



Nonconvex Nonlocal Tucker Decomposition for 3D Medical Image Super-Resolution

Huidi Jia^{1,2,3}, Xi'ai Chen^{1*}, Zhi Han^{1,2}, Baichen Liu^{1,2,3}, Tianhui Wen⁴ and Yandong Tang^{1,2}

¹ State Key Laboratory of Robotics, Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang, China, ² Institutes for Robotics and Intelligent Manufacturing, Chinese Academy of Sciences, Shenyang, China, ³ University of Chinese Academy of Sciences, Beijing, China, ⁴ School of Professional Studies, Columbia University, New York, NY, United States

Limited by hardware conditions, imaging devices, transmission efficiency, and other factors, high-resolution (HR) images cannot be obtained directly in clinical settings. It is expected to obtain HR images from low-resolution (LR) images for more detailed information. In this article, we propose a novel super-resolution model for single 3D medical images. In our model, nonlocal low-rank tensor Tucker decomposition is applied to exploit the nonlocal self-similarity prior knowledge of data. Different from the existing methods that use a convex optimization for tensor Tucker decomposition, we use a tensor folded-concave penalty to approximate a nonlocal low-rank tensor. Weighted 3D total variation (TV) is used to maintain the local smoothness across different dimensions. Extensive experiments show that our method outperforms some state-of-the-art (SOTA) methods on different kinds of medical images, including MRI data of the brain and prostate and CT data of the abdominal and dental.

OPEN ACCESS

Edited by:

Hancan Zhu,
Shaoxing University, China

Reviewed by:

Baojie Fan,
Nanjing University of Posts and
Telecommunications, China
Yao Wang,
Xi'an Jiaotong University, China

*Correspondence:

Xi'ai Chen
chenxiai@sia.cn

Received: 21 February 2022

Accepted: 14 March 2022

Published: 25 April 2022

Citation:

Jia H, Chen X, Han Z, Liu B, Wen T
and Tang Y (2022) Nonconvex
Nonlocal Tucker Decomposition for 3D
Medical Image Super-Resolution.
Front. Neuroinform. 16:880301.
doi: 10.3389/fninf.2022.880301

Keywords: 3D super-resolution, low rank tensor decomposition, nonlocal self-similarity, 3D total variation, medical image

1. INTRODUCTION

High-resolution (HR) medical images provide rich detailed information that is critical for accurate lesion segmentation, diagnosis, and treatment (Greenspan, 2009; Shi et al., 2010). Currently, the most widely used imaging techniques in clinical settings and research include magnetic resonance imaging (MRI) and computed tomography (CT). However, both MRI and CT have their limitations. The main limitation of MRI is the balance between image quality and scan time (Bustin et al., 2018). In general, it requires a long acquisition time to obtain an HR MRI with a high signal-to-noise ratio (SNR). This not only costs a lot but also affects patients' breathing and causes motion artifacts. Contrast-enhanced CT can show small lesions more clearly and blood flow in lesions. Due to the influence of a high radiation dose and contrast agent, contrast-enhanced CT scans are not allowed in many cases, such as in patients with hyperthyroidism and hypersensitivity (Marcelino et al., 2019). For instance, micro-CT (μ CT), which is applied to determine the 3D structure of teeth, has a higher resolution than cone beam computed tomography (CBCT), but it can only be used for extracted teeth due to the acquisition time and radiation (Hatvani et al., 2019). Super-resolution provides an efficient method for obtaining HR images from low-resolution (LR) images when acquisition conditions are limited. Therefore, super-resolution has become an important research issue in image processing and is widely applied in medical imaging (Zhang et al., 2012b; Shi et al., 2015; Hatvani et al., 2019; Qiu et al., 2021; Zhao et al., 2021; Zhu and Qiu, 2021).

Super-resolution image reconstruction is a highly ill-posed problem because it predicts multiple HR images from a given LR image. To solve this ill-posed problem, various super-resolution methods have been proposed. The most direct methods are based on interpolation (Duchon, 1979; Keys, 1981), such as nearest neighbor (NN) interpolation, bilinear interpolation, and bicubic interpolation. Interpolation-based methods are fast but not very accurate. Learning-based methods (Zhang et al., 2012a; Salvador, 2016; Lim et al., 2017; Liu et al., 2017) learn example patches from fixed and finite HR image sets. Thus, the performance of these methods highly depends on the learned HR image patches. Deep learning-based methods (Ledig et al., 2017; Zhao et al., 2019; Fan et al., 2020; Liu et al., 2021) have obtained outstanding performances in high-level image processing tasks, such as alignment, segmentation, and object detection. Recently, they also began to show their advantage in low-level image processing tasks. It is widely known that sufficient data are necessary for training an effective deep learning model. However, unlike natural images, medical databases for training are not always available due to some privacy regulations and laws.

As an important and powerful modeling tool, low-rank has attracted increasing attention in many fields, such as signal processing (E-Asim et al., 2020), image processing (Xu et al., 2015; Chen et al., 2020), and machine learning (Yi et al., 2021). Adding low-rank regularization can achieve a good reconstruction result in super-resolution (Shi et al., 2015; Yair and Michaeli, 2018). Liu et al. (2016) exploited the low-rank property of an image by reshaping multidimensional data to a matrix and applying a low-rank constraint to the matrix. Veganzones et al. (2016) took advantage of the low-rank property to propose two HSI super-resolution methods by local dictionary learning using end member induction algorithms. Although these low-rank matrix methods obtain good super-resolution results, they ignore considerable structural information of data in the process of reshaping an image to a 2D matrix.

Mathematically, MRI and CT images are multidimensional data with high spatial resolution and slice resolution. Tensors provide a natural way to represent multidimensional data. A tensor is a multiway array that can be viewed as a generalization of vectors and matrixes. For 3D medical images, there is a strong correlation between slices, and a 3rd-order tensor can be used to represent an image to avoid destroying the process of reshaping data into the matrix. It has been proven that tensors are a reasonable representation that can preserve the original structure of data and significantly improve the quality of reconstruction images (Liu et al., 2013; Yin et al., 2017; Li et al., 2018; Xie et al., 2018; Hatvani et al., 2019; Prevost et al., 2019). According to different decompositions of tensor, a series of low-rank tensor methods have emerged. Liu et al. (2013) extended low-rank matrix completion to the tensor case for tensor completion. Different from the matrix, the rank of a tensor is not clearly defined and the decomposition of the tensor is not unique. The CP (CANDECOMP/PARAFAC) decomposition (Carroll and Chang, 1970) of a tensor is a representation based on a sum of several rank-1 tensors. CP decomposition based methods require a small memory space, but these methods need

to predefine the rank, and the calculation of CP rank is an NP-hard problem (Kolda and Bader, 2009). Tucker decomposition (Tucker, 1966) represents a tensor as the product of a core tensor and several factor matrixes and minimizes the rank of the core tensor and the factor matrixes. The appearance of Tucker decomposition solves the calculation problem of CP rank. Previous methods (Chen et al., 2014; Dian et al., 2017) have proved that Tucker decomposition was effective and obtained satisfactory results in many fields. Li et al. (2017) applied tensor Tucker decomposition to the tensor completion problem and utilized the trace norm as a low-rank constraint to the factors of Tucker decomposition (Tucker, 1966). In low-rank structure learning, tensor norms, e.g., trace norms and nuclear norms, penalize large entries of vectors overly and usually introduce modeling bias (Leng et al., 2006). To correct the estimation bias of the convex tensor norms, unbiased folded-concave norms are considered. Hatvani et al. (2019) proposed a tensor factorization method for 3D super-resolution and applied it to dental CT. However, they did not consider the prior information of images.

As there are often many similar structures in human organs or tissues, we consider nonlocal self-similarity in our model. Nonlocal self-similarity is an important patch-based prior. This means that for a given patch in an image, some similar patches within the whole image can be found. In many tasks, such as denoising and recovery, nonlocal similarity-based methods (Dabov et al., 2007; Mairal et al., 2009; Wang et al., 2018) have demonstrated their effectiveness. Most Tucker decomposition-based methods (Chen et al., 2014; Dian et al., 2017; Li et al., 2017; Yair and Michaeli, 2018) only employ the global prior information in their models. To further exploit the low-rank prior hidden in the data and improve the reconstruction performance of super-resolution, nonlocal similarity in the tensor cubes is exploited.

Due to statistical uncertainty in physical measurements, inevitable noise is introduced in MRI and CT data. The noise in medical images will affect the clinical diagnosis accuracy and increase the difficulty of high-level tasks such as registration and segmentation (Zhang et al., 2015; Diwakar and Kumar, 2018). TV, which is defined as the integral of the absolute gradients of the image, is an effect regularization for suppressing noise and preserving the local spatial consistency of images. Shi et al. (2015) combined both global low-rank priori and two-dimensional TV regularization to obtain a higher resolution MR image. The two-dimensional TV only considers the local smoothness without utilizing the interframe smoothness. To take advantage of the spatial and interframe local smoothness simultaneously, weighted 3-dimensional TV (3D TV) is adopted in super-resolution reconstruction.

