City Research Online

City St George's, University of London

Citation: Yang, G., Ye, X., Slabaugh, G. G., Keegan, J., Mohiaddin, R. & Firman, D. (2016). Super-Resolved Enhancement of a Single Image and Its Application in Cardiac MRI. Paper presented at the 7th International Conference, ICISP 2016, 30 May - 01 Jun 2016, Quebec, Canada.

This is the accepted version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

Permanent repository link: https://openaccess.city.ac.uk/id/eprint/15471/

Copyright and Reuse: Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to City Research Online policy.

City Research Online: http://openaccess.city.ac.uk/ publications@citystgeorges.ac.uk

Super-Resolved Enhancement of a Single Image and Its Application in Cardiac MRI

Guang Yang[†1,2], Xujiong Ye[3], Greg Slabaugh[4], Jennifer Keegan[1,2], Raad Mohiaddin[1,2], and David Firmin[1,2]

[1] Cardiovascular MR Unit, Royal Brompton Hospital, SW3 6NP, London, UK
[2] National Heart & Lung Institute, Imperial College London, SW7 2AZ, London, UK
[3] School of Computer Science, University of Lincoln, LN6 7TS, Lincoln, UK
[4] Dept. of Computer Science, City University London, EC1V 0HB, London, UK

Abstract. Super-resolved image enhancement is of great importance in medical imaging. Conventional methods often require multiple low resolution (LR) images from different views of the same object or learning from large amount of training datasets to achieve success. However, in real clinical environments, these prerequisites are rarely fulfilled. In this paper, we present a self-learning based method to perform super-resolution (SR) from a single LR input. The mappings between the given LR image and its downsampled versions are modeled using support vector regression on features extracted from sparse coded dictionaries, coupled with dual-tree complex wavelet transform based denoising. We demonstrate the efficacy of our method in application of cardiac MRI enhancement. Both quantitative and qualitative results show that our SR method is able to preserve fine textural details that can be corrupted by noise, and therefore can maintain crucial diagnostic information.

1 Introduction

High resolution (HR) images are in demand for cardiac magnetic resonance imaging (CMRI). However, HR CMRI images are often costly to acquire. In particular, the quality and resolution of CMRI images can be limited dramatically by low signal to noise ratio, cardiac and respiratory motion, and restricted scanning time due to patient symptoms etc. In consequence, noise corrupted and low resolution (LR) images are produced that may reduce the visibility of vital pathological details and potentially compromise the clinical outcomes.

Instead of optimizing hardware settings (e.g., improved cardiac gating and respiratory navigation) and reducing scanning time (e.g., accelerated by parallel imaging and compressed sensing), image super-resolution (SR) provides an alternative solution to boost the perceptual quality of CMRI images in terms of the spatial resolution enhancement. Essentially, the goal of SR methods is to recover a HR image from a single or multiple LR images [1]. Comprehensive reviews on various SR methods can be found elsewhere [2, 3], and here we briefly review the most relevant publications. Existing SR algorithms can be broadly categorized into three classes including interpolation based, reconstruction based,

[†]Contact Email: g.yang@imperial.ac.uk. This work has been funded by NIHR.

and learning or example based methods [4]. Widely used interpolation based SR methods, which assume that images are spatially smooth, typically result in overly smoothed edges with ringing and jagged artifacts [4, 5]. Reconstruction based algorithms, which solve an inverse problem by recovering the HR image by fusing multiple LR images, are time-consuming and limited to an upsampling factor of two [4, 5, 6]. For learning or example based methods, the mapping function between LR and HR images (or their patches) is learned from a representative set of training image pairs. Once the mapping function is learned, it is applied to a single testing image to achieve SR. Despite numerous learning or example based methods (e.g., [7, 8, 9, 10]) claiming success for single-image SR, these methods hinge on the availability of training data of LR and HR image pairs; therefore, we define them as *pseudo-single-image* SR.

Compared to pseudo-single-image SR methods, Glasner et al. [1] proposed a self-similarity based single-image SR method using a training dataset that is directly established from the LR input, by exploiting patch redundancy among in-scale and cross-scale images in an image pyramid to enforce constraints for recovering the unknown HR image [4]. In this method, no extrinsic large training dataset is required as a priori, but the abundance of self-similar patches is crucial for successful SR recovery. Instead of searching for similar image patches, Yang and Wang [11] used support vector regression (SVR) to learn SR models from patches at different image scales. They also applied a sparse representation to extract effective image features for SVR to make their SR algorithm more computationally feasible for real applications, following the application of sparse representation originally proposed for solving the SR problem by Yang et al. [12]. More recently, Singh and Ahuja [13] added sub-band energy constraints for the self-similarity based SR method, and Huang et al. [14] enriched the dataset of self-similar patches by looking at their transformed exemplars. Although these studies have demonstrated promising results, the SR methods were applied on natural images, which are relatively clean. Therefore, the performance of these methods for noisy inputs, e.g., CMRI images, hasn't been demonstrated.

