



Original contribution

Applications of a deep learning method for anti-aliasing and super-resolution in MRI



Can Zhao^{a,*}, Muhan Shao^a, Aaron Carass^{a,b}, Hao Li^a, Blake E. Dewey^{a,e}, Lotta M. Ellingsen^{a,h}, Jonghye Woo^{c,d}, Michael A. Guttman^g, Ari M. Blitz^g, Maureen Stone^f, Peter A. Calabresi^g, Henry Halperin^g, Jerry L. Prince^{a,g}

^a Department of Electrical and Computer Engineering, Johns Hopkins University, Baltimore, MD, USA

^b Department of Computer Science, Johns Hopkins University, Baltimore, MD, USA

^c Department of Radiology, Harvard Medical School, Boston, MA, USA

^d Massachusetts General Hospital, Boston, MA, USA

^e Kirby Center for Functional Brain Imaging, Kennedy Krieger Institute, Baltimore, MD, USA

^f Department of Neural and Pain Sciences, University of Maryland, Baltimore, MD, USA

^g Johns Hopkins University School of Medicine, Baltimore, MD, USA

^h Department of Electrical and Computer Engineering, University of Iceland, Reykjavik, Iceland

ARTICLE INFO

Keywords:

Deep learning
MRI
Super-resolution
Aliasing
Segmentation
Reconstruction
SMORE

ABSTRACT

Magnetic resonance (MR) images with both high resolutions and high signal-to-noise ratios (SNRs) are desired in many clinical and research applications. However, acquiring such images takes a long time, which is both costly and susceptible to motion artifacts. Acquiring MR images with good in-plane resolution and poor through-plane resolution is a common strategy that saves imaging time, preserves SNR, and provides one viewpoint with good resolution in two directions. Unfortunately, this strategy also creates orthogonal viewpoints that have poor resolution in one direction and, for 2D MR acquisition protocols, also creates aliasing artifacts. A deep learning approach called SMORE that carries out both anti-aliasing and super-resolution on these types of acquisitions using no external atlas or exemplars has been previously reported but not extensively validated. This paper reviews the SMORE algorithm and then demonstrates its performance in four applications with the goal to demonstrate its potential for use in both research and clinical scenarios. It is first shown to improve the visualization of brain white matter lesions in FLAIR images acquired from multiple sclerosis patients. Then it is shown to improve the visualization of scarring in cardiac left ventricular remodeling after myocardial infarction. Third, its performance on multi-view images of the tongue is demonstrated and finally it is shown to improve performance in parcellation of the brain ventricular system. Both visual and selected quantitative metrics of resolution enhancement are demonstrated.

1. Introduction

In many clinical and research applications, MR images with high resolution and high signal-to-noise ratio (SNR) are desired. However, acquiring such MR images takes a long time, which is costly, lowers patient throughput, and increases both patient discomfort and motion artifacts. A common compromise in practice is to acquire MRI with good in-plane resolution and poor through-plane resolution. With a reasonable acquisition time, the resulting elongated voxels have good SNR and one viewpoint (the in-plane view) with good resolution in two directions that yields acceptable diagnostic quality. Though they are acceptable in diagnosis, these elongated voxels are not ideal for

automatic image processing software, which is generally focused on 3D analysis and needs isotropic resolution. A common first step in automatic analysis is to interpolate (nearest neighbor, linear, b-spline, zero-padding in k -space, etc.) the data into isotropic voxels. Unfortunately, since interpolation does not restore the high frequency information that was not acquired in the actual scan itself, the interpolated images have blurry edges in the through-plane direction. Also, for 2D MRI acquisitions the low sampling rate in the through-plane direction causes aliasing artifacts which appear as unnatural high-frequency textures that cannot be removed through interpolation.

To address the problem of low through-plane resolution, researchers have developed a number of super-resolution (SR) techniques,

* Corresponding author.

E-mail address: czhao20@jhu.edu (C. Zhao).

<https://doi.org/10.1016/j.mri.2019.05.038>

Received 29 December 2018; Received in revised form 25 May 2019; Accepted 26 May 2019

0730-725X/ © 2019 Elsevier Inc. All rights reserved.

including multi-image SR and single-image SR methods, that attempt to restore the missing and attenuated high frequency information. Multi-image SR methods reconstruct a high resolution (HR) image from the acquisition of multiple low-resolution (LR) images each with degraded resolution in a different direction [1,2], such that the methods can combine high-frequency information from multiple orientations. Although multi-image SR methods [1,2] can be effective, there are two major disadvantages. First, it takes extra time to acquire the multiple images, which is costly and inconvenient. Second, the images must be registered together, an operation that is prone to errors because of differences in both image resolution and geometric distortions (both caused by differences in the orientations of the acquired images). Single-image SR, on the other hand, only involves one acquired LR image, which solves the problem of acquisition time and focuses our attention on the SR algorithms. The most successful algorithms for single-image SR in MRI have been learning-based including sparse coding [3], random forests [4], and convolutional neural networks (CNNs) [5–7]. The basic strategy of these methods is to learn a mapping from LR atlas images (also called training images or exemplars) to HR atlas images and apply the learned mapping to the acquired LR subject images (also called testing images). Methods like sparse coding and random forests need a manually designed feature as the input for the mapping. On the other hand, CNNs take either images or patches directly as input, and learn the features automatically during training. Some CNNs are trained with aliased LR images [7,8]. These networks do anti-aliasing and super-resolution at the same time.

Many of these learning-based methods provide good results, especially the state-of-the-art methods CNNs [9,10]. In NTIRE CVPR 2017 [11,12] and PIRM ECCV 2018 super-resolution challenge [13], a variety of SR methods were evaluated. Among them, generative adversarial network (GAN)-based models provide results with better visual quality, while a CNN model called Enhanced Deep Residual Network (EDSR) [9] achieves the best accuracy, which is more important than visual quality for medical imaging. Although these SR methods provide far better results than interpolation, they have not been universally adopted in MRI for two major reasons. First, training data requires paired LR and HR MR images which can be difficult to obtain, primarily because high-SNR HR MR images take a long time to acquire and may suffer from motion artifacts. Second, to avoid overfitting in deep networks, the tissue contrast of external training images must closely match with the subject data, but this is difficult because MRI has no standardized intensity scale. Because of these two reasons, it is highly desirable for SR in MR to avoid external atlases.

SR methods that do not require external atlases have been developed in the past. Methods include total variation methods [14,15], non-local means [16,17], brain hallucination [18], and self super-resolution methods (SSR) [19–22]. These methods all have features that make them non-optimal. For example, total variation SR methods are essentially image enhancement methods designed to strengthen edges and suppress noise, where the parameters need to be carefully tuned to decide how much high frequency information to be restored. Non-local means SR assumes that small patches repeat themselves at different resolutions within the same image, which may not be true in medical images. Brain hallucination assumes that the LR/HR properties in a T1-weighted image are the same as those in a T2-weighted image, which may not be true. Self super-resolution methods, with the exception of [22], assume that images obey a spatial self-similarity wherein the LR/HR relationship at a coarse scale applies to that at a finer scale. The SR method presented in Jog et al. [22] is the first SR method to improve resolution in the through-plane direction using the higher resolution data that is already present in the in-plane directions. Our method is motivated by and improves upon this prior work, which we refer to as JogSSR in this paper.

We previously developed an SSR algorithm called Synthetic Multi-Orientation Resolution Enhancement (SMORE) [8,23]. SMORE does not use external training data, there are no parameters to tune, there is

no intensity smoothing or regularization, and the only pre-processing that is required is N4 intensity inhomogeneity correction [24]. SMORE takes advantage of the fact that the in-plane slices have high resolution data and can be used to extract paired LR/HR training data by down-sampling. For 3D MRI, SMORE uses the state-of-the-art network EDSR as the underlying SSR CNN [23]. For 2D MRI, SMORE uses a self anti-aliasing (SAA) deep CNN that precedes the SSR CNN, also using EDSR as the underlying CNN model [8]. We have already demonstrated that SMORE can give results with better SR accuracy than competing methods, including standard interpolation and JogSSR. Also Dewey et al. [25] have demonstrated that SMORE can be used to improve the performance of image synthesis and white matter lesion segmentation.

