Scientific Research Publishing

Performance Evaluation of Super-Resolution Methods Using Deep-Learning and Sparse-Coding for Improving the Image Quality of Magnified Images in Chest Radiographs

Kensuke Umehara, Junko Ota, Naoki Ishimaru, Shunsuke Ohno, Kentaro Okamoto, Takanori Suzuki, Takayuki Ishida

Department of Medical Physics and Engineering, Graduate School of Medicine, Osaka University, Suita, Japan
Email: kensuke.umehara@ieee.org, tishida@sahs.med.osaka-u.ac.jp

Abstract

Purpose: To detect small diagnostic signals such as lung nodules in chest radiographs, radiologists magnify a region-of-interest using linear interpolation methods. However, such methods tend to generate over-smoothed images with artifacts that can make interpretation difficult. The purpose of this study was to investigate the effectiveness of super-resolution methods for improving the image quality of magnified chest radiographs. **Materials and Methods**: A total of 247 chest X-rays were sampled from the JSRT database, then divided into 93 training cases with non-nodules and 154 test cases with lung nodules. We first trained two types of super-resolution methods, sparse-coding super-resolution (ScSR) and super-resolution convolutional neural network (SRCNN). With the trained super-resolution methods, the high-resolution image was then reconstructed using the super-resolution methods from a low-resolution image that was down-sampled from the original test image. We compared the image quality of the super-resolution methods and the linear interpolations (nearest neighbor and bilinear interpolations). For quantitative evaluation, we measured two image quality metrics: peak signal-to-noise ratio (PSNR) and structural similarity (SSIM). For comparative evaluation of the super-resolution methods, we measured the computation time per image. **Results**: The PSNRs and SSIMs for the ScSR and the SRCNN schemes were significantly higher than those of the linear interpolation methods ($p < 0.001$ or $p < 0.05$). The image quality differences between the super-resolution methods were not statistically significant. However, the SRCNN computation time was significantly faster than that of ScSR ($p < 0.001$). **Conclusion**: Super-resolution methods provide significantly better image quality than linear

interpolation methods for magnified chest radiograph images. Of the two tested schemes, the SRCNN scheme processed the images fastest; thus, SRCNN could be clinically superior for processing radiographs in terms of both image quality and processing speed.

Keywords

1. Introduction

Chest radiography is the most commonly performed diagnostic imaging technique for identifying various pulmonary diseases, including lung nodules, pneumonia, and pneumoconiosis. When radiologists need to verify small diagnostic signals such as lung nodules on an image, they enlarge the region-of-interest (ROI) using well-established linear interpolation methods. Such methods are commonly used for improving image resolution of a low-resolution image to generate a high-resolution image. However, the linear interpolation methods tend to generate over-smoothed images with aliasing, blur, and halo around the edges [1].

The single image super-resolution method is the post-processing approach for reconstructing a high-resolution image from a low-resolution image, and can greatly reduce artifacts resulting from linear interpolation methods. Recent super-resolution methods are example-based methods that learn the relationship between low-resolution and high-resolution image pairs. The sparse-coding super-resolution (ScSR) scheme [2] [3] is the archetypal example-based super-resolution method. Previous studies demonstrated the superiority of the ScSR method over conventional linear interpolation methods in the image quality of medical images [4] [5].

Deep-learning, also known as the deep convolutional neural network (DCNN), has recently attracted much attention in computer vision by demonstrating state-of-the-art performance in many image-based classification tasks [6] [7]. Moreover, DCNNs have been applied to image restorations such as denoising [8], inpainting [8], and deblurring [9]. The super-resolution convolutional neural network (SRCNN) [10] [11], which is an emerging deep-learning-based super-resolution method, has been proposed in computer vision. We previously demonstrated that the use of the SRCNN scheme has the potential to provide an effective approach for improving image resolution in chest radiographs [12]. However, few studies have investigated which super-resolution method is more suitable for clinical imaging applications, which require both fast processing speeds and high image quality.

In this paper, we applied and evaluated two types of super-resolution methods, *i.e.*, ScSR and SRCNN schemes for their ability to improve image quality

of magnified images in chest radiographs over that of linear interpolation. We then compared the two super-resolution methods in terms of processing speed by calculating the computation time per image.