In this article, we propose a super-resolution method for 3D medical images based on nonconvex nonlocal Tucker decomposition with weighted 3D total variation (NNTDTV). Different from most existing single image super-resolution methods, our method improves volume super-resolution, i.e., it not only improves the spatial resolution of images but also improves the slice resolution. The framework of our method is shown in **Figure 1**. The main contributions of this article are summarized as follows:

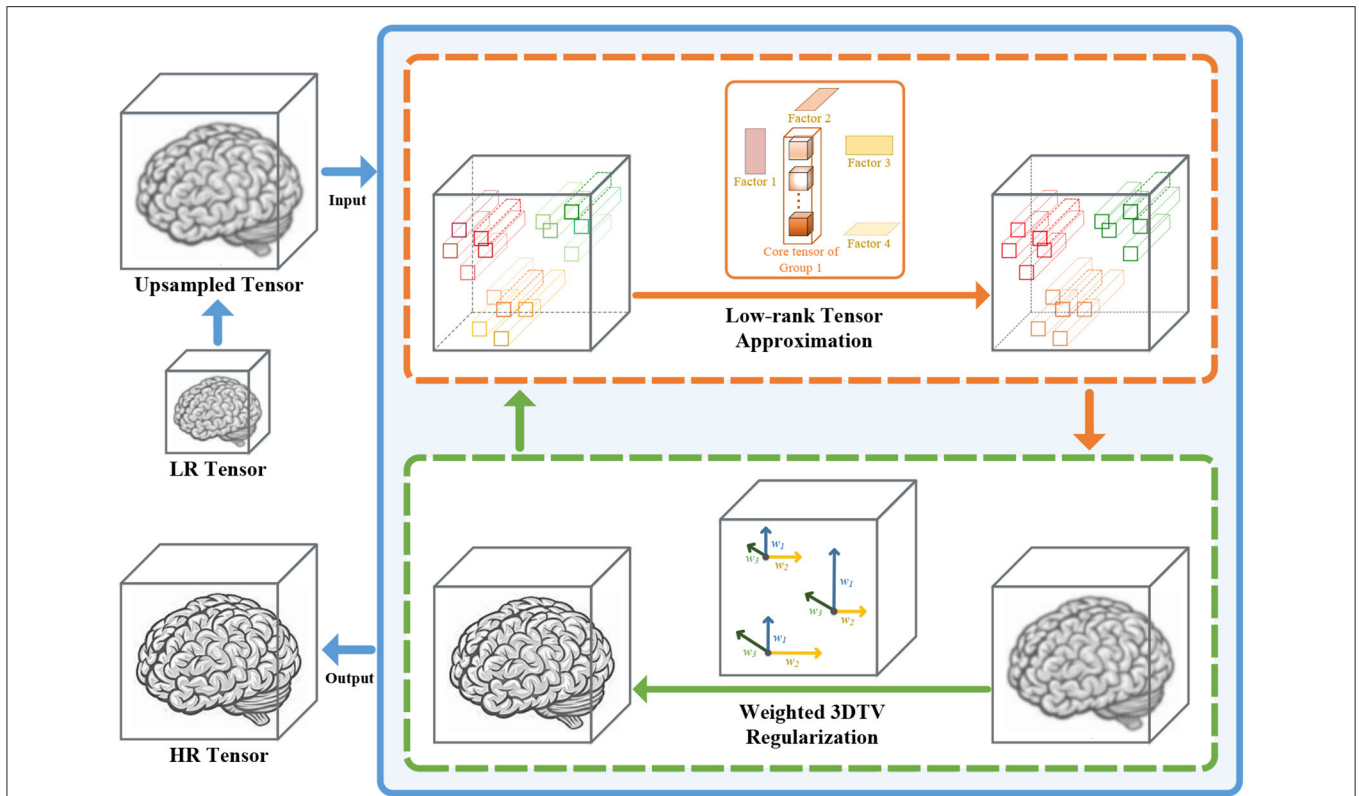


FIGURE 1 | The detailed framework of our method. The proposed method is divided into two parts, the low-rank regularization term, and the 3D total variation (TV) regularization term. The low-rank regularization term (the orange box) reconstructs the main information of the data by performing Tucker decomposition for the nonlocal similar self-similarity patches. The 3D TV regularization term (the green box) uses the local smoothness of the data to suppress the noise and keep the details.

1. We propose a nonconvex nonlocal Tucker decomposition-based super-resolution method to approximate HR medical images by utilizing the nonlocal similarity structure hidden in 3D medical images. A nonconvex optimization procedure is used to avoid estimation bias caused by the traditional convex procedure.
2. Weighted 3D total variation (TV) is used to exploit the smoothness of 3D medical images and suppress noise to some extent.
3. Extensive experiments show that, compared with the existing methods, our method has better performance and generality on different kinds of medical images. The reconstructed super-resolution images by our model have a clearer edge while maintaining local smoothness.

2. METHODS

2.1. Notation and Preliminaries

A tensor is a multidimensional array that can be viewed as a high-dimensional generalization of a matrix. We consider scalars as zero-order tensors and denote them by lowercase letters (x, y, z, \dots). Vectors and matrices are first-order and second-order tensors and denoted by bold lowercase letters ($\mathbf{x}, \mathbf{y}, \mathbf{z}, \dots$) and uppercase letters (X, Y, Z, \dots), respectively. A high-order tensor can be expressed as $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$, and

its element is x_{i_1, i_2, \dots, i_N} . The mode- n matricization or unfolding of a tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$ is the operation of reshaping a tensor into a matrix $X_{(n)} \in \mathbb{R}^{I_n \times (I_1 \dots I_{n-1} I_{n+1} \dots I_N)}$. The elements ($i_1, \dots, i_{n-1}, i_n, i_{n+1}, \dots, i_N$) of matrix $X_{(n)}$ satisfy:

$$j = 1 + \sum_{k=1, k \neq n}^N (i_k - 1) J_k \text{ with } J_k = \prod_{m=1, m \neq n}^{k-1} I_m. \quad (1)$$

For two tensors of the same size, their inner product is defined as:

$$\langle \mathcal{X}, \mathcal{Y} \rangle := \sum_{i_1, i_2, \dots, i_K} x_{i_1, i_2, \dots, i_K} \cdot y_{i_1, i_2, \dots, i_K}. \quad (2)$$

The corresponding Frobenius norm is defined as:

$$\|\mathcal{X}\|_F = \sqrt{\langle \mathcal{X}, \mathcal{X} \rangle}. \quad (3)$$

A tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$ and a matrix $Z \in \mathbb{R}^{J \times I_n}$'s mode- n product is:

$$\mathcal{Y} = \mathcal{X} \times_n Z \in \mathbb{R}^{I_1 \times \dots \times I_{n-1} \times J \times I_{n+1} \times \dots \times I_N}, \quad (4)$$

with entries

$$y_{i_1, \dots, i_{n-1}, j, i_{n+1}, \dots, i_N} = \sum_{i_n=1}^{I_n} (x_{i_1, \dots, i_N} z_{j, i_n}). \quad (5)$$

For a tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$, its Tucker decomposition is defined as:

$$\mathcal{X} = \sum_{s_1=1}^{S_1} \times \sum_{s_N=1}^{S_N} \mathcal{G}_{s_1 s_2 \dots s_N} [u_{s_1}^{(1)} \circ u_{s_2}^{(2)} \circ \dots \circ u_{s_N}^{(N)}] \quad (6)$$

$$= \mathcal{G} \times_1 U^{(1)} \times_2 U^{(2)} \dots \times_N U^{(N)}.$$

where $\mathcal{G} \in \mathbb{R}^{S_1 \times S_2 \times \dots \times S_N}$ denotes the core tensor and $\{U^{(n)}\}_{n=1}^N \in \mathbb{R}^{I_n \times S_n}$ denote the decomposition factors by Tucker decomposition of the tensor.

2.2. Problem Formulation

Affected by acquisition modality, motion blur, and noise, the observation model of an LR image can be expressed as:

$$\mathcal{Y} = DS\mathcal{X} + \varepsilon, \quad (7)$$

where $\mathcal{Y} \in \mathbb{R}^{w \times h \times S}$ denotes the observed LR image, D denotes the downsampling operator, S denotes the blurring operator, ε is the observation noise, and $\mathcal{X} \in \mathbb{R}^{W \times H \times S}$ is the HR image that we want to reconstruct. To reconstruct an HR image, we can estimate \mathcal{X} by minimizing

$$\min_{\mathcal{X}} \|DS\mathcal{X} - \mathcal{Y}\|_F^2. \quad (8)$$

According to Equation (8), recovering HR \mathcal{X} from \mathcal{Y} is an ill-posed inverse problem. This problem can be solved by introducing some regularization terms. Thus, we obtain the following cost function:

$$\min_{\mathcal{X}} \|DS\mathcal{X} - \mathcal{Y}\|_F^2 + \lambda R(\mathcal{X}), \quad (9)$$

where $\|DS\mathcal{X} - \mathcal{Y}\|_F^2$ is the fidelity term, $R(\mathcal{X})$ is the regularization term, and λ is a parameter used to balance the fidelity term and regularization term. The reconstruction effectiveness of the proposed model is closely related to the rationality of regularization terms.

2.3. Proposed Model

We use Weighted 3D TV and nonlocal low-rank terms as the regularization terms to approximate \mathcal{X} in our model. Therefore, the proposed model for the super-resolution task is formulated as follows:

$$\min_{\mathcal{X}} \|DS\mathcal{X} - \mathcal{Y}\|_F^2 + \lambda_{tv} \|\mathcal{X}\|_{3DTV} + \lambda_{rank} rank(\mathcal{P}_{\mathcal{X}}^k), \quad (10)$$

where $\|\cdot\|_{3DTV}$ is the 3D TV regularization and $rank(\cdot)$ denotes the low-rank terms. Here, $\mathcal{P}_{\mathcal{X}}^k$ denotes the k th group of patches, $k \in [1, K]$ and K is the number of groups, and the order of $\mathcal{P}_{\mathcal{X}}^k$ is N . λ_{tv} and λ_{rank} are the regularization parameters.

