Super-resolution of 3D Magnetic Resonance Images by random shifting and convolutional neural networks

Karl Thurnhofer-Hemsi*, Ezequiel López-Rubio*, Núria Roé-Vellvé†,
Enrique Domínguez* and Miguel A. Molina-Cabello*
*Department of Computer Languages and Computer Science
University of Málaga, Bulevar Louis Pasteur, 35, 29071 Málaga, Spain
†Molecular Imaging Unit, Centro de Investigaciones Médico-Sanitarias
General Foundation of the University of Málaga, C/ Marqués de Beccaria, 3, 29010 Málaga, Spain
Emails: {karlkhader,ezeqlr,enriqued,miguelangel}@lcc.uma.es*, nroe@fguma.es†

Abstract—Enhancing resolution is a permanent goal in magnetic resonance (MR) imaging, in order to keep improving diagnostic capability and registration methods. Super-resolution (SR) techniques are applied at the postprocessing stage, and their use and development have progressively increased during the last years. In particular, example-based methods have been mostly proposed in recent state-of-the-art works. In this paper, a combination of a deep-learning SR system and a random shifting technique to improve the quality of MR images is proposed, implemented and tested. The model was compared to four competitors: cubic spline interpolation, non-local means upsampling, low-rank total variation and a three-dimensional convolutional neural network trained with patches of HR brain images (SRCNN3D). The newly proposed method showed better results in Peak Signal-to-Noise Ratio, Structural Similarity index, and Bhattacharrya coefficient. Computation times were at the same level as those of these up-to-date methods. When applied to downsampled MR structural T1 images, the new method also yielded better qualitative results, both in the restored images and in the images of residuals.

Index Terms—Magnetic Resonance Imaging, super resolution, convolutional neural networks, supervised learning

I. INTRODUCTION

Estimation of high-resolution (HR) images from low-resolution (LR) images is a classical problem in image processing. This process, called super-resolution (SR), consists in using post-processing techniques to improve the image quality [1]. In the medical image processing subfield, despite the advances in acquisition technology, such as Computerized Tomography (CT), Positron Emission Tomography (PET), Magnetic Resonance Imaging (MRI) or combined modalities (e.g. SPECT/CT), obtaining an image at a desired resolution is not easy due to the limitations of physical imaging systems as well as quality factors such as noise, which is inherent in medical imaging, or blurring.

In the MRI typical clinical setting, data (both low- and high-resolution images of different types) is usually acquired with different voxel sizes for different MRI modalities. Moreover, in-plane resolution is usually higher than out-plane (i.e in the slice direction) one, producing non-isotropic voxel sizes

(i.e. rectangular voxels). The final resolution of the acquired images is normally limited by hardware, signal to noise ratio (SNR), limited acquisition time or movements of the patient. Therefore, in some cases, the acquired images have to be upsampled to decrease the voxel size and obtain higher-resolution images, which will be post-processed or analyzed.

Traditional techniques [2], [3], such as linear interpolation or spline-based methods, have been extensively used to increase the resolution of data. However, such techniques estimate new points assuming the homogeneity of regions, hence these interpolated images are typically blurred. Among different techniques [4], [5], example-based methods have been mostly proposed in recent state of the art works, either by exploiting internal similarities of the same image [6], [7], or learning from external training datasets [8]–[11].

The super-resolution convolutional neural network (SR-CNN) [12] is one of the most recent external example-based methods, where a deep CNN is proposed in order to learn the mapping function between low- and high-resolution images instead of learning dictionaries or manifolds to model the high-resolution space.

CNNs were inspired by the animal visual cortex and they are one of the first truly successful deep learning architectures, which have shown outstanding performance in processing images and videos. Nowadays, with the help of GPU-accelerated computing techniques, CNNs have been successfully applied to object recognition (e.g. handwriting, face, behavior...), recommender systems or image classification. Hundreds of papers have been published in the last years providing different types of deep neural networks in some selected areas [13], [14], particularly in medical image analysis [15], [16], where CNNs have become increasingly popular and widely applied.

In this work, a three dimensional convolutional neural network together with a random shifting technique are proposed to increase the resolution of MR images. The proposed method has been successfully applied to different brain MRI datasets and compared with state of the art SR algorithms. Results have shown the superior performance of the proposed method, even

in qualitative aspects.

The rest of paper is organized as follows. In the next section II, the proposed combination of a deep learning SR system and a random shifting technique is presented and both models (CNN and random shifting) explained. Comparative results are reported in the section III, where competitors and datasets are also described. Finally, the conclusions and future works are presented in section IV.

II. THE MODEL

In this section a combination of a deep-learning SR system and a random shifting technique to improve the quality of MR images is proposed. In Subsection II-A the SRCNN method employed for our model is described, and in Subsection II-B the shifting procedure to estimate the final HR image given a LR one is designed.

A. Convolutional neural network

The standard CNN based SR reconstruction is carried out in two steps:

- First, given a LR image $\mathbf{X}$, a spline interpolation I is performed in order to obtain a HR image $\mathbf{Z} = I(\mathbf{X})$.
- Second, a convolutional neural network is applied to restore the image.

The CNN is composed by three convolutional layers and a Rectified Linear Unit (ReLU) layer is applied in the first two filter responses. If we call g_1, g_2, g_3 for each of these operations, the net computes a HR image

$$g = g_3 \circ g_2 \circ g_1 \tag{1}$$

and the restoration is computed minimizing the Euclidean loss

$$\tilde{f} = arg\,min_g \sum ||\mathbf{X} - g(\mathbf{Z})||^2 \tag{2}$$

The first convolutional layer applies 64 filters of size $9 \times 9 \times 9$, 32 of size $1 \times 1 \times 1$ in the second, and one filter of size $5 \times 5 \times 5$ in the last layer. Overlapping patches are extracted from a set of HR reference images in order to have enough samples to train the network, and feature maps are computed. For each patch, a down-sampling and up-sampling is applied and a relationship is established to learn an end-to-end mapping between LR and HR images.

