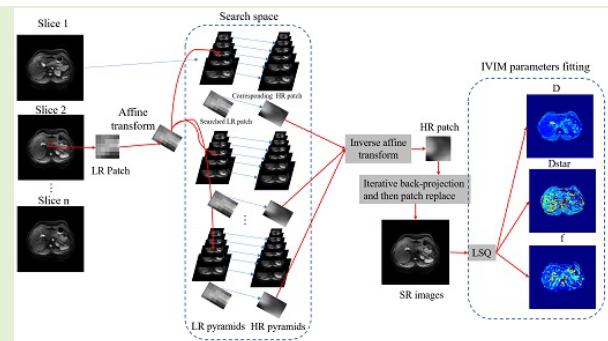


Super-Resolution of Intravoxel Incoherent Motion Imaging Based on Multisimilarity

Jiqing Huang, Lihui Wang^{ID}, Jin Qin, Yi Chen, Xinyu Cheng, and Yuemin Zhu

Abstract—Intravoxel incoherent motion imaging (IVIM) is a new magnetic resonance imaging (MRI) technique that can detect both diffusion and pseudo perfusion microscopic tissue information. However, it requires to acquire the diffusion-weighted (DW) images with multiple b-values, which leads to long acquisition time. To reduce the time, the spatial resolution of IVIM is usually not high enough to reflect the image details. To solve this problem, in this study, a multi-similarity block-based super-resolution reconstruction (SRR) method for IVIM was proposed. The fact that the image block similarity of DW images occurs not only across or within the images at different scales, but also the adjacent slices with the same b-values. The intra-slice and inter-slice similarity at multi-scales is defined as multi-similarity in this work. To fully explore the multi-similarity patches to assist SRR, a chain of Gaussian pyramids of a given image slice and the corresponding adjacent slices with the same b-value are established first, and then a random search strategy is used to find and identify the most similar blocks in low resolution (LR) image pyramids. Next, the matched blocks are mapped to high resolution (HR) image pyramids and then unwrapped by inverse affine transformation to replace original LR blocks. Finally, the super-resolution images are obtained through an iterative back-projection algorithm. To evaluate the performance of proposed method, we compared the brain and liver IVIM reconstruction results with several state-of-art methods. The results indicate that the multi-similarity-based method can achieve the best reconstruction results in both DW images and the corresponding IVIM parameter maps.

Index Terms—Intravoxel incoherent motion imaging, diffusion-weighted image, super-resolution reconstruction, gaussian pyramid, iterative back-projection.



I. INTRODUCTION

INTRAVOXEL incoherent motion (IVIM) is a promising magnetic resonance imaging (MRI) technique for investigating the pathophysiological properties of biological tissues from a microscopic view [1]. To map the IVIM parameters accurately, diffusion-weighted (DW) images with multiple b-values are usually required. Nevertheless, acquisition of these images increases the acquisition time, especially in

the case of high-spatial resolution data acquisitions. Long acquisition times induce many undesirable effects, such as motion blurring and unpleasant feelings for patients. Therefore, to reduce the acquisition time of IVIM is still a clinical problem that needs to be addressed. Generally, the number of b-values cannot be decreased to guarantee the accuracy of IVIM parameter mapping. Therefore, decreasing the spatial resolution of the DW images is an alternative option to accelerate the acquisition. However, low-spatial resolution images will blur the details of the structure and reduce the accuracy of the image analysis, thus hindering clinical diagnosis.

To deal with this dilemma, a potential image processing technique, known as super-resolution image reconstruction (SRIR), has been investigated extensively in recent years [2], [3]. SRIR attempts to predict the high-resolution (HR) image from a single image or a sequence of low-resolution (LR) images. This is a well-known ill-posed problem given that the image degradation process is unknown. Therefore, the reconstructed HR image cannot be determined uniquely. Numerous methods have been proposed to solve this problem, such as interpolation-based approaches, modeling-based methods, and learning-based methods etc., [4], [5].

Interpolation-based methods use a common and simple assumption based on which the predicted pixel value is

Manuscript received February 18, 2020; revised May 6, 2020; accepted May 8, 2020. Date of publication May 11, 2020; date of current version August 14, 2020. This work was supported in part by the National Nature Science Foundations of China under Grant 61661010 and Grant 61562009, in part by the Program PHC Yuanpei 2018 under Grant 41400TC, in part by the Funds for Talents of Guizhou University under Grant 2013(33), and in part by the Guizhou Science and Technology Plan Project under Grant Qiankehe [2018]5301. The associate editor coordinating the review of this article and approving it for publication was Dr. Ing. Emiliano Schena. (Corresponding author: Lihui Wang.)

Jiqing Huang, Lihui Wang, Jin Qin, Yi Chen, and Xinyu Cheng are with the Key Laboratory of Intelligent Medical Image Analysis and Precise Diagnosis of Guizhou Province, School of Computer Science and Technology, Guizhou University, Guiyang 550025, China (e-mail: jiqing-huang@outlook.com; wlh1984@gmail.com; qin_gs@163.com; yi-chen.gzu@outlook.com; xycheng@gzu.edu.cn).

Yuemin Zhu is with the University of Lyon, CNRS UMR 5220, Inserm U1206, INSA Lyon, Creatis, 69100 Villeurbanne, France (e-mail: zhu@creatis.insa-lyon.fr)

Digital Object Identifier 10.1109/JSEN.2020.2993873

always repeated along the contour-like textures and edges or surrounding them [6]–[8]. Although these methods are easy to implement within a short time, they often generate poor reconstruction results with block or blur effects owing to the simplicity of the posed hypothesis. Therefore, this type of SRIR technique has not been extensively used in recent years. To overcome the problem of interpolation, modeling-based reconstruction methods were proposed [9]–[12], which established a complex equation to describe the relationship between the LR and HR images, including image blurring, distortion, noise addition, and down-sampling. Based on these models, the SR(Super-resolution) image is iteratively reconstructed from an initial solution by solving an optimized problem with certain prior constraints. However, these types of methods are extremely sensitive to the previously selected priors and the initial solution. Accordingly, the reconstruction results cannot be guaranteed. To address these issues, many powerful learning-based methods, such as sparse coding [13]–[15] and dictionary learning [16]–[18] were proposed to exploit the complex mapping between low- and high-resolution images. The recent successful applications of machine learning and deep learning in the field of computer vision led to the formulation of SRIR using convolutional neural networks (CNN). This pioneering work was termed as super-resolution convolutional neural network (SRCNN) [18], it used a shallow convolutional neural network to learn the end-to-end mapping relationship between the LR and HR images. Inspired by this work, various deep learning based SRIR methods were proposed. Among them, the most typical ones are the very deep super-resolution (VDSR) [19], enhanced deep super-resolution (EDSR) [4], efficient sub-pixel convolutional neural network (ESPCN) [20], super-resolution generative adversarial networks (SRGAN) [21], fast super-resolution convolutional network (FSRCNN) [22], and deep back-projection network (DBPN) [23] etc.. Although the deep learning based SRIR methods achieve superior performances compared to conventional filtering techniques, their stringent requirements for the number of training data and the training instability limit their applications, especially for applications to medical images.

The inadequate training data encourages the use of several approaches to exploit internal patch similarities for SRIR. These methods are based on the statistical properties of the image, which assume that the image patches are repeated within and across different image scales [24], [25]. Searching the local or non-local patches that are similar to a given LR patch across different scales and using these similar patches to reconstruct HR patch is usually named as self-similarity-based method, which means that the similar patches were searched from the multiscale expressions of one image slice. However, in the case of IVIM, patch similarities are present not only in the intra-slice but also the inter-slices. Therefore, to fully exploit the similarities for further improving the SRIR accuracy of IVIM, we propose herein a multi-similarity based method, which constructs firstly a series of gaussian pyramids of given paired HR and LR image slices and the corresponding adjacent slices with the same b-value, and then a random search strategy is used to find

and identify the most similar blocks in the LR pyramids. After that, the selected similar blocks are mapped into the HR pyramids by an affine transformation. Since the mapped blocks in HR pyramids contain image details which are useful for the reconstruction, we implement the inverse affine transformation on the blocks in HR pyramids and use the resulted blocks to replace original LR blocks. Finally, the image obtained with the blocks replacing was considered as the initial estimate, which is further fine-tuned using an iterative back-projection algorithm.