2.3.1. 3DTV Regularization

Total variation is a well-known regularization approach for preserving the local spatial consistency of data in various image processing tasks. Traditional TV is designed to exploit the local

spatial smoothness prior to removing noise and retaining edges in images. Considering that 3D medical images can be seen as a 3rd-order tensor, smoothness is in three dimensions. Therefore, we introduce the weighted 3D TV to model the spatial and slice smoothness simultaneously. 3D TV is formulated as follows:

$$TV(\mathcal{X}) = \sum_{ijk} w_1 |x_{ijk} - x_{ij,k-1}| + w_2 |x_{ijk} - x_{i,j-1,k}| + w_3 |x_{ijk} - x_{i-1,jk}| \quad (11)$$

where x_{ijk} denotes the pixel at location (i, j) in the k th band, and $w_d (d = 1, 2, 3)$ denotes the weight along with different modes of \mathcal{X} . With constraints on spatial and slice dimensions, this weighted 3D TV remains piecewise smoothness in three dimensions.

2.3.2. Nonlocal Low-Rank Regularization

The spatial non-local similarity is one of the most important priors in image processing. For a given local patch in an image, we can find some other similar patches. To exploit this prior, we perform a series of operations on the original tensor. First, we separate image \mathcal{X} into a set of patches $\Omega = \{P_k \in \mathbb{R}^{b \times b \times B}\}_{k=1}^K$, where $b \times b$ is the size of the patch, B is the number of bands, and K is the number of patches with overlap. Second, for a given local patch, we find $d-1$ patches with the smallest Euclidean distance from it. Then, we stack them together with a local patch, forming a 4th-order tensor $\mathcal{P}_{\mathcal{X}}^k$ with size $b \times b \times B \times d$.

The global spectral correlation and spatial nonlocal similarity of images give the matched 4th-order tensor $\mathcal{P}_{\mathcal{X}}^k$ a good low-rank property. Furthermore, we perform Tucker decomposition for the nonlocal low-rank terms $\mathcal{P}_{\mathcal{X}}^k$ to exploit the low-rank characteristics and obtain:

$$\min_{U_k^{(n)}, \mathcal{G}_k} \lambda_1 \sum_{k=1}^K \sum_{n=1}^N \|U_k^{(n)}\|_{P_\lambda} + \lambda_2 \sum_{k=1}^K \|\mathcal{G}_k\|_F^2 \quad s.t. \quad \mathcal{P}_{\mathcal{X}}^k = \mathcal{G}_k \times_1 U_k^{(1)} \times_2 U_k^{(2)} \dots \times_4 U_k^{(N)}, \quad (12)$$

where \mathcal{G}_k and $U_k^{(n)}$ are the core tensor and decomposition factors, respectively, obtained by Tucker decomposition of $\mathcal{P}_{\mathcal{X}}^k$, and λ_1 and λ_2 are the regularization parameters.

The fold-concave penalty is a nonconvex norm. Different from the matrix nuclear norm that punishes the large singular value resulting in bias (Leng et al., 2006), the fold-concave penalty is theoretically proven to be an almost unbiased estimation of the rank (Fan and Li, 2001). For a given matrix X , we define its folded-concave norm as:

$$\|X\|_{P_\lambda} := \sum_{j=1}^r P_\lambda[\sigma_j(X)], \quad (13)$$

where r is the rank of X , $\sigma_j(X)$ is its j th singular value, and P_λ is called the folded-concave penalty function.

The minmax concave plus (MCP) penalty is one of the fold-concave penalties, and it has a simple form and excellent performance. In this article, we use the MCP penalty to exploit

the low-rank structure of the Tucker decomposition factors $U_k^{(n)}$ ($n = 1, 2, \dots, N$) of \mathcal{P}_{λ}^k . MCP penalty defined as follows:

$$P_{\lambda}(t) = \begin{cases} a\lambda^2/2, & \text{if } |t| \geq a\lambda, \\ \lambda |t| - \frac{t^2}{2a}, & \text{otherwise,} \end{cases} \quad (14)$$

where a is a constant. In addition, the last term in Equation (12) imposes the Frobenius norm on the core tensor \mathcal{G}_k to avoid overfitting.

In summary, our tensor super-resolution model is formulated as follows:

$$\begin{aligned} \min_{\mathcal{X}} & \|DS\mathcal{X} - \mathcal{Y}\|^2 + \lambda_{tv}\|\mathcal{X}\|_{3DTV} + \lambda_1 \sum_{k=1}^K \sum_{n=1}^N \|U_k^{(n)}\|_{P_{\lambda}} \\ & + \lambda_2 \sum_{k=1}^K \|\mathcal{G}_k\|_F^2, \\ \text{s.t.} & \mathcal{P}_{\mathcal{X}}^k = \mathcal{G}_k \times_1 U_k^{(1)} \times_2 U_k^{(2)} \dots \times_4 U_k^{(N)}. \end{aligned} \quad (15)$$

2.4. Optimization

For optimization purposes, we rewrite Equation (15) as:

$$\begin{aligned} \min_{\mathcal{X}} & \|DS\mathcal{X} - \mathcal{Y}\|^2 + \lambda_{tv}\|G_w(\mathcal{X})\|_1 + \lambda_1 \sum_{k=1}^K \sum_{n=1}^N \|U_k^{(n)}\|_{P_{\lambda}} \\ & + \lambda_2 \sum_{k=1}^K \|\mathcal{G}_k\|_F^2, \\ \text{s.t.} & \mathcal{P}_{\mathcal{X}}^k = \mathcal{G}_k \times_1 U_k^{(1)} \times_2 U_k^{(2)} \dots \times_4 U_k^{(N)}, \end{aligned} \quad (16)$$

where $\|G_w(\mathcal{X})\|_1 = [w_1 \times G_1(\cdot); w_2 \times G_2(\cdot); w_3 \times G_3(\cdot)]$ is the weighted three-dimensional difference operator and G_1, G_2 , and G_3 are the first-order difference operators in three dimensions.

We introduce some necessary auxiliary variables, including $\mathcal{X} = \mathcal{M}$ and $\mathcal{F} = G_w\mathcal{X}$, to split the interdependencies of the terms in Equation (16) and obtain:

$$\begin{aligned} \min_{\mathcal{X}} & \|DS\mathcal{X} - \mathcal{Y}\|^2 + \lambda_{tv}\|G_w(\mathcal{X})\|_1 + \lambda_1 \sum_{k=1}^K \sum_{n=1}^N \|U_k^{(n)}\|_{P_{\lambda}} \\ & + \lambda_2 \sum_{k=1}^K \|\mathcal{G}_k\|_F^2, \\ \text{s.t.} & \mathcal{P}_{\mathcal{X}}^k = \mathcal{G}_k \times_1 U_k^{(1)} \times_2 U_k^{(2)} \dots \times_4 U_k^{(N)}, \quad \mathcal{X} = \mathcal{M}, \\ & \mathcal{F} = G_w\mathcal{X}, \\ & \{V_k^{(n)} = U_k^{(n)}\}_{n=1}^N, \quad \{\mathcal{P}_{\mathcal{M}}^k = \mathcal{G}_k \times_1 U_k^{(1)} \times_2 U_k^{(2)} \dots \\ & \times_N U_k^{(N)}\}_{k=1}^K \end{aligned} \quad (17)$$

We solve the cost function Equation (17) by adopting the alternating direction method of multipliers (ADMM) algorithm

and obtain the object function as follows:

$$\begin{aligned} \mathcal{L} & (\mathcal{X}, \mathcal{F}, U_k^{(n)}, V_k^{(n)}, \mathcal{G}_k, \mathcal{M}, \mathcal{W}_1, \dots, \mathcal{W}_4) \\ & = \|DS\mathcal{X} - \mathcal{Y}\|_F^2 + \frac{w_1}{2} \left\| \mathcal{X} - \mathcal{M} + \frac{\mathcal{W}_1}{w_1} \right\|_F^2 + \lambda_{tv}\|\mathcal{F}\|_1 \\ & + \frac{w_2}{2} \left\| \mathcal{F} - G_w\mathcal{X} + \frac{\mathcal{W}_2}{w_2} \right\|_F^2 \\ & + \sum_{k=1}^K \left[\lambda_2 \|\mathcal{G}_k\|_F^2 + \sum_{n=1}^N \left(\lambda_1 \|U_k^{(n)}\|_{P_{\lambda}} \right) \right] \\ & + \sum_{k=1}^K \sum_{n=1}^N \left(\frac{w_3}{2} \left\| U_k^{(n)} - V_k^{(n)} + \frac{\mathcal{W}_3^{(n)}}{w_3} \right\|_F^2 \right) \\ & + \sum_{k=1}^K \left(\frac{w_4}{2} \left\| \mathcal{P}_{\mathcal{M}}^k - \mathcal{G}_k \times_1 V_k^{(1)} \times_2 V_k^{(2)} \dots \times_N V_k^{(N)} + \frac{\mathcal{W}_4^{(k)}}{w_4} \right\|_F^2 \right) \end{aligned} \quad (18)$$

where $\mathcal{W}_1, \mathcal{W}_2, \{\mathcal{W}_3^{(n)}\}_{n=1}^N$, and $\{\mathcal{W}_4^{(k)}\}_{k=1}^K$ are Lagrange multipliers. According to ADMM, we divide Equation (18) into Equation (6) subproblems and solve them by iteratively updating the variables. The following are the detailed variables updating procedures. The optimization procedure of the proposed model is shown in **Algorithm 1**.

