



Research article

Multi-modal brain MRI images enhancement based on framelet and local weights super-resolution

Yingying Xu, Songsong Dai, Haifeng Song, Lei Du and Ying Chen*

School of Electronics and Information Engineering, Taizhou University, Taizhou 318000, China

* **Correspondence:** Email: yichen222@tzc.edu.cn.

Abstract: Magnetic resonance (MR) image enhancement technology can reconstruct high-resolution image from a low-resolution image, which is of great significance for clinical application and scientific research. T1 weighting and T2 weighting are the two common magnetic resonance imaging modes, each of which has its own advantages, but the imaging time of T2 is much longer than that of T1. Related studies have shown that they have very similar anatomical structures in brain images, which can be utilized to enhance the resolution of low-resolution T2 images by using the edge information of high-resolution T1 images that can be rapidly imaged, so as to shorten the imaging time needed for T2 images. In order to overcome the inflexibility of traditional methods using fixed weights for interpolation and the inaccuracy of using gradient threshold to determine edge regions, we propose a new model based on previous studies on multi-contrast MR image enhancement. Our model uses framelet decomposition to finely separate the edge structure of the T2 brain image, and uses the local regression weights calculated from T1 image to construct a global interpolation matrix, so that our model can not only guide the edge reconstruction more accurately where the weights are shared, but also carry out collaborative global optimization for the remaining pixels and their interpolated weights. Experimental results on a set of simulated MR data and two sets of real MR images show that the enhanced images obtained by the proposed method are superior to the compared methods in terms of visual sharpness or qualitative indicators.

Keywords: magnetic resonance imaging; multi-modal; image enhancement; super-resolution; framelet transform

1. Introduction

Magnetic resonance imaging (MRI) is one of the most important medical imaging technology, it has the following advantages in clinical diagnosis: (1) no ionizing radiation damage; (2) can obtain the body part images that are difficult or inaccessible to other imaging technologies; (3) has better soft

tissue resolution than computed tomography (CT); (4) belongs to multi-parameter imaging, which can provide multi-modal diagnostic information. Overall, MRI images show clear anatomical structures and high contrast, and it is also possible to observe pathological tissues with small onset, which is very beneficial for early diagnosis and treatment of brain diseases [1, 2].

However, when MRI is used to collect images, due to the constraints of some conditions, such as the limited sampling time resulting in the use of a fast imaging method, the collected images are often low-resolution (LR) images. Compared with high-resolution (HR) images, LR images lose many edges and details of biological tissue structures, which is not conducive to subsequent clinical diagnosis. To overcome and compensate for the hardware limitations of MRI scanners, super-resolution reconstruction techniques can be used to recover HR images from the corresponding LR images [3,4]. Super-resolution reconstruction technology is an image enhancement method, which improves the spatial resolution of the image through a post-processing algorithm, and optimizes the image edge structure and details.

The super-resolution algorithm of MRI images can be roughly divided into single-image super-resolution and multi-contrast image super-resolution. The super-resolution algorithms of MRI images can be roughly divided into single-modal and multi-modal image super-resolution. Single-modal super-resolution can enhance image quality based on a single image or a sequence of images in the same modal [5]. Interpolation is the most traditional super-resolution reconstruction method, such as bicubic interpolation [6]. Due to its simple and efficient algorithm, it is widely used to improve the resolution of MRI. However, bicubic assumes that the image has regional smoothness in all regions, it damages the high-frequency component of the image and causes image blurring and edge distortion. One way to improve the interpolation effect is to use prior knowledge of the image. The new edge-directed interpolation (NEDI) [7] algorithm estimates the local covariance coefficients by using the statistical features of the LR image's edge structure, and uses the coefficients as the linear regression weights of adjacent pixels to interpolate the image. NEDI improves edge interpolation effects, but is more time-consuming than bicubic. Manjón et al. [8] proposed an upsampling model based on the nonlocal similarity of image patches. Wei and Ma [9] studied using the edge contrast information of LR images for interpolation, and proposed a contrast-guided interpolation (CGI) algorithm to enhance the sharpness of reconstructed edges through various directional filtering, which is much faster than NEDI. Shi et al. [10] and Tourbier et al. [11] introduced total variation regularizations to enhance the edge sharpness of the MR image. However, the prior structure information based on a single image is limited, and the enhancement effect of image super-resolution is also limited. Therefore, the image super-resolution method based on dictionary learning is gradually widely used in MR image enhancement. Rueda et al. [12] proposed an MR image super-resolution model based on the overcomplete dictionary and sparse representation. Zhang et al. [13] combined the sparse representation with nonlocal similarity and sparse derivative prior to improving the edge reconstruction. Jia et al. [14] and Jia et al. [15] train the overcomplete dictionary from in-plane MR images to reconstruct the isotropic high spatial resolution MR image. Meanwhile, Huang et al. [16] use weakly-supervised joint convolutional sparse coding to enhance the resolution of MR images based on a small number of well-registered images of different modalities. Recently, various convolutional networks based on deep learning methods have been widely used for MR image super-resolution [17–21], which essentially use deep neural networks to seek better nonlinear mapping from LR images to HR images. However, although these methods can achieve high scores in quantitative evaluation indicators, their common problem is that the models are less explicable and may produce meaningless image contents that do not exist originally [22, 23].

Multi-modal image super-resolution algorithm is established on the basis of a single-modal image super-resolution algorithm by introducing the structure information of different modal images as the prior knowledge of the reconstruction model. MR images of different modals can be obtained by using different MRI techniques or different software and hardware parameter settings when imaging the same object [24, 25]. Multi-modal MRI is increasingly being used in complex disease diagnosis and medical research experiments. Different imaging modes can provide specific information about biological tissues, both structural and functional. The synthesis of MR images of various modes is beneficial to more comprehensive observation of the imaging results. Multi-modal image super-resolution uses the structural correlation between different modals to improve the image resolution of one modal by using the information of another modal. Multi-modal images are usually presented as multi-contrast images, among which the most typical ones are T1-weighted and T2-weighted images. The signal generation mechanism of the multi-contrast image can be described by the following signal intensity model [26]:

$$SI \propto \rho(1 - 2e^{-TI/T1} + e^{-TR/T1})e^{-TE/T2}, \quad (1.1)$$

where the signal intensity SI is proportional to the proton density ρ and further depends on the pulse train parameters repetition time (TR), inversion time (TI) and echo time (TE). In formula (1.1), when TR is long and TE is short, the signal is roughly proportional to ρ . If TR and TE are both short, the signal is dominated by $\rho(1 - e^{-TR/T1})$, i.e., T1 weighting image is obtained [27]. When TR and TE are both long, the term is dominated by $\rho e^{-TE/T2}$, i.e., T2 weighting image is obtained. Where “long” and “short” are relative to the typical T1 and T2 constants.

