicated network architectures. Meanwhile, to improve discriminator capabilities and stabilize training dynamics, the researchers used a U-Net discriminator with spectral normalization regularizations. For most real-world images, Real-ESRGAN trained on synthetic data is able to increase details while reducing annoying artifacts.

2.3 The Evaluation Component

There exist two types of image quality evaluation methods: subjective and objective [14]. The former relies on the subjective perception of the experimenter, while the latter measures the image quality according to the quantitative index. The use of one or more image indicators creates a mathematical model, ensuring that the outcomes of objective evaluations are congruent with people's subjective sentiments. According to whether the

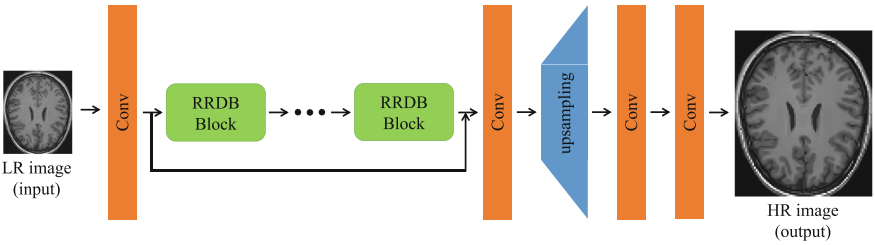

Fig. 3. The structure of Generative Network of Real-ESRGAN.

original image is referenced and the degree of reference, the objective quality evaluation can be divided into the following three types: (a) Full Reference (FR), (b) Reduced Reference (RR), (c) No Reference (NR) [15]. Since both FR method and RR method require the information from the original images as the reference for assessment, and the data used in this article is not HR MR neuroimages. Consequently, this research opts for the NR method [16], which does away with the need for reference images and relies on some information from the image to be reviewed for quality evaluation. Following that, in the evaluation component, numerous distinct NR indicators are integrated, including Brenner, SMD, SMD2, Vollath, Entropy, NIQE and so on

1. **Brenner**. The Brenner gradient function is the simplest gradient evaluation function. It simply calculates the square of the gray difference between two adjacent pixels, which is given by:

$$D\left(f^{Brenner}\right) = \sum_y \sum_x |f(x+2, y) - f(x, y)|^2 \tag{1}$$

where $D\left(f^{Brenner}\right)$ is the index calculated by the Brenner gradient function, and $f(x, y)$ represents the gray value of the pixel (x, y) corresponding to image f.

2. **SMD**. The Sum of Modulus of gray Difference (SMD) function takes the gray change as the basis for focus evaluation. SMD extracts the change size of the gray value of point (x, y) and its adjacent points by performing a differential operation on the gray level of the point (x, y) and its neighboring points, which is given by:

$$D\left(f^{SMD}\right) = \sum_y \sum_x (|f(x, y) - f(x, y-1)| + |f(x, y) - f(x+1, y)|) \tag{2}$$

where $D\left(f^{SMD}\right)$ is the index calculated by the Sum of Modulus of gray Difference function.

3. **SMD2**. The SMD2 function is the product of modulus of gray difference. That is, the two gray differences in each pixel field are multiplied and then accumulated pixel by pixel, which is given by:

$$D\left(f^{SMD2}\right) = \sum_y \sum_x |f(x, y) - f(x+1, y)| * |f(x, y) - f(x, y+1)| \tag{3}$$

where $D\left(f^{SMD2}\right)$ is the index calculated by the SMD2 function.

4. **Variance**. The Variance function may be utilized as the evaluation function since the sharply focused image has a bigger gray difference than the unfocused image, which is given by:

$$D\left(f^{Variance}\right) = \sum_y \sum_x |f(x, y) - \mu|^2 \tag{4}$$

where $D\left(f^{Variance}\right)$ is the index calculated by the Variance function.

5. **Vollath**. The Vollath function is based on the image cross-correlation function, which efficiently suppresses noise and reduces interference from impurities. The function is defined as follows:

$$D\left(f^{Vollath}\right) = \sum_y \sum_x f(x, y) * f(x + 1, y) - M * N * \mu \tag{5}$$

where $D\left(f^{Vollath}\right)$ is the index calculated by the Vollath function, and μ indicates the average gray value of the whole image, and M and N are the image's width and height, respectively.