Update \mathcal{X} . Fixing the other variables, we extract terms that contain \mathcal{X} in Equation (18). We update \mathcal{X} by minimizing:

$$\begin{aligned} \min_{\mathcal{X}} & \|DS\mathcal{X} - \mathcal{Y}\|_F^2 + \frac{w_1}{2} \left\| \mathcal{X} - \mathcal{M} + \frac{\mathcal{W}_1}{w_1} \right\|_F^2 + \frac{w_2}{2} \|\mathcal{F} \\ & - G_w\mathcal{X} + \frac{\mathcal{W}_2}{w_2} \right\|_F^2 \end{aligned} \quad (19)$$

The partial derivative of Equation (19) with respect to \mathcal{X} is as follows:

$$\begin{aligned} 2(DS)^T DS\mathcal{X} + w_1\mathcal{X} + w_2G_w'G_w\mathcal{X} & = 2(DS)^T\mathcal{Y} + w_1\mathcal{M} \\ & - \mathcal{W}_1 + w_2G_w'\mathcal{F} + \mathcal{W}_2G_w' \end{aligned} \quad (20)$$

where $(DS)^T$ denotes the transposes of DS and G_w' denotes the adjoint of G_w .

Update \mathcal{F} . Update \mathcal{F} by minimizing:

$$\begin{aligned} \min_{\mathcal{F}} & \lambda_{tv}\|\mathcal{F}\|_1 + \frac{w_2}{2} \left\| \mathcal{F} - G_w\mathcal{X} + \frac{\mathcal{W}_2}{w_2} \right\|_F^2 \\ & = \min_{\mathcal{F}} \sum_{n=1}^N \left(\lambda_{tv}\alpha_n \|\mathcal{F}_{(n)}\|_1 + \frac{w_2}{2} \left\| \mathcal{F} - G_w\mathcal{X} + \frac{\mathcal{W}_2}{w_2} \right\|_F^2 \right) \end{aligned} \quad (21)$$

Then, we have:

$$\mathcal{F} = \text{fold}_n \left[\text{Soft}_{\lambda_{tv}\alpha_i/w_2} \left(G_w\mathcal{X}_{(n)} - \frac{\mathcal{W}_2^{(i)}}{w_2} \right) \right] \quad (22)$$

where Soft denotes the soft-thresholding operator, which is defined as:

$$\text{Soft}_\Delta(x) = \begin{cases} x - \Delta, & \text{if } x > \Delta, \\ x + \Delta, & \text{if } x < -\Delta, \\ 0, & \text{otherwise.} \end{cases} \quad (23)$$

where $x \in \mathbb{R}$ and $\Delta > 0$.

Update $U_k^{(n)}$. Update $U_k^{(n)}$ by minimizing:

$$\min_U \lambda_1 \left\| U_k^{(n)} \right\|_{P_\lambda} + \frac{w_3}{2} \left\| U_k^{(n)} - V_k^{(n)} + \frac{\mathcal{W}_{3(k)}^n}{w_3} \right\|_F^2 \quad (24)$$

To solve Equation (24), we introduce the singular value shrinkage operator, which is defined as:

$$S_\tau(X) := U_X D_\tau \left(\sum_X \right) V_X^T, \quad (25)$$

where $X = U_X \sum_X V_X^T$ is the singular value decomposition of X . For a matrix A , $[D_\tau(A)]_{ij} = \text{sgn}(A_{ij})(|A_{ij}| - \tau)_+$. $U_k^{(n)}$ can be obtained by:

$$U_k^{(n)} = S_{\lambda_1 \lambda / w_3} \left(V_k^{(n)} + \frac{\mathcal{W}_{3(k)}^n}{w_3} \right) \quad (26)$$

Update $V_k^{(n)}$. Update $V_k^{(n)}$ by minimizing:

$$\min_V \sum_{n=1}^N \left(\frac{w_3}{2} \left\| U_k^{(n)} - V_k^{(n)} + \frac{\mathcal{W}_{3(k)}^n}{w_3} \right\|_F^2 + \frac{w_4}{2} \left\| \mathcal{P}_{\mathcal{M}}^k - \mathcal{G}_k \times_1 V_k^{(1)} \times_2 V_k^{(2)} \cdots \times_N V_k^{(N)} \frac{\mathcal{W}_{4(k)}^n}{w_4} \right\|_F^2 \right) \quad (27)$$

We obtain the solution of Equation (27) as follows:

$$V_k^{(n)} = \left[-W_{3(k)}^n + w_3 U_k^{(n)} + \left(W_{4(k)} + w_4 \mathcal{P}_{\mathcal{M}}^k \right) V_k^{(-n)} G_{(n)}^T \right] \times \left(w_4 I + w_4 G_{(n)} V_k^{(-n)T} V_k^{(-n)} G_{(n)}^T \right)^{-1} \quad (28)$$

Update \mathcal{G}_k . Update \mathcal{G} by minimizing:

$$\min_{\mathcal{G}} \lambda_2 \left\| \mathcal{G}_k \right\|_F^2 + \frac{w_4}{2} \left\| \mathcal{P}_{\mathcal{M}}^k - \mathcal{G}_k \times_1 V_k^{(1)} \times_2 V_k^{(2)} \cdots \times_N V_k^{(N)} + \frac{\mathcal{W}_{4(k)}}{w_4} \right\|_F^2 \quad (29)$$

We obtain the following solution of \mathcal{G}_k :

$$\text{vec}(\mathcal{G}) = [V^{(-n)T} V^{(-n)} \otimes w_4 V^{(n)T} V^{(n)} + \lambda_2 I]^{-1} \times \text{vec}[V^{(n)T} (\mathcal{W}_{4(k)} + w_4 \mathcal{P}_{\mathcal{M}}^k) V^{(-n)}] \quad (30)$$

Update \mathcal{M} . Fixing the other variables, we extract terms containing \mathcal{M} in Equation (18). We update \mathcal{M} by minimizing:

$$\min_{\mathcal{M}} \frac{w_1}{2} \left\| \mathcal{X} - \mathcal{M} + \frac{\mathcal{W}_1}{w_1} \right\|_F^2 + \sum_{k=1}^K \left(\frac{w_4}{2} \left\| \mathcal{P}_{\mathcal{M}}^k - \mathcal{G}_k \times_1 U_k^{(1)} \times_2 U_k^{(2)} \cdots \times_4 U_k^{(N)} + \frac{\mathcal{W}_{4(k)}}{w_4} \right\|_F^2 \right) \quad (31)$$

We obtain \mathcal{M} by updating $\mathcal{P}_{\mathcal{M}}^k$ and aggregating them:

$$\mathcal{M} = \sum_{k=1}^K \mathcal{P}_{\mathcal{M}}^k. \quad (32)$$

Here, the formula to update $\mathcal{P}_{\mathcal{M}}^k$ can be expressed as:

$$\min_{\mathcal{P}_{\mathcal{M}}^k} \frac{w_1}{2} \left\| \mathcal{P}_{\mathcal{X}}^k - \mathcal{P}_{\mathcal{M}}^k + \frac{\mathcal{W}_{1(k)}}{w_1} \right\|_F^2 + \sum_{k=1}^K \left(\frac{w_4}{2} \left\| \mathcal{P}_{\mathcal{M}}^k - \mathcal{G}_k \times_1 U_k^{(1)} \times_2 U_k^{(2)} \cdots \times_4 U_k^{(N)} + \frac{\mathcal{W}_{4(k)}}{w_4} \right\|_F^2 \right) \quad (33)$$

Then, we obtain the following closed-form solution of $\mathcal{P}_{\mathcal{M}}^k$:

$$\mathcal{P}_{\mathcal{M}}^k = \frac{w_4 \mathcal{C} + w_1 \mathcal{P}_{\mathcal{X}}^k + \mathcal{W}_{1(k)} - \mathcal{W}_{4(k)}}{w_4 + w_1} \quad (34)$$

where $\mathcal{C} = \mathcal{G}_k \times_1 U_k^{(1)} \times_2 U_k^{(2)} \cdots \times_4 U_k^{(N)}$.

Algorithm 1: NNTDTV.

Input: Low-resolution image \mathcal{Y}

Output: Reconstructed high-resolution image \mathcal{X}

- 1 **Initialize:** $\mathcal{X} = \text{Upsample}(\mathcal{Y})$, \mathcal{M} , $\mathcal{P}_{\mathcal{M}}^k$, \mathcal{G}_k , $U_k^{(n)} = V_k^{(n)}$;
 - 2 **for** $l=1, 2, \dots, L$ **do**
 - 3 Extract the set of $\mathcal{P}_{\mathcal{M}}^k$ and find similar patch group;
 - 4 Update \mathcal{X} via Equation (20);
 - 5 Update \mathcal{F} via Equation (22);
 - 6 **for each patch do**
 - 7 Update $U_k^{(n)}$ via Equation (26);
 - 8 Update $V_k^{(n)}$ via Equation (28);
 - 9 Update \mathcal{G}_k via Equation (30);
 - 10 Update $\mathcal{P}_{\mathcal{M}}^k$ via Equation (34);
 - 11 **end**
 - 12 Aggregate $\mathcal{P}_{\mathcal{M}}^k$ to reshape the $\mathcal{X}^{(l)}$;
 - 13 Update W_1, W_2, W_3, W_4 ;
 - 14 **end**
-

3. EXPERIMENTS

3.1. Database

We conduct extensive experiments to evaluate the effectiveness of our method compared with the SOTA methods. We adopt four

kinds of medical images including the synthetic brain MRI data selected from the BrainWeb database (Cocosco et al., 1997)¹, the real abdominal CT data from the NIH pancreas segmentation database (Roth et al., 2015)², the real prostate MRI data from the NCI-ISBI 2013 database,³ and dental image published by TF-SISR.

