A Multi-Cluster Random Forests-Based Approach to Super-Resolution of Abdominal CT Images Using Deep Neural Networks

Mahdieh Akbari
Department of Biomedical Engineering, Engineering Faculty, Shahed University
Tehran, Iran
mahdyeh.akbari@shahed.ac.ir

Amir Hossein Foruzan*
Department of Biomedical Engineering, Engineering Faculty, Shahed University
Tehran, Iran
a.foruzan@shahed.ac.ir

Yen-Wei Chen
Intelligent Image Processing Lab, College of Information Science and Engineering, Ritsumeikan University,
Shiga, Japan
chen@is.ritsumei.ac.jp

***Abstract*— To enhance the resolution of abdominal computed tomography slices, a super-resolution method based on random forests is proposed. We classify input LR patches using an Auto-Encoder network, and we employ several random forest machines for the estimation of HR data. The random forest machines are trained, and they are used to reconstruct HR patches in the test phase. Compared to the conventional K-means clustering and interpolation techniques, our results improved the performance of the SR algorithm by at least 0.02 and 5.7% using the SSIM and PSNR measures. While the conventional RF-based method generated unwanted spots in the HR image, the quality of our results outperforms the standard RF-based approach. Moreover, the proposed intensity profile follows that of the original image and preserves the edges better; however, the conventional technique generates false high-frequency oscillations.**

***Keywords*— random forests, deep neural networks, auto-encoder, super-resolution, CT image.**

I. Introduction

High-resolution (HR) medical images assist physicians in obtaining accurate diagnosis and treatment planning. An HR image reveals more details of a tumor, detects small tumors, and visualizes thin vessels as well. The whole hardware and software solutions to get HR data are called Super Resolution (SR) techniques. There are challenges for the acquisition of an HR image, including hardware limitations, scanning time, and radiation hazards. Therefore, software resolutions are the preferred approach by the researchers.

A super-resolution technique is considered as an image recovery method. The classical SR methods obtain the intensity of a lost pixel by an interpolation approach using the adjacent pixel values. Typical interpolation methods are bilinear, bicubic, and nearest neighbor methods. Based on the number of input images, software SR methods are divided into single-image and multiple-image approaches. Multiple-image SR procedures provide additional information and details to reconstruct an image from a sequence of images captured with sub-pixel resolution. Multiple acquisitions of a scene are not usually possible, especially in medical images. Therefore, a significant number of researches focuses on single-image-based approaches.

Concerning single-image SR methods, they are divided into two groups. The first group is the model-based approach that solves an inverse problem. The corresponding model recovers the original image from the observations made during the degradation process, including warping, blurring, down-sampling, and noise. If there is sufficient prior knowledge of the model, we can guarantee that the original image is recovered with good approximation [1].

The second group is machine learning-based algorithms that usually consist of two phases, training and testing. The high-frequency details are reconstructed from the original image as well from the external training dataset. In the training phase, high-resolution images and their corresponding low-resolution (LR) data are divided into smaller patches that are then used to train a supervised algorithm [2]-[3]. Typical machine learning algorithms are Random Forest (RF) and Dictionary-learning methods. Random Forest is a tool with an ensemble of decision trees for classification or regression purposes. RFs are used in different branches of machine vision, such as object recognition, data classification, and single-image super-resolution [4].

Li *et al.* [5] used the feature augmented random forest and dramatically improved the results compared to conventional super-resolution RF methods. By combining random forest and coupled dictionary learning, Gu *et al.* [6] reduced noise and artifacts, increase computational speed, and image quality. Schulter *et al.* [7] proposed a fast and accurate SR algorithm using RFs. Lu *et al.* proposed a super-resolution method based on learning-based weighted RFs with nonlocal similarity structure from an external dataset [8].

In this paper, we propose a learning-based SR algorithm using random forests. Since the outcomes of RF-based super-resolution algorithms are remarkable, we employ random forests to reconstruct an HR patch from an LR input. Based on recent researches, the accuracy of learning-based techniques depends on the similarity of the training data [9]. The latest trend is employing more than one machine to learn training patterns to improve the performance of these algorithms [8]. Current approaches utilize standard features, including the variance of an image or its gradient to categorize patches. Here, we use a Deep Neural Network (DNN) to classify LR image patches without supervision that is the main novelty of the proposed method.

The remaining of the paper is organized as follows. In Section II, we give details on the proposed SR algorithm. The results and discussions are given in Section III. Section IV concludes the paper and offers future works.