In this paper, we use four applications to demonstrate the potential of SMORE in both research and clinical scenarios. The first application is on T2 FLAIR MR brain images acquired from multiple sclerosis (MS) patients. MS is an auto-immune disease in which myelin, the protective coating of nerves, is damaged and can be visualized as hyperintense lesions in FLAIR images. We show that visualization of MS lesions using SMORE is better than that obtained using cubic b-spline interpolation and JogSSR. The second application is on cardiac MRI where we explore the visualization of myocardial scarring from cardiac left ventricular remodeling after myocardial infarction. Characterizing such scarring is important factor in assessing the long-term clinical outcome after myocardial infarction [26] and it is challenging due to the competing requirements of high-resolution imaging and rapid scanning due to cardiac motion and breathing. We show improved visualization of such scars when using SMORE.

The third application of SMORE is on multi-orientation MR images of the tongue in tongue tumor patients. Because of the involuntary requirement to swallow during lengthy MR scans, acquisition times are very limited—< 3 min—in tongue imaging. A previous approach to obtaining super-resolution in the tongue used a computational combination of axial, sagittal, and coronal image stacks, each obtained in a separate stationary phase and registered together [2]. We demonstrate how the use of SMORE on a single acquisition is comparable to the result of combining three acquisitions. The fourth application of SMORE is on brain ventricle labeling in subjects with normal pressure hydrocephalus (NPH). NPH is a brain disorder usually caused by disruption of the cerebrospinal fluid (CSF) flow, leading to ventricle expansion and brain distortion. Having accurate parcellation of the ventricular system into its sub-compartments could potentially help in diagnosis and surgical planning in NPH patients [27]. Both visual and selected quantitative metrics of resolution enhancement are demonstrated.

In this paper, we make three important contributions about the implementation and utility of SMORE. First, we give a complete explanation of the method (previously described more briefly in conference publications [8,23]) in Section 2.1. Second, in order to show the versatility of SMORE, we present results on four MRI datasets from different pulse sequences and different organs, with three of them being real acquired LR MR datasets. Finally, we demonstrate that the proposed SR algorithm yields improvements not only in apparent image quality but, in the fourth experiment, show quantitative improvements when SMORE is applied as a preprocessing step for a segmentation task.

2. Material and methods

2.1. Overview of SMORE

Fig. 1 shows the workflow of SMORE for MRI acquired with 3D and 2D protocols assuming that axial slices are in-plane slices. For 3D MRI, SMORE(3D) simulates LR axial slices from HR axial slices by applying a filter $h(x)$ consisting of a rect filter as well as an anti-ringing filter in k -space that yields the through-plane resolution, and then uses the paired LR/HR data to train a self super-resolution (SSR) network. This trained network is applied on LR coronal slices to improve through-plane

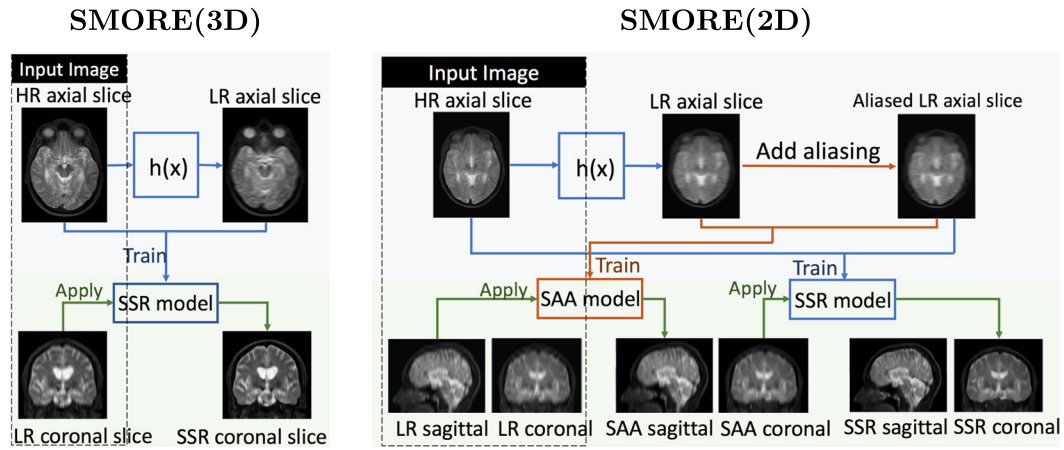


Fig. 1. Overview of SMORE. Workflow of SMORE for MRI acquired with 3D protocols and 2D protocols, referred as SMORE(3D) and SMORE(2D). They are simplified version of algorithms described in our previous conference papers [8,23].

resolution, resulting in a HR volume. Details can be found in [23], with a modification that the anti-ringing filter is changed to a Fermi filter to better mimic the behavior in scanners. SMORE(2D) uses the same general concept as SMORE(3D), but adds a self anti-aliasing (SAA) network trained with aliased axial slices. The aliased slices are created by first applying the filter $h(x)$, which in this case mimics the through-plane slice profile, and then a downsampling/upsampling sequence that produces aliasing at the same level as that found in the through-plane direction. We first apply the trained SAA network on sagittal slices to remove aliasing in the sagittal plane. We then apply the trained SSR network on the coronal plane to both remove aliasing in the coronal plane and improve through-plane resolution, resulting in an anti-aliased HR volume. Details can be found in [8]. For both SMORE(3D) and SMORE(2D), we only apply the trained networks in one orientation instead of two (or more) as described in our previous conference papers [8,23]. This reduces computation time from 20 min to 15 min for SMORE(3D), and from 35 min to 25 min for SMORE(2D) on a Tesla K40 GPU, with only a minor impact on performance. Also we omit SAA network and directly apply SSR network to the LR image to further reduce time cost from 25 min to 15 min for SMORE(2D) if the ratio r between through-plane and in-plane resolution is < 3 , since the aliasing is empirically not severe in this case.

The SAA and SSR neural networks currently used in SMORE are both implemented using the state-of-the-art super-resolution EDSR network [9]. In this paper, we implement patch-wise training with randomly extracted 32×32 patches. Training on small patches restricts the effect receptive field [28] to avoid structural specificity so that this network can better preserve pathology. It also reduces spatial correlation of the training data, which can accelerate convergence in theory [29]. To reduce training time, the networks are fine-tuned from pre-trained models that were trained from arbitrary data. When applying the trained networks, we apply them to entire coronal or sagittal slices (depending on whether it is SAA or SSR) rather than just 32×32 patches. This is possible since EDSR is a fully convolutional network (FCN) which allows an arbitrary input size [30].

2.2. Application to visual enhancement for MS lesions

In this experiment, we test whether super-resolved T2 FLAIR MR images can give better visualization of white matter lesions in the brain than the acquired images. The T2 Flair MR images were acquired from multiple sclerosis (MS) subjects using a Philips Achieva 3 T scanner with a 2D protocol and the following parameters: $0.828 \times 0.828 \times 4.4$ mm, TE = 68 ms, TR = 11 s, TI = 2.8 s, flip angle = 90° , turbo factor = 17, acquisition time = 2 m56 s. We performed cubic b-spline interpolation, JogSSR [22], and SMORE(2D) on

the data using a $0.828 \times 0.828 \times 0.828$ mm digital grid. We show a visual comparison on the regions of white matter lesions in axial, sagittal, and coronal slices for the three methods. We also plot 1D intensity profiles of the three methods across selected paths through different lesions.