The BrainWeb database contains a set of synthetic brain MRI data volumes (image size of $181 \times 217 \times 181$ and spatial resolution of $1 \text{ mm}^3 \times 1 \text{ mm}^3 \times 1 \text{ mm}^3$) produced by an MRI simulator. The NIH pancreas segmentation database contains 82 real contrast-enhanced abdominal CT volumes (image size of $256 \times 256 \times B$, where $B \in [90, 233]$ and spatial resolution from $0.5 \text{ mm}^3 \times 0.5 \text{ mm}^3 \times 0.5 \text{ mm}^3$ to $1 \text{ mm}^3 \times 1 \text{ mm}^3 \times 1 \text{ mm}^3$). The NCI-ISBI 2013 database contains 30 prostate MRI cases (image size of $384 \times 384 \times B$, where $B \in [113, 217]$ and spatial resolution from $0.6 \text{ mm}^3 \times 0.6 \text{ mm}^3 \times 3.6 \text{ mm}^3$ to $0.625 \text{ mm}^3 \times 0.625 \text{ mm}^3 \times 4 \text{ mm}^3$). The dental image published by the TF-SISR contains a set of dental CT data (image size of $282 \times 266 \times 392$). The CBCT image was obtained with the Carestream 81003D system. The linewidth resolution of the CBCT machine was $500 \mu\text{m}$ and the volumes have a voxel size of $80 \times 80 \times 80 \mu\text{m}^3$. The μCT image was obtained with a Quantum FX system from Perkin Elmer, with a linewidth resolution of $50 \mu\text{m}$, and voxel size of $40 \times 40 \times 40 \mu\text{m}^3$.

3.2. Experimental Settings

Experiments on the BrainWeb database, NIH pancreas segmentation database, NCI-ISBI 2013 database, and LR images are obtained by applying Gaussian blurring with a blur kernel of 1 voxel wide and downsampling the original HR images with a factor of 2, similar to LRTV (Shi et al., 2015), and the original HR images are used as ground truth (GT).

In the experiment with dental images published by the TF-SISR method, our method can be validated on the real-world database. We use the CBCT as input and the μCT as GT. We follow the TF-SISR setting, including the downsampling and blurring kernel type, to ensure integrity and fairness. To evaluate the reconstruction effect, we compare the reconstruction results of each method with μCT images acquired from the same sample.

There are four important parameters in our model including λ_{tv} , λ_1 , λ_2 , and d . λ_{tv} , λ_1 , and λ_2 control the balance of 3DTV and nonlocal low-rank regularization terms, and d is the number of similar patches that form the fourth-order tensor \mathcal{P}_{λ}^k . We select 20 serial representative slices from the BrainWeb database for parameters analysis and selection. In **Figure 2**, we show the relationship between PSNR and the regularization parameters λ_{tv} , λ_1 , and λ_2 in Equation (15) with the other parameters fixed at optimal values. The change of λ_{tv} has the biggest influence on PSNR of results, while the change of λ_2 has the least influence. It can be seen that when λ_{tv} , λ_1 , and λ_2 vary within a wide range, PSNR can reach a high value. We observed

similar behavior in other cases. The value of λ_{tv} and λ_1 and λ_2 can be chosen in $[5, 20]$ and $[1, 10]$, respectively. In all experiments, the parameters are set as follows: the regularization parameters λ_{tv} , λ_1 , and λ_2 are fixed to 10, 5, and 20, respectively. **Figure 2D** shows the PSNR gains vs. the number of similar patches. The PSNR of reconstruction results becomes stable when the number of patches is larger than 10. We set the number of similar patches as 15 in our experiment. Great experimental results show that the set of parameters has good universality for different types of data. For LRTV and TF-SISR, we adjust the parameters for each database to optimize experimental results.

3.3. Quantitative Comparison

To evaluate the performance of our method on medical images, we compare our method with the nearest neighbor (NN) interpolation method, two state-of-the-art (SOTA) methods for the medical image 3D super-resolution including the low-rank and total variation regularizations (LRTV) method (Shi et al., 2015) and the tensor factorization single image super-resolution (TF-SISR) method (Hatvani et al., 2019). In 3D medical image super-resolution methods, LRTV and TF-SISR achieve competitive results. We use two quantitative picture quality indices to evaluate the quality of the reconstructed image, i.e., the peak signal-to-noise ratio (PSNR) and the structural similarity index (SSIM). It is known that higher values of PSNR and SSIM indicate better performances.

Table 1 shows the quantitative results of the BrainWeb database, NIH pancreas segmentation database, NCI-ISBI 2013 database, and dental image. The bolding values in table means the best results. These results show that our reconstruction method in terms of PSNR and SSIM outperforms the existing method in different types of medical images. Through **Table 1**, we can find that our model achieves the highest MPSNR value than those of other methods. The quantitative results by TF-SISR are relatively low in the first three databases. This shows that it is not enough to only use tensor factorization to recover images with complex structures, and it is necessary to add more priors. This is also demonstrated by our quantitative and visual results. Our model makes full use of the 3DTV and the nonlocal similarity to suppress noise and preserve tiny details, thus achieving the best results. For a further detailed analysis of the super-resolution results, we randomly select a set of data from the BrainWeb database, the NIH pancreas segmentation database, and the NCI-ISBI 2013 database and show the SSIM and PSNR values of each slice in **Figure 3**. Our method obtains much higher values of SSIM and PSNR than other methods for almost every slice.

3.4. Visual Quality Comparison

In this section, we demonstrate the visual results of each method on four databases. For the results of the BrainWeb databases, a typical slice of the coronal, sagittal, and axial views is shown in **Figure 4**, and a zoom in of the frontal region in the sagittal view is also provided. **Figures 5, 6** show a set of representative reconstruction results and the corresponding details from the

¹The BrainWeb database can be downloaded at: <https://brainweb.bic.mni.mcgill.ca/brainweb/>.

²The NIH pancreas segmentation database can be downloaded at: <https://wiki.cancerimagingarchive.net/display/Public/Pancreas-CT>.

³The NCI-ISBI 2013 database can be downloaded at: <https://liuquande.github.io/SAML/>.

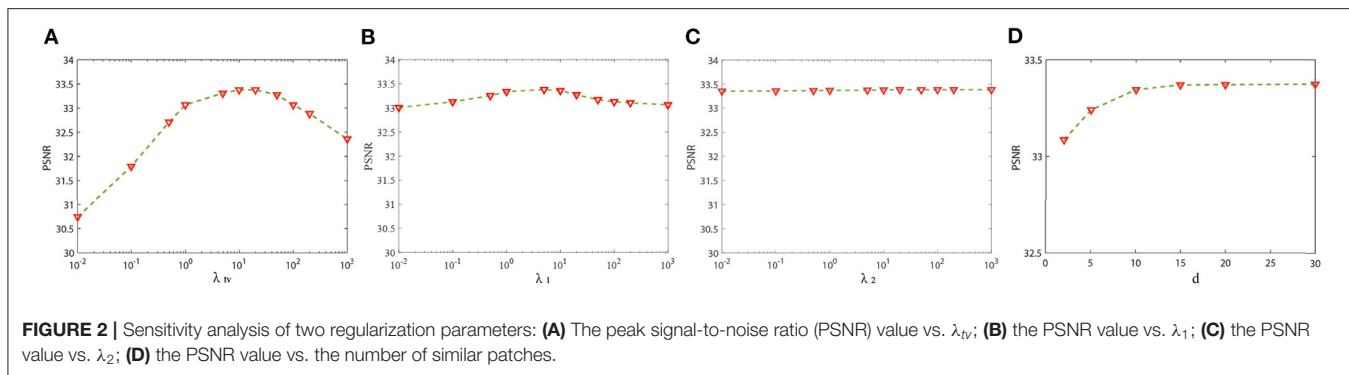
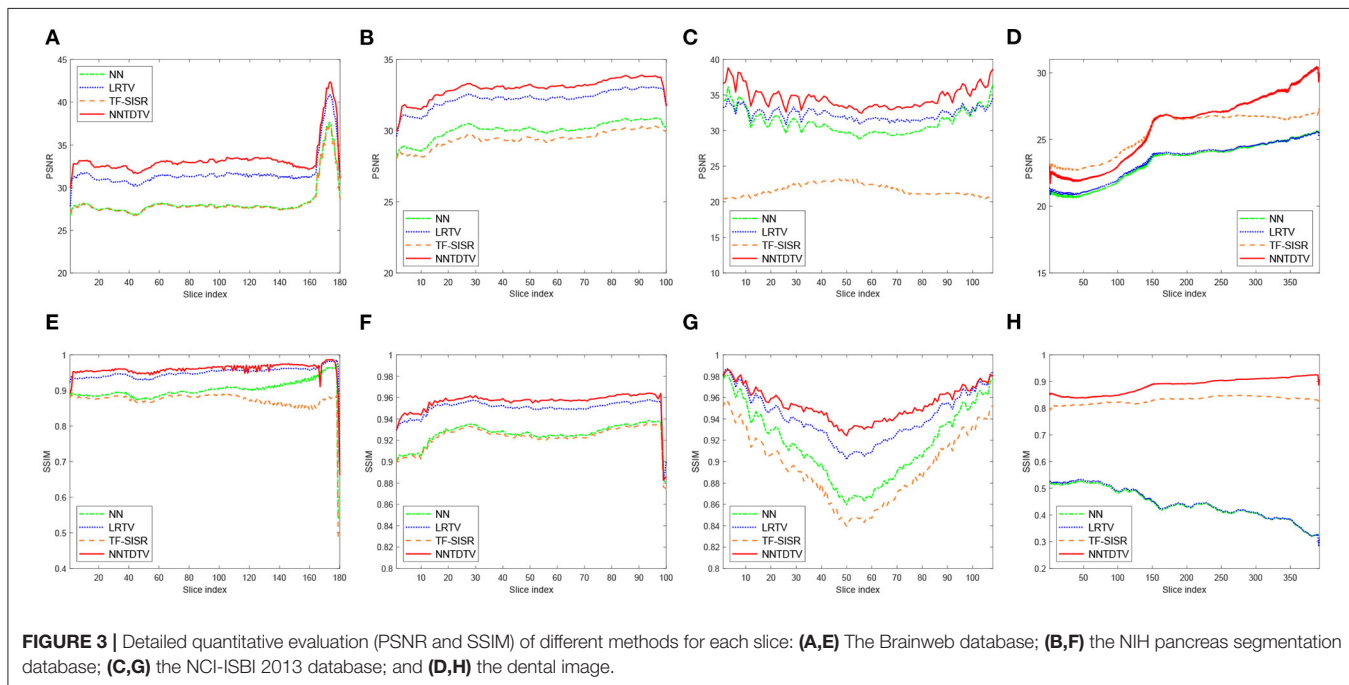


TABLE 1 | Quantitative results [peak signal-to-noise ratio (PSNR), structural 254 similarity index (SSIM)] by different methods on four medical databases.