The rest of the manuscript is organized as follows: the SRIR method, which is based on multi-similarity, is detailed in Section II. The segmented least-square-based IVIM parameter fitting method is introduced in Section III. Additionally, the evaluation criteria, experimental results, the discussion, and conclusions are listed in Sections IV, V, VI, and VII, respectively.

II. METHODS

To fully exploit the information in the neighboring slices, the multi-similarity instead of the self-similarity-based algorithm is proposed for the SRIR of IVIM. Firstly, we expand the searching space in a multi-similarity pyramids to identify available and useful patches from several neighboring slices acquired with the same b-value. Then, we propose a correlation evaluation approach between the target, matching patches, and the searching strategy to adapt the increment of the searching space. Depending on the designed cost function and searching strategy, paired candidate patches with affine transformation parameters for each point are then obtained. Finally, transformed HR patches are used to progressively reconstruct and restore details of HR images. The detailed algorithms are given as follows.

A. Construct Multi-Similarity Pyramidal Space

Given several DW image slices (LR) acquired with multiple b-values, we first group them according to the b-value,

$$I_n = \{i_{n,m} | m = 1, 2, \dots, M, n = 1, 2, \dots, N\} \quad (1)$$

where N and M represent the total number of b-values and slices respectively, n and m are their index.

For each image set I_n , the multi-similarity space including a LR pyramidal set ($D = \{D_n^l\}$) and a HR pyramid set ($U = \{U_n^l\}$) is constructed, as illustrated in Fig.1 and formulated in Eq.(2).

$$\begin{cases} D_n^l = \{d_{n,m}^l | m = 1, 2, \dots, M, l = 1, 2, \dots, \text{bottom}\} \\ U_n^l = \{u_{n,m}^l | m = 1, 2, \dots, M, l = 1, 2, \dots, \text{bottom}\} \end{cases} \quad n = 1, 2, 3, \dots, N \quad (2)$$

where l indicates the level of pyramids. As l increases, the size of the images also increases. As an initial, $u_{n,m}^{\text{bottom}}$ is set as HR image, and the other HR pyramid layer $u_{n,m}^l$ ($l = 1, \dots, \text{bottom}-1$) is obtained by down-sampling $u_{n,m}^{l+1}$ ($l = 1, \dots, \text{bottom}-1$) using bicubic interpolation with a scaling factor η . For the LR pyramids, the bottom layer $d_{n,m}^{\text{bottom}}$ is obtained by up-sampling $u_{n,m}^{\text{bottom}-1}$ through a bicubic interpolation with a magnification factor η , and the other LR pyramid

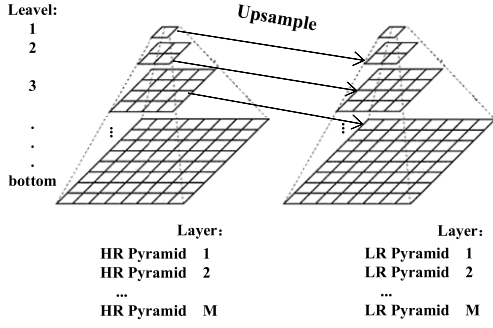


Fig. 1. Multiple HR-LR pyramid space.

layers were obtained through the same down-sampling process as that in HR pyramids. In order to reconstruct the SR image, we need to find the similar patches from both HR and LR pyramids of several neighboring slices to replace the initial estimated SR patches. In this process, the key points are how to define the similar patches and how to find them.

B. SR Reconstruction Based on Multi-Similarity

To construct a SR image from a series of HR and LR image pyramids, we estimated firstly the initial state of SR image $\hat{u}_{n,m,0}^{bottom+1}$ by up-sampling $u_{n,m}^{bottom}$ using a bicubic interpolation with a magnification factor η . Then, we segmented the SR image $\hat{u}_{n,m,0}^{bottom+1}$ into several patches with fixed size. The final reconstruction of SR image is realized by finding the proper patches to replace each patch in the initial SR image $\hat{u}_{n,m,0}^{bottom+1}$.

To achieve this, one image patch $P(t_i)$ in HR image $u_{n,m}^{bottom}$ was selected randomly, the center of which locates at t_i . In the LR pyramids of the adjacent layers among $d_{n,m-L_n}^{bottom+1}$ to $d_{n,m+L_n}^{bottom+1}$, where L_n is the number of adjacent slices, we search the closest matching patch $Q(t_i)$. Considering that the motion of patient may cause distortions of the matching block, an affine transformation T_i is applied to the target image block $P(t_i)$ to extract the similarity information effectively. In this study, affine transformation consists of similarity, rotations, and shear transformations, expressed as:

$$T_i = S(\theta_i^s, \theta_i^\phi) \times A(\theta_i^\alpha, \theta_i^\beta) \quad (3)$$

where $S(\theta_i^s, \theta_i^\phi)$ includes the both similarity and rotation transformation, formulated as:

$$\begin{aligned} S(\theta_i^s, \theta_i^\phi) &= \theta_i^s R(\theta_i^\phi) \\ &= \theta_i^s \begin{bmatrix} \cos(\theta_i^\phi) & -\sin(\theta_i^\phi) \\ \sin(\theta_i^\phi) & \cos(\theta_i^\phi) \end{bmatrix} \end{aligned} \quad (4)$$

where θ_i^s is the scale factor and $R(\theta_i^\phi)$ is a two-dimensional (2D) rotation matrix with a rotation angle of θ_i^ϕ . In Eq.(3), $A(\theta_i^\alpha, \theta_i^\beta)$ represents the shear transformation,

$$A(\theta_i^\alpha, \theta_i^\beta) = \begin{bmatrix} 1 & \theta_i^\alpha \\ \theta_i^\beta & 1 \end{bmatrix} \quad (5)$$

where $\theta_i^\alpha, \theta_i^\beta$ are the shear factors along the X-axis and Y-axis respectively.

The parameters of the affine transformation T_i can be defined as $\theta_i = (\theta_i^s, \theta_i^\phi, \theta_i^\alpha, \theta_i^\beta)$. After its application to the target patch, the cost function for searching the best matching patches is formulated as

$$C(t_i, \theta_i) = \min_{\{\theta_i\}} \sum_{i \in \Omega} E_{\text{res}}(t_i, \theta_i) + E_{\text{scale}}(t_i, \theta_i) \quad (6)$$

where Ω denotes the set of all pixel in $u_{n,m}^{bottom+1}$ which is the HR pyramid of the n^{th} b-value, m^{th} slice, and $bottom + 1$ level, E_{res} is the mean-squared error (MSE) between the target and matching image patches, defined as

$$E_{\text{res}}(t_i, \theta_i) = \|W_i(P(t_i) - Q(t_i, \theta_i))\|_2^2 \quad (7)$$

where W_i is the Gaussian weight kernel with a variance of 3, used for weighting each block based on its contribution to the similarity cost function, and $Q(t_i, \theta_i)$ is the matched image block in $d_{m-L_n:m+L_n}^l$, $l = 1, 2, \dots, \text{bottom}-1$, which is transformed by T_i . To avoid a trivial solution, in which the matching blocks are more likely to be found in the close level space, a scale loss function $E_{\text{scale}}(t_i, \theta_i)$ is introduced,

$$E_{\text{scale}}(t_i, \theta_i) = \lambda_{\text{scale}} \max(0, \alpha - \text{Scale}(T_i)) \quad (8)$$

where α is the scaling factor indicating the difference in the pyramidal level between the target and searching patches, and $\text{Scale}(T_i)$ is the estimation function of the affine transformation T_i . According to ref [25], it is defined as

$$\text{Scale}(T_i) = \sqrt{\det(T_i)} \quad (9)$$

where $\det(\cdot)$ denotes the value of the determinant.