2. Materials and Methods

2.1. Materials

A total of 247 chest radiographs were sampled from the JSRT Database, which is an open-access database created by the Japanese Society of Radiological Technology [13]. The database contained 154 cases with lung nodules and 93 cases with non-nodules. The 247 cases were divided into a training dataset comprised of the 93 cases without lung nodules, and a test dataset of the 154 cases with lung nodules.

2.2. Sparse-Coding Super-Resolution (ScSR)

Figure 1 shows an overview of the sparse-coding super-resolution (ScSR) [2] [3] scheme that we used in this study. The ScSR scheme can be divided into a training phase and a testing phase. In the training phase, two types of dictionaries, D_l and D_h, were learned from (and comprised of) the low- and high-resolution image patches, respectively, to optimize the over-complete dictionaries. The sparsest representation of a patch y of the low-resolution image can be defined as:

$$\min \|\alpha\|_0 \quad \text{s.t.} \quad \|FD_l\alpha - Fy\|_2^2 \leq \varepsilon, \tag{1}$$

where F is a feature extraction operator including four 1-D high-pass filters and α is a vector of coefficients of a sparse linear combination. However, Equation (1) is non-deterministic polynomial time-hard (NP-hard), so as long as the desired vector of coefficients α is sufficiently sparse, they can be efficiently recovered by instead minimizing the l^1-norm, as follows:

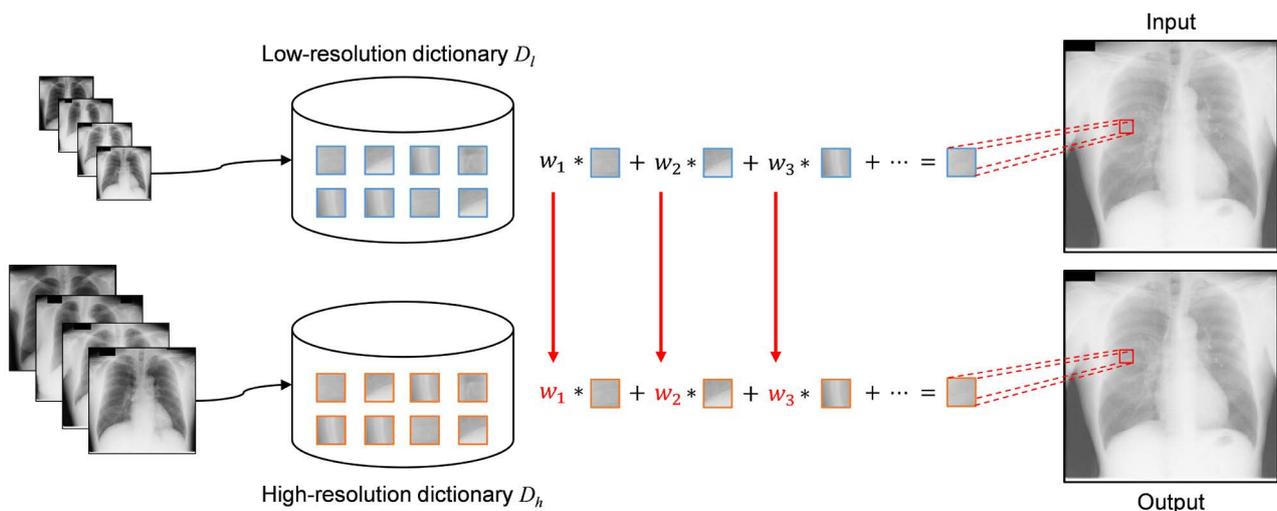$$\min \|\alpha\|_1 \quad \text{s.t.} \quad \|FD_l\alpha - Fy\|_2^2 \leq \varepsilon. \tag{2}$$

Figure 1. Overview of the sparse-coding super-resolution (ScSR) scheme.

In the testing phase, each patch of the low-resolution inputs was searched as a sparse representation of the low-resolution dictionary, represented as down-sampled image patches of high-resolution ones. The vector of coefficients of a representation of low-resolution patches α^* which is the coefficient corresponding with α, was used to generate the high-resolution output. Finally, the high-resolution output x can be reconstructed as follows:

$$x = D_h \alpha^*. \tag{3}$$

2.3. Super-Resolution Convolutional Neural Network (SRCNN)

Figure 2 shows an overview of the super-resolution convolutional neural network (SRCNN) [10] [11] scheme that we used in this study. The SRCNN scheme also has a training and testing phase; these used the same training and testing datasets, respectively, as described for the ScSR scheme. The testing phase consisted of a high-resolution image reconstructed from a low-resolution input image using the trained SRCNN model.