Database		NN	LRTV	TF-SISR	NNTDTV
BrainWeb	PSNR (dB)	28.3336	31.7784	28.2847	33.3313
	SSIM	0.8996	0.9492	0.8955	0.9593
NIH pancreas	PSNR (dB)	30.6488	32.6053	27.2705	33.2831
	SSIM	0.9039	0.9242	0.8935	0.9289
NCI-ISBI 2013	PSNR (dB)	29.6478	31.2985	26.6026	33.2587
	SSIM	0.8632	0.9019	0.8582	0.9239
Dental	PSNR (dB)	23.3410	23.4459	25.5183	25.9325
	SSIM	0.4418	0.4449	0.8310	0.8851

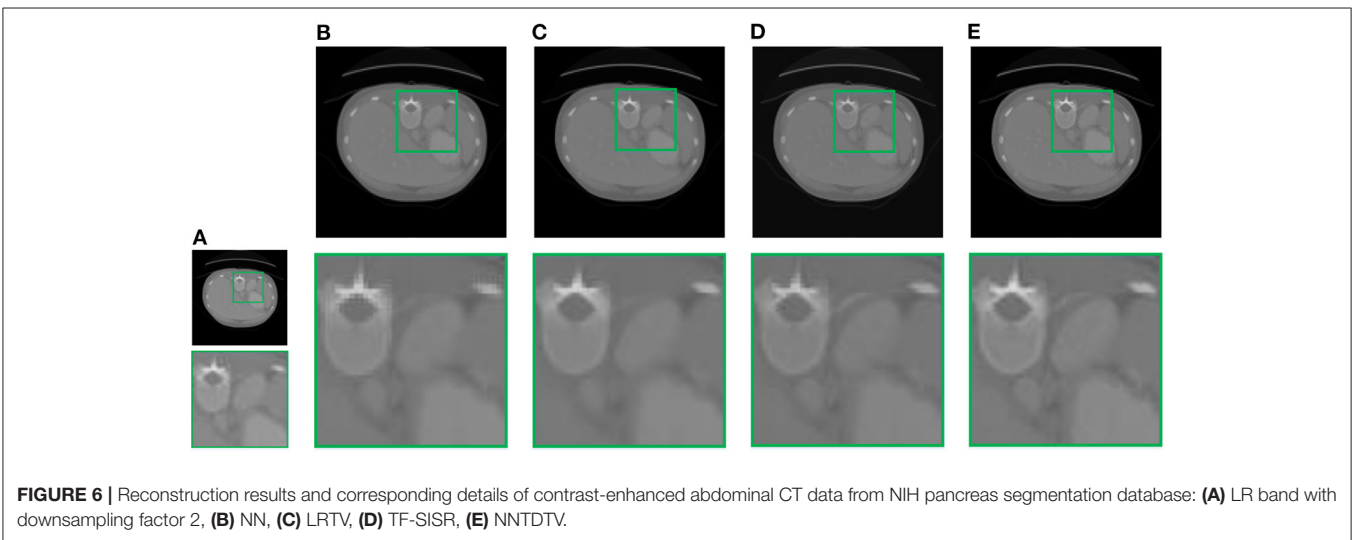
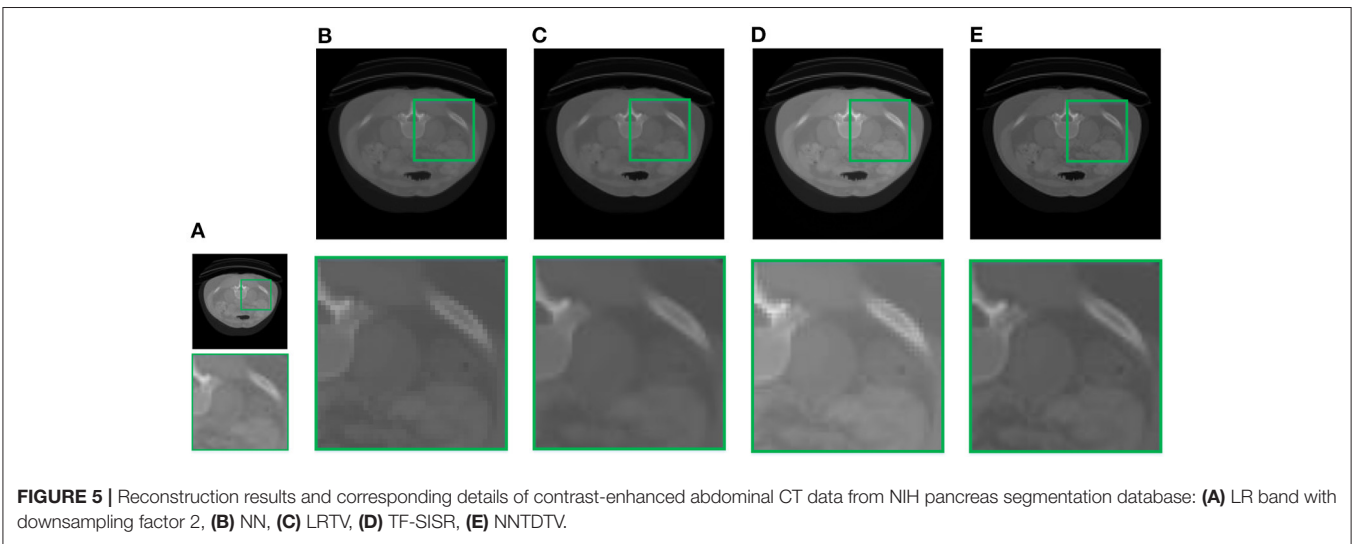
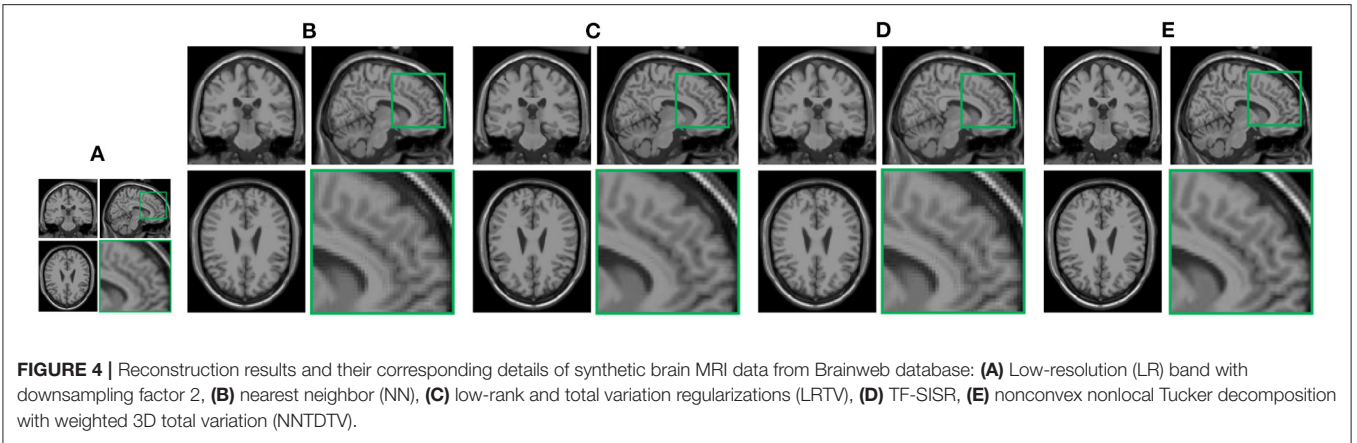
The best results are shown in bold.