For the proposed model described in the next section, we use this net as a base, which is called SRCNN3D and is implemented in [17].

B. Random shifting model

Given an LR image $\mathbf{X}$, the SRCNN model learns an approximation $\tilde{f}(\mathbf{X})$ of the underlying HR image $f(\mathbf{X})$:

$$f(\mathbf{X}) = \tilde{f}(\mathbf{X}) + \tilde{\epsilon}(\mathbf{X}) \tag{3}$$

where $\tilde{\epsilon}(\mathbf{X})$ is the approximation error. Shifted versions LR image can be considered as inputs to the SRCNN, where $\mathbf{a} \in [0, K) \times [0, K) \times [0, K)$ is a shift vector and K is the window size (in voxels) of the first layer of the SRCNN. Therefore a

different approximation $\bar{f}_\mathbf{a}(\mathbf{X})$ of the underlying HR image f is obtained for each shift vector $\mathbf{a}$:

$$\bar{f}_\mathbf{a}(\mathbf{X}) = \tilde{f}(\mathbf{X} \ominus \mathbf{a}) \oplus \lambda\mathbf{a} \tag{4}$$

$$\bar{\epsilon}_\mathbf{a}(\mathbf{X}) = \tilde{\epsilon}(\mathbf{X} \ominus \mathbf{a}) \oplus \lambda\mathbf{a} \tag{5}$$

$$f(\mathbf{X}) = \left(\tilde{f}(\mathbf{X} \ominus \mathbf{a}) + \tilde{\epsilon}(\mathbf{X} \ominus \mathbf{a}) \right) \oplus \lambda\mathbf{a} =$$

$$\left(\tilde{f}(\mathbf{X} \ominus \mathbf{a}) \oplus \lambda\mathbf{a} \right) + (\tilde{\epsilon}(\mathbf{X} \ominus \mathbf{a}) \oplus \lambda\mathbf{a}) =$$

$$\bar{f}_\mathbf{a}(\mathbf{X}) + \bar{\epsilon}_\mathbf{a}(\mathbf{X}) \tag{6}$$

where $\ominus$ and $\oplus$ stand for the image left shift and right shift operators, respectively, and λ is the super-resolution zoom factor. If $\mathbf{a}$ is regarded as a random variable, then a consensus approximation can be obtained as the expectation of $\bar{f}_\mathbf{a}(\mathbf{X})$ with respect to $\mathbf{a}$, $E_\mathbf{a}\left[\bar{f}_\mathbf{a}(\mathbf{X})\right]$. An estimation of such expectation can be readily obtained by averaging M individual approximations:

$$\hat{f}(\mathbf{X}) = \frac{1}{M} \sum_{m=1}^{M} \bar{f}_{\mathbf{a}_m}(\mathbf{X}) \tag{7}$$

where the $\mathbf{a}_j$ are uniformly drawn at random from $[0, K) \times [0, K) \times [0, K)$. From (7) and the strong law of large numbers, it follows that $\hat{f}(\mathbf{X})$ converges almost surely to $E_\mathbf{a}\left[\bar{f}_\mathbf{a}(\mathbf{X})\right]$ as the number of samples M tends to infinity:

$$\hat{f}(\mathbf{X}) \to E_\mathbf{a}\left[\bar{f}_\mathbf{a}(\mathbf{X})\right] \text{ when } M \to \infty \tag{8}$$

Now, it is reasonable to think that $E_\mathbf{a}\left[\bar{f}_\mathbf{a}(\mathbf{X})\right]$ is a better approximation to $f(\mathbf{X})$ than most of the individual approximations $\bar{f}_\mathbf{a}(\mathbf{X})$, since $E_\mathbf{a}\left[\bar{f}_\mathbf{a}(\mathbf{X})\right]$ does not depend on a particular shift $\mathbf{a}$ of the LR image, while the individual approximations do. Hence we propose to use $\hat{f}(\mathbf{X})$ as an approximator to $f(\mathbf{X})$. In practice the $\mathbf{a}_j$ are constrained to be integer vectors, so that fractional shifts are not necessary, since fractional shifts would pose a difficult problem themselves.

The proposed super-resolution algorithm is as follows:

1) Given M, compute the following set of vectors:

$$\mathcal{S}_M = \{\mathbf{a}_m\}_{m \in \{1,...,M\}} \tag{9}$$

$$\mathbf{a}_m = (a_1^{(m)}, a_2^{(m)}, a_3^{(m)}) \in \mathbb{R}^3 \tag{10}$$

where a_i is a random value drawn from the uniform distribution on the set $\{0, ..., K-1\}$, for each $i \in \{1, 2, 3\}$.

2) For an input LR image $\mathbf{X}$, compute M circularly image shifts:

$$\{\mathbf{X} \ominus \mathbf{a}_m, \mathbf{a}_m \in \mathcal{S}_M\} \tag{11}$$

3) Apply the CNN to obtain a set of HR images:

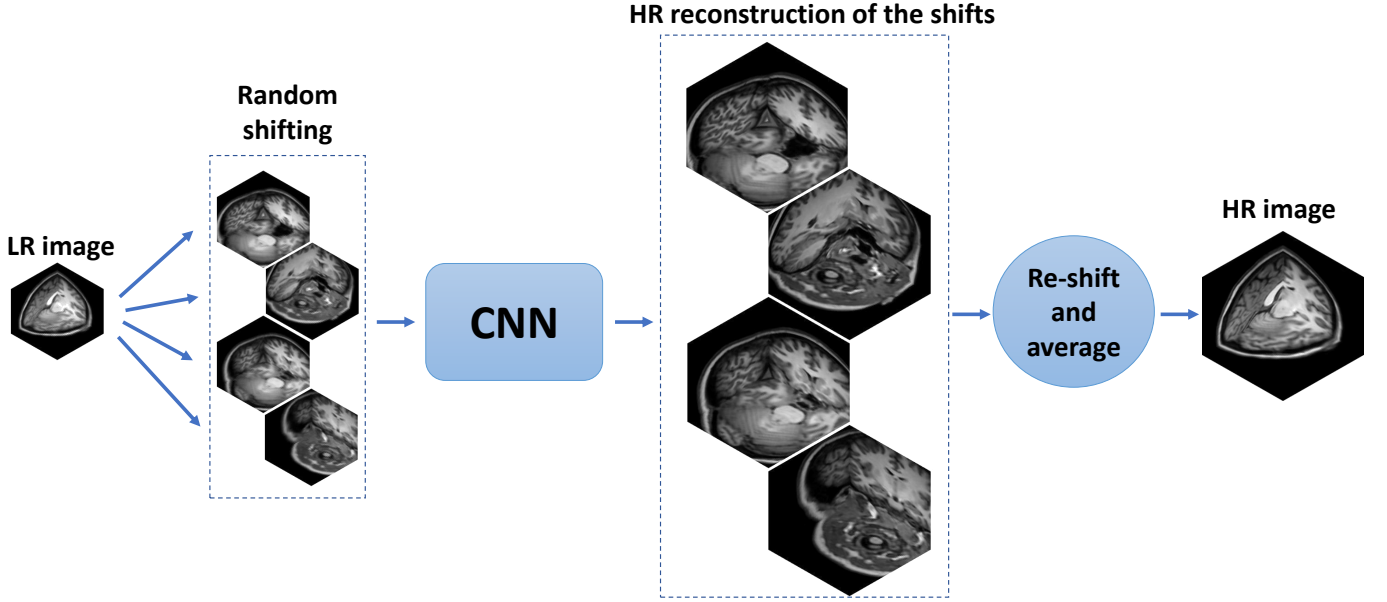$$\{\tilde{f}(\mathbf{X} \ominus \mathbf{a}_m), \mathbf{a}_m \in \mathcal{S}_M\} \tag{12}$$

Fig. 1. Scheme of the proposed algorithm.

4) Recompose the images considering the super-resolution zoom factor $\{\bar{f}_{\mathbf{a}_m}(\mathbf{X}), \mathbf{a}_m \in \mathcal{S}_M\}$ and compute the final restored HR image $\hat{f}(\mathbf{X})$ following Eq. (7).

A schematic depiction of the operation of our algorithm is shown in Fig. 1.

III. EXPERIMENTAL RESULTS

The experiments we have carried out and the obtained results are reported in this section. The LR image generation, software and hardware that we have used, and the selected performance metrics for comparison between methods are specified in Subsection III-A. Then, the MR image datasets are described in III-B. The set of tuned parameters are in Subsection III-C and the descriptions of the competing algorithms in Subsection III-D. At the end, in Subsection III-E we report the results of the performance tests.

A. Methods

Input LR images for each method were obtained from the original HR images following this procedure: first, HR images were cropped in relation to the zoom factor to be applied. Then, a three-dimensional Gaussian filter was applied, with standard deviation equal to 1. Finally, cubic interpolation was used to obtain the LR image, through the *imresize3* function of Matlab (default parameters).

The SRCNN3D method was used as the superresolution convolutional neural network since it has demonstrated a considerable effectiveness compared with other state-of-art methods. We employed a pre-trained model which is available in the project. The CNN was trained over 470000 iterations, using momentum of 0.9, learning rate of 0.0001 and batch size of 256. Stochastic gradient descent was used for model optimization.

The comparison experiments have been carried out on a 64-bit Personal Computer with an eight-core Intel i7 3.60GHz CPU, 32 GB RAM and standard hardware, and using Matlab R2017b. A Nvidia GTX 1080 GPU was used for the image testing of the CNN methods.

Three quality measures were used to evaluate the proposed method:

- *Peak Signal-to-Noise Ratio (PSNR)*, measured in (decibels) dB, which is commonly used in medical image processing (higher is better).
- *Structural Similarity index (SSIM)* [18], which focuses on structural similarities between images (higher is better):

$$SSIM(x,y) = \frac{(2\mu_x\mu_y)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (13)$$

where μ_x and μ_y are the mean value of images x and y, σ_x and σ_y are the standard deviation of images x and y, σ_{xy} is the covariance of x and y, $c_1 = (k_1L)^2$ and $c_2 = (k_2L)^2$ (default values were used: $L = 1$ is the dynamic range, $k_1 = 0.01$ and $k_2 = 0.03$).

- *Bhattacharrya coefficient (BC)* [19], which measures the closeness of the two discrete pixel probability distributions P and $\hat{P}$ corresponding to the ground truth (GT) and modeled images with values in the range $[0, 255]$:

$$BC = \sum_{j=0}^{255} P(j)\hat{P}(j) \quad (14)$$

where $BC \in [0, 1]$ and higher is better.