Combining Eqs. (7) and (8), it can be inferred that if the scale of the searched block is large, $E_{\text{scale}}(t_i, \theta_i)$ is invalid, otherwise $E_{\text{scale}}(t_i, \theta_i)$ is activated when the scale of the searched block is smaller than η . Searching for similar blocks at large scales can keep more image details and texture information to ensure good reconstruction quality.

By minimizing the above cost function, the similar patches and the optimize affine transformation parameters can be determined. Although the multi-similarity space can guarantee the SR reconstruction accuracy, it increases the searching space and therefore increases the computation complexity. To seep up the match searching, we optimized the searching strategy as described in the following subsection.

C. Optimization Based on Similarity Patch Searching

To solve the nearest-neighboring field estimation problem and avoid exhaustive searching in LR pyramids, the patch match algorithm is introduced. The main idea of which is that the similar patch often matches in the localized neighborhood. By utilizing a priori knowledge in terms of the image pixel value continuity, the computation time is significantly reduced. The process of patch match includes three important steps, namely the initialization, propagation, and random searching, as illustrated in Fig. 2. In which, A and B represent HR and LR images respectively. Suppose that the orange square represents the target image block P , the red and green squares are the randomly selected surrounding blocks, noted as P_1 and P_2 . The centers of P , P_1 , and P_2 are marked as t, t_1 and t_2

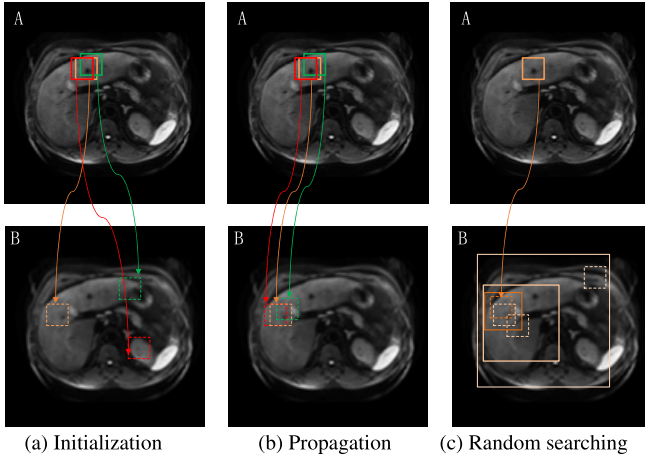


Fig. 2. Main concept of patch-match algorithm. A and B represent HR and LR images respectively. In subfigure (a), solid orange square in HR image represents the target block, and the red and green solid square are randomly selected surrounding blocks. The dotted squares in LR image are randomly generated, in which, the one with the minimum cost is determined as the initial matched patch. In subfigure (b), propagating the blocks in LR images along the 4 neighboring directions to update the match block. In subfigure (c), the approximate optimal matching block is searched in a gradually reduced window.

respectively. During the initialization, the shifts off , off_1 and off_2 , are assigned randomly. Additionally, $t + off$, $t_1 + off_1$ and $t_2 + off_2$ are the central pixels of the matches Q , Q_1 , and Q_2 , demonstrated with the dotted boxes. The central pixel of block P is then updated in the propagation step. Comparing the similarity between the target block P and the candidates Q , Q_1 , and Q_2 , the offset with the smallest cost function will become the new off . This step definitely needs multiple iterations to obtain the optimal solution. Therefore, more matched blocks in the region centered at $t + off$ with a radius of r are searched to find the optimal matching patches. To speed up the search process and ensure the matching accuracy, the radius r is gradually reduced by a certain exponential order until the radius is unity in each iteration. Finally, the sub-region of SRR is completed by replacing the searched image block Q with P .

In the initialization of the traditional block matching algorithms, the candidate center offsets are randomly generated. This stochastic process easily causes the search results to fall into a local optimum, and thus costs much time for searching. Therefore, we use the following initialization process to promote the local similarity of images and add the characteristics of affine transformation to increase the diversity of the search space. The initial locations of patches s_i in the LR pyramid coincide with the center positions of the target block t_i in the HR image. This means that the offset is set to be zero and the affine transformation is initialized as a diagonal matrix. During the propagation, the matching image blocks are respectively moved by one pixel in four neighboring directions, thereby yielding four adjoining matching patches. It is assumed that the affine transform parameters for the center patch and its four neighbors are θ_{ij} , $j = 0, 1, 2, 3, 4$. The matching patch which is the closest to the target patch should be that can minimize the cost function $C(t_i, \theta_{ij})$. The corresponding parameters of

the affine transformation is propagated to θ_i , expressed as,

$$\theta_i = \arg \min_j \{C(t_i, \theta_i) | j = 0, 1, 2, 3, 4\} \quad (10)$$

After the affine transformation $T(\theta_i)$ of the initial position s_i , it is difficult to guarantee that the transformed block is optimal. To deal with this issue, we modified the traditional patch matching method in terms of random search, which not only randomly selects the similarity patches within a certain area but also perturbs the parameters of the affine transformation.

$$\begin{cases} s_{i,iter}^x = s_i^x + r \times \mu^{iter} \times rand(0, 1) \\ s_{i,iter}^y = s_i^y + r \times \mu^{iter} \times rand(0, 1) \\ s_{i,iter}^m = randi(m - L_n, m + L_n) \\ \theta_{i,iter}^s = scale_{pyramid}(s_{i,iter}^m) + rand(0, 1)/iter \\ \theta_{i,iter}^\phi = \phi \times rand(0, 1)/iter \\ \theta_{i,iter}^\alpha = \kappa \times rand(0, 1)/iter \\ \theta_{i,iter}^\beta = \kappa \times rand(0, 1)/iter \end{cases} \quad (11)$$

where r is an initial window radius, μ denotes the attenuation coefficient of the window size, which is less than 1 to reduce the search range of the image. In addition, $iter$ records the number of iterations, $s_{i,iter}^x$ and $s_{i,iter}^y$ represent the center position of the image windows in the x -direction and y -direction respectively for $iter^{th}$ iteration. $s_{i,iter}^m$ indicates the slice index of the image windows for $iter^{th}$ iteration. $\theta_{i,iter}^s$, $\theta_{i,iter}^\phi$, $\theta_{i,iter}^\alpha$, and $\theta_{i,iter}^\beta$ are the scale, rotation angle, and shear parameters of the affine transformation at the $iter^{th}$ iteration. In addition, $\theta_{i,iter}^s$ is decided by the pyramidal scale of the matching patch $scale_{pyramid}(s_{i,iter}^m)$,

$$scale_{pyramid}(s_{i,iter}^m) = \begin{cases} 1/\alpha, & \text{when } s_{i,iter}^m < m \\ 1, & \text{when } s_{i,iter}^m \geq m \end{cases} \quad (12)$$

where, α is a step-down-sampling parameter of the pyramid. As the number of iterations increases, the parameters tend to be stable and the matched patches remain almost unchanged so that the rotation and shear parameters of the affine transformation gradually approach zero. However, the value of the scale parameter depends on the level of the pyramid that the matching patches located at. Thus, at each iteration, $\theta_{i,iter}^\phi$ is obtained randomly based on the uniform distribution on a closed interval $[0, \phi/iter]$, $\theta_{i,iter}^\alpha$ and $\theta_{i,iter}^\beta$ are obtained randomly from the uniform distribution within a closed interval $[0, \kappa/iter]$, where ϕ is the maximum rotational angle of the affine transformation, and κ is the maximum shear parameter of the shear transformation. In Eqs.(11), the function $rand(a, b)$ is used to generate a random fraction that is uniformly distributed within the closed interval $[a, b]$ and $randi(a, b)$ is used to generate a random integer that is uniformly distributed in the closed interval $[a, b]$. When a new matching patch candidate is generated, if it is closer to the target patches compared to the previous matching patches, it replaces the former; otherwise it maintains the original one. After updating the matching patches, the propagation and random searching are continued until the search window size is unity, or until the stop criteria is satisfied. The optimized patch matching flow chart is shown in algorithm 1.