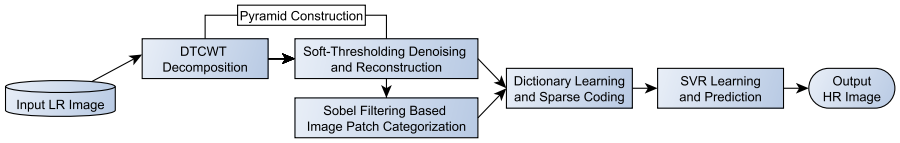

Fig. 1. Flowchart of our approach including both learning and prediction.

In this paper, we propose an SVR based SR method that learns a sparse representation as robust and effective features across various down-sampled and denoised versions (using dual-tree complex wavelet transform, i.e., DTCWT) of a single input LR image. Once the SVR has been trained, we can apply the best model to predict the final SR image from a single LR image (Figure 1). Compared to the previous learning based methods [7, 8, 9], our method does not require construction of paired LR and HR training datasets. Compared to Glasner et al. [1] (Figure 2(a)) and more recent work by Singh and Ahuja [13], our method does not assume self-similarity of image patches. In contrast to the

SR methods [11], [15], in which SVR was originally applied, we use DTCWT to suppress the noise, and we hypothesize that our method is more applicable for medical imaging applications [16], in particular for the CMRI images.

2 Method

The overall workflow of our method is summarized in Figure 1, and details of each step are described below.

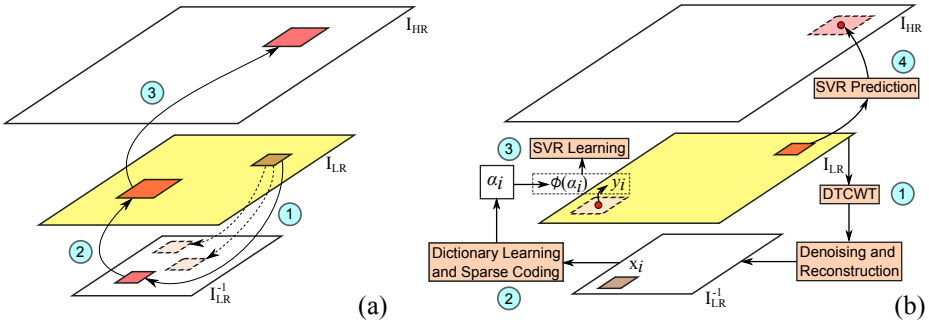

Fig. 2. Schematic diagram of (a) conventional self-similarity based SR method and (b) our self-learning based SR approach. Panel (a): In step 1, the input LR image I_{LR} (yellow) is downsampled by a factor of two recursively into several lower levels, and here only one downsampled level (I_{LR}^{-1}) is shown for illustration purposes. Then k nearest neighbors of the example patch (brown patch P_{LR} in I_{LR}) are located in the downsampled level. For simplicity, we show the case where the most similar patch (red patch in I_{LR}^{-1}) is used (i.e., $k = 1$). In step 2, the corresponding patch at the same location (red patch in I_{LR}) is used as the HR predictor (notice the larger size). In step 3, this patch is copied and pasted in the HR image (I_{HR}). In the case of $k > 1$, multiple predictors are averaged and then pasted; Panel (b): Four major steps of our method and details are described in Sections 2.1-2.3. Instead of downsampling using bicubic interpolation, DTCWT based denoising and reconstruction was applied to construct the image pyramid.

2.1 Denoising and Pyramid Construction Using DTCWT

Most previous self-similarity or self-learning methods (e.g., [1, 11, 13]) used bicubic interpolation to construct the image pyramid, but this can be susceptible to noise. Intuitively, discrete wavelet transform (DWT) based denoising has an intrinsic multiscale structure of decomposition and reconstruction, and therefore can be suitable for the pyramid construction task while effectively suppressing the noise. In addition, the dual-tree complex wavelet transform (DTCWT) overcomes two shortcomings of the widely used conventional DWT: lack of shift-invariance and lack of directional selectivity [17]. Regarding memory usage, the DTCWT is an over-complete transform with a redundancy factor of four for 2D image processing while DWT is non-redundant. An alternative method to obtain the orientation selective sub-bands is using steerable pyramid decomposition [13], which is over-complete by a factor of 4k/3 (where k denotes the number of orientation sub-bands) [18]. Regarding computational complexity, the DTCWT is linear $\mathcal{O}(\mathrm{N})$ where N is the number of input pixels. In contrast, the steerable pyramid decomposition has a complexity of $\mathcal{O}(\mathrm{N} \log \mathrm{N})$. By considering both memory and computational complexity, instead of using steerable pyramid decomposition to overcome limitations of DWT, we decomposed the original image using DTCWT. To accomplish the denoising and image pyramid construction, each level of the pyramid was reconstructed using a soft-thresholding scheme [17], which we used to train the SVR (step 1 in Figure 2(b)).