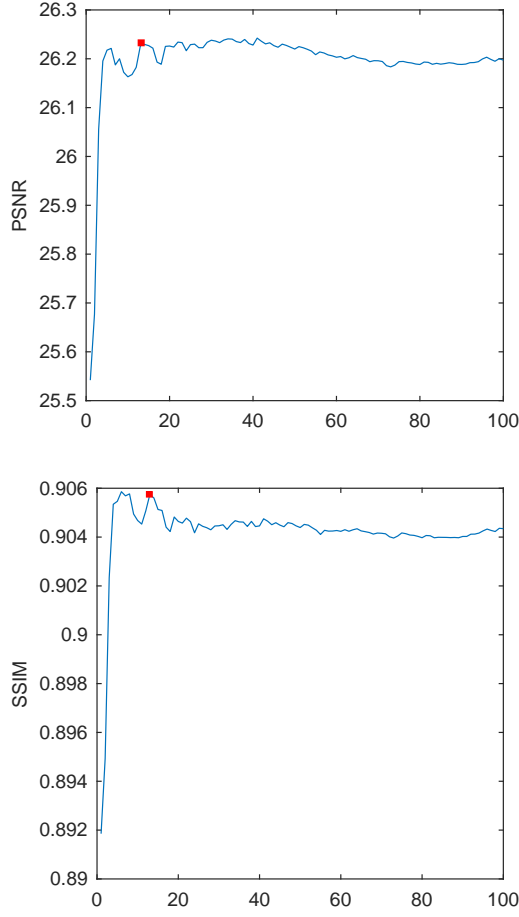| Parameter	Value
Window size K	25
Number of shifts M	13

Fig. 2. Evolution of the MSE and SSIM metrics when varying the number of shifts employed for the reconstruction (higher is better). Image 1 of the Kirby 21 dataset, image 80 of disc1 of the OASIS dataset and T1 noiseless image of Brainweb were used for the analysis.

Also, CPU time was measured for each method. In the case of our method, as it is based on the execution of a CNN as many times as shifted images we have generated, it could be parallelized in different GPU's. However, we measured the processing time sequentially with only one GPU.

In addition to this, the results were examined from a qualitative point of view. For that purpose, residual images r for each method were computed as the difference between the original HR image h and the super-resolved one s:

$$r = h - s \qquad (15)$$

The optimum residual image should be the darkest one, as the differences must be close to zero. However, for a better discrimination between methods, these residuals were subtracted to a positive constant, so the images are displayed in gray.

B. Datasets

Images obtained from different datasets were considered in order to evaluate the results of the SR algorithms. Two T1-weighted MRI images of the Kirby 21 (images 10 and 11) [8]. These data were acquired using a 3-T MR scanner with a $1.0 \times 1.0 \times 1.2 \, mm^3$ voxel resolution over an field-of-view (FOV) of $240 \times 204 \times 256 \, mm$ acquired in the sagittal plane. Two T1 images of the OASIS dataset (images 1 and 2 of the cross-sectional data) [20] . Data were acquired on a 1.5-T Vision scanner with a $1.0 \times 1.0 \times 1.25 \, mm^3$ voxel resolution over an FOV of $256 \times 256 \, mm$. One image of the IBSR public dataset [21]. It is named IBSR_07, it has image size $256 \times 256 \times 128$, with $1.5 \times 1.0 \times 1.0 \, mm^3$ voxel resolution. Finally, a T1-weighted image was acquired at CIMES using a 3-T MR scanner with a $0.93 \times 0.93 \times 1.0 \, mm^3$ voxel resolution over an FOV of $256 \times 256 \, mm$.

C. Parameter selection

Apart from the parameters defined for the SRCNN model, which were set at their default values, our model needs to define two parameters: the window size K and the number of shifts M. For this tuning, three images different from those selected for the experiments were used: image 1 of the Kirby 21 dataset, image 80 of the OASIS dataset and a normal brain T1 image of the Brainweb[1] simulated database (slice thickness $1 \, mm$, 0% noise level and $RF = 0$). Mean PSNR and SSIM of these images were computed and the optimum parameters are reported in TABLE I. The evolution of these metrics as a function of the number of shifts employed in the reconstruction is displayed in Fig. 2. It can be seen that there is a stabilization as the number of shifts increases. A red square indicates the selected value after considering both measures.

D. Competitors

The proposed method has been compared with other four SR algorithms:

- *Spline*: cubic spline interpolation as implemented in Matlab (Mathworks Inc.).
- *NLMU* (non-local means upsampling) [6]: it is a data patch-based reconstruction.
- *LRTV* (low-rank total variation) [22]: low-rank regularization and total variation techniques were used to integrate both local and global information for image reconstruction.
- *SRCNN3D* [17]: three-dimensional convolutional neural network trained with patches of HR brain images.

The first three have been implemented in Matlab, but the convolutional network developed by *SRCNN3D*, which it is used by our proposal, has been developed using Caffe package [11] on a Python framework.
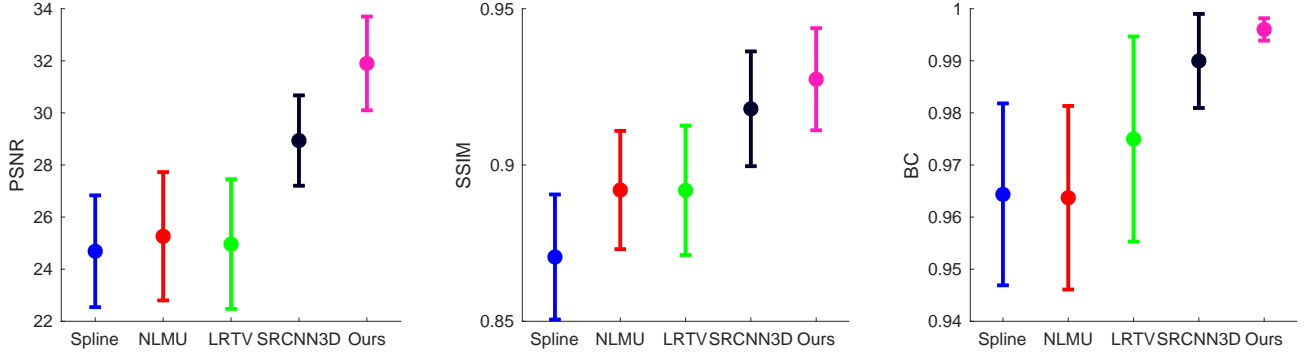
---

[1]http://mouldy.bic.mni.mcgill.ca/brainweb/

Fig. 3. Comparison of the MSE, SSIM and BC for the five methods. Mean and standard deviation of the results for all the test images are displayed, using $\lambda = 2$.