Algorithm 1 Modified Patch-Matching Algorithm

Input: Target reconstruction image, LR pyramid
Output: Parameters of matching patches
Initialize matching patches parameters s_i, θ_i
Exterior layer (traverse all pixel points of the image, $i = 1, 2, 3, \dots$):
Inner layer (iteratively update matching patches parameters, $\text{iter} = 1, 2, 3, \dots$):
Propagate parameters using an affine transformation:

$$\theta_i = \arg \min_j \{C(t_i, \theta_i) | j = 0, 1, 2, 3, 4\}$$

Update matching patch parameters according to Eqs.(11)
Calculate the smallest cost function $C(t_i, \theta_{i, \text{iter}})$

$$\theta_i = \min(C(t_i, \theta_{i, \text{iter}}), C(t_i, \theta_{i, \text{iter}-1}))$$

D. Evaluation Criteria

To quantitatively evaluate the ability of the proposed SRR method in the DW image restoration and IVIM parameters estimation, we use peak signal-to-noise ratio (PSNR), signal-to-noise ratio (SNR), and Riesz-transform-based feature similarity metric (RFSIM) to evaluate the quality of the reconstructed DW images, and use average of the residual error (ARE) to evaluate the quality of the reconstructed IVIM parameters.

PSNR and SNR are extensively used for the measurement of the reconstruction accuracy, and RFSIM is used for the evaluation of the structure and feature similarity between the reconstructed and reference images. PSNR is defined as

$$\text{PSNR} = 10 \log_{10} \left(\frac{\max(G)^2}{\text{MSE}} \right) \quad (13)$$

where

$$\text{MSE} = \sum_{t_i^x=1}^{N_x} \sum_{t_i^y=1}^{N_y} (R(t_i^x, t_i^y) - G(t_i^x, t_i^y))^2 \quad (14)$$

in which G are the reference images, R refers to the reconstructed images, N_x and N_y stand are the dimensions of the images. For MRI images, especially for IVIM applications, SNR is one of the most important factors that influences the IVIM parameter fitting, defined as

$$\text{SNR} = 10 \log_{10} \left(\frac{\text{Power}}{\text{MSE}} \right) \quad (15)$$

where Power is given by

$$\text{Power} = \sum_{t_i^x=1}^{N_x} \sum_{t_i^y=1}^{N_y} (G(t_i^x, t_i^y))^2 \quad (16)$$

RFSIM is computed by comparing the feature maps at key locations marked between the reference and the reconstructed images. Let f_g and f_r represent the Riesz-transform features of G and R . Both features have five dimensions. The similarity between the two feature maps f_g and f_r is defined as

$$d_i(t_i^x, t_i^y) = \frac{2 f_{gi}(t_i^x, t_i^y) \cdot f_{ri}(t_i^x, t_i^y) + c}{f_{gi}^2(t_i^x, t_i^y) + f_{ri}^2(t_i^x, t_i^y) + c} \quad (17)$$

where c is a small constant value. The RFSIM can then be computed as

$$\text{RFSIM} = \prod_{i=1}^5 \frac{\sum_{t_i^x=1}^{N_x} \sum_{t_i^y=1}^{N_y} d_i(t_i^x, t_i^y) \cdot M(t_i^x, t_i^y)}{\sum_{t_i^x=1}^{N_x} \sum_{t_i^y=1}^{N_y} M(t_i^x, t_i^y)} \quad (18)$$

where $M(\cdot)$ indicate the Riesz-transform features at certain pixel more than 0.

Average of the residual error (ARE) is used to compare the difference between the reference and reconstructed IVIM parameters, formulated as

$$\text{ARE} = \sum_{t_i^x=1}^{N_x} \sum_{t_i^y=1}^{N_y} |G(t_i^x, t_i^y) - R(t_i^x, t_i^y)| / \sum_{t_i^x=1}^{N_x} \sum_{t_i^y=1}^{N_y} 1 \quad (19)$$

III. EXPERIMENTAL RESULTS**A. Data Description and Experimental Setup**

Two IVIM datasets were used for evaluating the SRR performance, one was for liver and the other was for brain. The Liver dataset contained 26 DW images of two patients, which were acquired with a 3 T Philips MRI scanner using the following parameters: TR/TE = 2000/53 ms, matrix size = 128×128 , slice thickness = 8 mm, slice spacing = 8 mm, and b value = 0, 10, 20, 40, 60, 80, 100, 200, 300, 400, 600, 800 s/mm². The brain dataset contained 20 DW images of two patients, which were acquired with a 3 T GE MRI scanner. The acquisition parameters are as follows: TR/TE = 2800/77 ms, matrix size = 128×128 , slice thickness = 5 mm, slice spacing = 6.5 mm, and b value = 0, 10, 20, 30, 40, 50, 70, 100, 200, 400, 600, 800, 1000 s/mm².

To validate the superiority of the proposed method, we compared our proposed method with a variety of state-of-art methods, including self-similarity, SRCNN and EDSR in terms of the visual quality and quantitative metrics. After the verification, the SR images were obtained by upsampling directly the acquired images, subsequently, the IVIM fitting model was applied on the reconstructed SR DW images to evaluate the change of IVIM parameters. In this work, we use segmented least squares (SLSQ) method to fit the IVIM parameters, including the diffusivity (D), pseudo-diffusivity (D^*), and perfusion fraction (f) [26].

As the SRCNN and EDSR models were well trained, we simply changed the input layer to adopt medical image data. Therefore, the detailed parameters of the deep neural network are the same as that used in the original works [4], [25], [27]. The hyperparameters that used in the multi-similarity and self-similarity methods are listed in Table I.

In Table I, L_n is the number of adjacent slices, η describes the scaling factor of Gaussian pyramids, patchsize denotes the size of the reconstruction block that we used (e.g., 3×3 blocks in the liver dataset and 7×7 blocks in the brain dataset, γ is the size of the searching window, κ and ϕ denote the small disturbances of target patches in angle and shear, and μ represents the attenuation coefficient for search windows.

TABLE I
MAJOR PARAMETER SETTINGS IN SIMILARITY-BASED METHODS

	L_n	η	$patchsize$	γ	κ	ϕ	μ
Self-similarity	-	$\sqrt[3]{1/2}$	3 or 7	$\frac{Nx}{4}$	$\frac{\pi}{4}$	0.05	0.5
Multi-similarity	2	$\sqrt[3]{1/2}$	3 or 7	$\frac{Ny}{4}$	$\frac{\pi}{4}$	0.05	0.5

B. HR Reconstruction Results for DW Images

We downsampled the original image as the input image and use different methods to reconstruct. In this case, these reconstructed images are visually compared with the reference images. The input image is simply an enlargement of the down-sampled image to help visualization. Fig.3 shows only the normalized reconstructions for a middle slice of DW images for different organs and patients and the corresponding residuals with the normalized ground truth. The color bar represents the residual between all reconstructed image and ground truth which is multiplied by five to improve the visual effect. Comparison of these residuals proves that the multi-similarity method provides the best SRR in both the smooth and edge regions. As to the deep learning-based reconstruction results, the edge reconstruction is not desirable.

In addition to the visual assessment, we used PSNR, SNR, and RFSIM to evaluate the SRR performance, as shown by the box plots in Fig. 4. These groups of figures reflect the statistical information of all the slices. The two ends of the dashed line represent the minimum and maximum values. The lid and bottom parts of the box plots are the quartile values, while the middle line of the box part is the mean value of the qualification index. We found that, our method was significantly improved compared to other methods for almost all the samples in terms of quantitative metrics. This indicates that the multi-similarity method not only provides the best visual effect but also maintains the original pixel value. Compared with the EDSR method, the multi-similarity-based algorithm has improved the PSNR, SNR, and RFSIM, by at least 17%, 31%, and 27% respectively. Compared with the SRCNN method, the proposed algorithm has improved the PSNR, SNR, and RFSIM, by at least 7%, 9%, and 5% respectively, compared to the original self-similarity method, our method can increase the PSNR, SNR and RFSIM by 1%, 2%, and 1% respectively.