II. The Proposed Method

A. *The outline of our method*

The steps involved in the proposed multi-cluster RF super-resolution method are shown in Figure 1. Our SR method includes training and test phases. After reading a CT slice of the abdomen region, it is divided into non-overlapping patches, and their LR and HR versions are prepared. In the training phase, we employ a DNN to cluster the input patches into six groups by an unsupervised scheme. There is a single random forest corresponding to each cluster that is trained by its corresponding image patches. In the test phase, input patches are clustered by the DNN, and the random forests reconstruct them. The decision of the number of clusters was performed heuristically.

B. *Preprocessing*

A Gaussian filter, with a variance of 0.5, smoothes an input LR image. Then, we reduce the dimensionality of the slice by bicubic interpolation with a factor of 0.5 to prepare the LR data. The whole slice is divided into non-overlapping patches of the size of 5×5 pixels.

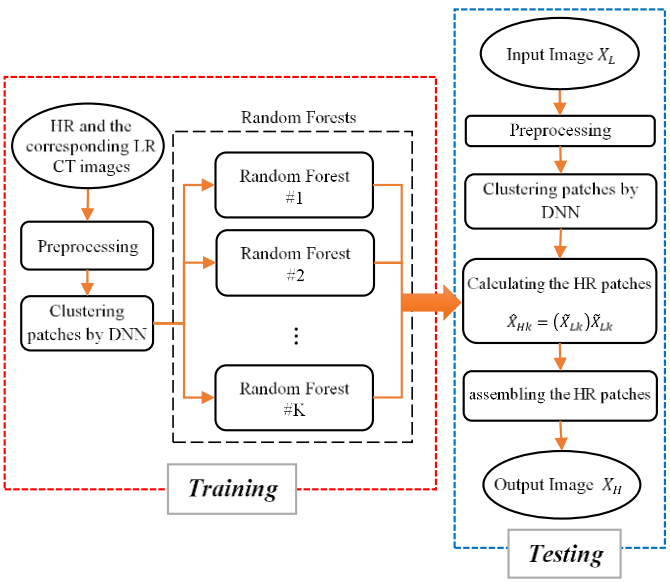

Figure 1. Flowchart of the multi-clustering RF-based super-resolution.

C. *Clustering patches based on deep neural network*

Deep neural networks are an emerging field of machine learning with diverse applications, including clustering, segmentation, registration, data augmentation, and regression [10]. In the case of clustering, DNNs can do the job both supervised and with no supervision. In the supervised case, labels of the data are given to the network together with input data. When clustering is performed with no supervision, the network itself decides on the labels of the data. Concerning medical images, it is a very tedious, boring, and difficult task to label small patches. We divide image patches into six groups using an Auto-Encoder (AE). The architecture of the network is 25 x 500 x 500 x 2000 x 6 (Figure 2). We represent a 2D 5 x 5 patch by a 25 x 1 vector, and it is normalized so that the range of the intensities is [-1, 1]. Compared to the conventional clustering techniques used in recent researches, the AE does not require any feature extraction step, and the network itself considers both low-level and high-level features simultaneously. If there is a network with m layers and ni neurons in the ith layer, the output of the jth neuron is obtained by Equation (1).

$$u_{ji} = \sigma(\mathcal{W}_{ji} \cdot \mathcal{U}_i + b_j) \tag{1}$$

In Equation (1), $\mathcal{W}_{ji}$ and b_j are the weight vector and bias corresponding to the jth neuron of the ith layer. The weights and the biases are initialized randomly (for example, by normal distribution) and optimized through an iteration scheme. The input data is passed through the network layers.

D. *Multiple Random Forest for Super-Resolution*

Random forests are ensembles of decision trees, classification trees, or regression trees. To predict response data by an RF, we must follow the decision tree from the root node down to a leaf node. It forms the tree structure by sequentially dividing the input data into two separate sub-spaces while restricting the maximum depth to ξ_{max}. The left and right child nodes are labeled as 0 and 1, respectively.

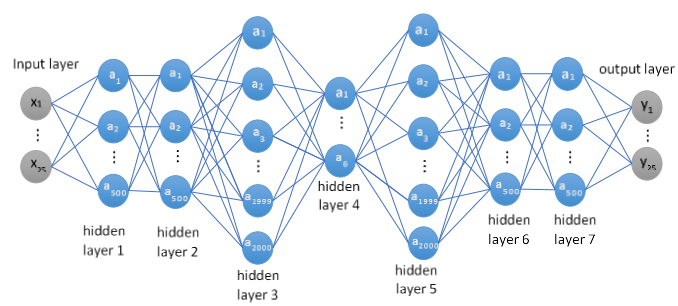

Figure 2. The architecture of the clustering DNN.