2.3. Application to visual enhancement of scarring in cardiac left ventricular remodeling

In this experiment, we test whether super-resolved images can give better visualization of the scarring caused by left ventricular remodeling after myocardial infarction than the acquired images. We acquired two T1-weighted MR images from an infarcted pig, each with a different through-plane resolution. One image, which serves as the HR reference image, was acquired with resolution equal to $1.1 \times 1.1 \times 2.2$ mm, and then it was sinc interpolated on the scanner (by zero padding in k -space) to $1.1 \times 1.1 \times 1.1$ m. The other image was acquired with resolution equal to $1.1 \times 1.1 \times 5$ mm. Both of these images were acquired with a 3D protocol, inversion time = 300 ms, flip angle = 25° , TR = 5.4 ms, TE = 2.5 ms, and GRAPPA acceleration factor $R = 2$. The HR reference image has a segmented centric phase-encoding order with 12 k -space segments per imaging window (heart beat), while the LR subject image has 16 k -space segments.

In our experiment, we performed sinc interpolation, JogSSR, and SMORE(3D) on the $1.1 \times 1.1 \times 5.0$ mm data using a $1.1 \times 1.1 \times 1.1$ mm digital grid. These images were then rigidly registered to the reference image for comparison. We are interested in the regions of thinning layer of midwall scar between the endocardial and epicardial layers of normal myocardium and the thin layer of normal myocardium between the scar and epicardial fat. These two regions of interest are cropped and zoomed to show the details.

2.4. Application to multi-view reconstruction

In this experiment, we test whether a super-resolved image from a single acquisition can give a comparable result to a multi-view super-resolution image reconstructed from three acquisitions. MR images of the tongue were acquired from normal speakers and subjects who had tongue cancer surgically resected (glossectomy). Scans were performed on a Siemens 3.0 T Tim Treo system using an eight-channel head and neck coil. A T2-weighted Turbo Spin Echo sequence with an echo train length of 12, TE = 62 ms, and TR = 2500 ms was used. The field-of-view (FOV) was 240×240 mm with a resolution of 256×256 . Each dataset contained a sagittal, coronal, and axial stack of images containing the tongue and surrounding structures. The image size for the high-resolution MRI was $256 \times 256 \times z$ (z ranges from 10 to 24) with

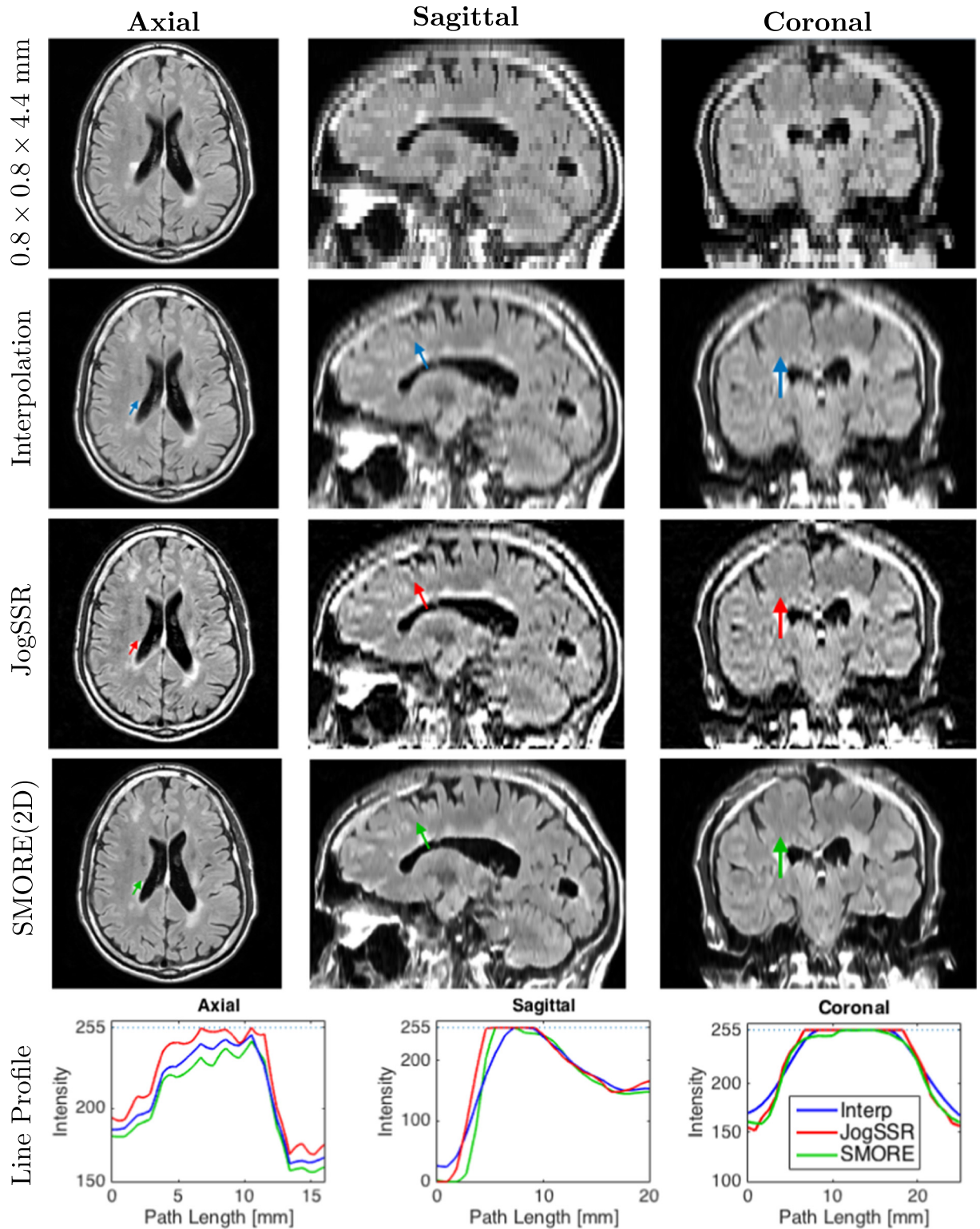


Fig. 2. T2 Flair MRI from an MS subject: Axial, sagittal, and coronal views of the acquired $0.828 \times 0.828 \times 4.4$ mm image, and the reconstructed volumes with $0.828 \times 0.828 \times 0.828$ mm digital grid through cubic b-spline interpolation, JogSSR, and SMORE(2D). In each view, we pick a path across lesions, shown as colored arrows in the images, and plot the line profiles of the three methods in the same plot on the bottom of each view.

0.9375×0.9375 mm in-plane resolution and 3 mm slice thickness. The datasets were acquired at a rest position and the subjects were required to remain still for 1.5–3 min for each orientation. For each subject, the three axial, sagittal, and coronal acquisitions were interpolated onto a $0.9375 \times 0.9375 \times 0.9375$ mm digital grid and N4 corrected [24].

We applied both JogSSR and SMORE(3D) on single acquisitions to

compare to the multi-view super-resolution reconstruction. The multi-view reconstruction algorithm we used for comparison is an improved version of the algorithm described in Woo et al. [2]. This approach takes three interpolated image volumes, aligns them using ANTs affine registration [31] and SyN deformable registration [32], and then uses a Markov random field image restoration algorithm (with edge

enhancement) to reconstruct a single HR volume.

2.5. Application to brain ventricle parcellation

This experiment demonstrates the effect of super-resolution on brain ventricle parcellation and labeling using the Ventricle Parcellation Network (VParNet) described in Shao et al. [33]. In particular, we test whether super-resolved images can give better VParNet results than images from either interpolation or JogSSR. The data for this experiment are from an NPH data set containing 95 T1-w MPRAGE MRIs (age range: 26–90 years with mean age of 44.54 years). They were acquired on a 3 T Siemens scanner with scanner parameters: TR = 2110 ms, TE = 3.24 ms, FA = 8°, TI = 1100 ms, and voxel size of $0.859 \times 0.859 \times 0.9$ mm. There are also 15 healthy controls from the Open Access Series on Imaging Studies (OASIS) dataset involved in this experiment. All the MRIs were interpolated to a $0.8 \times 0.8 \times 0.8$ mm digital grid, and then pre-processed using N4-bias correction [24], rigid registration to MNI 152 atlas space [34], and skull-stripping [35].