These multi-contrast images have similar anatomical structures due to the same proton density values in corresponding spatial positions, but the difference in signal intensity leads to the difference in contrast. Studies have shown that the structural information shared between multi-contrast images can be used to guide image enhancement and reconstruction tasks such as MRI super-resolution. For example, T2-weighted images can clearly display the edges of many kinds of pathological tissues, but the imaging time is long. T1-weighted images have the advantage of showing normal brain tissue structure, requiring less imaging time, but the lesions are often not as clear as T2-weighted images. In order to solve the contradiction between imaging time and image quality, a compromise method is to collect all the signals of T1-weighted images while only part of the signals of T2-weighted images. In this way, the T2-weighted image obtained by fast imaging has a low resolution, and then the high-resolution T1 image is used as the reference image for super-resolution reconstruction to obtain the high-resolution T2-weighted image.

Rousseau [28, 29] constructed a super-resolution model for MR images by taking advantage of the similar structure of multi-mode images, and verified the superiority of using different contrast image information for super-resolution compared with single-mode super-resolution. Manjón et al. [30] further combined the prior information of the image to be reconstructed and improved the iterative algorithm. Jafari-Khouzani [31] proposed a nonlocal means method based on multiple modal features to further improve the edge sharpness of MR images after super-resolution. Lu et al. [32] hypothesized that T2-weighted LR image and T1-weighted HR image have similar geometric structures in manifold space and proposed a super-resolution approach via manifold regularized sparse learning. Zheng et al. [33] proved that the local regression weights used for interpolation are very similar among T1-weighted and T2-weighted images, and proposed a super-resolution method for T2 images by using the local weights learned from paired T1 images. Then, Zheng et al. [34] established an associated model of gradient

value between different MRI contrast images, and proposed a new super-resolution method by using gradient information of the HR reference image that has another contrast. In general, the multi-modal image super-resolution reconstruction algorithm makes full use of the local structure similarity of MRI multi-modal image, and introduces the structure information of high-resolution reference image with different modes (different contrast) in the reconstruction process, which makes the reconstruction result more ideal.

This paper proposed a new model for multi-modal MR image super-resolution based on the work of [33]. Considering that in the problem of T2 image super-resolution by using the structural information of T1 image, the method proposed in Zheng et al. [33] interpolate T2 using the coefficient matrix calculated from self-defined image edges of T1 image, which can be improved in two aspects: first, it may be inaccurate to define the image edge by gradient threshold; second, the interpolation coefficient matrix can be fine-tuned to better match the internal structural association of T1 and T2 image pairs. Therefore, this paper constructs a modified model of bivariate optimization based on framelet transformation. We use framelet transform to decompose the super-resolution target T2 image into high-frequency and low frequency parts, so that the main energy and edge information of the image can be treated differently during the reconstruction. The advantage of our method lies in that it can more accurately emphasize the weight borrowing of image edges by separating them more finely, and at the same time we carry out global optimization and fine-tuning of the interpolation weights obtained according to T1 image, so as to achieve better super-resolution effects.

2. Proposed Approach

In this section, we first introduce the framelet transform and regression weights used in the proposed method, then construct our super-resolution model and present the optimization algorithm.

2.1. Framelet transform and image representation

Framelet transform is a mathematical method that can transform images in space domain and frequency domain [35]. To construct framelets, one starts from a compactly supported refinable function (a scaling function) $\psi \in L^2(\mathbb{R})$ with a refinement mask (low-pass filter) $\xi_0 \in L^2(\mathbb{Z})$ satisfying a refinement equation

$$\psi(x) = \sum_{l \in \mathbb{Z}} \xi_0(l) \psi(2x - l). \quad (2.1)$$

Then, for the given compactly supported refinable function, a tight framelet system can be constructed by finding an appropriate set of framelets $\Phi = \{\phi^1, \dots, \phi^r\} \subset L^2(\mathbb{R})$. Let $\{\xi_1, \dots, \xi_r\} \subset L^2(\mathbb{Z})$ be a set of framelet masks (high-pass filters), then the framelets are defined as

$$\phi^j = \sum_{l \in \mathbb{Z}} \xi_j(l) \psi(2x - l), \quad j = 1, \dots, r. \quad (2.2)$$

Thus, the construction of framelets Φ amounts to design a tight framelet filter bank $\{\xi_0, \xi_1, \dots, \xi_r\}$ [36,37]. The unitary extension principle (UEP) in [38] gives the condition for $\chi(\Phi)$ to form as a tight frame system, i.e., the filter bank $\{\xi_0, \xi_1, \dots, \xi_r\}$ satisfies

$$\zeta_{\xi_0}(\omega) \overline{\zeta_{\xi_0}(\omega + \lambda\pi)} + \sum_{j=1}^r \zeta_{\xi_j}(\omega) \overline{\zeta_{\xi_j}(\omega + \lambda\pi)} = \delta(\lambda), \quad \lambda = 0, 1, \quad (2.3)$$

for almost all $\omega \in \mathbb{R}$. Here $\zeta_\xi(\omega) = \sum_l \xi(l)e^{il\omega}$ and $\delta(\lambda)$ is a delta function. Based on the UEP, a piecewise linear B-spline can be used as the refinable function ψ . The refinement mask is $\xi_0 = [\frac{1}{4}, \frac{1}{2}, \frac{1}{4}]$, and the two corresponding high-pass filters are

$$\xi_1 = [-\frac{1}{4}, \frac{1}{2}, -\frac{1}{4}], \quad \xi_2 = [-\frac{\sqrt{2}}{4}, 0, -\frac{\sqrt{2}}{4}]. \quad (2.4)$$

We usually refer to the process of image transformation from spatial domain to frequency domain as framelet decomposition, and vice versa as framelet reconstruction. In the numerical scheme of framelet transform, we use \mathcal{F} to represent framelet decomposition operator, it decompose the image into low frequency and high frequency coefficients, and can be written as

$$\mathcal{F} = \begin{bmatrix} \mathcal{F}_0 \\ \mathcal{F}_1 \end{bmatrix}, \quad (2.5)$$

where \mathcal{F}_0 denotes the low-pass filter operator that can retain the main information of the image, and \mathcal{F}_1 consists of remaining band-pass and high-pass filter operators that can separate the edge information of multiple directions of the image. Based on the unitary extension principle (UEP) [38], the following equation holds

$$\mathcal{F}^T \mathcal{F} = \mathcal{F}_0^T \mathcal{F}_0 + \mathcal{F}_1^T \mathcal{F}_1 = \mathcal{I}, \quad (2.6)$$

where \mathcal{F}^T and \mathcal{I} are the inverse framelet transform and equivalent transform, respectively.