C. HR Reconstruction Results for IVIM Parameters

To show that the SR can also help improve the accuracy of the parameter fitting, we used groups of DWI which were reconstructed with different methods to fit the IVIM parameters. For simplicity, we illustrated only the parameter fitting map of the liver 1 dataset in Fig. 5. The fitting parameter mapping verified that the proposed method outperforms the other methods. To further validate this point, the quantitative evaluation metric ARE of the fitted IVIM parameters obtained with different methods for all the samples was given in Fig.6. It is shown that our proposed method yields more accurate

IVIM parameters. The ARE obtained by EDSR are the biggest, SRCNN is better than EDSR but worse than self-similarity method. Compared with self-similarity method, the proposed multi-similarity method is better for reconstructing the perfusion parameter D^* and perfusion fraction f .

D. SR Reconstruction Results

To improve the spatial resolution in the real application, we directly magnify the original acquired images using the same parameters, and the results were given in Fig.7. As demonstrated, the resolution is improved, which allows the visualization of details and edges, and yields accurate pixel values and IVIM parameter maps, as illustrated in Fig.8.

IV. DISCUSSION

IVIM has potential to screen several diseases at microscopic view. However, for clinical IVIM acquisitions, higher b values required to be used, which resulted in usually DW images with low SNR. The SNR loss limits the spatial resolution of the clinical IVIM acquisitions. Currently, numerous SRR research have been conducted on natural and medical images but few efforts have focused on IVIM [28]. In this work, we presented a type of integrated method for DW image SRR and IVIM parameter estimation. For the SRR of IVIM, the reconstruction accuracy for both DW image intensity and the corresponding parameter maps is the key issue for following precise diagnosis. In this approach, more attention was paid to the local reconstruction accuracy and SNR for IVIM SR reconstruction, and the results were evaluated on a pointwise manner.

For the DW image reconstruction, as illustrated in Fig. 3, it was observed that the residuals between the reconstructed images and the ground-truth declined from left to right. The learning-based methods have large residuals in the boundary regions, this phenomenon can also be found in the nature image SRR [29]–[31], while the similarity-patch-based methods have small and uniform residuals. In the case of the SRCNN method, there is only one convolution layer that is used to extract local features, so it is difficult to learn detailed features in a global sense. At the same time, the cost function of SRCNN was the MSE, which led to a larger edge penalty, and caused the solution to attain the average value of all the possible solutions. In the case of the EDSR, the deeper network partially overcame the shortcomings of the simplified SRCNN method, but it is easy to cause an overfitting problem and results in a reconstruction instability. This explains why the learning-based SRR methods for DW images are not well as expected. The self-similarity-based method can explore more similar patches at multiple scales by constructing the multilevel Gaussian pyramidal model, the combination of such similar patches could provide different details for SRR, and therefore promote the reconstruction accuracy. Furthermore, the combination of multiple loss functions used in the self-similarity based method balances the contradiction between the edge protection and smoothing area uniformity, which explains why it can achieve results with comparable or better quality than those of the deep-learning networks.

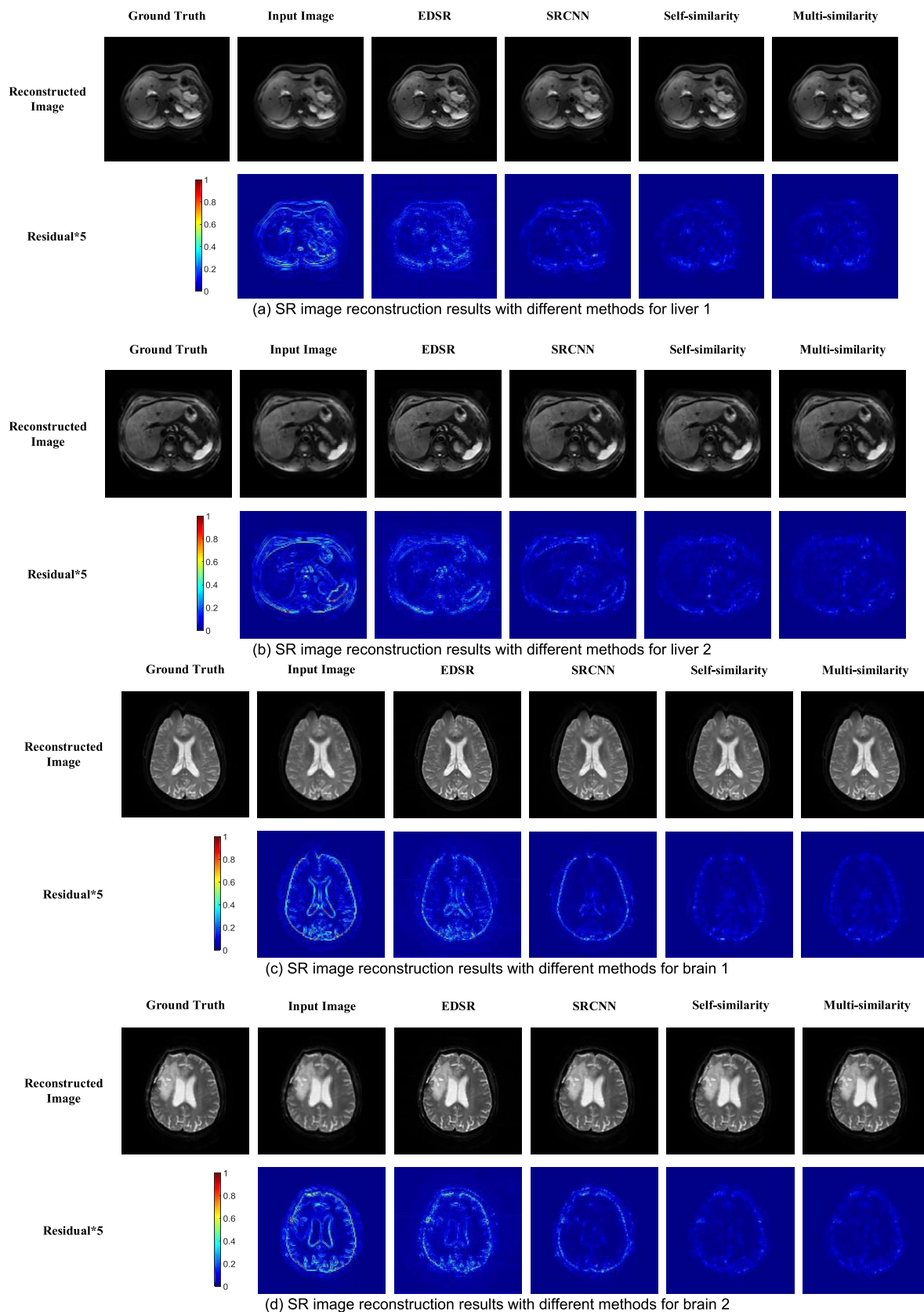


Fig. 3. SR reconstruction results and the corresponding residual maps compared to ground-truth with different methods for different DW images.

Compared with the self-similarity method, the multi-similarity algorithm proposed in this work investigates much more similar patches by establishing multiple groups of Gaussian

pyramids, in which, not only the similarity between the different pyramid scales was considered, but also the similarity between the adjacent slices was taken into account. The extra

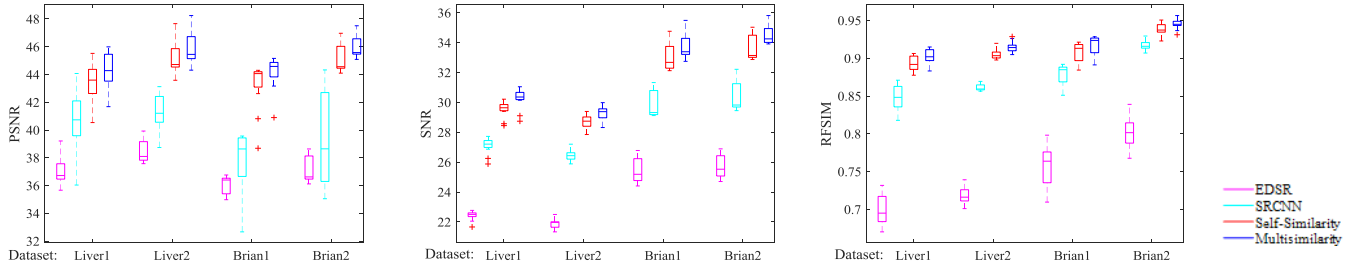


Fig. 4. Quantitative evaluation of DW images reconstructed with different methods. Five horizontal lines from top to bottom of each box indicate the maximum, third quartile, median, first quartile, minimum, respectively.