E. Results

A comparison between the evaluated methods from a quantitative point of view is summarized in Fig. 3. The graphics show the mean and standard deviation for each method in error bars, for the five tested images, using a super-resolution zoom factor of 2. Traditional methods, i.e., *Spline*, *NLMU* and *LRTV*, are clearly worse than CNN methods, both in mean and standard deviation values. There are slight differences in PSNR, but we can appreciate that the *LRTV* method overcomes the other two in SSIM and BC. An improvement of all the metrics can be seen when comparing our proposal with *SRCNN3D*. PSNR increases from 29 dB up close to 32 dB, and SSIM is also slightly improved, which indicates that the local brain structures are restored in a better way. It is important to recall that the most representative metric for this kind of images is usually SSIM, because it focuses on the morphology of the brain, and not so much on the intensity values, which, in a real application, are not so relevant for the HR image. Moreover, the high value (very close to 1) and small standard deviation we have obtained in BC are also remarkable, so we can ensure that the restored image is close to the original HR image.

The mean processing time required for the execution of each method is displayed in Fig. 4. The *Spline* method has the best performance, but, as we has explained before, the restored image lacks sufficient quality. *NLMU* and *LRTV* are in the opposite situation for time performance, as they need a lot of time to process a single image and are also very dependent on its dimensions. The higher the number of voxels, the more time they use to generate the HR image. Learning methods do not suffer too much for this reason. *SRCNN3D* processes an image in around 8 seconds. Even though our method is not the fastest one, it should be considered that a single GPU was employed when using it. If N GPUs were simultaneously employed, the total time would decrease linearly with N. In the present conditions, our method takes around 53 seconds, which is less still than *NLUM* and *LRTV*, while providing the best quantitative performance.

A three-dimensional view of the Kirby 21 (image 10) restoration is shown in Fig. 5. Slices of each plane are dis-
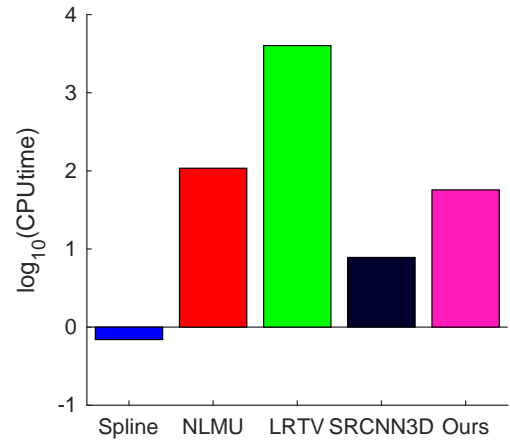

Fig. 4. Comparison of the CPU time for the five methods. Mean of the results for all the test images are displayed, in a logarithmic scale, using $\lambda = 2$.

played to evaluate the performance of each method from a qualitative point of view. First, it is clear that both *Spline* and *NLMU* generate a smoother HR image, which does not correspond with the target image. This may indicate that they depend a lot on the grade of smoothness of the input LR image. *LRTV* is also slightly different. The CNN methods yield similar results in this case.

The quantitative outcomes for the OASIS images are detailed in TABLE II. In some cases the differences are small, but it is clear that our method yields better measures in both PSNR and SSIM, which combined with the really high values of BC, demonstrate that our proposal is effective.

In Fig. 6 we can observe the notable performance of our proposal, where the real T1 image from CIMES was processed. If we focus on the residual images, we can extract two different conclusions. First, *Ours* obtains the most uniform gray residual, which indicates that the restored image has the most similar voxel intensities with respect to the original HR image. Second, if we compare internal parts of the brain where the gray intensities are different, we see that the residual
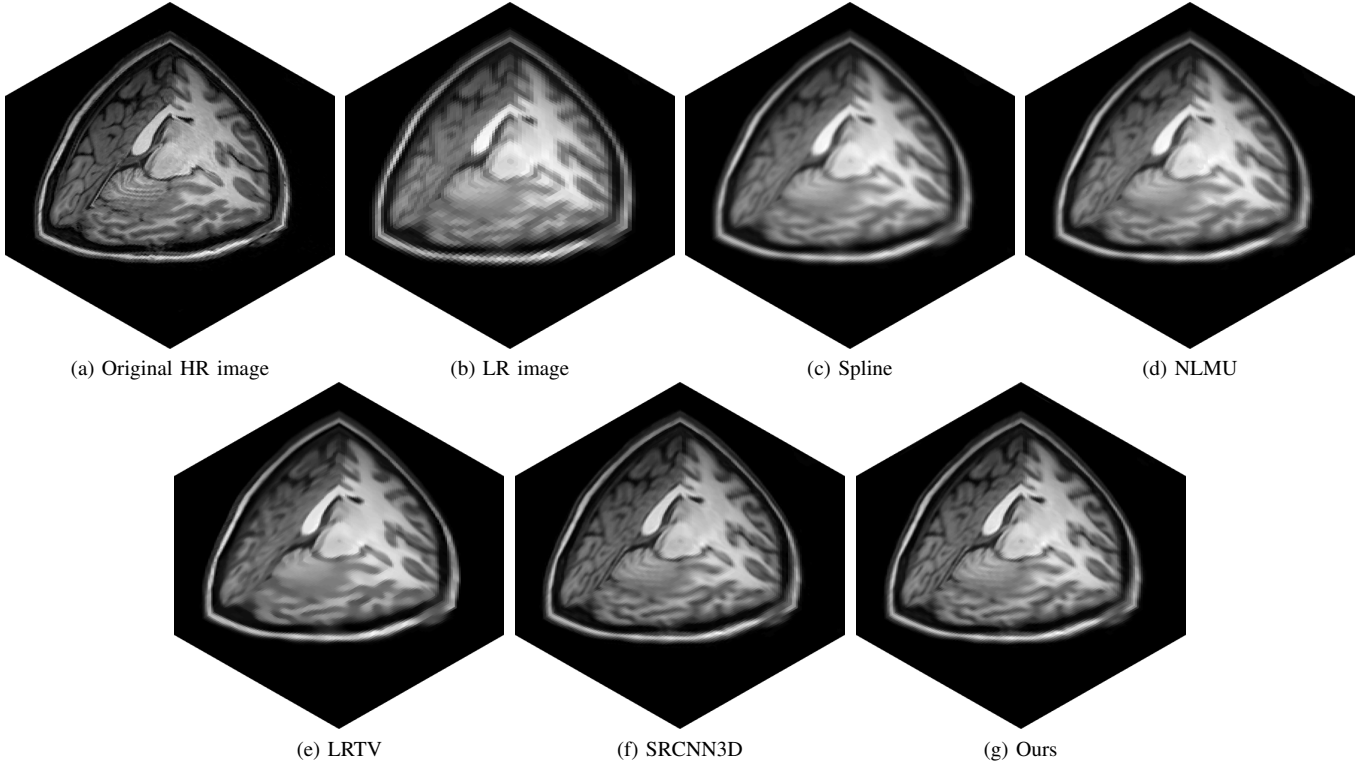
Fig. 5. Qualitative results for KKI2009-10-MPRAGE T1-weighted image for each method, applied with zoom factor 2. Three-dimensional images are shown, where the XY plane corresponds to a slice of the axial view, XZ to a slice of the sagittal view and YZ to a slice of the coronal view.