The framelet decomposition coefficients of an MRI image \mathbf{Y} is given by

$$\mathcal{F} \mathbf{Y} = \begin{bmatrix} \mathcal{F}_0 \mathbf{Y} \\ \mathcal{F}_1 \mathbf{Y} \end{bmatrix}, \quad (2.7)$$

where $\mathcal{F}_0 \mathbf{Y}$ is the approximation coefficients, and $\mathcal{F}_1 \mathbf{Y}$ contains detail coefficients. Accordingly, framelet reconstruction, as the inverse process of decomposition, can be expressed as

$$\mathcal{F}^T \mathbf{Y} = \begin{bmatrix} \mathcal{F}_0^T \mathbf{Y} \\ \mathcal{F}_1^T \mathbf{Y} \end{bmatrix}. \quad (2.8)$$

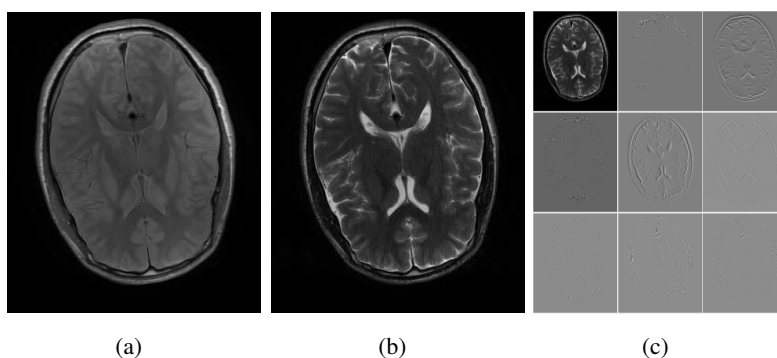


Figure 1. Schematic diagram of multimodal brain MR image and its framelet decomposition. (a) T1-weighted MR image, (b) T2-weighted MR image, (c) one level framelet decomposition results of T2 image.

Figure 1(c) shows an intuitive example of framelet decomposition on image (b), as we can see that Figure 1(c) contains nine subgraphs after one level decomposition, in which the upper left corner is approximation coefficient and the remaining eight are detail coefficients (we scale the display based on the range of pixel values in each detail subimage for better visual effects).

2.2. Shared weights in multi-modal MR images

Taking MRI images of T1 and T2 modes as an example, Zheng et al. [33] proved that based on NEDI that the local regression weight of one high-resolution image could be used as interpolation coefficient to amplify low-resolution images with different contrast but the same structure in another mode, that is, high-resolution T1 and T2 images shared similar local regression weights. The core idea of [33] is to estimate the regression weight from the local edge position of the high-resolution T1 image, and then perform a weighted summation of the four adjacent pixels in the corresponding position in the low-resolution T2 image with the regression weight to obtain a new pixel, which is the interpolation pixel for the target T2 image.

As described in [33], the regression weights $\mathbf{b}_i \in \mathbb{R}^{4 \times 1}$ estimated from the position of the i -th edge pixel of T1 image is used to interpolate T2 image, and the new pixel y_i generated in T2 image can be solved by the following formula:

$$y_i = \mathbf{b}_i^T \mathbf{a}_i, \quad (2.9)$$

where $\mathbf{a}_i \in \mathbb{R}^{4 \times 1}$ contains the four nearest pixels of T2 interpolation image along the diagonal direction at the i -th pixel position. That is, [33] proved that in the image edge region, a newly inserted pixel y_i in T2 image can be represented by a linear combination of its four diagonal neighborhood pixels, while the T1 image with the same structure as T2 can provide the linear regression weights used for interpolation.

2.3. The proposed model

As mentioned above, the purpose of our model is to recover high-resolution T2-weighted image from low-resolution T2-weighted image \mathbf{L}_{T2} with the help of high-resolution T1-weighted image. Let's denote \mathbf{X} the image magnified by the traditional interpolation method which is the same size as \mathbf{H}_{T1} , and \mathbf{Y} as the super-resolution result of our method. We can combine the regression weight vectors of all pixels and write them in matrix form, the initialized regression weighting matrix is denoted as $\mathbf{B} \in \mathbb{R}^{P \times P}$, here $P = M \times N$. When the coordinates of the interpolated pixel is (m, n) , then the $(m \times n)$ -th row of \mathbf{B} is the corresponding regression coefficients of $[b_1, b_2, b_3, b_4]$, and their column coordinates are the pixel numbers of the four diagonal neighborhoods of the current pixel. To allow interpolation weights can be updated and adjusted, we introduce additional variables $\mathbf{W} \in \mathbb{R}^{P \times P}$, which is the fine-tuning regression weighting matrix to be solved. Combined with framelet decomposition, our super-resolution model is constructed as the following optimization problem:

$$\begin{aligned} \min_{\mathbf{Y}, \mathbf{W}} \quad & \frac{1}{2} \|\mathbf{D}\mathbf{Y} - \mathbf{L}_{T2}\|_{\mathbb{F}}^2 + \frac{\alpha}{2} \|\mathcal{F}_0 \mathbf{Y} - \mathcal{F}_0 \mathbf{W}\mathbf{X}\|_{\mathbb{F}}^2 \\ & + \frac{\beta}{2} \|\mathcal{F}_1 \mathbf{Y} - \mathcal{F}_1 \mathbf{W}\mathbf{X}\|_{\mathbb{F}}^2 + \frac{\gamma}{2} \|\mathbf{B} - \mathbf{W}\|_{\mathbb{F}}^2, \end{aligned} \quad (2.10)$$

where $\|\mathbf{D}\mathbf{Y} - \mathbf{L}_{T2}\|_{\mathbb{F}}^2$ is the data fidelity term, and D represents a downsampling operator. Parameters α, β and γ are all non-negative, which are used to balance the constraint terms of super-resolution images and regression weights respectively.

An overview of the proposed approach is summarized in Figure 2. First, the weighting matrix is calculated based on [33] for T1-weighted HR image. Second, the T2-weighted LR image is pre-interpolated by the bicubic method. Then, the HR image of interest and its weighted reference image are decomposed by framelet transformation into high-frequency and low-frequency coefficients. Finally, the HR image and the weight matrix are optimized collaboratively based on Eq (2.10).

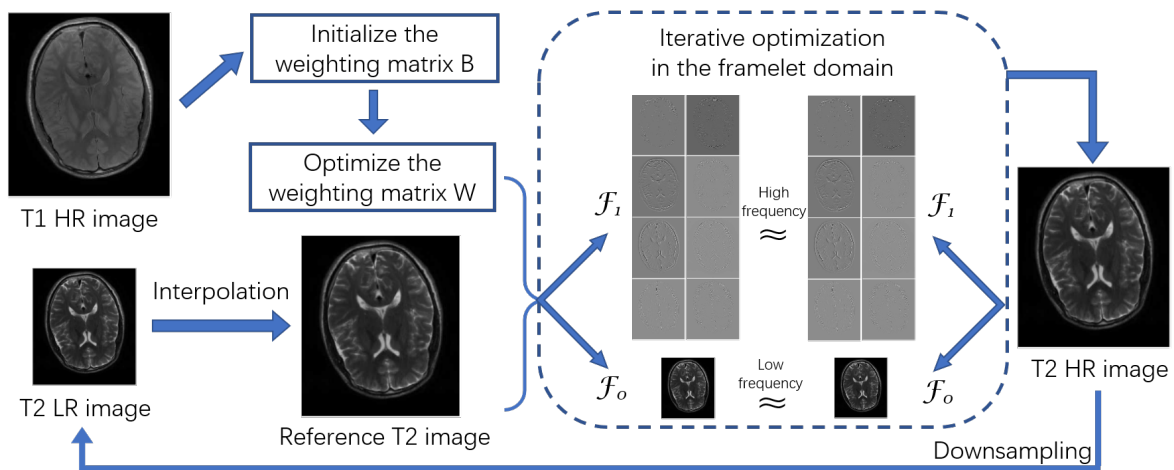


Figure 2. Block diagram of the proposed method.