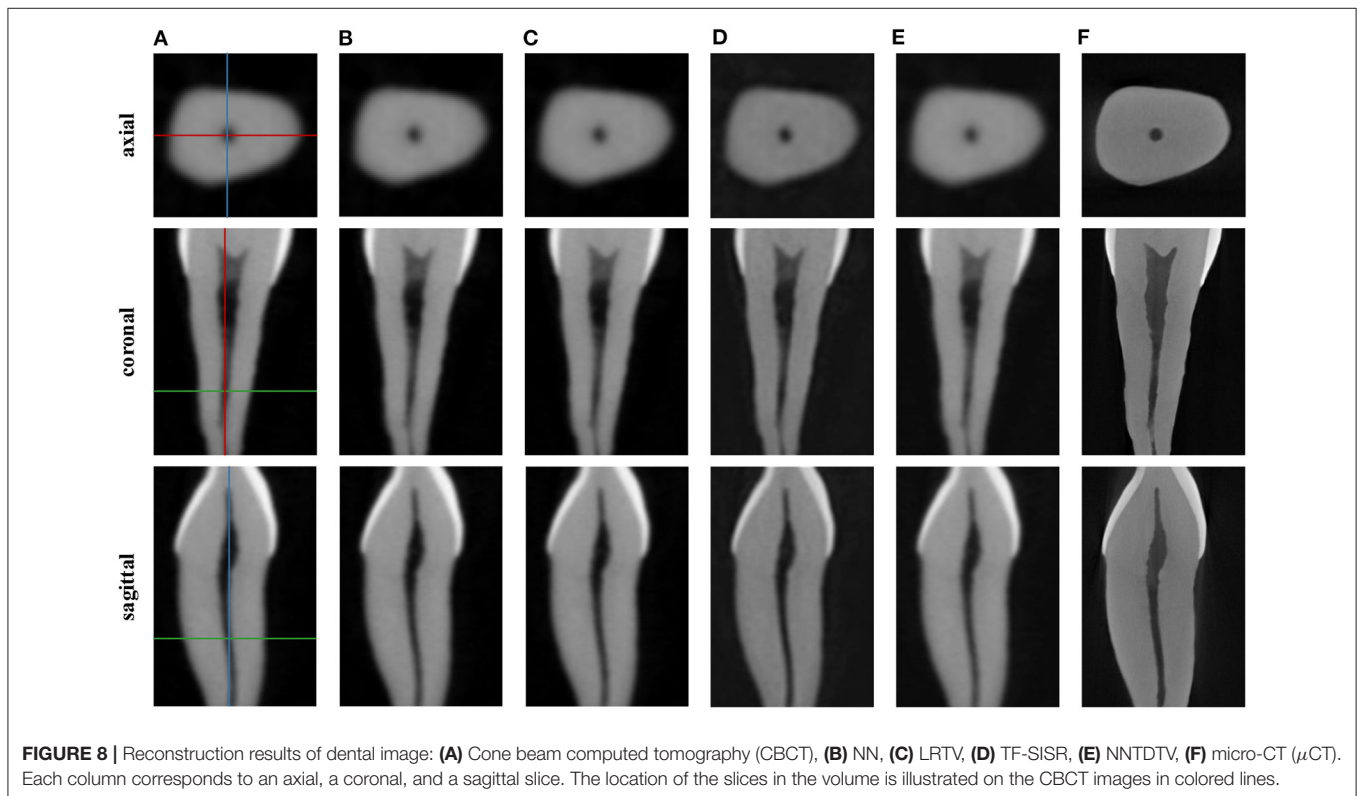
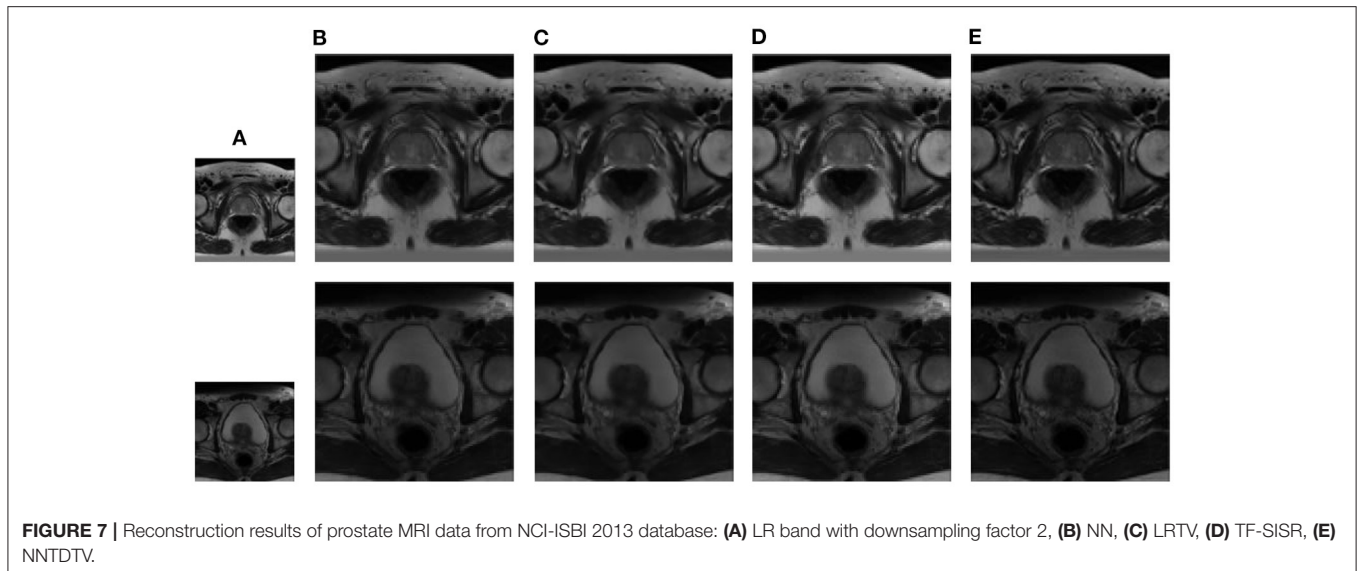


NIH pancreas segmentation database, respectively. **Figure 7** shows two sets of results from the NCI-ISBI 2013 database. As shown in the figures, the reconstruction results of NN and TF-SISR have obvious serrated edges while the results of LRTV are too smooth to keep the details clear. Compared with them, our method not only preserves the exact color

but also reconstructs clearer edges while maintaining the local smoothness.

Notably, the TF-SISR method obtains a relatively good performance only on dental images. This may be because the TF-SISR method only utilizes tensor factorization to recover images without any prior information. Therefore, when the





structure of an image is complex, the reconstruction effect is not satisfactory. The TF-SISR method was proposed for 3D image super-resolution and applied to dental images. In this part of the experiment, we follow the experimental setting of the TF-SISR method. We show the PSNR and SSIM in **Table 1** and the corresponding axial, coronal, and sagittal slices in **Figure 8**. Our method is not only superior to other methods with various types of medical images but also achieves a slightly better effect than TF-SISR with the dental

image. This shows that our method has good performance and generalization.

4. DISCUSSION

Super-resolution reconstruction provides an economical and effective solution to improve the resolution of CT and MRI at the software level. However, there are few super-resolution methods designed for 3D medical images. We proposed a

single image super-resolution method based on tensor low-rank decomposition that combines the local smoothness and nonlocal similarity and apply to 3D CT and MRI data. Our method shows better quantitative results and visual quality compared with existing 3D medical super-resolution methods. We maintain the local smoothness while keeping the clear edges of images.

This study has certain limitations. First, nonlocal similarity prior makes full use of the structural information of images and improves the effect of super-resolution reconstruction. However, it increases computational costs. In the future study, we will explore more efficient priors to include in our framework. We will further consider the global spatial correlation in all directions of data by constructing reasonable tensors and then performing low-rank constraints. We will combine the low-rank prior data with the deep learning method to obtain accurate data features and achieve better super-resolution reconstruction results. Second, although our super-resolution model can theoretically be used for images of any dimension, we only verify it on 3D images in this article. In the future, the experiments will be conducted on diverse data.

5. CONCLUSION

In this article, we proposed a new super-resolution method based on nonlocal Tucker tensor decomposition and 3D TV regularization. Nonlocal Tucker tensor decomposition fully exploits the spatial and inter-frame low-rank information to reconstruct the data. The 3D TV regularization term retains the local smoothness of spatial and spectral information, thus enhancing the detailed information. Extensive experimental studies on four different kinds of medical data, including MRI data of the brain and prostate and CT data of the abdominal and dental, validated our method compared with the SOTA methods.

REFERENCES

- Bustin, A., Voilliot, D., Menini, A., Felblinger, J., De Chillou, C., Burschka, D., et al. (2018). Isotropic Reconstruction of MR Images Using 3D Patch-Based Self-Similarity Learning. *IEEE Trans. Med. Imag.* 37, 1932–1942. doi: 10.1007/BF02310791
- Carroll, J. D., and Chang, J. J. (1970). Analysis of individual differences in multidimensional scaling via an n-way generalization of "Eckart-Young" decomposition. *Psychometrika* 35, 283–319.
- Chen, W., Zhu, X., Sun, R., He, J., Li, R., Shen, X., and Yu, B. (2020). "Tensor low-rank reconstruction for semantic segmentation," in *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics LNCS)*, Vol. 12362 (Cham), 52–69.
- Chen, Y. L., Hsu, C. T., and Liao, H. Y. M. (2014). Simultaneous tensor decomposition and completion using factor priors. *IEEE Trans. Pattern Anal. Mach. Intell.* 36, 577–591. doi: 10.1109/TPAMI.2013.164
- Cocosco, C. A., Kollokian, V., Kwan, R. K., and Evans, A. C. (1997). Brain web: online interface to a 3D MRI simulated brain database. *Neuroimage* 5, 528–537.
- Dabov, K., Foi, A., and Egiazarian, K. (2007). Video denoising by sparse 3D transform-domain collaborative filtering. *Eur. Signal Process. Conf.* 16, 145–149.
- Dian, R., Fang, L., and Li, S. (2017). "Hyperspectral image super-resolution via non-local sparse tensor factorization," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.* vol. 2017 (Honolulu, HI), 3862–3871.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found at: BrainWeb, <https://brainweb.bic.mni.mcgill.ca/brainweb/>.

ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent from the patients/participants or legal guardian/next of kin was not required to participate in this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

XC and ZH contributed to the conception and design of the study. HJ wrote this manuscript and participated in the whole experiment process. BL wrote sections of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

FUNDING

This study was supported in part by the National Natural Science Foundation of China under Grant 61903358, 61873259, and 61821005, in part by the Youth Innovation Promotion Association of the Chinese Academy of Sciences under Grant 2022196 and Y202051, National Science Foundation of Liaoning Province under Grant 2021-BS-023, and in part by the National Key Research and Development Program of China under Grant 2020YFB1313400.