To achieve a mapping between the LR and HR patches, we use multiple random forest machine. We provide image patches form the training LR and HR dataset, and we represent the whole images by X_L and X_H matrices, respectively. The columns of the matrices, i.e. x_L and x_H include the LR and HR patch vectors, respectively, that have a size of 25 x 1. The estimation of an HR patch can be considered as a local regression problem of the corresponding LR patch, and it can be described by Equation (2).

$$X_H = W(X_L).X_L \tag{2}$$

where $W(.)$ denotes the local linear regression function. The learning error is formulated as Equation (3).

$$arg \min_{W(x_L)} \sum_{n=1}^{N} \|x_L^n - W(x_L^n).x_L^n\|_2^2 \tag{3}$$

By solving Equation (3), the linear regression function is obtained, and the super-resolution random forest is established.

The linear regression function expresses the dependencies between LR and HR patches. When the random forest is trained, the recovered SR image is estimated by Equation (4).

$$x_H = m(x_L) = W(x_L).x_L = \frac{1}{T}\sum_{t=1}^{T} m_{l(t)}(x_L) \qquad (4)$$

where T denotes the number of the decision tree, and $m_{l(t)}$ and $l(t)$ denote the local linear regression function and the leaf node of the tree t generated by x_L, respectively.

III. Results and Discussions

A. Experimental Setup

We performed several experiments on medical data for the evaluation of our algorithm. The dataset consisted of 30 abdominal CT scans in portal-vein (70-80 seconds after the injection of contrast agent). There were lesions in the images. The dataset belonged to Mr Run Shaw Hospital, Hangzhou, China [11].

We implemented the proposed method on an Intel(R) Core™ i7-4710HQ CPU @ 2.50GHz and 12GB of RAM and NVIDIA GeForce GTX 850M. The preprocessing and random forests were implemented in MATLAB R2018b environment and clustering the patches were coded in Python 3.7 using TensorFlow and Keras backend libraries.

Table 1. The (mean ± STD) of PSNR and SSIM values for 40 slices of abdominal CT images.

	Bicubic	**RF-based super-resolution using k-means clustering**	**Proposed method**
PSNR (dB)	31.95 ± 4.52	30.97 ± 6.99	**37.60 ± 3.64**
SSIM	0.87 ± 0.09	0.96 ± 0.02	**0.98 ± 0.01**

A. Evaluation Metrics

We evaluated the reconstructed SR images by Peak Signal to Noise Ratio (PSNR) and the Structural Similarity Index Measurement (SSIM) defined in Equations (5)-(7).

$$PSNR = \log\left(\frac{255^2}{MSE}\right) \qquad (5)$$

$$MSE = \frac{\left(\sum_{j=1}^{m}\sum_{i=1}^{n}\left(I_{org}(i,j) - I_{rec}(i,j)\right)^2\right)}{m \times n} \qquad (6)$$

Where I_{org} and I_{rec} are the original and the reconstructed images respectively, m and n are the height and width of the image, and MSE is the mean square error. The higher the $PSNR$ value is, the closer the reconstructed image will be to the original image. Small values of the $PSNR$ indicate that the reconstructed image has lost more high-frequency information. The SSIM is defined in Equation (6).

$$SSIM(x,y) = \frac{(2u_x u_y + C_1)(2\sigma_{xy} + C_2)}{(u_x^2 + u_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \qquad (7)$$

where u_x , u_y , σ_x and σ_y are the mean and standard deviation of x,y, respectively; σ_{xy} is the covariation of x,y; C_1, C_2 are constant that prevents division by zero and we set them to one [2]. The SSIM is used to measure the similarity between two images. The SSIM value ranges from 0 to 1.

B. Quantitative Results

Quantitative evaluation of our method included comparisons with similar state-of-the-art researches that employed random forests to estimate HR patches and conventional interpolation techniques. The results shown in Table 1 reveals that the improvements in image quality achieved by our method are considerable. Compared to the Bicubic interpolation and reconstruction based on K-means clustering, the achievement is 5.7% and 6.6% using PSNR, and 0.11 and 0.02% using SSIM, respectively. While the performance of our method is superior to the two mentioned techniques using PSNR and SSIM metrics, the SSIM results are of more importance. PSNR is a measure that is conventionally used in the assessment of SR algorithms. It is only based on the intensity difference of the original and reconstructed images and; therefore, it is sensitive to noise. However, SSIM is a more reliable index. According to Table 1, the proposed clustering approach improved the RF-based super-resolution algorithm by 0.02 using the SSIM measure. Moreover, the lower value of the standard deviation of our method reveals its stability compared to other techniques. Since the performance of an SR algorithm is different in natural and clinical images, we did not evaluate the performance of the proposed method using natural images.