In particular, the motivation for this model is derived from our experimental observations. By comparing the differences between the ground-truth high-resolution image and the interpolation enhanced image using the method of [33], we find that the different values of the high-frequency coefficients is significantly smaller than the low-frequency coefficients. Figure 3(a) displays the average pixel errors of the super-resolution reconstruction result and the HR image on the framelet decomposition coefficients, in which, the first 9 histograms correspond to 9 decomposition subgraphs, and the 10th histograms shows the average error of the 8 detail graphs.

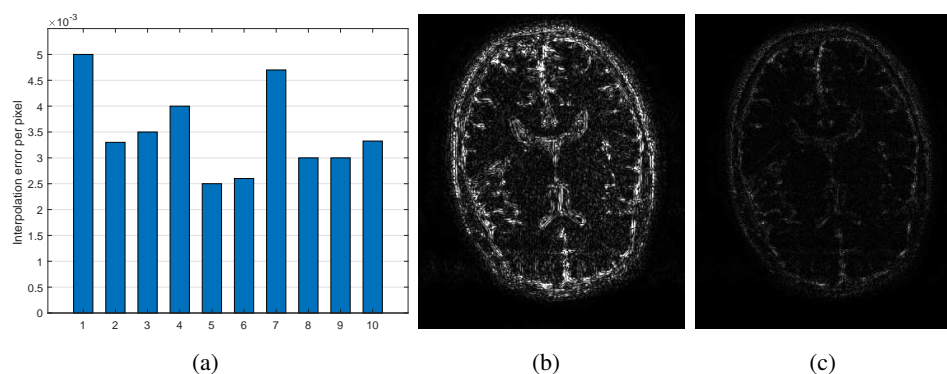


Figure 3. Comparison of similarity between the original HR image and the interpolated image of [33] on the framelet decomposition results. (a) the reconstruction error of framelet coefficients, (b) the difference image of approximation image, (c) the average difference image of eight detail images.

We also illustrate the absolute value of such differences as difference maps in Figure 3(b) and (c), the

former shows the reconstruction errors on the low-frequency coefficients between the Figure 1(b) and its interpolated version obtained by [33], and the latter shows that on the high-frequency coefficients, darker pixel means a smaller difference. Therefore, this phenomenon prompted us to use wavelet decomposition to separate the super-resolution enhanced image into approximate and detailed components, that is, in the model, through parameter adjustment, more emphasis is placed on the fidelity of high-frequency detail components, while more freedom is given to the low-frequency components of the image.

To sum up, different from [33], which only reconstructed the edge pixels of the enlarged LR image, our model considers the global pixel reconstruction and uses framelet decomposition to more finely separate the edge structure of the enlarged LR image in multiple directions. In addition, the regression weights can also be globally fine-tuned in optimization to achieve better results.

2.4. Numerical algorithm

Our model can be efficiently solved by the alternating direction method of multipliers (ADMM) [39]. We introduce an auxiliary variable \mathbf{V} , and let $\mathbf{V} = \mathbf{W}\mathbf{X}$, the optimization problem (2.10) is equivalently written in the following form:

$$\begin{aligned} \min_{\mathbf{Y}, \mathbf{W}, \mathbf{V}} \quad & \frac{1}{2} \|\mathbf{D}\mathbf{Y} - \mathbf{L}_{T2}\|_{\mathbb{F}}^2 + \frac{\alpha}{2} \|\mathcal{F}_0\mathbf{Y} - \mathcal{F}_0\mathbf{V}\|_{\mathbb{F}}^2 \\ & + \frac{\beta}{2} \|\mathcal{F}_1\mathbf{Y} - \mathcal{F}_1\mathbf{V}\|_{\mathbb{F}}^2 + \frac{\gamma}{2} \|\mathbf{B} - \mathbf{W}\|_{\mathbb{F}}^2 \\ & + \frac{\rho}{2} \|\mathbf{V} - \mathbf{W}\mathbf{X}\|_{\mathbb{F}}^2 + \langle \mathbf{G}, \mathbf{V} - \mathbf{W}\mathbf{X} \rangle, \end{aligned} \quad (2.11)$$

where \mathbf{G} is the matrix of the Lagrange multipliers, and ρ is the penalty parameter.

To solve the super-resolution image \mathbf{Y} , take the first variation of Eq (2.11) with respect to \mathbf{Y} equal to 0:

$$D^T(\mathbf{D}\mathbf{Y} - \mathbf{L}_{T2}) + \alpha\mathcal{F}_0^T(\mathcal{F}_0\mathbf{Y} - \mathcal{F}_0\mathbf{V}) + \beta\mathcal{F}_1^T(\mathcal{F}_1\mathbf{Y} - \mathcal{F}_1\mathbf{V}) = 0, \quad (2.12)$$

which actually lead to

$$(\mathbf{D}^T\mathbf{D} + \mathbf{K})\mathbf{Y} = \mathbf{D}^T\mathbf{L}_{T2} + \mathbf{K}\mathbf{V}, \quad (2.13)$$

where \mathbf{K} is an operator which is defined as

$$\mathbf{K} = \alpha\mathcal{F}_0^T\mathcal{F}_0 + \beta\mathcal{F}_1^T\mathcal{F}_1. \quad (2.14)$$

We have the following property for \mathbf{K} : assume ξ_0 is the corresponding low-pass filter of \mathcal{F}_0 , \mathcal{T} is the fast Fourier transform (FFT), and $\mathcal{T}(\cdot)^T$ is the complex conjugate of $\mathcal{T}(\cdot)$, then, $\mathcal{T}(\mathbf{K}) = (\alpha - \beta)\mathcal{T}(\xi_0)^T\mathcal{T}(\xi_0) + \beta$. Thanks to Eq (2.6), we deduce that

$$\mathbf{K} = \alpha\mathcal{F}_0^T\mathcal{F}_0 + \beta\mathcal{F}_1^T\mathcal{F}_1 = (\alpha - \beta)\mathcal{F}_0^T\mathcal{F}_0 + \beta\mathbf{I}. \quad (2.15)$$

Then we have

$$\mathcal{T}(\mathbf{K}) = (\alpha - \beta)\mathcal{T}(\mathcal{F}_0^T\mathcal{F}_0) + \beta = (\alpha - \beta)\mathcal{T}(\mathcal{F}_0)^T\mathcal{T}(\mathcal{F}_0) + \beta. \quad (2.16)$$

In numerical scheme, $\mathcal{T}(\mathcal{F}_0)$ is essentially the $\mathcal{T}(\xi_0)$. Therefore,

$$\mathcal{T}(\mathbf{K}) = (\alpha - \beta)\mathcal{T}(\xi_0)^T\mathcal{T}(\xi_0) + \beta. \quad (2.17)$$

Using FFT, we can obtain the closed form solution of \mathbf{Y} from Eq (2.12) directly,

$$\mathbf{Y} = \mathcal{T}^{-1} \left(\frac{\mathcal{T}(K) \odot \mathcal{T}(\mathbf{V}) + \mathcal{T}(D^T \mathbf{L}_{T2})}{\mathcal{T}(K) + \mathcal{T}(D^T D)} \right), \quad (2.18)$$

where \odot is dot product operator, and \mathcal{T}^{-1} denotes the inverse FFT.