2.2 Extracting Features Using Image Sparse Representation

Directly working on pixels can be time-consuming and infeasible for real clinical applications. In this study, a sparse image representation [19] was learned, representing robust and effective features for SVR that can be formulated as an optimization problem, $\min \frac{1}{n} \sum_{i=1}^{n} ||D\alpha_i - \mathrm{x}_i||_2^2 + \lambda||\alpha_i||_1, i = 1, \ldots, n$, in which x_i is the vectorized image patch (patch size was set to $P = 5 \times 5$ in this study), and is denoted as $\mathrm{x} = \mathrm{vectorize}(\mathrm{P_{LR}}) - \mathrm{mean}(\mathrm{P_{LR}}) \in \mathbb{R}^{P \times 1}$, and the subtraction of the mean value of each LR patch ($\mathrm{mean}(\mathrm{P_{LR}})$) indicates the learning of local pixel value variations instead of learning absolute pixel values. Here n is the number of training patches and D is the over-complete dictionary to be learned, α_i is the corresponding sparse coefficient vector, and λ is the Lagrange multiplier term that regularizes the model sparsity (l_1-norm term) and the l_2-norm based residual. According to previous research [11], better SR performance can be achieved by patch categorization, i.e., clustering patches into low and high spatial frequency ones and learning their dictionaries separately. We applied a Sobel filter in this study, such that any patch including an edge derived from the Sobel filter is considered to have high spatial frequency. In doing this, two dictionaries (D_l and D_h) for low and high spatial frequency patches have been learned separately. Accordingly, we used the corresponding sparse coefficient vectors α_l and α_h as the features for the SVR models (step 2 in Figure 2(b)).

2.3 Support Vector Regression

Finally, support vector regression (in this study we used ν-SVR) [20], which can fit the data in a high dimensional feature space without assuming the data distribution, was applied to model the relationship between the input sparse coefficient vector and the associated SR pixel value. In the training procedure, the SVR solves the following optimization problem,

$$\begin{aligned} \min \frac{1}{2}||\omega||^2 + \rho\left(\nu\epsilon + \frac{1}{n}\sum_{i=1}^{n}(\xi_i + \xi_i^*)\right) \\ \text{subject to } y_i - \langle\omega, \Phi(\alpha_i)\rangle - b \leq \epsilon + \xi_i \\ \langle\omega, \Phi(\alpha_i)\rangle + b - y_i \leq \epsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0,\ \epsilon \geq 0,\ i, \ldots, n \end{aligned} \tag{1}$$

in which y_i denotes the associated target variable, i.e., pixel value at the center of the patch considered in the HR image ($\mathrm{mean}(\mathrm{P_{LR}})$ is also subtracted from each corresponding y_i). $\Phi(\alpha_i)$ is the sparse image patch features in the transformed space; thus, $\Phi(\alpha_i)$ and y_i form the feature and target variable pairs for training. And ω is the norm vector of the nonlinear mapping function to be learned. In addition, ρ is the tradeoff between the generalization and the upper and lower bounds of training errors ξ_i and ξ_i^* subject to a margin of ϵ, and ν controls the amount of support vectors in the resulting model. $K(\alpha_i, \alpha_j) \equiv \Phi(\alpha_i)^T\Phi(\alpha_j)$ is called the kernel function. In this study, we used a nonlinear Gaussian Radial Basis Function (RBF) kernel $K(\alpha_i, \alpha_j) = \exp(-\gamma||\alpha_i - \alpha_j||^2)$ with scaling-factor, $\gamma > 0$, to map feature vectors into a nonlinear transformed feature space. SVR

parameters (ρ and γ) were estimated using grid search with cross-validation. Furthermore, SVR model selection was achieved according to a minimum-error-rate classification rule based on Bayesian decision theory [11], and the trained best SVR model was applied to predict the final SR output for a test LR input (steps 3 and 4 in Figure 2(b)).