VParNet was trained to parcellate the ventricular system of the human brain into its four cavities: the left and right lateral ventricles (LLV and RLv), and the third and the fourth ventricles. It was trained on 25 NPH subjects and 15 healthy controls (not involved in the evaluations). In the original experiment of Shao et al. [33], the remaining 70 NPH subjects were used for testing. In this experiment, we downsampled the 70 NPH subject images first so that we could study the impact of super-resolution. In order to remove the impact of pre-processing, we downsampled the 70 pre-processed test datasets instead of the raw datasets. In particular, we downsampled the data to a resolution of $0.8 \times 0.8 \times 0.8r$ mm following a 2D acquisition protocol, where r is the through-plane to in-plane resolution ratio. The slice number of the HR images happens to be a prime number. Since the downsampled images must have integer slice numbers, the downsample ratio r which is also the ratio between the slice number of HR images and downsampled images must be non-integers. In the experiment, we choose ratio r of 1.50625, 2.41, 3.765625, 4.82, 6.025. The downsampled images have voxel length ($0.8r$ mm) in z -axis of 1.205 mm, 1.928 mm, 3.0125 mm, 3.856 mm, 4.82 mm. To apply VParNet, which was trained on $0.8 \times 0.8 \times 0.8$ mm images, to these downsampled images, we used cubic b-spline interpolation, JogSSR, and SMORE(2D) to produce images on a $0.8 \times 0.8 \times 0.8$ mm digital grid. These images were then used in the same trained VParNet to yield ventricular parcellation results.

The HR NPH images have physical resolution in z -axis of 0.9 mm. We used them as ground truth and evaluated the accuracy of super-resolved images using the Structural Similarity Index (SSIM) and the Peak Signal to Noise Ratio (PSNR) within brain masks. As for the ventricle parcellation performance, we evaluated the automated parcellation results using manual delineations. We computed Dice coefficients [36] to evaluate the parcellation accuracy of the same network on different super-resolved and interpolated images. By comparing the parcellation accuracy, we can evaluate how much improvement we get from SMORE(2D) compared with interpolation.

3. Results

3.1. Application to visual enhancement for brain white matter lesions

Fig. 2 shows an example of T2 FLAIR images reconstructed from the acquired resolution of $0.828 \times 0.828 \times 4.4$ mm input image onto a $0.828 \times 0.828 \times 0.828$ mm digital grid using cubic-b-spline interpolation, JogSSR, and SMORE(2D). On these images, MS lesions appear as bright regions in the brain's white matter. We can see that both JogSSR and SMORE(2D) give sharper edges than interpolation and the SMORE(2D) result looks more realistic than JogSSR. This is in part because JogSSR does not carry out anti-aliasing, which allows aliasing artifacts, which are seen in the original and interpolated images, to remain. We

note that in Fig. 2, SMORE also enhances resolution in the axial slice slightly, which is originally 0.828×0.828 mm HR. Although we apply super-resolution in the through-plane, structures like edges that pass through-plane slices obliquely also gets enhanced, permitting in-plane edges to also be enhanced.

Aside from the visual impression of performance differences gleaned from looking at the images directly, we also examined selected intensity profiles within the images. Each reconstructed image in Fig. 2 contains a small colored arrow. These arrows depict the line segment and direction over which intensity profiles shown in the bottom row of the figure are extracted. For example, the three colored arrows in the axial images of the first column yield the profiles on the bottom right graph. These axial profiles show that other than some differences in overall intensity, the resolutions of the methods appear to be very similar. This is to be expected since the axial image already has good resolution. The profiles through the ventricle and lesion in the sagittal orientation, however, show significant differences. Both super-resolution approaches show a steeper edge than the interpolated image (although the JogSSR result is inexplicably shifted relative to the true position of the edge). The profiles from the lesion in the coronal images show a similar property—steeper edges from the super-resolution approaches. Overall, the selected intensity profiles suggest resolution enhancement from both SMORE(2D) and JogSSR.

3.2. Application to visual enhancement for scarring in cardiac left ventricular remodeling

Cardiac imaging data that were acquired with a resolution of $1.1 \times 1.1 \times 5.0$ mm are shown in Fig. 3 after application of interpolation (zero-padding in k -space), JogSSR, and SMORE(3D) applied on a $1.1 \times 1.1 \times 1.1$ mm digital grid. A thin layer of the midwall scar between the endocardium and epicardium of normal myocardium appears as bright strip in the magenta boxes. A thin layer of normal myocardium between scar and epicardial fat appears as dark strip in the cyan boxes. They are zoomed in to show the details below the short-axis (SAX) slices with acquired resolution of 1.1×1.1 mm and long-axis (LAX) slices with originally acquired resolution of 1.1×5 mm for the first three columns, or 1.1×2.2 mm for the column of “HR ref.”. Each zoomed box contains a colored arrow which depicts a line segment. The corresponding line profiles are shown on the bottom.

As seen in the long-axis (LAX) images and zoomed regions, the borders between normal myocardium, enhanced scar and blood are significantly clearer in SMORE(3D) compared with JogSSR and interpolation. The intensity profile of SMORE(3D), the green line shown in the magenta box marked “LAX (1)”, very closely matches that of the HR reference image. For the short-axis (SAX(1) and SAX(2)), the resolution was already high and there is less to be gained. Nevertheless, it is apparent that the image clarity is slightly improved by SMORE(3D) while faithfully representing the patterns from the input images.

Furthermore, we computed the SSIM and PSNR between each method and HR reference image. The SSIM for interpolation, JogSSR, and SMORE results are 0.5070, 0.4770, and 0.5146, correspondingly. The PSNR for interpolation, JogSSR, and SMORE are 25.8816, 24.4142, and 25.3002, correspondingly. SMORE gives best SSIM, yet worse PSNR than sinc interpolation. Note that the registration cannot be perfect among different sets of cardiac images, due to motion or changing physiological state. When computing SSIM, images are prefiltered. Therefore, SSIM is less sensitive to image distortion. On the other hand, PSNR is a measure of noise level. SMORE does not consider noise reduction. This might be the reason why the SMORE result has better SSIM but worse PSNR. Also this evaluation is done on only one pair of LR/HR data, and is not very informative statistically.