6. **Entropy**. The Entropy function, which is based on statistical characteristics, is a useful metric for assessing the richness of visual data. According to information theory, the information quantity of an image is measured by the information entropy of the image, which is given by:

$$D\left(f^{Entropy}\right) = -\sum_{I=0}^{L-1} P_i \ln(P_i) \tag{6}$$

where $D\left(f^{Entropy}\right)$ is the index calculated by the Entropy function, and P_i is the probability of pixels with gray value I in the image, and L is the total number of gray levels (usually 256).

7. **NIQE**. The Natural Image Quality Evaluator (NIQE) [17] is the evaluation index of ECCV's PIRM competition in 2018. The NIQE first extracts the region of interest from the image when extracting the statistical features of the image. It stems from the fact that the human eye prefers to judge the image quality by the clearer part of the image. When using the collected 36 features for image quality evaluation, the Multivariate Gaussian Model (MVG) is used to fit, which is given by:

$$f_X(x_1, ..., x_k) = \frac{1}{(2\pi)^{k/2}|\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(x - v)^T \Sigma^{-1}(x - v)\right) \tag{7}$$

where $(x_1, ..., x_k)$ are the 36 features collected, v and Σ are the mean and covariance matrix of MVG, which can be obtained by maximum likelihood estimation.

The value of NIQE is obtained by calculating the distance of the MVG parameters between the natural image and the distorted image:

$$D\left(f^{NIQE}\right) = D(v_1, v_2, \Sigma_1, \Sigma_2) = \sqrt{(v_1 - v_2)^T \left(\frac{\Sigma_1 + \Sigma_2}{2}\right)^{-1} (v_1 - v_2)} \tag{8}$$

where $D\left(f^{NIQE}\right)$ is the index calculated by the NIQE.

3 Results and Discussions

3.1 MR Neuroimage Acquisition and Preprocessing

In this work, two datasets with the accession numbers of ds001553 and ds001555 are obtained from the OpenNeuro platform (https://openneuro.org). Each study was performed by three people with no known history of neurological diseases (5 females; age $= 25 \pm 5$ years). All participants gave informed consent in compliance with a protocol approved by the Institutional Review Board of the National Institute of Mental Health in Bethesda, MD, USA. For 3T images, its accession number is ds001553. 372 coronal slices were obtained in a General Electric 3T MRI scanner using image parameters: TR $= 7240$ ms, TE $= 2.7$ ms, TI $= 725$ ms, Flip angle $= 12°$, and resolution $= 1 \times 1 \times 1$ mm^3. For 7T images, its accession number is ds001555. 354 coronal slices were obtained in a Siemens 7T MRI scanner equipped with a 32-element receive coil using image parameters: TR $= 3000$ ms, TE $= 3.88$ ms, TI $= 1500$ ms, Flip angle $= 6°$, and resolution $= 1 \times 1 \times 1$ mm^3. As for spatial alignment, we registered the corresponding 3T and 7T images using Matlab SPM12 toolbox, to minimize the possible global distortions. To do so, all the images were linearly aligned to the MNI space by using an individual template [18]. After one subject's MRI data is aligned to the MNI standard space, the NIfTI file is converted to 78 images of PNG format by slicing. Finally, we obtained 234 normalized images from three subjects, with respect to the 3T scale, and the same number of normalized images at the 7T scale. As individual differences are not concerned about the core point in this paper, the average effectiveness of all super-resolution images corresponding to different scale is evaluated respectively in the following evaluation phase.

3.2 The Super-Resolution Results of the MR Neuroimaging

From Fig. 4, it can be seen that the resolution of MR neuroimages has been greatly increased by using the proposed SR-MRI framework.

In this framework, the super-resolution images using the bicubic interpolation algorithm have relatively higher ambiguity compared with the results from the SRCNN and Real-ESRGAN methods, not only on the edge, but also in the region. The image rebuilt using the SRCNN model has somewhat enhanced intra-regional resolution, however it still falls short of the image recovered by the Real-ESRGAN model. By using the Real-ESRGAN reconstruction, the images are clearest at the junction of white matter and gray matter, and the edge, fineness and texture characteristics are recovered to a large extent. Therefore, the Real-ESRGAN method is embedded into the current framework with stronger recommendation. In the next section, we compare the performance of super-resolution results from both 3T and 7T MRI using quality evaluation of no reference.