TABLE II
RESULTS OBTAINED FOR EACH METHOD FOR THE TESTED OASIS IMAGES (HIGHER IS BETTER FOR PSNR, SSIM AND BC)

IMAGE 1	PSNR	SSIM	BC	CPU time (sec.)
Spline	26.7531	0.8885	0.9658	0.6459
NLMU	27.8348	0.9108	0.9653	80.3417
LRTV	27.8079	0.9156	0.9882	3172.3286
SRCNN3D	29.0975	0.9225	**0.9926**	**0.2463**
Ours	**29.9088**	**0.9277**	0.9925	5.4251
IMAGE 2	PSNR	SSIM	BC	CPU time (sec.)
Spline	25.4974	0.9064	0.9679	0.6305
NLMU	26.3516	0.9271	0.9674	83.1330
LRTV	25.9279	0.9311	0.9876	3241.5940
SRCNN3D	28.6044	0.9370	0.9943	**0.2265**
Ours	**29.9607**	**0.9432**	**0.9949**	5.9557

images of the competing methods still show those differences, while our method does not. This implies that the other methods could distort the original brain structures and voxel intensities.

A graphical depiction of a super-resolution zoom factor 3 with IBSR_07 image is shown in Fig. 7. *Spline* still remains the worst reconstruction algorithm. Not too many differences can be appreciated between the other methods apart from the gray scale intensities in some regions of the brain. We can assure that *Ours* also works well with different zoom factors.

IV. CONCLUSION

In this paper, a method for magnetic resonance image super-resolution is presented. It is based on the combination of two different methodologies. Low-resolution images are processed through a convolutional neural network to perform an image restoration in order to obtain a high-resolution image. The quality of the restored images is increased by applying a random shifting model to the input images and then recomposing them into a consensus. A variety of images of different datasets were used to evaluate the efficiency of
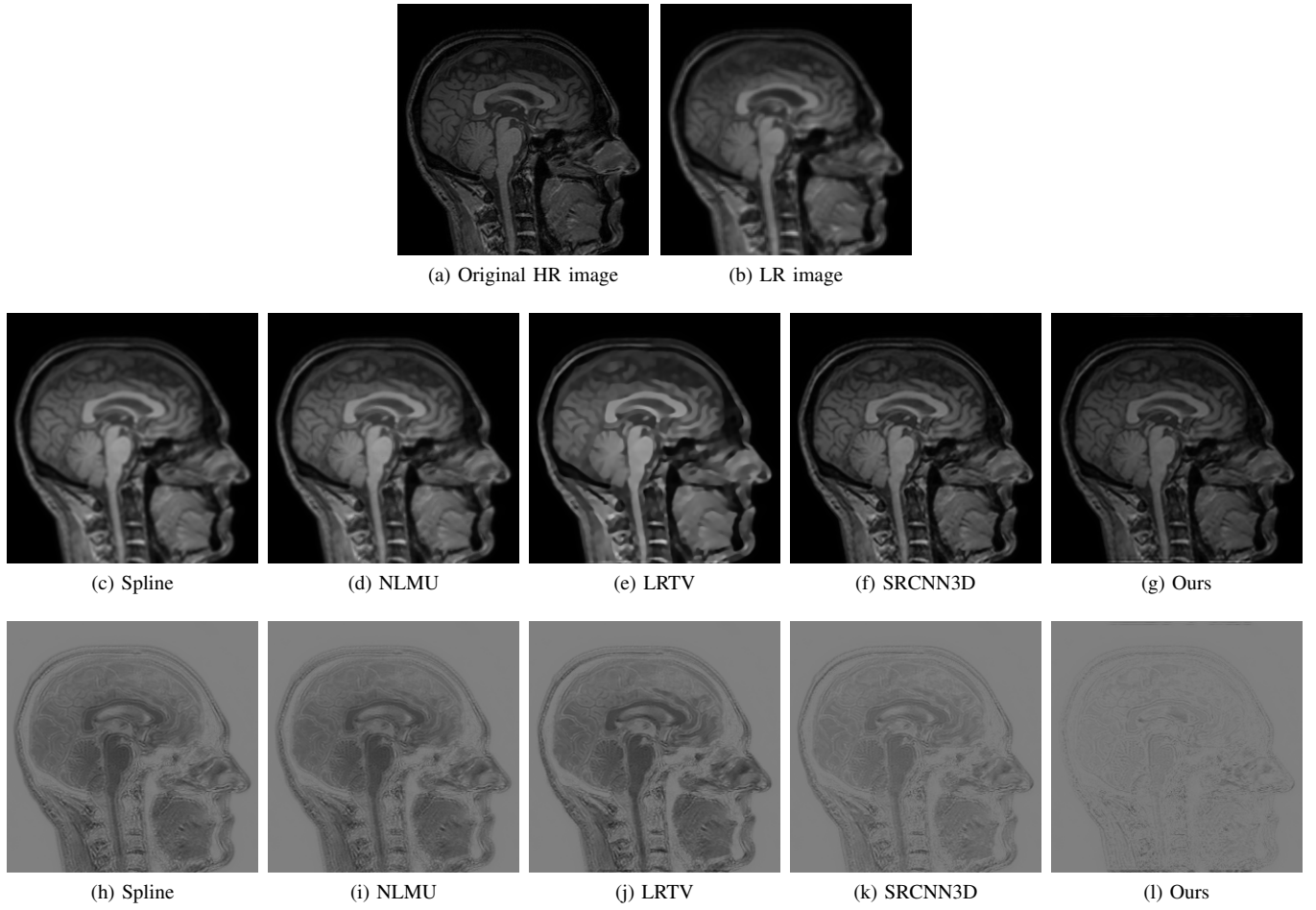
Fig. 6. Qualitative results for the T1-weighted image from CIMES for each method, applied with zoom factor 2. The second row shows the image reconstructed by each algorithm and the third row shows residual images between the reconstructed and the original HR images.