3.3. Application to multi-frame reconstruction

Tongue data with 0.9375×0.9375 mm in-plane resolution and

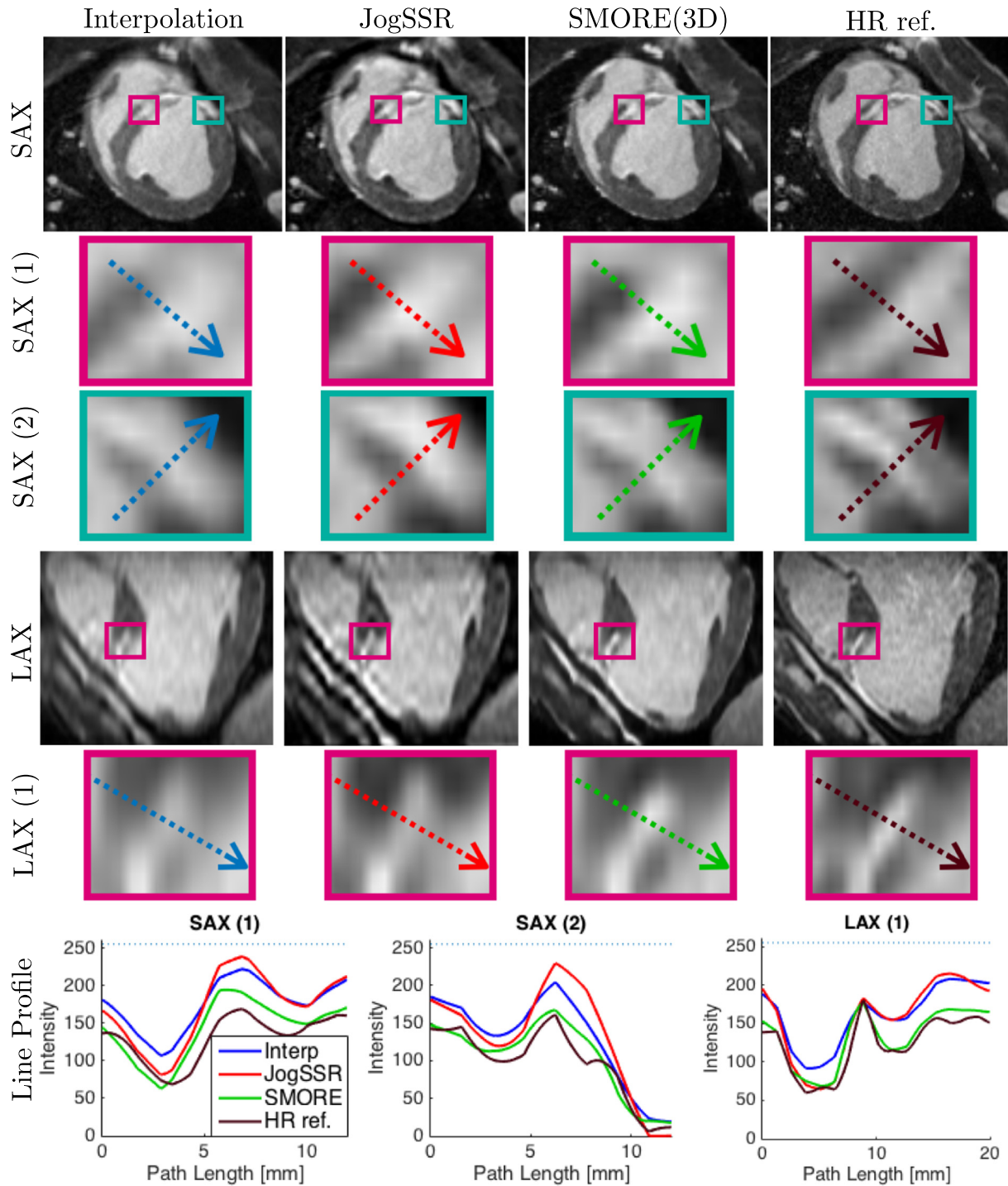


Fig. 3. Late gadolinium enhancement (LGE) from an infarct swine subject: Short-axis (SAX) and long-axis (LAX) views arranged in columns using $1.1 \times 1.1 \times 1.1$ mm digital grid: output of 1) sinc-interpolation, 2) JogSSR and 3) SMORE(3D) for the subject LR image acquired at $1.1 \times 1.1 \times 5$ mm; 4) sinc-interpolated HR reference image for comparison acquired at $1.1 \times 1.1 \times 2.2$ mm. SAX(1) and LAX(1) boxes contain a thinning layer of enhanced midwall scar between *endo-* and *epi* layers of normal myocardium (hypo-intense). SAX(2) boxes contain a thin layer of normal myocardium (hypo-intense) between scar and epicardial fat (both hyper-intense). In each box, we pick a path across the region of interest, shown as colored arrows, and plot the profiles in the last row. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3 mm through-plane resolution were acquired with in-plane view of axial, sagittal, and coronal. They are shown in Fig. 4 on a $0.9375 \times 0.9375 \times 0.9375$ mm digital grid after both cubic b-spline interpolation (blue boxes) and JogSSR (red boxes), and SMORE(2D) (green boxes). The in-plane views are only shown for interpolation since they are already HR slices. In through-plane views, SMORE(2D) always gives visually better results than both interpolation

and JogSSR. In particular, we can see the edges are sharper in SMORE (2D) and no artificial structures are created.

A comparison of interpolation and SMORE(2D) (where each used only the coronal image) and the multi-view reconstruction (which used all three images) is shown in Fig. 5. The arrows point out at the bright pathology region—i.e., scar tissue formed after removing a tumor. We can see that SMORE has visually better resolution than the interpolated

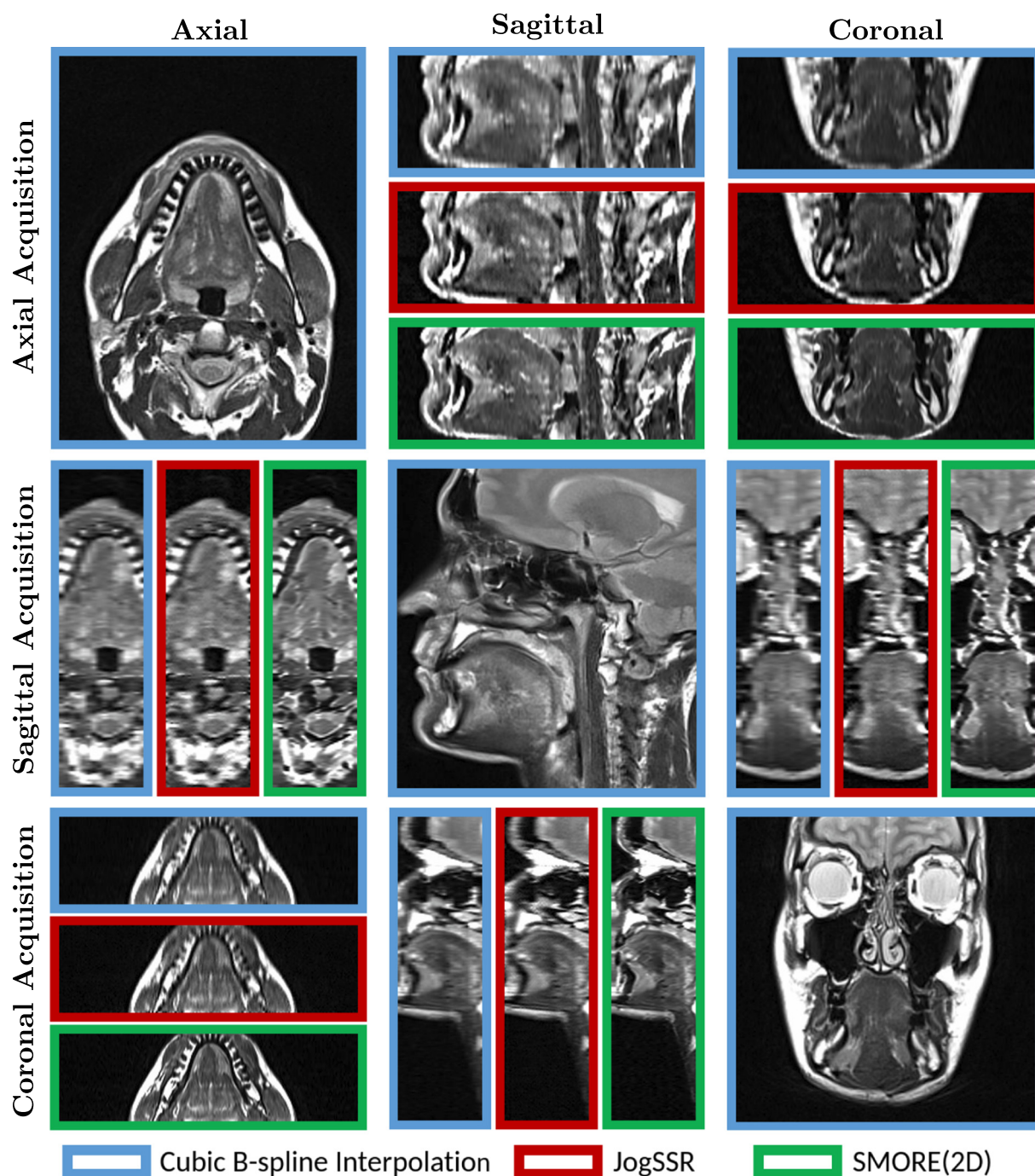


Fig. 4. T2w MRI from a tongue tumor subject: Axial, Sagittal, and Coronal views of the three acquisitions in Axial, Sagittal, and Coronal planes (not registered). We show the through-plane views of the resolved volumes with isotropic digital resolution that result from cubic b-spline interpolation (blue boxes), JogSSR (red boxes), and our SMORE(2D) (green boxes). The in-plane views are only shown with interpolation results since they are already HR slices. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

image, but several places within the multi-view reconstruction have visually better detail. On the other hand, the pathology region in the multi-view reconstruction appears to be somewhat degraded in appearance over both the SMORE(2D) and interpolation result. We believe that this loss of features may be caused by regional mis-registration between the three acquisitions.