2.4 Experimental Settings and Performance Measure

First, ex-vivo swine heart MR imaging data were acquired on a Siemens Skyra 3T MRI system (Erlangen, Germany). The subject was scanned using a 4-element receive coil setup and a 3D fast low-angle shot sequence (3D FLASH, TR/TE = 50 ms/5.36 ms, flip angle = 35°, voxel resolution = 0.6 mm^3, and reconstructed into 0.3 mm^3 isotropically). Second, in-vivo patient data were acquired on a Siemens Avanto 1.5T MRI system. Ten subjects were scanned using a navigator-gated inversion-prepared 3D segmented gradient echo sequence with late gadolinium enhancement (TR/TE = 5.2 ms/2.3 ms, flip angle = 20°, voxel resolution = 1.5x1.5x4 mm^3, and reconstructed into 0.7x0.7x2 mm^3) [21].

For the ex-vivo data, we extracted the central slice from the axial and saggital views respectively in order to test the efficacy of our method on both in-plane and out-of-plane slices. The original resolution is 400x400 pixels for both slices, and used as HR ground truth for SR results evaluation. Inputs were synthesized by downsampling (by a factor of 2 and 4 denoted as x2 and x4 resulted in 200x200 and 100x100 inputs respectively), and degrading (by various additive noises [(σ=10,15,20,and 25) and run 10 times for each noise level]). Evaluations have been done qualitatively by visual inspection and quantitatively using peak signal to noise ratio (PSNR) and mean squared error (MSE) against the HR ground truth. The performance of our method was compared with bicubic interpolation with standard low-pass Gaussian filter based denoising and the self-learning based SR approach (SLSR) proposed in [11]. For the in-vivo data, we randomly chose 40 ROIs (each of 200x200 pixels) from 10 patient cases. For these 40 noise corrupted LR slices, we have no HR ground truth available. In addition to visual inspection of x2 SR, we compared the line profiles extracted from the results of three SR methods, namely, bicubic (with standard low-pass Gaussian filter based denoising), SLSR, and ours, and estimated the remaining noise in the SR results using the method described in [22].

3 Results

Figure 3(a) shows the synthesized LR input by a downsampling factor of two. The added noise ($\sigma = 15$) is shown in Figure 3(b). Figure 3(c) is the zoomed-in ROI of the HR ground truth. Compared to the results of using bicubic interpolation and SLSR method (Figure 3(d) and (e)), our SR result (Figure 3(f)) is much more homogeneous in the myocardium region. In addition, green arrows pointed at three example fine structures that are much clearer in our SR result. Figure 3(g)-(i) are the results recovered from the synthesized LR input by an x4 downsampling with additive noise ($\sigma = 15$). Pink arrows pointed at textures that have been preserved using our SR method (Figure 3(i)) while missing in the bicubic interpolation and SLSR method (Figure 3(g) and (h)). Tables 1 and

2 show the PSNR and MSE with various additive noises, and the standard deviations were calculated by running the random noise simulation 10 times at each noise level ($\sigma = 10, 15, 20,$ and 25). In addition, our SR results (both PSNR and MSE in Figure 4) showed significant improvement compared to the results of bicubic interpolation and SLSR (statistical significances were given by two-sample Wilcoxon rank-sum test between the results of each two SR methods with significance level of $p < 0.05$). There is no significant difference between bicubic interpolation and SLSR when the noise level is high (e.g., when $\sigma = 25$).