Meanwhile, to solve \mathbf{V} , similar to \mathbf{Y} , take the first variation of Eq (2.11) with respect to \mathbf{V} equal to 0, we can obtain

$$(S + \rho)\mathbf{V} = S\mathbf{Y} + \rho\mathbf{W}\mathbf{X} - \mathbf{G}, \quad (2.19)$$

where S is an operator which is defined as

$$S = \mathcal{F}_0^T \mathcal{F}_0 + \alpha \mathcal{F}_1^T \mathcal{F}_1. \quad (2.20)$$

Using FFT, we can obtain the closed form solution of \mathbf{V} from Eq (2.19) directly,

$$\mathbf{V} = \mathcal{T}^{-1} \left(\frac{\mathcal{T}(S) \odot \mathcal{T}(\mathbf{Y}) + \rho \mathcal{T}(\mathbf{W}\mathbf{X} - \mathbf{G})}{\mathcal{T}(S) + \rho} \right). \quad (2.21)$$

To solve the regression weighting matrix \mathbf{W} , take the first variation of Eq (2.11) with respect to \mathbf{W} equal to 0, we can obtain

$$\mathbf{W} = (\beta + \rho \mathbf{X}\mathbf{X}^T)^{-1} (\beta \mathbf{B} + \rho \mathbf{V} + \mathbf{G}\mathbf{X}). \quad (2.22)$$

The last step is to update the Lagrange multiplier \mathbf{G} using the fastest descent method, and the update formula is as follows:

$$\mathbf{G} = \mathbf{G} + \rho(\mathbf{V} - \mathbf{W}\mathbf{X}). \quad (2.23)$$

Through the iterative update of the above variables, we can get the optimized super-resolution enhanced image \mathbf{Y} . The stop rule for our algorithm is when the number of iterations t reaches the maximum t_{max} or the variable \mathbf{Y} reaches the convergence condition $\|\mathbf{Y}^{(t)} - \mathbf{Y}^{(t-1)}\|_F / \|\mathbf{Y}^{(t)}\|_F < \epsilon$. We set $t_{max} = 300$ and $\epsilon = 10^{-4}$ in the following experiments.

Finally, we give some discussions for the computational complexity of the proposed algorithm. The computation of framelet transform \mathcal{F} is essentially the convolution, thus the complexity of $\mathcal{F}\mathbf{Y}$ is $O(P \log N)$ when using FFT. Besides, since the downsampling operator D and the T2 reference HR image \mathbf{X} are known, the associated calculations can be computed in advance outside of the iterative loop. The computational complexities of the four variables $\{\mathbf{Y}, \mathbf{V}, \mathbf{W}, \mathbf{G}\}$ that need to be optimized and updated in the iterations are $O(P \log N + P)$, $O(P \log N + P)$, $O(P)$ and $O(P)$, respectively. Thus, the overall complexity is $O(P \log N + P)$.

3. Experiments and analysis

We reiterate that the purpose of this paper is to achieve super-resolution enhancement of low-resolution T2 images by using high-resolution T1 images that can be quickly imaged as prior information, so as to reduce the time required for multi-modal imaging of the same target. In the experiments, we use accurately registered T1 and T2 image pairs as ground-truth, and a simulated data pair and two real data pairs of brain MRI images are used to verify the effectiveness of the proposed method. The synthetic T1 and T2 images shown in Figure 5 are generated from the BrainWeb [40] using default parameters and

have a spatial resolution of 256×256 pixels. The first realistic T1 (TR = 2000 ms, TE = 9.7 ms) and T2 (TR = 5000 ms, TE = 97 ms) images that are shown in Figure 6 are acquired by a 3T Siemens Trio Tim MRI scanner (FOV = 230×187 mm², slice thickness = 5.0 mm) and have a spatial resolution of 384×324 pixels. The second realistic T1 (TR = 700 ms, TE = 15 ms) and T2 (TR = 3500 ms, TE = 90 ms) images that are shown in Figure 7 are acquired by a Philips 3T MRI scanner (FOV = 260×260 mm², slice thickness = 4.0 mm) and have a spatial resolution of 512×512 pixels. Besides, in order to simulate low-resolution T2 image, we first smooth the high-resolution T2 image with 3×3 Gaussian filter with standard deviation 0.5 and then down-sample it by a factor of two, then the high-resolution T2 image are used as the ground-truth to quantitatively evaluate the performance of our method. These practices are the same as described in [33].

As usual, we use Peak Signal-to-Noise Ratio (PSNR) and structural similarity (SSIM) [41] to objectively evaluate the image quality of super-resolution results. PSNR and SSIM reflect the global similarity between super-resolution enhanced images and the ground-truth from the perspective of image pixels and image structures respectively. The higher the PSNR value, the closer the reconstructed pixel value is to the pixel value in the HR source image. The higher the SSIM value, the more accurate the reconstructed edge. That is to say, the larger the two values are, the better the super-resolution effect is. In the following experiments, we compare the proposed method with the four methods: bicubic interpolation [6], NEDI [7], CGI [9] and Zheng et al. [33], which have similar technical principles, and demonstrate the effectiveness of the proposed approach from both subjective visual effects and objective index evaluations.

For parameter settings, we adopt grid searches to adjust the parameters α , β , and γ in Eq (2.10) to make PSNR and SSIM reach the best, while the barrier parameter ρ is updated iteratively to speed up the convergence (following closely Boyd et al. [42]). We found that for the three datasets, selecting the optimal parameters from a fixed parameter set $\Psi = [0.01, 0.05, 0.1, 0.3, 0.5, 1, 1.5, 3, 5, 10]$ was enough to demonstrate the effectiveness of the proposed method. Taking the Siemens MR image as an example, we observe the influence of model parameters on experimental results according to the numerical changes of quantitative indicators under different parameters. Figure 4 shows the stability of a single parameter's value change on PSNR and SSIM when the other two parameters are fixed as optimal. It can be seen that the experimental results are the least sensitive to γ , and desirable results can be achieved as long as α and β are within the appropriate range, and the influence trends of parameter changes on PSNR and SSIM are basically the same. As for other comparison methods, we also adjusted them to obtain the optimal effects.

The super-resolution visual effects of the three experimental datasets are shown in Figures 5, 6 and 7, respectively. In order to get a better look at the edge enhancement quality of the super-resolution, we give a magnified view of the local detail in the second line of each set of figures.

As we can see from Figure 5, the image enhanced by bicubic interpolation and NEDI algorithm has edge fracture, artifact and blur, while the overall reconstruction quality of CGI is better than bicubic and NEDI, but the recovery effect is not good enough for small structures. The method proposed in [33] is an improvement on the basis of CGI, so the edge structure of its reconstructed image is more complete. On the whole, through the visual contrast of local amplification, we can clearly observe that the super-resolution image obtained by our method has the clearest edges and is closest to the HR source image in terms of detail structures. For experiments on real MR images, we can get similar observations in the Philips image comparison in Figure 6, while the visual contrast in Figure 7 is less obvious.