3.4. Application to brain ventricle parcellation

Example images from an NPH subject, all reconstructed on a $0.8 \times 0.8 \times 0.8$ mm digital grid, are shown in Fig. 6. The LR image

depicted using cubic-bspline interpolation has resolution $0.8 \times 0.8 \times 3.856$ mm LR while the ground truth image has resolution $0.859 \times 0.859 \times 0.9$ mm. The JogSSR and SMORE(2D) results are also shown. To reveal more detail, the second row shows zoomed images of the 4th ventricle, where the zoomed region is shown using blue boxes in the first row. The VParNet [33] parcellations as well as the manually delineated label of the 4th ventricle are shown using purple voxels on the third row. Visually, of all the results derived from the LR data, SMORE(2D) gives the best super-resolution and parcellation results. In particular, the VParNet parcellation on the SMORE(2D) result is very close to the VParNet on the HR image.

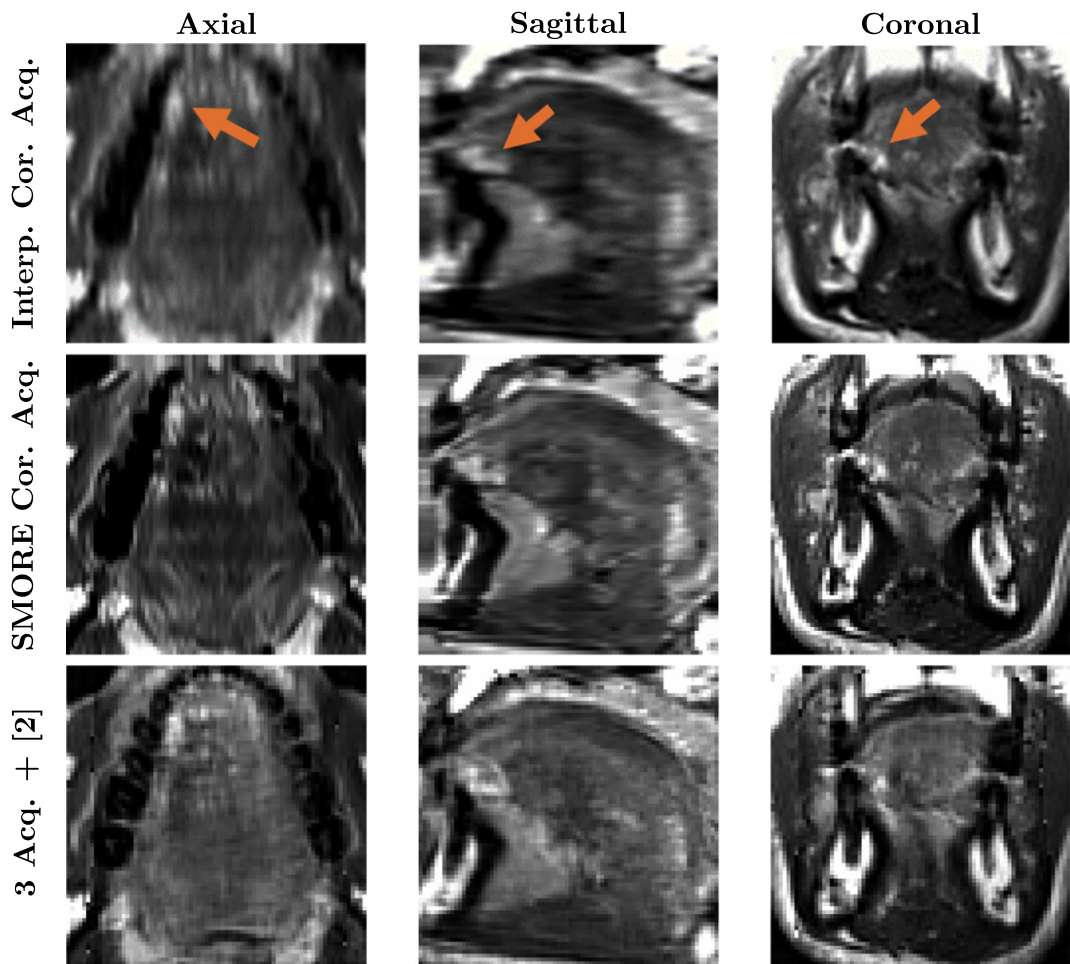


Fig. 5. Comparison between SMORE(2D) and multi-view reconstruction for a tongue tumor subject: Axial, Sagittal, and Coronal views of the tongue region in cubic b-spline interpolation and SMORE(2D) results for a single coronal acquisition, and multi-view reconstructed image [2] using three acquisitions. The arrows point out the bright looking scar tissue from a removed tumor.

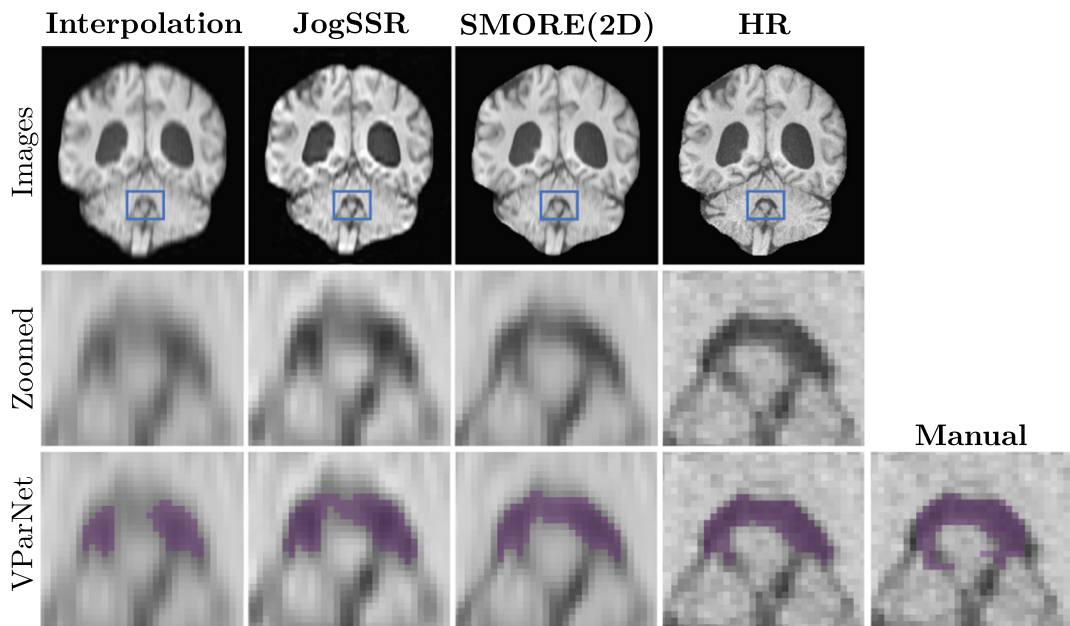


Fig. 6. Coronal views of brain ventricle parcellation on an NPH subject: The volumes with digital resolution of $0.8 \times 0.8 \times 0.8$ mm that resolved from $0.8 \times 0.8 \times 3.856$ mm LR image using cubic-b-spline interpolation, JogSSR, SMORE(2D), and the interpolated $0.8 \times 0.8 \times 0.9$ mm HR image. The patches in blue boxes are zoomed in the second row to show details of the 4th ventricle. The last row shows the VParNet [33] parcellation results and the manual labeling for the 4th ventricle. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 1

Evaluation of brain ventricle parcellation on 70 NPH subjects. We performed paired two-tail Wilcoxon signed-rank tests between interpolation, JogSSR, and SMORE. The ‘*’ ($p < 0.005$) and ‘†’ ($p < 0.05$) indicate the method is significantly better than the other two methods. For the column of ‘HR’, they indicate that HR images give significantly better parcellation results than resolved images from the three methods with thickness of 1.205 mm.