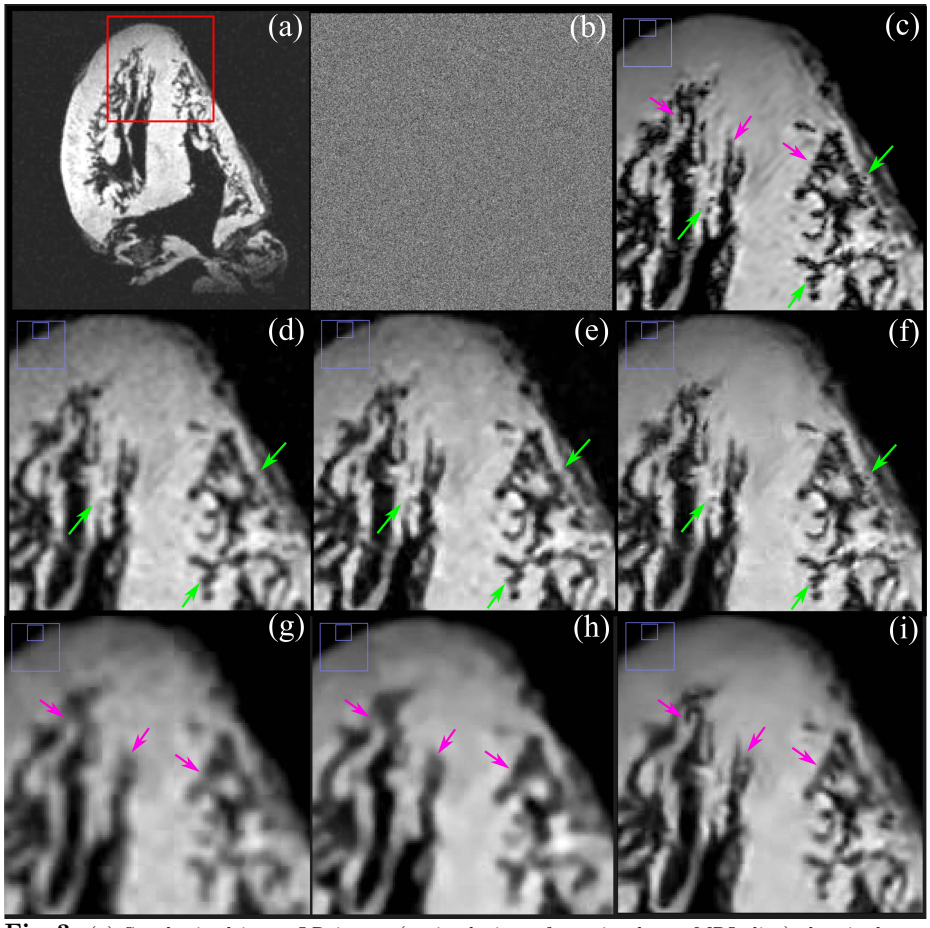

Fig. 3. (a) Synthesized input LR image (sagittal view of a swine heart MRI slice) that is down-sampled (x2) and noise corrupted (red box showing the ROI; contrast has been altered for better visualization of the background noise); (b) Random noise ($\sigma = 15$) added to the original HR image; (c) ROI of the original HR image without blurring or noise (ground truth); Panels (d)-(f): x2 SR results; (d) ROI of the bicubic interpolation result; (e) ROI of the SLSR result; (f) ROI of our result; Green arrows pointed three example fine structures that have been super-resolved using our method, but much more blurred using bicubic interpolation and SLSR method; Panels (g)-(i): x4 SR results; (g) ROI of the bicubic interpolation result; (h) ROI of the SLSR result; (i) ROI of our result; Pink arrows point to three example textures that have been preserved using our method, but which are more vague in the results of using bicubic interpolation or the SLSR method.

Figure 5(a)-(c) shows the x2 SR results for one ROI of an in-vivo late gadolinium enhancement scan. Our SR result (Figure 5(c)) shows noise suppression in the blood pool region of the left atrium, and therefore less confounding artifacts with similar intensities as the real late gadolinium enhancement of the myocardial wall. In addition, line profiles (Figure 5(d) and (e)) across the SR results show that bicubic interpolation and SLSR method tend to be much noisier. Furthermore, the noise level estimations have been performed according to [22], and our method obtained a lower level of remaining noise (Table 3 shows that $\sigma = 32.04 \pm 4.17$ compared to $\sigma = 35.76 \pm 3.67$ obtained by bicubic interpolation and $\sigma = 40.17 \pm 4.89$ obtained by SLSR method, and noise level of the original LR input is $\sigma = 53.52 \pm 5.48$).

Table 1. PSNR of x2 and x4 SR on ex-vivo cardiac MRI data from different views. For $\sigma > 0$ we run 10 times for each noise level to obtain the mean±std.

	PSNR				
	N0	**N10**	**N15**	**N20**	**N25**
Axial x2					
Bicubic	**33.84**	**33.22±0.07**	**32.37±0.04**	**31.61±0.10**	**30.74±0.09**
SLSR	**36.01**	**35.28±0.11**	**33.96±0.14**	**32.54±0.13**	**31.35±0.18**
Ours	**39.81**	**37.54±0.13**	**35.00±0.13**	**33.14±0.10**	**31.70±0.16**
Sagittal x2					
Bicubic	**31.96**	**31.30±0.02**	**30.60±0.06**	**29.77±0.03**	**28.96±0.07**
SLSR	**33.47**	**32.54±0.04**	**31.40±0.09**	**30.21±0.13**	**28.96±0.26**
Ours	**36.63**	**34.88±0.09**	**32.46±0.10**	**30.69±0.10**	**29.31±0.11**
Axial x4					
Bicubic	**26.80**	**26.66±0.01**	**26.47±0.02**	**26.23±0.03**	**25.96±0.04**
SLSR	**27.86**	**27.76±0.05**	**27.51±0.05**	**27.19±0.05**	**26.82±0.08**
Ours	**30.36**	**30.31±0.02**	**29.91±0.05**	**29.02±0.07**	**27.71±0.09**
Sagittal x4					
Bicubic	**25.53**	**25.37±0.01**	**25.19±0.01**	**24.93±0.02**	**24.66±0.02**
SLSR	**26.40**	**26.34±0.04**	**26.07±0.04**	**25.74±0.02**	**25.62±0.04**
Ours	**28.39**	**28.36±0.04**	**28.08±0.04**	**27.30±0.05**	**26.21±0.21**

Table 2. MSE of x2 and x4 SR on ex-vivo cardiac MRI data from different views. For $\sigma > 0$ we run 10 times for each noise level to obtain the mean±std.