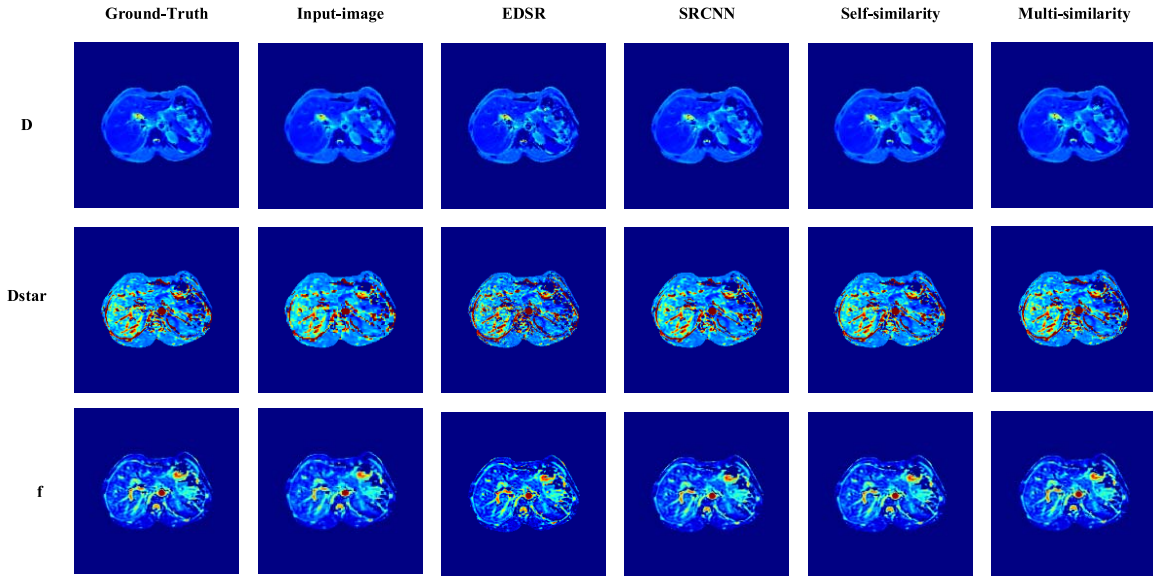


Fig. 5. Reconstructed IVIM parameter maps with different methods for liver 1, including diffusion coefficients (D), perfusion coefficients (D^*), and perfusion fraction (f).

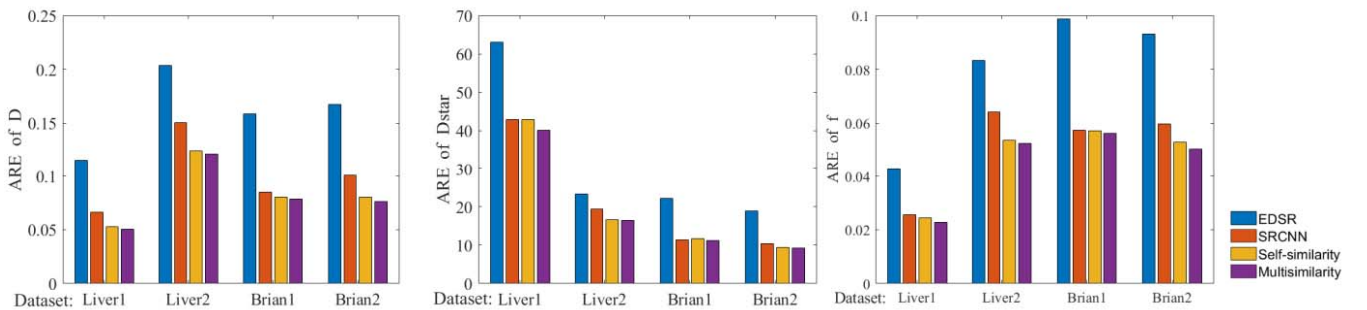


Fig. 6. Average residual values (ARE) of the IVIM parameters reconstructed with different methods.

information provided by the adjacent slices can improve the reconstruction details.

For the IVIM parameters, as illustrated in Fig 5 and Fig 6, we found that the ARE maps of D , D^* and f obtained by EDSR is much larger than the similarity-based methods. Since these IVIM parameters are calculated with the DW images acquired at different b values, if the DW images were not reconstructed well at some b values, the corresponding IVIM parameters would experience a big bias. The Fig.5 and Fig.6 demonstrate that the DW image quality reconstructed with EDSR model at different b values are not stable. As we mentioned above, when the b value increases, the SNR of

the acquired DW images becomes lower, since the learning-based method is very sensitive to the noise, consequently, the DW image reconstruction results at high b values are not guaranteed. However, the similarity-based methods overcome this kind of problem. Instead of learning the corresponding relationship between the LR and HR images, the similarity-based methods fully explore the useful information from LR images to reconstruct HR images, more similar patches are explored, more accurate the HR image reconstruction.

Although the proposed multi-similarity-based IVIM SRR method achieves the best SRR results for both DW images and the IVIM parameters, there are still some limitations. Firstly,