In Figure 3, typical results for 10 test slices are shown. Except for the case#1, our method achieved better performance compared to a K-means based SR approach.

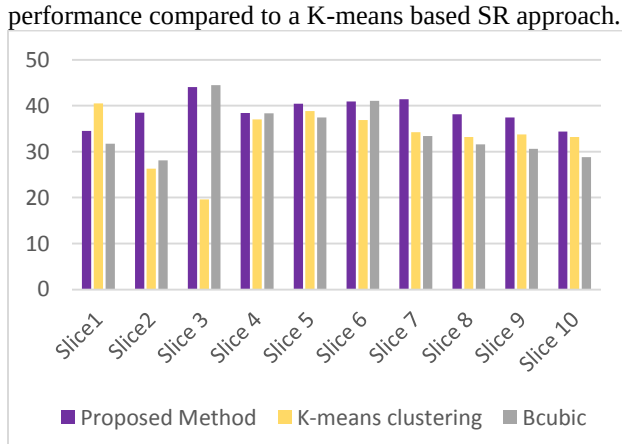

Figure 3. The proposed method often yields better results than the other method.

C. Visual Inspection

For the quantitative evaluation of our method, a typical HR and LR slice (Figure 4a and 4b, respectively), the reconstructed image by conventional RF technique (Figure 4c),

and the reconstructed image by the proposed method (Figure 4d) are shown in Figure 4. As can be seen in Figure 4c, HR patches obtained by the conventional RF technique, which uses K-means clustering, contains some white points that are not seen in the original HR data (Figure 4a). Since odd intensities in a medical image may represent abnormality, the quality of the mentioned HR image may be rejected by a physician. However, there is not such an anomaly in our results. Therefore, the outcome of our method has the potential to be employed in medical diagnosis.

A compelling tool for the assessment of the quality of a reconstructed HR image is to survey intensity profiles neighboring object boundaries. The intensity profiles of the original HR data together with the reconstructed HR images by the conventional RF technique and the proposed method are shown in Figure 5. As the black and red curves reveal, the intensity profile of our approach is more similar to the original profile, while the conventional method produces noisy high-frequency data.

To compare the clusters obtained by the K-means algorithm and the AE network, we show typical data corresponding to each of the clusters in Figure 6. For a better comparison, we prepared 20×20 patches instead of 5×5. Qualitative inspection depicts that the patches of each class are more alike in the AE-based approach compared to the K-means method. This observation proves that the enhancement of the clustering output improves the overall output of the SR method.

We used a limited number of images in the training phase due to hardware limitations, but as the number of training data is increased, the improvement in the output is boosted as well.

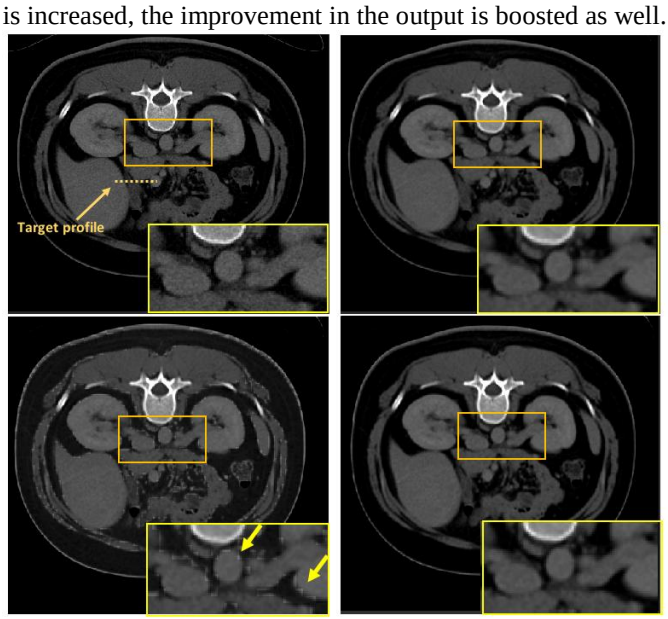

Figure 4. (a) Typical input HR Image; (b) Input LR Image (c) Reconstructed image using RFs based on K-means clustering; (d) Reconstructed image using the proposed method.