	MSE				
	N0	**N10**	**N15**	**N20**	**N25**
Axial x2					
Bicubic	**26.91**	**31.76±0.13**	**37.88±0.11**	**46.21±0.50**	**56.88±0.43**
SLSR	**17.34**	**20.79±0.21**	**28.12±0.37**	**39.22±0.43**	**54.55±1.04**
Ours	**6.85**	**12.11±0.16**	**21.41±0.30**	**33.23±0.56**	**46.60±0.70**
Sagittal x2					
Bicubic	**32.79**	**38.23±0.16**	**45.00±0.36**	**54.40±0.30**	**65.99±0.40**
SLSR	**23.18**	**28.68±0.26**	**37.55±0.28**	**50.09±0.63**	**68.00±0.78**
Ours	**11.19**	**16.86±0.19**	**29.51±0.41**	**44.22±0.71**	**61.38±0.70**
Axial x4					
Bicubic	**135.78**	**140.46±0.42**	**146.53±0.51**	**154.86±0.99**	**164.91±1.38**
SLSR	**106.35**	**109.02±1.25**	**115.27±1.25**	**124.12±1.41**	**135.40±2.43**
Ours	**59.91**	**60.60±0.33**	**66.32±0.76**	**81.56±1.31**	**110.33±1.98**
Sagittal x4					
Bicubic	**144.20**	**149.61±0.28**	**155.90±0.51**	**165.51±0.77**	**176.35±0.62**
SLSR	**118.01**	**119.72±0.98**	**127.25±1.10**	**137.29±1.68**	**152.61±1.85**
Ours	**74.59**	**75.17±0.68**	**80.27±0.71**	**95.97±1.10**	**126.70±1.32**

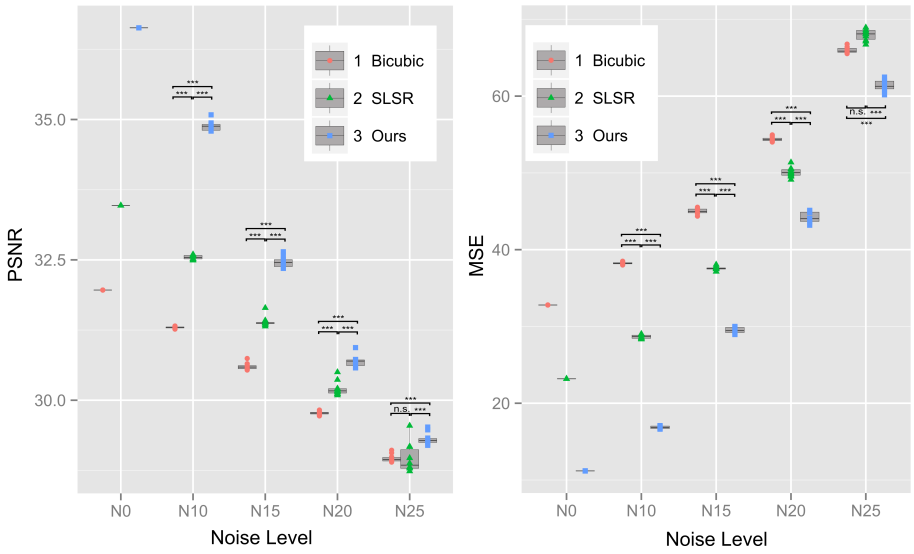

Fig. 4. Left: Boxplot of the PSNR (Sagittal x2 SR); Right: Boxplot of the MSE (Sagittal x2 SR). Statistical significant between each two groups are showing above the boxplot (*** stands for $p < 0.05$ and n.s. means no significant difference between two groups).