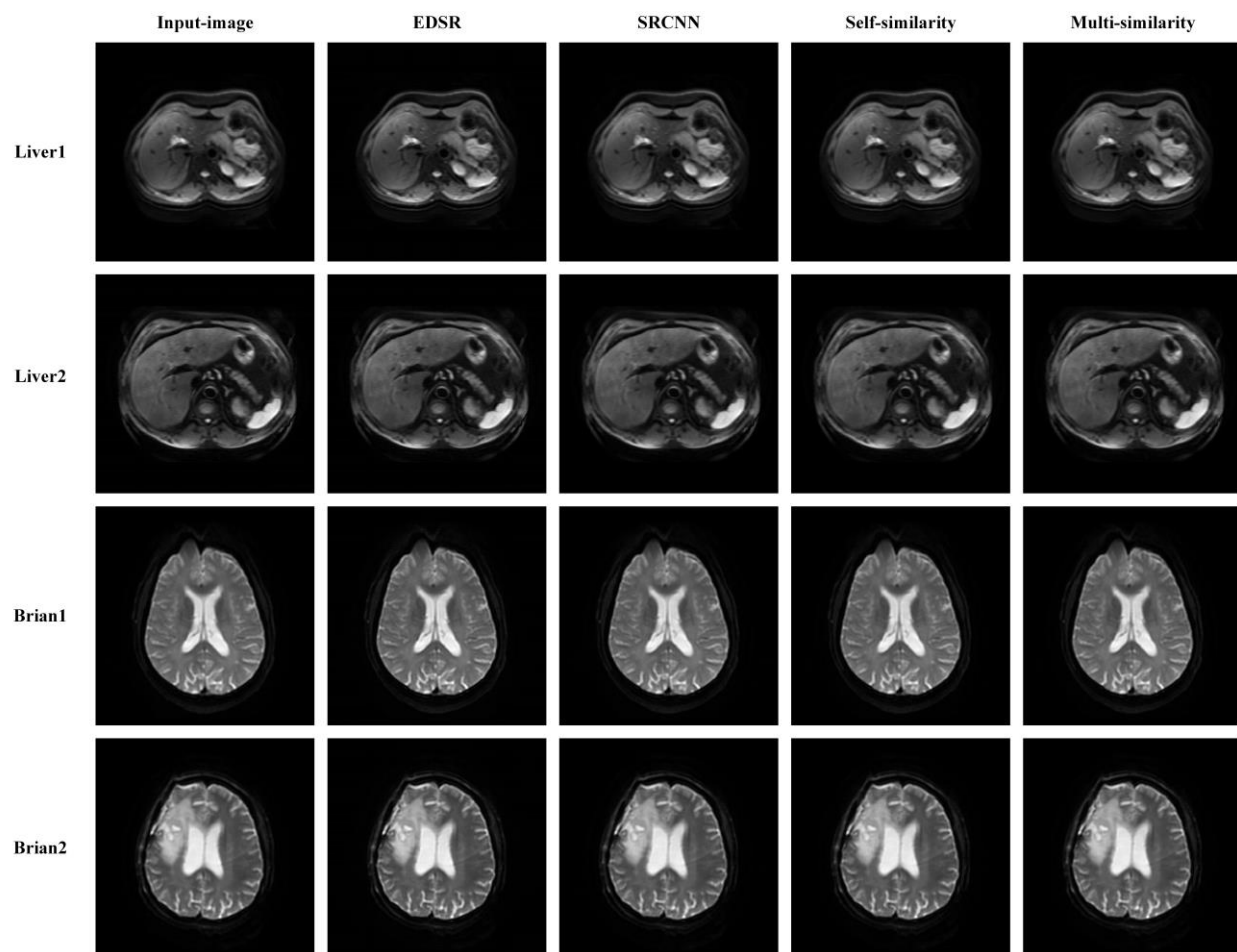


Fig. 7. Overall SRR of original images.

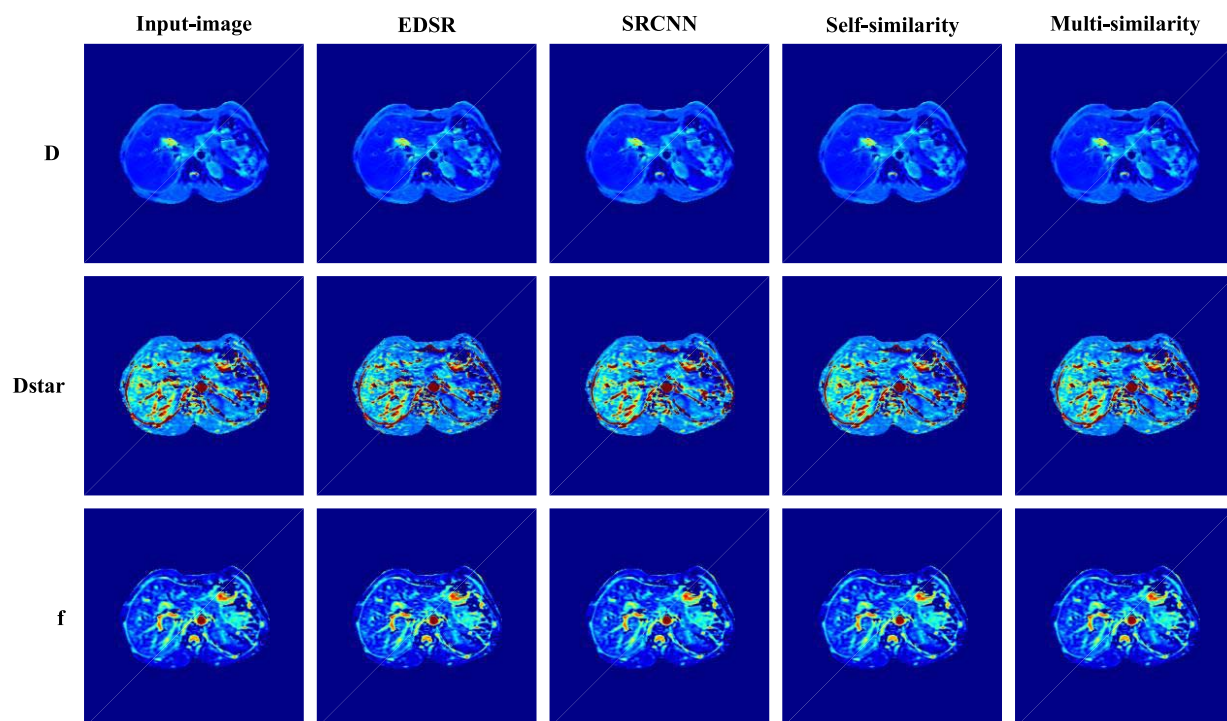


Fig. 8. Visualization of the parameters of the liver dataset in SRR of high-resolution (HR) images.

the proposed method was vulnerable to the window size, and we chose different window sizes for different datasets. Thus, the self-adaptation strategy constitutes important future work. By building more delicate multi-size patch pyramids, or by

designing new self-adaptation search strategies, the speed and precision—both in terms of the DW images and IVIM parameters—will be potentially improved. Secondly, the SRR implemented in this work intended to increase in-plane resolution, how to increase the interplane resolution and reconstruct more slices will be interesting. Finally, the accuracy of the IVIM parameters was also affected by the fitting methods. Thus, finding a better fitting method is still an important issue for comparing the IVIM SRR results, especially to evaluate the clinical value of the SRR methods.

V. CONCLUSION

We proposed a multi-similarity-based IVIM image SR reconstruction method by expanding the searching space among the Gaussian pyramids of several adjacent DW image slices. The multi-similarity patch for each location was found in this space and was used to get SR reconstruction patches. To promote the reconstruction speed, we modified also the searching strategy by introducing the perturbation of affine transformation parameters. The experimental results demonstrated that the proposed method achieved the better results compared with the state-of-art deep learning models and self-similarity-based methods. In addition, in terms of the IVIM parameter maps, the proposed methods outperform also the other methods.