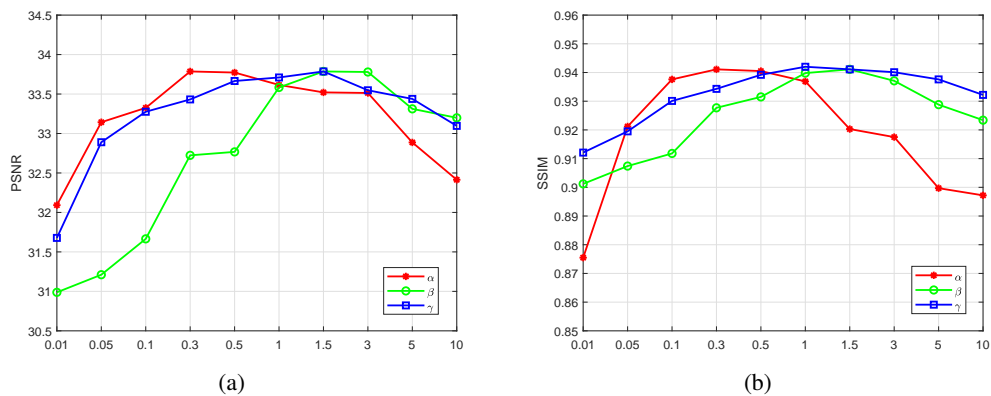


Figure 4. The stability test of three parameters on two evaluation indicators (PSNR and SSIM) takes Siemens MR image shown in Figure 6 as an example.

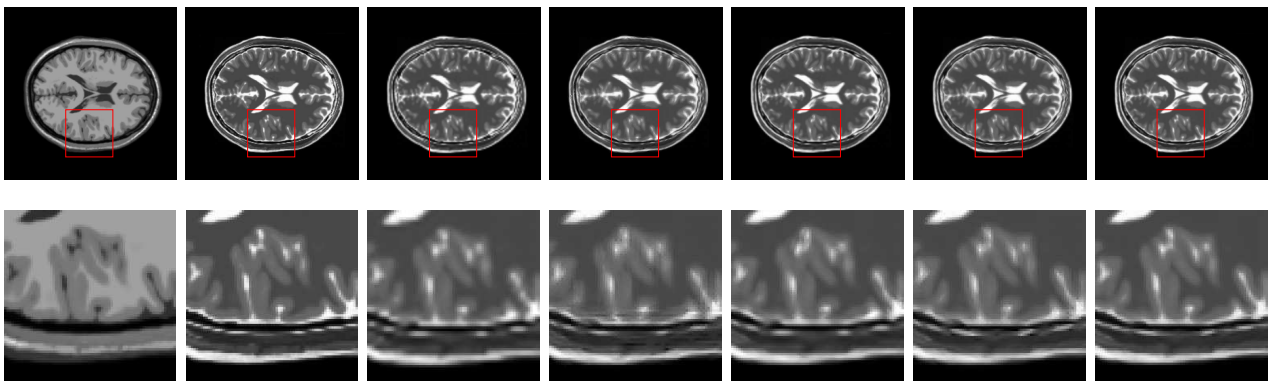


Figure 5. Super-resolution results of BrainWeb MR data. The first row from left to right: T1 HR image; T2 HR image; the result of bicubic interpolation; NEDI; CGI; method of [33]; our method. The second row from left to right: the magnification of the red rectangle in the corresponding each example.

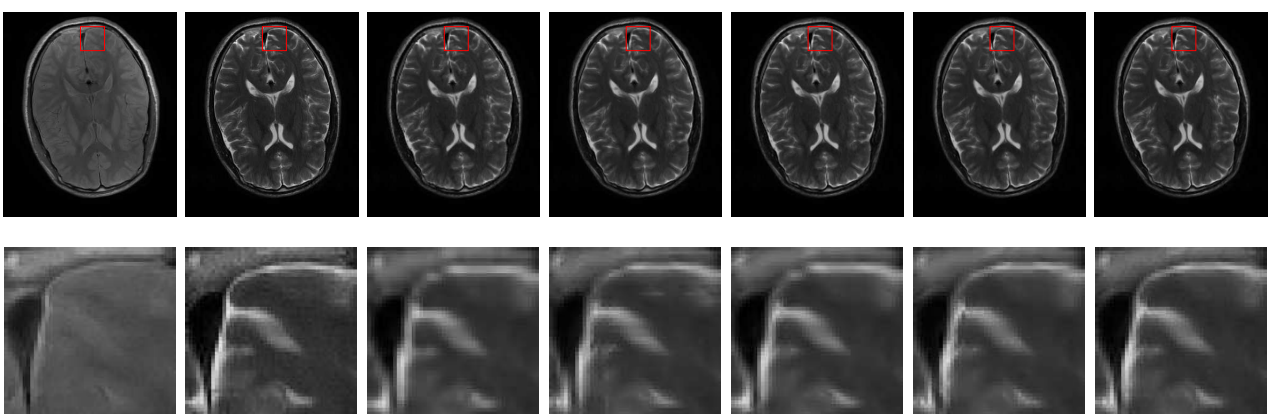


Figure 6. Super-resolution results of Siemens MR data. The first row from left to right: T1 HR image; T2 HR image; the result of bicubic interpolation; NEDI; CGI; method of [33]; our method. The second row from left to right: the magnification of the red rectangle in the corresponding each example.

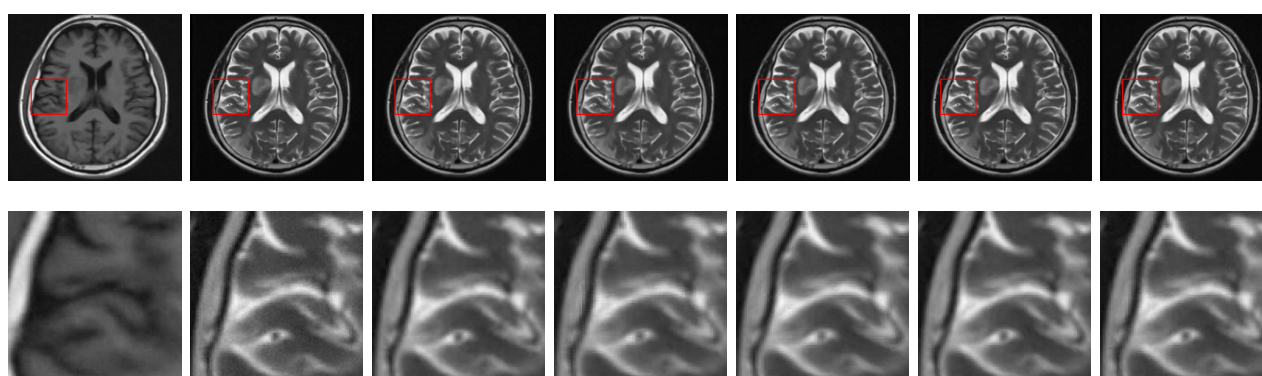


Figure 7. Super-resolution results of Philips MR data. The first row from left to right: T1 HR image; T2 HR image; the result of bicubic interpolation; NEDI; CGI; method of [33]; our method. The second row from left to right: the magnification of the red rectangle in the corresponding each example.