- Diwakar, M., and Kumar, M. (2018). A review on CT image noise and its denoising. *Biomed. Signal Process. Control* 42, 73–88. doi: 10.1016/j.bspc.2018.01.010
- Duchon, C. E. (1979). Lanczos Filtering in One and Two Dimensions. *J. Appl. Meteorol.* 18, 1016–1022.
- E-Asim, F., De Almeida, A. L., Haardt, M., Cavalcante, C. C., and Nossek, J. A. (2020). Rank-one detector for kronecker-structured constant modulus constellations. *IEEE Signal Process. Lett.* 27, 1420–1424. doi: 10.1109/LSP.2020.3010133
- Fan, J., and Li, R. (2001). Variable selection via nonconcave penalized likelihood and its oracle properties. *J. Am. Stat. Assoc.* 96, 1348–1360. doi: 10.1198/016214501753382273
- Fan, S., Bian, Y., Chen, H., Kang, Y., Yang, Q., and Tan, T. (2020). Unsupervised cerebrovascular segmentation of tof-mra images based on deep neural network and hidden markov random field model. *Front. Neuroinf.* 13, 77. doi: 10.3389/fninf.2019.00077
- Greenspan, H. (2009). Super-resolution in medical imaging. *Comput. J.* 52, 43–63. doi: 10.1093/comjnl/bxm075
- Hatvani, J., Basarab, A., Tournet, J. Y., Gyongy, M., and Kouame, D. (2019). A tensor factorization method for 3-D super resolution with application to dental CT. *IEEE Trans. Med. Imag.* 38, 1524–1531. doi: 10.1109/TMI.2018.2883517
- Keys, R. G. (1981). Cubic convolution interpolation for digital image processing. *IEEE Trans. Acoust.* 29, 1153–1160.
- Kolda, T. G., and Bader, B. W. (2009). Tensor decompositions and applications. *SIAM Rev.* 51, 455–500. doi: 10.1137/07070111X

- Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., et al. (2017). "Photo-realistic single image super-resolution using a generative adversarial network," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2017 (Honolulu, HI), 105–114.
- Leng, C., Lin, Y., and Wahba, G. (2006). A note on the Lasso and related procedures in model selection. *Stat. Sin.* 16, 1273–1284. Available online at: <http://www.jstor.org/stable/24307787>
- Li, S., Dian, R., Fang, L., and Bioucas-Dias, J. M. (2018). Fusing hyperspectral and multispectral images via coupled sparse tensor factorization. *IEEE Trans. Image Process.* 27, 4118–4130. doi: 10.1109/TIP.2018.2836307
- Li, X., Ye, Y., and Xu, X. (2017). "Low-rank tensor completion with total variation for visual data inpainting," in *Proc. AAAI Conf. Artif. Intell.*, vol. 2017 (San Francisco, CA), 2210–2216.
- Lim, B., Son, S., Kim, H., Nah, S., and Lee, K. M. (2017). "Enhanced deep residual networks for single image super-resolution," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Work.*, vol. 2017 (Honolulu, HI), 1132–1140.
- Liu, C., Fang, F., Xu, Y., and Shen, C. (2016). "Single image super-resolution based on nonlocal sparse and low-rank regularization," in *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics LNCS)*, Vol. 9810 (Cham), 251–261.
- Liu, J., Musialski, P., Wonka, P., and Ye, J. (2013). Tensor completion for estimating missing values in visual data. *IEEE Trans. Pattern Anal. Mach. Intell.* 35, 208–220. doi: 10.1109/TPAMI.2012.39
- Liu, J., Yang, W., Zhang, X., and Guo, Z. (2017). Retrieval compensated group structured sparsity for image super-resolution. *IEEE Trans. Multimed.* 19, 302–316. doi: 10.1109/TMM.2016.2614427
- Liu, L., Zhang, J., Wang, J.-x., Xiong, S., and Zhang, H. (2021). Co-optimization learning network for mri segmentation of ischemic penumbra tissues. *Front. Neuroinf.* 15, 782262. doi: 10.3389/fninf.2021.782262
- Mairal, J., Bach, F., Ponce, J., Sapiro, G., and Zisserman, A. (2009). "Non-local sparse models for image restoration," in *Proc. IEEE Int. Conf. Comput. Vis.*, vol. 2009 (Kyoto), 2272–2279.
- Marcelino, J., Carvalho, S., Duarte, F. C., Costa, A. C., and Barbosa, M. P. (2019). *Adverse Reactions To Iodinated Contrast Media*, vol. 27, 9–20. doi: 10.1055/s-0033-1348885
- Prevost, C., Usevich, K., Comon, P., and Brie, D. (2019). "Coupled tensor low-rank multilinear approximation for hyperspectral super-resolution," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, vol. 2019 (Brighton), 5536–5540.
- Qiu, D., Cheng, Y., and Wang, X. (2021). Gradual back-projection residual attention network for magnetic resonance image super-resolution. *Comput. Methods Programs Biomed.* 208, 106252. doi: 10.1016/j.cmpb.2021.106252
- Roth, H. R., Lu, L., Farag, A., Shin, H. C., Liu, J., Turkbey, E. B., et al. (2015). "Deeporgan: multi-level deep convolutional networks for automated pancreas segmentation," in *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)* Vol. 9349 (Cham), 556–564.
- Salvador, J. (2016). *Example-Based Super Resolution*. doi: 10.1016/c2015-0-06719-3
- Shi, F., Cheng, J., Wang, L., Yap, P. T., and Shen, D. (2015). LRTV: MR image super-resolution with low-rank and total variation regularizations. *IEEE Trans. Med. Imaging* 34, 2459–2466. doi: 10.1109/TMI.2015.2437894
- Shi, F., Fan, Y., Tang, S., Gilmore, J. H., Lin, W., and Shen, D. (2010). Neonatal brain image segmentation in longitudinal MRI studies. *Neuroimage* 49, 391–400. doi: 10.1016/j.neuroimage.2009.07.066
- Tucker, L. R. (1966). Some mathematical notes on three-mode factor analysis. *Psychometrika* 31, 279–311. doi: 10.1007/BF02289464
- Veganzones, M. A., Simões, M., Licciardi, G., Yokoya, N., Bioucas-Dias, J. M., and Chanussot, J. (2016). Hyperspectral super-resolution of locally low rank images from complementary multisource data. *IEEE Trans. Image Process.* 25, 274–288. doi: 10.1109/TIP.2015.2496263
- Wang, X., Girshick, R., Gupta, A., and He, K. (2018). "Non-local neural networks," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2018 (Salt Lake City, UT), 7794–7803.
- Xie, Q., Zhao, Q., Meng, D., and Xu, Z. (2018). Kronecker-basis-representation based tensor sparsity and its applications to tensor recovery. *IEEE Trans. Pattern Anal. Mach. Intell.* 40, 1888–1902. doi: 10.1109/TPAMI.2017.2734888
- Xu, Y., Hao, R., Yin, W., and Su, Z. (2015). Parallel matrix factorization for low-rank tensor completion. *Inverse Probl. Imag.* 9, 601–624. doi: 10.3934/ipi.2015.9.601
- Yair, N., and Michaeli, T. (2018). "Multi-scale weighted nuclear norm image restoration," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2018 (Salt Lake City, UT), 3165–3174.
- Yi, Z., Liu, Y., Zhao, Y., Xiao, L., Leong, A. T., Feng, Y., et al. (2021). Joint calibrationless reconstruction of highly undersampled multicontrast MR datasets using a low-rank Hankel tensor completion framework. *Magn. Reson. Med.* 85, 3256–3271. doi: 10.1002/mrm.28674
- Yin, X. X., Hadjiloucas, S., Chen, J. H., Zhang, Y., Wu, J. L., and Su, M. Y. (2017). Tensor based multichannel reconstruction for breast tumours identification from DCE-MRIs. *PLoS ONE* 12, e0172111. doi: 10.1371/journal.pone.0172111
- Zhang, K., Gao, X., Tao, D., and Li, X. (2012a). "Multi-scale dictionary for single image super-resolution," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2012 (Providence, RI), 1114–1121.
- Zhang, X., Xu, Z., Jia, N., Yang, W., Feng, Q., Chen, W., and Feng, Y. (2015). Denoising of 3D magnetic resonance images by using higher-order singular value decomposition. *Med. Image Anal.* 19, 75–86. doi: 10.1016/j.media.2014.08.004
- Zhang, Y., Wu, G., Yap, P. T., Feng, Q., Lian, J., Chen, W., and Shen, D. (2012b). "Reconstruction of super-resolution lung 4D-CT using patch-based sparse representation," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.* (Providence, RI), 925–931.
- Zhao, C., Dewey, B. E., Pham, D. L., Calabresi, P. A., Reich, D. S., and Prince, J. L. (2021). SMORE: a self-supervised anti-aliasing and super-resolution algorithm for MRI using deep learning. *IEEE Trans. Med. Imag.* 40, 805–817. doi: 10.1109/TMI.2020.3037187
- Zhao, J., Xiong, R., Xu, J., Wu, F., and Huang, T. (2019). "Learning a deep convolutional network for subband image denoising," in *Proc. IEEE Int. Conf. Multimed. Expo.* vol. 2019 (Shanghai), 1420–1425.
- Zhu, D., and Qiu, D. (2021). Residual dense network for medical magnetic resonance images super-resolution. *Comput. Methods Programs Biomed.* 209, 106330. doi: 10.1016/j.cmpb.2021.106330

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Jia, Chen, Han, Liu, Wen and Tang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.