REFERENCES

- [1] Y. T. Li, J.-P. Cercueil, J. Yuan, W. Chen, R. Loffroy, Y. Xiáng, and J. Wáng, "Liver intravoxel incoherent motion (IVIM) magnetic resonance imaging: A comprehensive review of published data on normal values and applications for fibrosis and tumor evaluation," *Quantum Imag. Med. Surg.*, vol. 7, no. 1, p. 59, 2017.
- [2] B. Scherrer, A. Gholipour, and S. K. Warfield, "Super-resolution in diffusion-weighted imaging," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assisted Interv.*, 2011, pp. 124–132.
- [3] N. M. H. Elsaid and Y.-C. Wu, "Super-resolution diffusion tensor imaging using SRCNN: A feasibility study," in *Proc. 41st Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2019, pp. 2830–2834.
- [4] B. Lim, S. Son, H. Kim, S. Nah, and K. M. Lee, "Enhanced deep residual networks for single image super-resolution," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jul. 2017, pp. 136–144.
- [5] L. M. Delfin, R. P. Elias, H. de Jesús Ochoa Domínguez, and O. O. Villegas, "Auto-regularized gradients of adaptive interpolation for MRI super-resolution," *J. Signal Process. Syst.*, vol. 91, no. 8, pp. 885–898, Aug. 2019.
- [6] J. V. Manjón, P. Coupé, A. Buades, V. Fonov, D. L. Collins, and M. Robles, "Non-local MRI upsampling," *Med. Image Anal.*, vol. 14, no. 6, pp. 784–792, Dec. 2010.
- [7] F. Yang, Y.-M. Zhu, I. E. Magnin, J.-H. Luo, P. Croisille, and P. B. Kingsley, "Feature-based interpolation of diffusion tensor fields and application to human cardiac DT-MRI," *Med. Image Anal.*, vol. 16, no. 2, pp. 459–481, Feb. 2012.
- [8] S. Yu, R. Zhang, S. Wu, J. Hu, and Y. Xie, "An edge-directed interpolation method for fetal spine MR images," *Biomed. Eng. OnLine*, vol. 12, no. 1, p. 102, 2013.
- [9] F. Shi, J. Cheng, L. Wang, P.-T. Yap, and D. Shen, "LRTV: MR image super-resolution with low-rank and total variation regularizations," *IEEE Trans. Med. Imag.*, vol. 34, no. 12, pp. 2459–2466, Dec. 2015.
- [10] J. T. Hsu, C. C. Yen, C. C. Li, M. Sun, B. Tian, and M. Kaygusuz, "Application of wavelet-based POCS super-resolution for cardiovascular MRI image enhancement," in *Proc. 3rd Int. Conf. Image Graph. (ICIG)*, 2004, pp. 572–575.
- [11] S. Tourbier, X. Bresson, P. Hagmann, J.-P. Thiran, R. Meuli, and M. B. Cuadra, "An efficient total variation algorithm for super-resolution in fetal brain MRI with adaptive regularization," *NeuroImage*, vol. 118, pp. 584–597, Sep. 2015.
- [12] X. Zhang, E. Y. Lam, E. X. Wu, and K. K. Y. Wong, "Application of Tikhonov regularization to super-resolution reconstruction of brain MRI images," in *Medical Imaging and Informatics*. Berlin, Germany: Springer, 2008, pp. 51–56.
- [13] J. Yang, J. Wright, T. S. Huang, and Y. Ma, "Image super-resolution via sparse representation," *IEEE Trans. Image Process.*, vol. 19, no. 11, pp. 2861–2873, Nov. 2010.
- [14] W. Dong, L. Zhang, G. Shi, and X. Wu, "Image deblurring and super-resolution by adaptive sparse domain selection and adaptive regularization," *IEEE Trans. Image Process.*, vol. 20, no. 7, pp. 1838–1857, 2011.
- [15] R. Timofte, V. De Smet, and L. Van Gool, "A+: Adjusted anchored neighborhood regression for fast super-resolution," in *Proc. Asian Conf. Comput. Vis.*, 2014, pp. 111–126.
- [16] Y.-H. Wang, J. Qiao, J.-B. Li, P. Fu, S.-C. Chu, and J. F. Roddick, "Sparse representation-based MRI super-resolution reconstruction," *Measurement*, vol. 47, pp. 946–953, Jan. 2014.
- [17] K. Srinivasan, A. Sharma, and A. Ankur, "Group sparse based super-resolution of magnetic resonance images for superior lesion diagnosis," in *Proc. 1st Int. Conf. Med. Health Informat. (ICMHI)*, 2017, pp. 95–100.
- [18] J. Ota, K. Umehara, N. Ishimaru, and T. Ishida, "Application of sparse-coding super-resolution to 16-bit DICOM images for improving the image resolution in MRI," *Open J. Med. Imag.*, vol. 7, no. 4, p. 144, 2017.
- [19] J. Kim, J. K. Lee, and K. M. Lee, "Accurate image super-resolution using very deep convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 1646–1654.
- [20] W. Shi *et al.*, "Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 1874–1883.
- [21] C. Ledig *et al.*, "Photo-realistic single image super-resolution using a generative adversarial network," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 4681–4690.
- [22] C. Dong, C. C. Loy, and X. Tang, "Accelerating the super-resolution convolutional neural network," in *Proc. Eur. Conf. Comput. Vis.*, 2016, pp. 391–407.
- [23] M. Haris, G. Shakhnarovich, and N. Ukita, "Deep back-projection networks for super-resolution," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 1664–1673.
- [24] S. Jain *et al.*, "Patch-based super-resolution of MR spectroscopic images: Application to multiple sclerosis," *Frontiers Neurosci.*, vol. 11, p. 13, Jan. 2017.
- [25] J.-B. Huang, A. Singh, and N. Ahuja, "Single image super-resolution from transformed self-exemplars," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 5197–5206.
- [26] C. Ye *et al.*, "Estimation of intravoxel incoherent motion parameters using low b-values," *PLoS ONE*, vol. 14, no. 2, Feb. 2019, Art. no. e0211911.
- [27] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 2, pp. 295–307, Feb. 2016.
- [28] M. Delbany *et al.*, "One-millimeter isotropic breast diffusion-weighted imaging: Evaluation of a superresolution strategy in terms of signal-to-noise ratio, sharpness and apparent diffusion coefficient," *Magn. Reson. Med.*, vol. 81, no. 4, pp. 2588–2599, Apr. 2019.
- [29] J. Salvador, *Example-Based Super Resolution*. New York, NY, USA: Academic, 2016.
- [30] H. Wang *et al.*, "Deep learning enables cross-modality super-resolution in fluorescence microscopy," *Nature Methods*, vol. 16, no. 1, pp. 103–110, Jan. 2019.
- [31] W.-S. Lai, J.-B. Huang, N. Ahuja, and M.-H. Yang, "Deep Laplacian pyramid networks for fast and accurate super-resolution," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 624–632.



Jiqing Huang was born in Anhui, China, in 1994. He received the M.S. degree in computer science and technology from Guizhou University in 2019. He is currently pursuing the Ph.D. degree with INSA, LYON, Sciences-Health Interdisciplinary Doctoral School, France. His research interests include medical imaging, diffusion-weighted magnetic resonance imaging, image processing, and machine learning.



Lihui Wang was born in Jilin, China, in 1982. She received the B.S. degree in electronic science and technology and the M.S. degree in optics from Harbin Engineering University, Harbin, in 2006 and 2008, respectively, and the Ph.D. degree in image and system from INSA Lyon, Lyon, France, in 2013. From 2013 to 2017, she was an Assistant Professor with the Computer Science and Technology Department, Guizhou University, Guiyang. Since 2017, she has been an Associate Professor with the Key Lab of Intelligent Medical Image Analysis and Precise Diagnosis, Guiyang, China. She is the author of more than 30 articles. Her research interests include medical imaging, image processing, pattern recognition, and machine learning.



Jin Qin was born in Guizhou, China, in 1978. He received the B.S. and M.S. degrees in computer science and software from Guizhou University, China, in 2000 and 2003, respectively, and the Ph.D. degree in control science and engineering from the University of Science and Technology, Beijing, China, in 2012. Since 2010, he has been an Associate Professor with the College of Computer Science and Technology, Guizhou University. His research interests include computational intelligence and machine learning.



Yi Chen was born in Guizhou, China, in 1996. She received the M.S. degree in computer science and technology from Guizhou University in 2019.

Since 2019, she was an Assistant Lecturer with the Department of Computer Science and Technology, Guizhou University, and the Key Laboratory of Intelligent Medical Image Analysis and Precise Diagnosis, Guiyang, China. Her research interests include medical imaging, image processing and machine learning.



Xinyu Cheng was born in Guizhou, China, in 1978. He received the B.E. degree in computer application from Guizhou University, Guiyang, in 2007, and the M.E. degree in computer application from the Guizhou University of Technology, Guiyang, in 2011. From 2001 to 2004, he was a Teaching Assistant with the Department of Computer Engineering, Guizhou University of Technology. Since 2005, he has been an Associate Professor with the Computer Science and Technology Department, Guizhou University. He has been a member of the Key Lab of Intelligent Medical Image Analysis and Precise Diagnosis, Guiyang, since 2017. He is the author of more than 20 articles. His research interests include machine learning, computer vision, software engineering, and network communication.



Yuemin Zhu received the B.Sc. degree in telecommunication from the Huazhong University of Science and Technology (HUST), China, in 1982, the M.Sc. and Ph.D. degrees from INSA, Lyon, France, in 1984 and 1988, respectively, and the Habilitation à Diriger des Recherches, France, in 1993.

He is a permanent Research Director of the Centre National de la Recherche Scientifique (CNRS), France. His current research interests include image representation, reconstruction, correction, denoising, registration, segmentation, visualization, and fusion. His current research projects focus on human cardiac diffusion tensor imaging and human cardiac polarize light imaging. He has so far published more than 260 research articles, including 105 journal articles, two book chapters, 12 invited conference keynotes, and 147 conference proceedings. Due to his research contribution, he has received twice the French Prize of Scientific Excellence (Prime d'Excellence Scientifique), in 2010 and 2014, respectively. He is also an Editor of the *Journal of Medical Engineering* and invited reviewer of more than 20 international journals.