the algorithm, obtaining successful results. Quantitative results show that it overcomes other state-of-art methods, enhancing measures as SSIM, which indicates that the brain structures are not distorted. Our proposal restores the images qualitatively well for different zoom factors, avoiding over-smoothing.

REFERENCES

[1] E. López-Rubio, "Superresolution from a single noisy image by the median filter transform," *SIAM Journal on Imaging Sciences*, vol. 9, no. 1, pp. 82–115, 2016.

[2] T. M. Lehmann, C. Gonner, and K. Spitzer, "Survey: interpolation methods in medical image processing," *IEEE Transactions on Medical Imaging*, vol. 18, no. 11, pp. 1049–1075, 1999.

[3] P. Thevenaz, T. Blu, and M. Unser, "Interpolation revisited [medical images application]," *IEEE Transactions on Medical Imaging*, vol. 19, no. 7, pp. 739–758, 2000.

[4] A. J. Shah and S. B. Gupta, "Image super resolution-a survey," in *2012 1st International Conference on Emerging Technology Trends in Electronics, Communication Networking*, 2012, pp. 1–6.

[5] C. S. Balure and M. R. Kini, "A survey – super resolution techniques for multiple, single, and stereo images," in *2014 Fifth International Symposium on Electronic System Design*, 2014, pp. 215–216.

[6] J. V. Manjón, P. Coupé, A. Buades, V. Fonov, D. L. Collins, and M. Robles, "Non-local MRI upsampling." *Medical Image Analysis*, vol. 14, no. 6, pp. 784–792, 2010.