Metrics	Thickness	Interp.	JogSSR	SMORE	HR(0.9 mm)
SSIM	1.205 mm	0.9494	0.9507	0.9726*	1*
	1.928 mm	0.9013	0.9106	0.9389*	
	3.0125 mm	0.8290	0.8400	0.8893*	
	3.856 mm	0.7677	0.7812	0.8387*	
	4.82 mm	0.7003	0.7170	0.7817*	
PSNR	1.205 mm	35.0407	34.0472	39.5053*	–
	1.928 mm	31.9321	30.6444	35.7429*	
	3.0125 mm	29.2785	27.4384	31.9878*	
	3.856 mm	27.7562	25.7118	29.6050*	
	4.82 mm	26.4127	24.2377	28.1593*	
Dice(RLV)	1.205 mm	0.9704	0.9705	0.9712†	0.9715*
	1.928 mm	0.9678	0.9690	0.9706*	
	3.0125 mm	0.9610	0.9635	0.9693*	
	3.856 mm	0.9527	0.9578	0.9648*	
	4.82 mm	0.9405	0.9498	0.9629*	
Dice(LLV)	1.205 mm	0.9710	0.9709	0.9715†	0.9717†
	1.928 mm	0.9690	0.9693	0.9710*	
	3.0125 mm	0.9638	0.9641	0.9699*	
	3.856 mm	0.9571	0.9585	0.9663*	
	4.82 mm	0.9469	0.9510	0.9638*	
Dice(3rd)	1.205 mm	0.9149	0.9149	0.9163†	0.9174*
	1.928 mm	0.9095	0.9097	0.9141*	
	3.0125 mm	0.8945	0.8940	0.9073*	
	3.856 mm	0.8779	0.8761	0.8937*	
	4.82 mm	0.8560	0.8545	0.8832*	
Dice(4th)	1.205 mm	0.8954	0.8941	0.8973*	0.8983*
	1.928 mm	0.8891	0.8851	0.8947*	
	3.0125 mm	0.8741	0.8657	0.8878*	
	3.856 mm	0.8550	0.8463	0.8753*	
	4.82 mm	0.8254	0.8216	0.8629*	
Dice(whole)	1.205 mm	0.9690	0.9690	0.9696†	0.9699*
	1.928 mm	0.9665	0.9672	0.9690*	
	3.0125 mm	0.9602	0.9614	0.9675*	
	3.856 mm	0.9524	0.9552	0.9632*	
	4.82 mm	0.9408	0.9470	0.9607*	

We also evaluated these results quantitatively. The mean values of SSIM and PSNR are shown in Table 1. Paired two-tail Wilcoxon signed-rank tests [37] were performed between results from interpolation, JogSSR, and SMORE(2D). The ‘*’ ($p < 0.005$) and ‘†’ ($p < 0.05$) indicate the method is significantly better than the other two methods. We can see that the SR results from SMORE(2D) are always significantly more accurate than interpolation and JogSSR, with SSIM and PSNR as accuracy metrics. For images with thickness of 1.205 mm, the significance is weak, indicating that the improvement is not dramatic. For images with thicker slices, the significance is strong, and therefore the improvement is large. The Dice coefficient of the parcellation results of the four cavities (RLV, LLV, 3rd, 4th) and the whole ventricular system are also shown in Table 1. From the table, we can find that for example, VParNet on SMORE(2D) results of thickness 4.82 mm is better than interpolation results of 3.856 mm, while the later needs 56.25% longer scanning time. It shows the potential of reducing scanning time by using SMORE.

We wondered whether the parcellation results from HR images are significantly better than SMORE(2D) from images with thickness 1.205 mm. In the column of ‘HR’ of Table 1, we show results of paired two-tail Wilcoxon signed-rank tests between HR and the results from the three methods. The ‘*’ and ‘†’ indicate that HR images give significantly better evaluation values than SMORE(2D) for thickness of 1.205 mm. The strong significance always holds, except for the Dice of LLV for which only weak significance holds. It shows that acquiring HR images with adequate SNR gives better parcellation results than LR images,

even with SMORE(2D) applied to improve spatial resolution. However, if the acquired data are already limited to be anisotropic LR, which is common in clinical and research, SMORE(2D) can give better parcellation than interpolation.

4. Discussion and conclusions

This paper demonstrates the applications of a self super-resolution (SSR) algorithm, SMORE, on four different MRI datasets, and shows the improved MRI resolution both visually and quantitatively. The methodology of SMORE has been introduced in our previous conference papers [8,23]. While in this paper, we make important contributions about the implementation and utility of SMORE, which are useful and necessary to appreciate that SMORE can be reliably and widely used in practice. First, this paper provides a complete explanation of SMORE, while previous conference papers have been incomplete due to their required brevity. Second, this paper shows the application of SMORE on MR images produced from different pulse sequences and contrasts, in different organs. To our best knowledge, no other published deep-learning SR method has been demonstrated for improvement of such diverse MRI data sets without training data. Finally, we have demonstrated how SMORE can improve segmentation accuracy, a result which shows there are quantifiable improvements from using SMORE, in addition to the visual improvement in image quality.

In this paper, we demonstrate the applications of SMORE in real world scenarios for MR images. First of all, we consider an important distinction between general SR on natural images and SR on real acquired MR images. Although the general SR problem has been discussed a lot in computer vision application, the common SR problem setting requires well-established LR/HR paired external training data. In contrast to natural images, such external training data is much more difficult to obtain for MR images. In this paper, we develop SMORE to be a SSR algorithm, which requires no external training data; in other words, what it needs is only the input subject image itself. This makes SMORE more applicable in real world scenarios. Second, SMORE is developed for a common type of MRI acquisition which has high in-plane resolution but low through-plane resolution (thick slices). This type of MRI is widely acquired in clinical and research applications. Third, the four experiments in this paper are performed on four different MRI datasets, with three out of them being real acquired LR datasets. From a visual comparison, we find that SMORE enhances edges but does not create structures out of nothing; this reduces the risk of wrongly altering anatomical structures. Finally, the experiment in Section 3.4 shows that SMORE is not only visually appealing, it also gives quantitative improvements on SSIM and PSNR. More importantly, applying SMORE as a preprocessing step significantly improves ventricular segmentation accuracy on this brain MRI dataset. Furthermore, we note that sometimes lower resolution images processed with SMORE yield better segmentation results than those from higher resolution images processed with interpolation. This suggests that SMORE post-processing may allow shorter scan times.

The SMORE method has limitations. First, we use EDSR as the SR deep network in SMORE, since it was evaluated as one of the state-of-the-art SR architecture by extensive comparisons [11,13]. We note that these evaluations were performed on natural images, not MR images, and it is possible that there might be another SR network that performs better on MR images. Fortunately, EDSR can be easily replaced under the framework of SMORE if a better SR network becomes available. Second, the best resolution that can be achieved by SMORE is limited to the in-plane resolution. Because of this, for example, there is no way to use SMORE to enhance images that have been acquired with isotropic resolution. Also, SMORE does not consider cases where resolution differs in three orientations. Third, SMORE does not address motion artifacts. Finally, both SMORE(2D) and SMORE(3D) require knowledge of $h(x)$, the point spread function (or slice profile), which may not be known accurately in some cases. Future work should address these

issues.

In conclusion, SMORE produces results that are not only visually appealing, but also more accurate than interpolation. More importantly, applying it as a preprocessing step can improve segmentation accuracy. All these SMORE results were obtained without collecting any external training data. This makes SMORE a useful preprocessing step in many MRI analysis tasks.

Acknowledgements

This research was supported in part by NIH grant R01NS082347 and by the National Multiple Sclerosis Society grant RG-1601-07180. This research project was conducted using computational resources at the Maryland Advanced Research Computing Center (MARCC). Thanks to Dan Zhu for valuable discussion on MRI acquisition models.