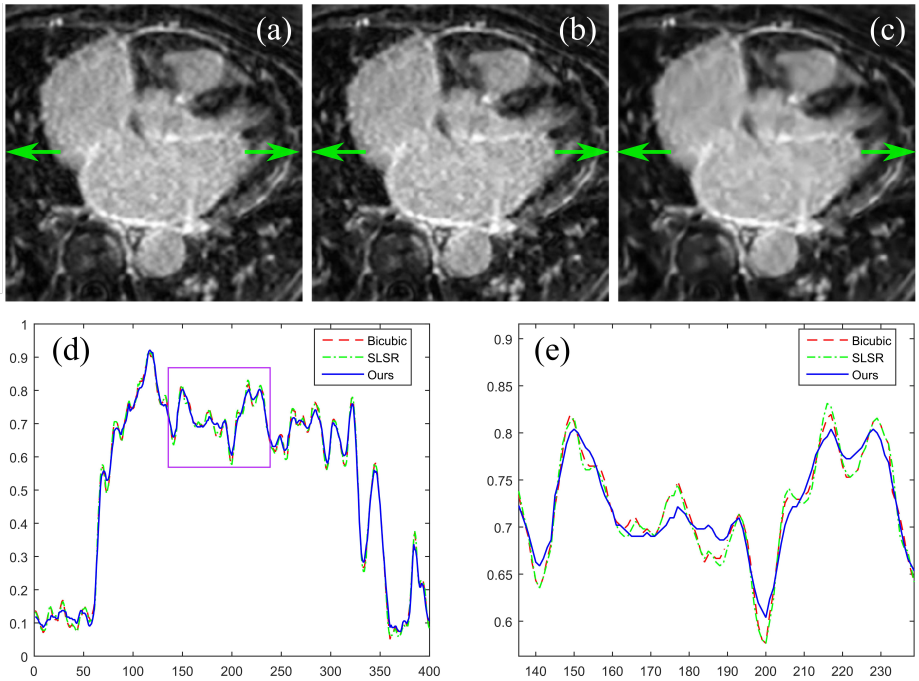

Fig. 5. Panels (a)-(c): x2 SR results of an in-vivo late gadolinium enhancement scan using bicubic interpolation, SLSR method and our approach respectively; (d) Horizontal line profiles through the green arrows of the SR results; (e) Zoomed-in line profiles of the pink region in (d).

4 Discussion and Conclusion

The objective of this study was to develop an image post-processing method to super-resolve LR cardiac MR images from single input while effectively suppressing the noise. Results of this work provide compelling evidence that our single image SR method, which is coupled with multiscale DTCWT based denoising, is capable of reconstructing superior HR results with significantly lower remaining noise compared with conventional bicubic interpolation and a state-of-the-art self-learning based SR approach, i.e., SLSR.

Kernel based single image SR methods (example kernels including linear, bicubic, and Lanczos kernel) are the most widely used in clinical applications due to their efficiency. Results of our experiments on ex-vivo cardiac MR images showed that bicubic kernel based interpolation obtained lower PSNR and higher MSE than SLSR or our methods when the noise level $\sigma \leq 20$ (Tables 1 and 2). This can be attributed to the fact that bicubic interpolation makes strong assumptions on spatial smoothness of images that can result in blurred edges, loss of fine textures, and averaged noisy pixels with their less noisy neighbors (Figure 3(d) and (g)).

Table 3. Noise estimation of in-vivo late gadolinium enhancement data for randomly selected 40 ROIs to obtain the mean±std.

	Original LR Image	**Bicubic**	**SLSR**	**Ours**
Noise Level	**53.52±5.48**	**35.76±3.67**	**40.17±4.89**	**32.04±4.17**

Self-similarity based methods and later proposed self-learning based methods utilize a multiscale decomposition of the LR image to understand the relationship between its current scale and its lower resolution versions. Then the HR output is reconstructed either using prediction from candidates selected by k-nearest neighbors method (self-similarity based) or SVR (self-learning based). To the best of our knowledge, previously published self-similarity or self-learning based methods (e.g., [1, 11, 13]) are demonstrated on natural images, which are assumed with less or no noise contamination. Results of the SLSR method tested on ex-vivo data with various additive noises showed that for the noise level $\sigma \leq 20$ SLSR outperformed bicubic interpolation with significant differences (Figure 3(e) and (h) and Tables 1 and 2). This may be attributed to the fact that SVR learns local pixel value variations nonlinearly, while negative lobes on the bicubic kernel tend to create ringing artifacts. When $\sigma = 25$ SLSR performed similar to bicubic interpolation (no significant difference) demonstrating that SLSR results can be affected by noise significantly. Tellingly, with noise suppression in each level of the multiscale pyramid our method showed significant improvement of both PSNR and MSE even for the dense noise contamination with high magnification factor (Tables 1 and 2). Both line profiles and remaining noise estimation further confirmed the merits of our SR method (Figure 5(d) and (e) and Table 3).