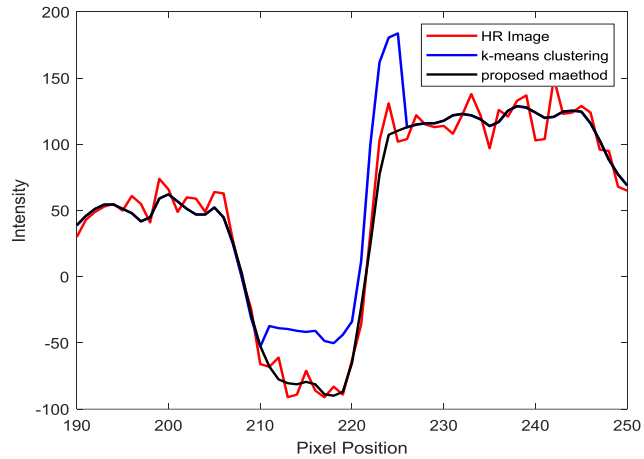

Figure 5. Intensity profiles neighboring an object boundary shown in Figure 4. The red curve: the original CT image. The blue curve: the reconstructed CT image using the conventional RF technique. The black curve: the reconstructed CT image using the proposed method.

IV. Conclusions and Future Works

In the paper, we proposed a novel SR method based on random forests for the reconstruction of abdominal CT images. We improved the clustering of the input patches and estimated HR patches of each cluster using individual random forests. The achievements of our method are significant, both qualitatively and using conventional metrics.

In the future, we decide to employ a much larger dataset and examine the effect of clustering on other learning-based SR algorithms.

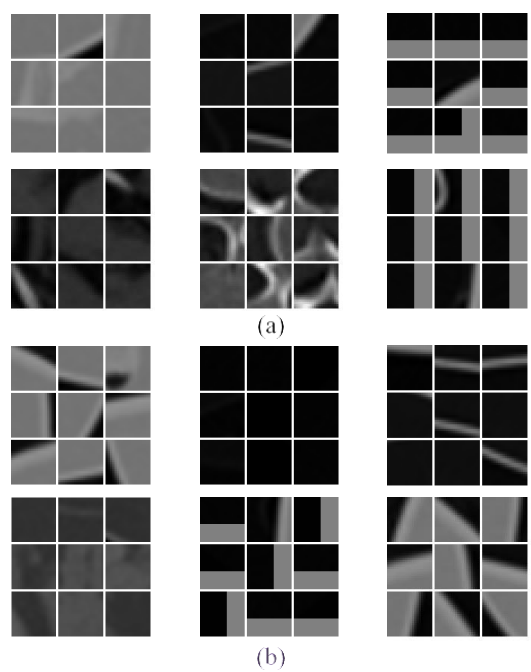

Figure 6. Clustered image patches with (a) K-means method, and (b) DNN-based technique.

REFERENCES

[1] Saafin, W., et al. "Image super-resolution from compressed sensing observations". in *2015 IEEE International Conference on Image Processing (ICIP)*. 2015. IEEE.

[2] Li, C.-m., et al., "Compressed Sensing, Pseudodictionary-Based, Superresolution Reconstruction". *Journal of Sensors., 2016.*

[3] Vishnukumar, S. and M. Wilscy, "Single image super-resolution based on compressive sensing and improved TV minimization sparse recovery". *Optics Communications.*,vol. 404: p. 80-93. 2017.

[4] Breiman, L., Random forests. Machine learning, 2001, vol.45(1): p. 5-32.

[5] Li, H., K.-M. Lam, and M. Wang, "Image super-resolution via feature-augmented random forest". *Signal Processing: Image Communication.*, vol.72: p. 25-34. 2019.

[6] Gu, P., et al., "Low-Dose Computed Tomography Image Super-Resolution Reconstruction via Random Forests". *Sensors.*, vol.19(1): p. 207. *2019.*

[7] Schulter, S., C. Leistner, and H. Bischof. "Fast and accurate image upscaling with super-resolution forests". in Proceedings of the *IEEE Conference on Computer Vision and Pattern Recognition.* 2015.

[8] Lu, Z., C. Wu, and X. Yu, "Learning Weighted Forest and Similar Structure for Image Super-Resolution". *Applied Sciences*, vol.9(3): p. 543. 2019.

[9] Zhang, K., et al., "Learning local dictionaries and similarity structures for single image super-resolution" . *Signal Processing.*, **142**: p. 231-243. 2018.

[10] Lee, J.-G., et al., Deep learning in medical imaging: general overview. Korean journal of *radiology*, vol.18(4): p. 570-584. 2017.

[11] Xu, Y., et al., Texture-specific bag of visual words model and spatial cone matching-based method for the retrieval of focal liver lesions using multiphase contrast-enhanced CT images. *International journal of computer assisted radiology and surgery.*, vol.13(1): p. 151-164. 2018.