References

- [1] Plenge E, Poot DH, Bernsen M, Kotek G, Houston G, Wielopolski P, et al. Super-resolution reconstruction in MRI: better images faster? Medical imaging 2012: image processing. Vol. 8314. International Society for Optics and Photonics; 2012. p. 83143V.
- [2] Woo J, Murano EZ, Stone M, Prince JL. Reconstruction of high-resolution tongue volumes from MRI. IEEE Transactions on Biomedical Engineering 2012;59(12):3511–24.
- [3] Wang Y-H, Qiao J, Li J-B, Fu P, Chu S-C, Roddick JF. Sparse representation-based MRI super-resolution reconstruction. Measurement 2014;47:946–53.
- [4] Alexander DC, Zikic D, Zhang J, Zhang H, Criminisi A. Image quality transfer via random forest regression: applications in diffusion MRI. International conference on medical image computing and computer-assisted intervention. Springer; 2014. p. 225–32.
- [5] Chen Y, Xie Y, Zhou Z, Shi F, Christodoulou Ag, Li D. Brain MRI super resolution using 3D deep densely connected neural networks. 15th IEEE international symposium on, IEEE. vol. 2018. 2018.
- [6] Mahapatra D, Bozorgtabar B, Hewavitharanage S, Garnavi R. Image super resolution using generative adversarial networks and local saliency maps for retinal image analysis. International conference on medical image computing and computer-assisted intervention. Springer; 2017. p. 382–90.
- [7] C.-H. Pham, A. Ducournau, R. Fablet, F. Rousseau, Brain MRI super-resolution using deep 3D convolutional networks, in: 2017 IEEE 14th international symposium on biomedical imaging (ISBI 2017), IEEE, vol. 2017, pp. 197–200.
- [8] Zhao C, Carass A, Dewey BE, Woo J, Oh J, Calabresi PA, et al. A deep learning based anti-aliasing self super-resolution algorithm for MRI. International conference on medical image computing and computer-assisted intervention. Springer; 2018. p. 100–8.
- [9] Lim B, Son S, Kim H, Nah S, Lee KM. Enhanced deep residual networks for single image super-resolution. The IEEE conference on computer vision and pattern recognition (CVPR) workshops. vol. 1. 2017. p. 4.
- [10] Ledig C, Theis L, Huszar F, Caballero J, Cunningham A, Acosta A, et al. Photo-realistic single image super-resolution using a generative adversarial network. vol. 2. CVPR; 2017. p. 4.
- [11] Timofte R, Agustsson E, Van Gool L, Yang M-H, Zhang L, Lim B, et al. Ntire 2017 challenge on single image super-resolution: methods and results. Computer vision and pattern recognition workshops (CVPRW), 2017 IEEE conference on. IEEE; 2017. p. 1110–21.
- [12] Agustsson E, Timofte R. Ntire 2017 challenge on single image super-resolution: dataset and study. The IEEE conference on computer vision and pattern recognition (CVPR) workshops. vol. 3. 2017. p. 2.
- [13] Blau Y, Mechrez R, Timofte R, Michaeli T, Zelnik-Manor L. PIRM challenge on perceptual image super-resolution. 2018. [arXiv preprint arXiv:1809.07517].
- [14] Tourbier S, Bresson X, Hagmann P, Thiran J-P, Meuli R, Cuadra MB. An efficient total variation algorithm for super-resolution in fetal brain MRI with adaptive regularization. Neuro Image 2015;118:584–97.
- [15] Shi F, Cheng J, Wang L, Yap P-T, Shen D. LRTV: MR image super-resolution with low-rank and total variation regularizations. IEEE Trans Med Imaging 2015;34(12):2459–66.
- [16] Manjón JV, Coupé P, Buades A, Fonov V, Collins DL, Robles M. Non-local MRI upsampling. Med Image Anal 2010;14(6):784–92.
- [17] Zhang K, Gao X, Tao D, Li X. Single image super-resolution with non-local means and steering kernel regression. IEEE Trans Image Process 2012;21(11):4544–56.
- [18] Rousseau F. Alzheimers disease neuroimaging initiative, a non-local approach for image super-resolution using intermodality priors. Med Image Anal 2010;14(4):594–605.
- [19] Huang J-B, Singh A, Ahuja N. Single image super-resolution from transformed self-exemplars. Proceedings of the IEEE conference on computer vision and pattern recognition. 2015. p. 5197–206.
- [20] Yang G, Ye X, Slabaugh G, Keegan J, Mohiaddin R, Firmin D. Combined self-learning based single-image super-resolution and dual-tree complex wavelet transform denoising for medical images. Medical imaging 2016: Image processing. vol. 9784. International Society for Optics and Photonics; 2016. p. 97840L.
- [21] Weigert M, Royer L, Jug F, Myers G. Isotropic reconstruction of 3D fluorescence microscopy images using convolutional neural networks. International conference on medical image computing and computer-assisted intervention, springer. 2017. p. 126–34.
- [22] Jog A, Carass A, Prince JL. Self super-resolution for magnetic resonance images. International conference on medical image computing and computer-assisted intervention. Springer; 2016. p. 553–60.
- [23] C. Zhao, A. Carass, B. E. Dewey, J. L. Prince, Self super-resolution for magnetic resonance images using deep networks, in: Biomedical imaging (ISBI 2018), 2018 IEEE 15th international symposium on, IEEE, 2018, pp. 365–368.
- [24] Tustison NJ, Avants BB, Cook PA, Zheng Y, Egan A, Yushkevich PA, et al. N4ITK: improved N3 bias correction. IEEE Trans Med Imaging 2010;29(6):1310–20.
- [25] Dewey BE, Zhao C, Carass A, Oh J, Calabresi PA, van Zijl PC, et al. Deep harmonization of inconsistent MR data for consistent volume segmentation. International workshop on simulation and synthesis in medical imaging. Springer; 2018. p. 20–30.
- [26] Bruder O, Wagner A, Jensen CJ, Schneider S, Ong P, Kispert E-M, et al. Myocardial scar visualized by cardiovascular magnetic resonance imaging predicts major adverse events in patients with hypertrophic cardiomyopathy. J Am Coll Cardiol 2010;56(11):875–87.
- [27] Yamada S, Tsuchiya K, Bradley W, Law M, Winkler M, Borzage M, et al. Current and emerging MR imaging techniques for the diagnosis and management of CSF flow disorders: a review of phase-contrast and time-spatial labeling inversion pulse. Am J Neuroradiol 2015;36(4):623–30.
- [28] Luo W, Li Y, Urtasun R, Zemel R. Understanding the effective receptive field in deep convolutional neural networks. Advances in neural information processing systems. 2016. p. 4898–906.
- [29] LeCun YA, Bottou L, Orr GB, Müller K-R. Efficient backprop. Neural networks: Tricks of the trade. Springer; 2012. p. 9–48.
- [30] Pan T, Wang B, Ding G, Yong J-H. Fully convolutional neural networks with full-scale-features for semantic segmentation. AAAI. 2017. p. 4240–6.
- [31] Avants Brian B, Tustison Nick, Song Gang. Advanced normalization tools (ANTS). Insight J. 2009;2:1–35.
- [32] Avants BB, Epstein CL, Grossman M, Gee JC. Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. Med Image Anal 2008;12(1):26–41.
- [33] Shao M, Han S, Carass A, Li X, Blitz AM, Prince JL, et al. Shortcomings of ventricle segmentation using deep convolutional networks. Understanding and interpreting machine learning in medical image computing applications. Springer; 2018. p. 79–86.
- [34] Fonov VS, Evans AC, McKinstry RC, Almli C, Collins D. Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. Neuro Image 2009;47:S102.
- [35] Roy S, Butman JA, Pham DL. Alzheimers disease neuroimaging initiative, robust skull stripping using multiple MR image contrasts insensitive to pathology. Neuro Image 2017;146:132–47.
- [36] Dice LR. Measures of the amount of ecologic association between species. Ecology 1945;26(3):297–302.
- [37] Wilcoxon F. Individual comparisons by ranking methods. Biometrics 1945;1(6):80–3.