Although previous research demonstrated the success of using paired LR and HR training datasets to super-resolve a given new LR image [9], in this study we emphasize the practical infeasibility of this class of methods for CMRI images where acquiring HR training data can require prohibitively long scan times. For

clinical CMRI studies, total scanning time per patient is already long (typically 60 minutes) and this can not easily be extended for HR data acquisition without having a significant input on patient throughput. A method, which doesn't require the use of HR training data, is therefore highly desirable. Moreover, there are potential hazards include painful peripheral nerve stimulation for prolonged scanning time. We are aware that our proof of concept study may have two limitations. First, there are several parameters that can be tuned in our framework, e.g., the thresholds for the calculated gradient magnitude in Sobel filtering and DTCWT denoising. Currently they are based on trial and error; however, we envisage that future assessment of the robustness of the SR framework will include parameter perturbation. Second, we explicitly assumed an additive Gaussian noise for simplicity, but more complex noise modeling and suppression could be easily plugged-in to the current framework. In summary, our SR framework, which couples SVR and DTCWT, can achieve promising HR CMRI results while effectively suppressing the noise.

References

[1] Glasner, D., et al.: Super-resolution from a single image. ICCV, pp. 349–356 (2009)
[2] Van Reeth, E., et al.: Super-resolution in magnetic resonance imaging: A review. Concepts in Magnetic Resonance Part A, Vol. 40(6), pp. 306–325 (2012)
[3] Nasrollahi, K. and Moeslund, T.B.: Super-resolution: A comprehensive survey. Mach. Vis. Appl., Vol. 25(6), pp. 1423–1468 (2014)
[4] Yang, C.Y., et al.: Exploiting self-similarities for single frame super-resolution. ACCV., pp. 497–510 (2010)
[5] Kangi, L., et al.: Self-learning-based low-quality single image super-resolution. MMSP., pp. 1–6 (2013)
[6] Lin, Z.C. and Shum, H.Y.: Fundamental limits of reconstruction-based super-resolution algorithms under local translation. IEEE Trans. Pattern Anal. Mach. Intell., Vol. 26(1), pp. 83–97 (2004)
[7] Chang, H., et al.: Super-resolution through neighbor embedding. CVPR., pp. 1–8 (2014)
[8] Yang, J., et al.: Coupled dictionary training for image super-resolution. IEEE Trans. Image Process., Vol. 21(8), pp. 3467–3478 (2012)
[9] Trinh, D.H., et al.: Novel example-based method for super-resolution and denoising of medical images. IEEE Trans. Image Process., Vol. 23(4), pp. 1882–1895 (2014)
[10] Hidane, M., et al.: Super-resolution from a low-and partial high-resolution image pair. ICIP., pp. 2145–2149 (2014)
[11] Yang , M.C. and Wang, Y.C.F.: A self-learning approach to single image super-resolution. IEEE Trans. Multimedia., Vol. 15(3), pp. 498–508 (2013)
[12] Yang, J., et al.: Image Super-resolution via sparse representation. IEEE Trans. Image Process., Vol. 19(11), pp. 2861–2873 (2010)
[13] Singh, A. and Ahuja, N.: Sub-band energy constraints for self-similarity based super-resolution. ICPR, pp. 4447–4452 (2014)
[14] Huang, J., et al.: Single image super-resolution from transformed self-exemplars. CVPR, pp. 5197–5206 (2015)
[15] Ni, K.S. and Nguyen, T.Q.: Image super-resolution using support vector regression. IEEE Trans. Image Process., Vol. 16(6), pp. 1596–1610 (2007)
[16] Yang, G., et al.: Combined self-learning based single-image super-resolution and dual-tree complex wavelet transform denoising for medical images, In press (2016)
[17] Selesnick, I.W., et al.: The dual-tree complex wavelet transform. Signal Process. Mag. IEEE., Vol. 22(6), pp. 123–151 (2005)
[18] Cohen, R.: Signal denoising using wavelets. Technical Report, Department of Electrical Engineering Technion, Israel Institute of Technology, Haifa (2012)
[19] Mairal, J., et al.: Online learning for matrix factorization and sparse coding. J. Mach. Learn. Res., Vol. 11, pp. 19–60 (2010)
[20] Chang, C. and Lin, C.: LIBSVM: a library for support vctor machines. ACM Trans. Intell. Syst. Technol., Vol. 2(5), pp. 1–39 (2011)
[21] Keegan, J., et al.: Improved respiratory efficiency of 3D late gadolinium enhancement imaging using the continuously adaptive windowing strategy (CLAWS). Magn. Reson. Med., Vol. 71(3), pp. 1064–1074 (2014)
[22] Liu, X., et al.: Single-image noise level estimation for blind denoising. IEEE Trans. Image Process., Vol. 22(12), pp. 5226–5237 (2013)