3.3 Quantitative Analysis Based on No-Reference Indicators

As shown in Table 1, the evaluation value of each index is given by averaging all values from a group of images (234 slices). Apart from the NIQE index of the seven indexes in this article, the greater an index value is, the higher an image quality has.

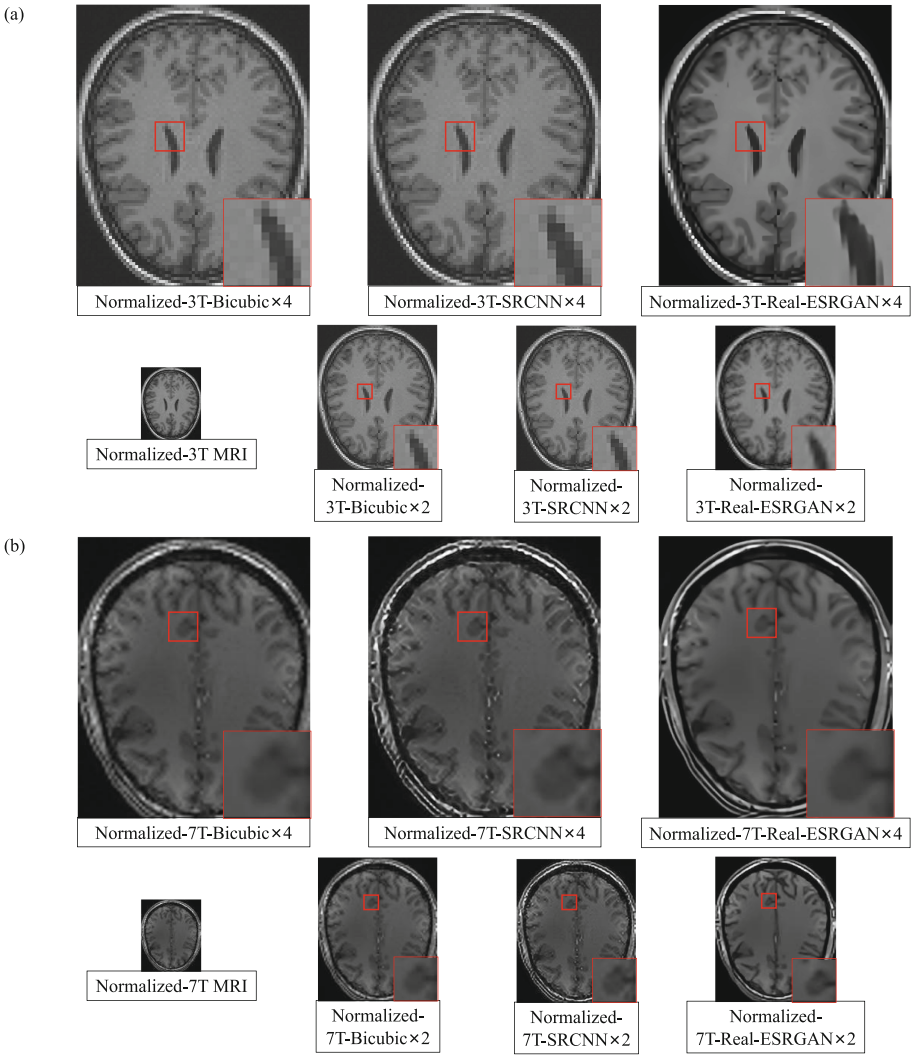

Fig. 4. The reconstructed MRI results based on the proposed SR-MRI framework. (a) shows the reconstructed results based on the normalized-3T MR neuroimages; (b) shows the reconstructed results based on the normalized-7T MR neuroimages. ($\times 2$ is the two times magnification for super-resolution reconstructed images, and $\times 4$ is the four times magnification.)