For objective comparison, PSNR and SSIM values obtained by different methods on the three sets of MR images are listed in Table 1. The PSNR values obtained by our method on BrainWeb data and Siemens data are more than 0.6 values higher than the suboptimal method, and the SSIM value is about 0.006 value higher than that of the suboptimal method. Therefore, we can directly observe the differences in the reconstructed edges from the detailed magnification of Figures 5 and 6. For Philips data, however, the proposed method only achieved a PSNR of about 0.1 and SSIM of about 0.001 higher than the suboptimal method, so the visual contrast differences between the results of each method in Figure 7 are not obvious by macroscopic observation.

Table 1. Super-resolution results of different methods on three sets of MR images.

Images		Bicubic [6]	NEDI [7]	CGI [9]	[33]	Our
BrainWeb	PSNR	23.5458	27.5649	28.2868	27.6197	28.9182
	SSIM	0.9100	0.9558	0.9659	0.9606	0.9723
Siemens	PSNR	29.3926	32.6016	33.0875	33.1506	33.7861
	SSIM	0.8986	0.9282	0.9341	0.9345	0.9411
Philips	PSNR	30.7899	31.9606	32.9059	33.7545	33.8562
	SSIM	0.9150	0.9141	0.9237	0.9261	0.9270

Table 2. The running time in seconds for different methods on three sets of MR images.

Images	Bicubic [6]	NEDI [7]	CGI [9]	[33]	Our
BrainWeb	0.0018	2.4389	0.0495	2.5089	4.8772
Siemens	0.0021	4.7875	0.0984	4.8831	9.6236
Philips	0.0029	11.7684	0.2432	11.9107	23.0549

Besides, all experimental algorithms are implemented with MATLAB on a personal computer with CPU 2.20 GHz and 16 GB memory. The computation time of different methods are shown in Table 2.

As can be seen from the table, the proposed method is relatively time-consuming.

To sum up, by comparing two quantitative indicators between different algorithms, we can conclude that subjective visual contrast is basically consistent with objective quantitative evaluation. When the visual contrast difference is not obvious, we rely on quantitative indicators to judge the effectiveness of the algorithm.

4. Conclusions

In this paper, a super-resolution model for multi-modal MR images is proposed for T1-weighted and T2-weighted image pairs. In order to overcome the inaccuracy of edge localization and fixed interpolation weights of the earlier methods, we use framelet decomposition to finely separate the image edges and detail components, then construct a global interpolation weighting matrix, and uses an iterative algorithm to co-optimize the super-resolution image and interpolation weights. Multiple experimental results show that the proposed method can effectively improve the visual effects and evaluation indexes of the enhanced images. In future work, we will further explore the prior information that can accurately describe the intrinsic correlation of multimodal images, so as to construct more effective image resolution enhancement models. In addition, we will also consider the combination of frequency domain priori and deep learning models to improve the interpretability of the super-resolution results.

Acknowledgments

This work was supported in part by the Zhejiang Provincial Natural Science Foundation of China under Grant No. LQ21F020001, and LQ21A010001, and in part by the National Natural Science Foundation of China under Grant 62101375, and 62006168.

Conflict of interest

The authors declare that there is no conflict of interest.

References

1. L. Hua, Y. Gu, X. Gu, J. Xue, T. Ni, A novel brain MRI image segmentation method using an improved multi-view fuzzy c-means clustering algorithm, *Front. Neurosci.*, **15** (2021), 662674. <https://doi.org/10.3389/fnins.2021.662674>
2. A. Hatamizadeh, V. Nath, Y. Tang, D. Yang, H. R. Roth, D. Xu, Swin UNETR: Swin transformers for semantic segmentation of brain tumors in MRI images, in *International MICCAI Brainlesion Workshop*, Springer, (2022), 272–284. https://doi.org/10.1007/978-3-031-08999-2_22
3. H. Greenspan, Super-resolution in medical imaging, *Comput. J.*, **52** (2009), 43–63. <https://doi.org/10.1093/comjnl/bxm075>
4. D. Qiu, Y. Cheng, X. Wang, Gradual back-projection residual attention network for magnetic resonance image super-resolution, *Comput. Meth. Prog. Bio.*, **208** (2021), 106252. <https://doi.org/10.1016/j.cmpb.2021.106252>

5. L. Wang, H. Zhu, Z. He, Y. Jia, J. Du, Adjacent slices feature transformer network for single anisotropic 3D brain MRI image super-resolution, *Biomed. Signal Proces.*, **72** (2022), 103339. <https://doi.org/10.1016/j.bspc.2021.103339>
6. R. Keys, Cubic convolution interpolation for digital image processing, in *IEEE Transactions on Acoustics, Speech and Signal Processing*, IEEE: Piscataway, (1981), 1153–1160. <https://doi.org/10.1109/TASSP.1981.1163711>
7. X. Li, M. Orchard, New edge-directed interpolation, *IEEE T. Image Process.*, **10** (2001), 1521–1527. <https://doi.org/10.1109/83.951537>
8. J. Manjón, P. Coupé, A. Buades, V. Fonov, D. L. Collins, M. Robles, Non-local MRI upsampling, *Med. Image Anal.*, **14** (2010), 784–792. <https://doi.org/10.1016/j.media.2010.05.010>
9. Z. Wei, K.-K. Ma, Contrast-guided image interpolation, *IEEE T. Image Process.*, **22** (2013), 4271–4285. <https://doi.org/10.1109/TIP.2013.2271849>
10. F. Shi, J. Cheng, L. Wang, P. T. Yap, D. Shen, LRTV: MR image super-resolution with low-rank and total variation regularizations, *IEEE T. Med. Imaging*, **34** (2015), 2459–2466. <https://doi.org/10.1109/TMI.2015.2437894>
11. S. Tourbier, X. Bresson, P. Hagmann, J. P. Thiran, R. Meuli, M. B. Cuadra, An efficient total variation algorithm for super-resolution in fetal brain MRI with adaptive regularization, *NeuroImage*, **118** (2015), 584–597. <https://doi.org/10.1016/j.neuroimage.2015.06.018>
12. A. Rueda, N. Malpica, E. Romero, Single-image super-resolution of brain MR images using overcomplete dictionaries, *Med. Image Anal.*, **17** (2013), 113–132. <https://doi.org/10.1016/j.media.2012.09.003>
13. D. Zhang, J. He, Y. Zhao, M. Du, MR image super-resolution reconstruction using sparse representation, nonlocal similarity and sparse derivative prior, *Comput. Biol. Med.*, **58** (2015), 130–145. <https://doi.org/10.1016/j.compbiomed.2014.12.023>
14. Y. Jia, Z. He, A. Gholipour, S. K. Warfield, Single anisotropic 3-D MR image upsampling via overcomplete dictionary trained from in-plane high resolution slices, *IEEE J. Biomed. Health*, **20** (2016), 1552–1561. <https://doi.org/10.1109/JBHI.2015.2470682>
15. Y. Jia, A. Gholipour, Z. He, S. K. Warfield, A new sparse representation framework for reconstruction of an isotropic high spatial resolution MR volume from orthogonal anisotropic resolution scans, *IEEE T. Med. Imaging*, **36** (2017), 1182–1193. <https://doi.org/10.1109/TMI.2017.2656907>
16. Y. Huang, L. Shao, A. F. Frangi, Simultaneous super-resolution and cross-modality synthesis of 3d medical images using weakly-supervised joint convolutional sparse coding, in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, (2017), 6070–6079.
17. Y. Chen, Y. Xie, Z. Zhou, F. Shi, A. G. Christodoulou, D. Li, Brain MRI super resolution using 3D deep densely connected neural networks, in *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, IEEE, 2018, 739–742. <https://doi.org/10.1109/ISBI.2018.8363679>
18. J. Shi, Z. Li, S. Ying, C. Wang, Q. Liu, Q. Zhang, et al., MR image super-resolution via wide residual networks with fixed skip connection, *IEEE J. Biomed. Health*, **23** (2019), 1129–1140. <https://doi.org/10.1109/JBHI.2018.2843819>