[7] H. Zheng, X. Qu, Z. Bai, Y. Liu, D. Guo, J. Dong, X. Peng, and Z. Chen, "Multi-contrast brain magnetic resonance image super-resolution using the local weight similarity," *BMC Medical Imaging*, vol. 17, no. 1, pp. 1–13, 2017.
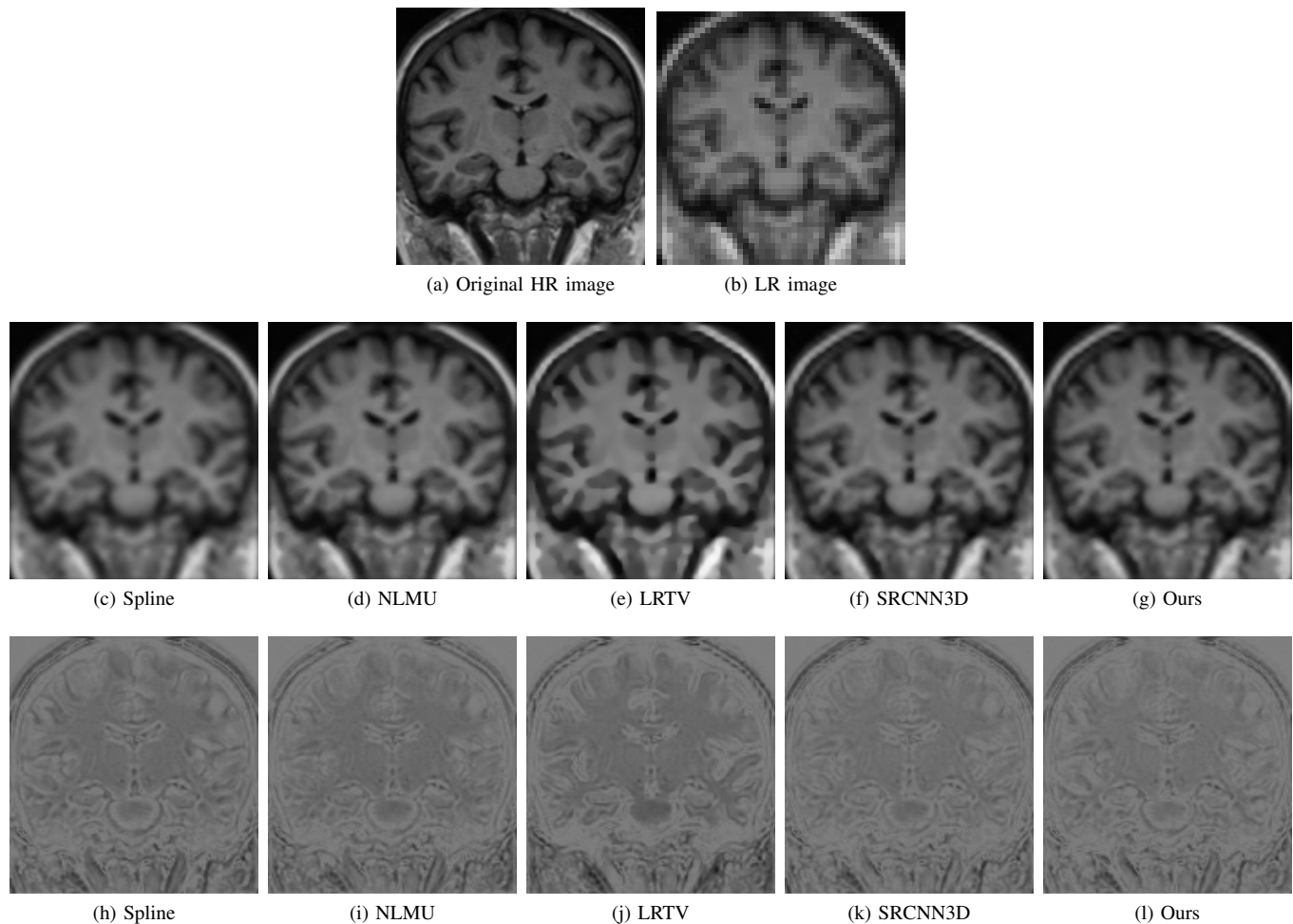
(a) Original HR image   (b) LR image

(c) Spline   (d) NLMU   (e) LRTV   (f) SRCNN3D   (g) Ours

(h) Spline   (i) NLMU   (j) LRTV   (k) SRCNN3D   (l) Ours

Fig. 7. Qualitative results for $IBSR\_07$ image for each method, applied with zoom factor 3. Second row shows the reconstructed image by each algorithm and third row shows residual images between the reconstructed and the original HR image.

[8] A. Rueda, N. Malpica, and E. Romero, "Single-image super-resolution of brain MR images using overcomplete dictionaries," *Medical Image Analysis*, vol. 17, no. 1, pp. 113–132, 2013.

[9] D. H. Trinh, M. Luong, F. Dibos, J. M. Rocchisani, C. D. Pham, and T. Q. Nguyen, "Novel example-based method for super-resolution and denoising of medical images," *IEEE Transactions on Image Processing*, vol. 23, no. 4, pp. 1882–1895, 2014.

[10] D. Zhang, J. He, Y. Zhao, and M. Du, "MR image super-resolution reconstruction using sparse representation, nonlocal similarity and sparse derivative prior," *Computers in Biology and Medicine*, vol. 58, pp. 130–145, 2015.

[11] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, and T. Darrell, "Caffe: Convolutional architecture for fast feature embedding," *arXiv preprint arXiv:1408.5093*, 2014.

[12] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 2, pp. 295–307, Feb 2016.

[13] W. Liu, Z. Wang, X. Liu, N. Zeng, Y. Liu, and F. E. Alsaadi, "A survey of deep neural network architectures and their applications," *Neurocomputing*, vol. 234, no. November 2016, pp. 11–26, 2017.

[14] M. T. McCann, K. H. Jin, and M. Unser, "Convolutional neural networks for inverse problems in imaging: A review," *IEEE Signal Processing Magazine*, vol. 34, no. 6, pp. 85–95, Nov 2017.

[15] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. van der Laak, B. van Ginneken, and C. I. Sánchez, "A survey on deep learning in medical image analysis," *Medical Image Analysis*, vol. 42, no. December 2012, pp. 60–88, 2017.

[16] L. Xiang, Y. Qiao, D. Nie, L. An, W. Lin, Q. Wang, and D. Shen, "Deep auto-context convolutional neural networks for standard-dose pet image estimation from low-dose pet/mri," *Neurocomputing*, vol. 267, pp. 406 – 416, 2017.

[17] C. H. Pham, A. Ducournau, R. Fablet, and F. Rousseau, "Brain mri super-resolution using deep 3d convolutional networks," in *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*, April 2017, pp. 197–200.

[18] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.

[19] A. Bhattacharyya, "On a measure of divergence between two multinomial populations," *Sankhy: The Indian Journal of Statistics (1933-1960)*, vol. 7, no. 4, pp. 401–406, 1946. [Online]. Available: http://www.jstor.org/stable/25047882

[20] D. S. Marcus, T. H. Wang, J. Parker, J. G. Csernansky, J. C. Morris, and R. L. Buckner, "Open access series of imaging studies (oasis): Cross-sectional MRI data in young, middle aged, nondemented, and demented older adults," *J. Cognitive Neuroscience*, vol. 19, no. 9, pp. 1498–1507, Sep. 2007. [Online]. Available: http://dx.doi.org/10.1162/jocn.2007.19.9.1498

[21] A. J. Worth, "MGH CMA internet brain segmentation repository (IBSR)," http://www.cma.mgh.harvard.edu/ibsr/, 2010.

[22] 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 Transactions on Medical Imaging*, vol. 34, no. 12, pp. 2459–2466, Dec 2015.