In Table 1, it is found that all indicators of 7T MRI are superior than those of 3T MRI. When comparing the image quality of three alternative processing algorithms in this SR-MRI framework, the Real-ESRGAN method outperforms the SRCNN and bicubic interpolation methods. Meanwhile, the four times magnification produces a superior visual quality than the two times magnification. In addition, to evaluate the enhanced scale before and after running this SR-MRI framework, the difference of indictors is calculated between both 3T and 7T normalized MR neuroimaging and the

Table 1. Comparison of super-resolution results between 3T and 7T MRI based on quality evaluation of no reference. ($*$ in $D(f^*)$ is Brenner, SMD, SMD2, Variance, Vollath, Entropy, and NIQE, respectively. I: Normalized 3T; II: Normalized 7T; III: 3T-Bic ×2; IV: 7T-Bicx2; V: 3T-SRCNN×2; VI: 7T-SRCNNx2; VII: 3T-Real-ESRGAN×2; VIII: 7T-Real-ESRGANx2; IX: 3T-Bic×4; X: 7T-Bicx4; XI: 3T-SRCNN×4; XII: 7T-SRCNNx4; XIII: 3T-Real-ESRGAN ×4; XIV: 7T-Real-ESRGANx4.)

Index / Image	$D(f^{Brenner})$ /10^6	$D(f^{SMD})$ /10^5	$D(f^{SMD2})$ /10^6	$D(f^{Variance})$ /10^7	$D(f^{Vollath})$ /10^7	$D(f^{Entropy})$ /10^0	$D(f^{NIQE})$ /10^0
Normalized 3T (I)	5.81	1.35	1.99	0.58	0.40	4.18	19.79
Normalized 7T (II)	7.50	1.87	3.04	0.75	0.51	4.34	15.22
$D(f^{*,II})-D(f^{*,I})$	**1.71**	**0.53**	**1.05**	**0.18**	**0.11**	**0.16**	**-4.57**
3T-Bic×2 (III)	8.71	2.56	1.60	2.07	1.90	4.20	12.08
7T-Bicx2 (IV)	13.30	3.44	2.25	2.63	2.39	4.33	11.03
$D(f^{*,IIV})-D(f^{*,III})$	**4.58**	**0.87**	**0.64**	**0.56**	**0.49**	**0.13**	**-1.05**
3T-SRCNN×2 (V)	21.55	4.05	5.52	2.60	2.14	4.25	13.21
7T-SRCNNx2 (VI)	38.07	5.79	8.43	3.55	2.73	4.43	11.56
$D(f^{*,VI})-D(f^{*,V})$	**16.52**	**1.74**	**2.92**	**0.94**	**0.58**	**0.18**	**-1.65**
3T-Real-ESRGAN×2 (VII)	16.46	3.43	3.65	3.04	2.72	4.30	11.54
7T-Real-ESRGANx2 (VIII)	28.72	4.71	5.77	3.95	3.38	4.45	10.80
$D(f^{*,VIII})-D(f^{*,VII})$	**12.25**	**1.27**	**2.12**	**0.91**	**0.65**	**0.15**	**-0.74**
3T-Bic×4 (IX)	10.56	5.07	1.65	7.72	7.52	4.23	11.45
7T-Bicx4 (X)	16.27	6.74	2.30	9.42	9.15	4.34	10.91
$D(f^{*,X})-D(f^{*,IX})$	**5.71**	**1.67**	**0.64**	**1.70**	**1.63**	**0.11**	**-0.54**
3T-SRCNN×4 (XI)	34.68	8.08	7.09	9.90	9.32	4.27	9.28
7T-SRCNNx4 (XII)	63.29	11.41	11.20	12.87	11.90	4.45	9.14
$D(f^{*,XII})-D(f^{*,XI})$	**28.60**	**3.33**	**4.12**	**2.97**	**2.53**	**0.18**	**-0.14**
3T-Real-ESRGAN×4 (XIII)	51.21	9.15	12.8	11.21	10.25	4.31	8.78
7T-Real-ESRGANx4 (XIV)	55.28	9.74	12.97	14.16	13.30	4.48	7.52
$D(f^{*,XIV})-D(f^{*,XIII})$	**3.97**	**0.59**	**0.17**	**2.95**	**3.05**	**0.17**	**-1.26**

super-resolution reconstructed images, respectively. On the one hand, by comparing with the NIQE difference of $|D(f^{*,II}) - D(f^{*,I})| = 4.57$ calculated by the normalized 3T and 7T images, the NIQE difference after the super-resolution reconstruction processes is reduced obviously. On the other hand, the difference calculated by the other six indicators is increased after running this framework. It is found that the SR-MRI framework can enhance the performance of images, and even the enhanced 3T-MRI could achieve the similar effect from the 7T-MRI machine.