19. X. Zhao, X. Hu, Y. Liao, T. He, T. Zhang, X. Zou, et al., Accurate MR image super-resolution via lightweight lateral inhibition network, *Comput. Vis. Image Und.*, **201** (2020), 103075. <https://doi.org/10.1016/j.cviu.2020.103075>
20. M. Jiang, M. Zhi, L. Wei, X. Yang, J. Zhang, Y. Li, et al., FA-GAN: Fused attentive generative adversarial networks for MRI image super-resolution, *Comput. Med. Imag. Grap.*, **92** (2021), 101969. <https://doi.org/10.1016/j.compmedimag.2021.101969>
21. H. Song, W. Yang, GSCCTL: A general semi-supervised scene classification method for remote sensing images based on clustering and transfer learning, *Int. J. Remote Sens.*, **2021** (2021), 1–25. <https://doi.org/10.1080/01431161.2021.2019851>
22. Z. Wang, J. Chen, S. C. Hoi, Deep learning for image super-resolution: A survey, *IEEE T. Pattern Anal.*, **43** (2021), 3365–3387. <https://doi.org/10.1109/TPAMI.2020.2982166>
23. Y. Li, B. Sixou, F. Peyrin, A review of the deep learning methods for medical images super resolution problems, *IRBM*, **42** (2021), 120–133. <https://doi.org/10.1016/j.irbm.2020.08.004>
24. Q. Lyu, H. Shan, C. Steber, C. Helis, C. Whitlow, M. Chan, et al., Multi-contrast super-resolution MRI through a progressive network, *IEEE Trans. Med. Imag.*, **39** (2020), 2738–2749. <https://doi.org/10.1109/TMI.2020.2974858>
25. Q. Lyu, H. Shan, G. Wang, MRI super-resolution with ensemble learning and complementary priors, *IEEE Trans. Comput. Imag.*, **6** (2020), 615–624. <https://doi.org/10.1109/TCI.2020.2964201>
26. B. M. Dale, M. A. Brown, R. C. Semelka, *MRI: Basic Principles and Applications*. John Wiley & Sons, 2015.
27. S. Rathore, A. Abdulkadir, C. Davatzikos, Analysis of MRI data in diagnostic neuroradiology, *Annu. Rev. Biomed. Data Sci.*, **3** (2020), 365–390. <https://doi.org/https://doi.org/10.1146/annurev-biodatasci-022620-015538>
28. F. Rousseau, Brain hallucination, in *European Conference on Computer Vision*, Springer, (2008), 497–508. https://doi.org/10.1007/978-3-540-88682-2_38
29. F. Rousseau, A non-local approach for image super-resolution using intermodality priors, *Med. Image Anal.*, **14** (2010), 594–605. <https://doi.org/10.1016/j.media.2010.04.005>
30. J. V. Manjón, P. Coupé, A. Buades, D. L. Collins, M. Robles, MRI superresolution using self-similarity and image priors, *Int. J. Biomed. Imag.*, **2010** (2010), 1–11. <https://doi.org/10.1155/2010/425891>
31. K. Jafari-Khouzani, MRI upsampling using feature-based nonlocal means approach, *IEEE Trans. Med. Imag.*, **33** (2014), 1969–1985. <https://doi.org/10.1109/TMI.2014.2329271>
32. X. Lu, Z. Huang, Y. Yuan, MR image super-resolution via manifold regularized sparse learning, *Neurocomputing*, **162** (2015), 96–104. <https://doi.org/10.1016/j.neucom.2015.03.065>
33. H. Zheng, X. Qu, Z. Bai, Y. Liu, D. Guo, J. Dong, et al., Multi-contrast brain magnetic resonance image super-resolution using the local weight similarity, *BMC Med. Imag.*, **17** (2017). <https://doi.org/10.1186/s12880-016-0176-2>
34. H. Zheng, K. Zeng, D. Guo, J. Ying, Y. Yang, X. Peng, et al., Multi-contrast brain MRI image super-resolution with gradient-guided edge enhancement, *IEEE Access*, **6** (2018), 57856–57867. <https://doi.org/10.1109/ACCESS.2018.2873484>

35. Y. R. Li, R. H. Chan, L. Shen, X. Zhuang, Regularization with multilevel non-stationary tight framelets for image restoration, *Appl. Comput. Harmon. Anal.*, **53** (2021), 332–348. <https://doi.org/10.1016/j.acha.2021.03.003>
36. Y. R. Li, R. H. Chan, L. Shen, Y. C. Hsu, W.-Y. Isaac Tseng, An adaptive directional haar framelet-based reconstruction algorithm for parallel magnetic resonance imaging, *SIAM J. Imag. Sci.*, **9** (2016), 794–821. <https://doi.org/10.1137/15M1033964>
37. Y. R. Li, L. Shen, X. Zhuang, A tailor-made 3-dimensional directional haar semi-tight framelet for pMRI reconstruction, *Appl. Comput. Harmon. Anal.*, **60** (2022), 446–470. <https://doi.org/10.1016/j.acha.2022.04.003>
38. A. Ron, Z. Shen, Affine systems in $l_2(\mathbb{R}^d)$: The analysis of the analysis operator, *J. Funct. Anal.*, **148** (1997), 408–447. <https://doi.org/10.1006/jfan.1996.3079>
39. E. Esser, *Primal Dual Algorithms for Convex Models and Applications to Image Restoration, Registration and Nonlocal Inpainting*, PhD thesis, University of California in Los Angeles, 2010.
40. C. A. Cocosco, V. Kollokian, K. S. Kwan, A. C. Evans, I. Centre, Brain Web: Online interface to a 3D MRI simulated brain database, *NeuroImage*, **5** (1997).
41. Z. Wang, A. C. Bovik, H. R. Sheikh, E. P. Simoncelli, Image quality assessment: from error visibility to structural similarity, *IEEE Trans. Image Process.*, **13** (2004), 600–612. <https://doi.org/10.1109/TIP.2003.819861>
42. S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein, Distributed optimization and statistical learning via the alternating direction method of multipliers, *Foundations and Trends® in Machine Learning*, **3** (2011), 1–122. <https://doi.org/10.1561/22000000016>



AIMS Press

©2023 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)