4 Conclusion

In this paper, by developing a novel super-resolution framework, namely SR-MRI, combined with the pre-trained SRCNN and Real-ESRGAN models, we apply the image super-resolution reconstruction technology based on deep neural network to enhance the

MR neuroimages. The effectiveness of the deep super-resolution model is verified by comparing multi-scale MRI data from 3T to 7T. It has been shown that when the 3T MR neuroimages are processed using the current framework, the image quality is increased considerably, and the effect from a 7T MRI machine could be achieved similarly. In the future, the further research is needed to integrate more super-resolution techniques into this framework and apply it to more potential scenarios towards accelerating the clinical practice.

References

1. Kuai, H., Zhong, N.: The extensible data-brain model: architecture, applications and directions. J. Comput. Sci. 101103 (2020)
2. Kuai, H., Zhong, N., Chen, J., Yang, Y.: Multi-source brain computing with systematic fusion for smart health. Inform. Fus. **75**, 150–167 (2021)
3. Weiskopf, N., Edwards, L.J., Helms, G.: Quantitative magnetic resonance imaging of brain anatomy and in vivo histology. Nat. Rev. Phys. **3**(8), 570–588 (2021)
4. Stucht, D., Danishad, K.A., Schulze, P.: Highest resolution in vivo human brain MRI using prospective motion correction. PLoS ONE **10**(7), e0133921 (2015)
5. Yang, W., Zhang, X., Tian, Y., Wang, W.: Deep learning for single image super-resolution: a brief review. IEEE Trans. Multimedia **21**(12), 3106–3121 (2019)
6. Liu, C., Wu, X., Yu, X., Tang, Y.: Fusing multi-scale information in convolution network for MR image super-resolution reconstruction. Biomed. Eng. Online **17**(1), 1–23 (2018)
7. Shi, J., Liu, Q., Wang, C., Zhang, Q.: Super-resolution reconstruction of MR image with a novel residual learning network algorithm. Phys. Med. Biol. **63**(8) (2018)
8. Qiu, D., Zheng, L., Zhu, J., Huang, D.: Multiple improved residual networks for medical image super-resolution. Futur. Gener. Comput. Syst. **116**, 200–208 (2021)
9. Lv, J., Zhu, J., Yang, G.: Which GAN? A comparative study of generative adversarial network-based fast MRI reconstruction. Phil. Trans. R. Soc. A **379**(2200), 20200203 (2021)
10. Zhong, N., Bradshaw, J.M., Liu, J., Taylor, J.G.: Brain informatics. IEEE Intell. Syst. **26**(5), 16–21 (2011)
11. Tavares, V., Prata, D., Ferreira, H.A.: Comparing SPM12 and CAT12 segmentation pipelines: a brain tissue volume-based age and Alzheimer's disease study. J. Neurosci. Methods **334**, 108565 (2020)
12. Dong, C., Loy, C.C., He, K.: Learning a deep convolutional network for image super-resolution. In: Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T. (eds.) European Conference on Computer Vision, LNIP, vol. 8692, pp. 184–199. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-10593-2_13
13. Wang, X., Xie, L., Dong, C., Shan, Y.: Real-esrgan: Training real-world blind super-resolution with pure synthetic data. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 1905–1914 (2021)
14. Mohammadi, P., Ebrahimi-Moghadam, A., Shirani, S.: Subjective and objective quality assessment of image: a survey. Majlesi J. Electric. Eng. **9**(1) (2014)
15. Yang, J., Zhao, Y., Liu, J., Jiang, B.: No reference quality assessment for screen content images using stacked autoencoders in pictorial and textual regions. IEEE Trans. Cybern. **99**, 1–13 (2020)
16. Hu, S., Wang, X., Wu, H., Luan, X.: Unified diagnosis framework for automated nuclear cataract grading based on smartphone slit-lamp images. IEEE Access **8**, 174169–174178 (2020)

17. Mittal, A., Soundararajan, R., Bovik, A.C.: Making a “completely blind” image quality analyzer. IEEE Signal Process. Lett. **20**(3), 209–212 (2012)
18. Holmes, C.J., Hoge, R., Collins, L., Woods, R.: Enhancement of MR images using registration for signal averaging. J. Comput. Assist. Tomogr. **22**(2), 324–333 (1998)