The SRCNN method can be divided into three parts: patch extraction and representation, non-linear mapping, and reconstruction. Patch extraction and representation refers to the first layer, which extracts patches from the low-resolution input image. The operation of the first layer is as follows:

$$F_1(\mathbf{Y}) = \max\left(0, W_1 * \mathbf{Y} + B_1\right), \tag{4}$$

where F, $\mathbf{Y}$, W_1, and B_1 represent the mapping function, the bicubic interpolated low-resolution image, the filters, and the biases, respectively.

Non-linear mapping refers to the middle layer, which maps the feature vectors non-linearly to another set of feature vectors, the high-resolution features. The operation of the middle layer is as follows:

$$F_2(\mathbf{Y}) = \max\left(0, W_2 * F_1(\mathbf{Y}) + B_2\right). \tag{5}$$

Reconstruction aggregates these high-resolution features to generate the final high-resolution image. The operation of the last layer is as follows:
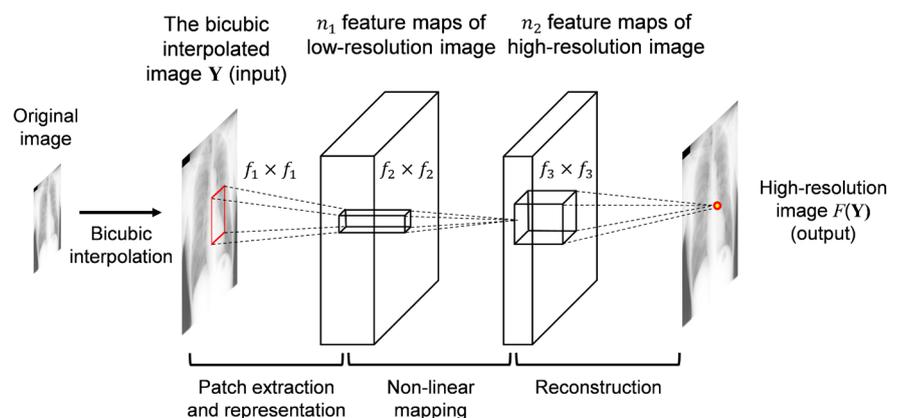

Figure 2. Overview of the super-resolution convolutional neural network (SRCNN) scheme.

$$F(\mathbf{Y}) = W_3 * F_2(\mathbf{Y}) + B_3. \tag{6}$$

2.4. Experimental Procedures

Figure 3 shows an overview of the evaluation scheme. The evaluation of super-resolution imaging is difficult because super-resolution methods estimate a high-resolution image from a low-resolution image; thus, there is no "correct" high-resolution image. Therefore, we performed an image-restoration experiment using the down-sampled original test image. Such an experiment provides a method for assessing whether the resulting high-resolution image was correctly restored or not relative to the original ROI image.

A total of 154 ROIs (matrix size: 320 × 320 pixels) centered on the nodules were cropped from each original test image. We first generated two types of low-resolution images by down-sampling. The matrix sizes of the resulting low-resolution images were 160 × 160 pixels and 80 × 80 pixels, respectively. Next, we reconstructed the high-resolution images from the down-sampled low-resolution image using the super-resolution methods to magnify by 2X or for 4X, respectively. Thus, the matrix size of the resulting high-resolution image was the same as that of the original ROI image (320 × 320 pixels). For comparative evaluation of the super-resolution and linear interpolation methods, we performed the same experiment using nearest neighbor and bilinear interpolations. Finally, we measured two image quality metrics, the peak signal-to-noise ratio (PSNR) [14] and structural similarity (SSIM) [15], using an original ROI image as the reference image. These metrics are widely used to measure image restoration quality objectively. PSNR measures image quality based on the pixel difference between two images. SSIM measures the similarity between two images to assess the perceptual image quality.

For comparative evaluation of processing speed of the super-resolution methods, we measured the computation time per image using our standard-performance computer (CPU: Intel® Core i7-4770S 3.1 GHz, RAM: 8 GB).

2.5. Statistical Analysis

The statistical significance of the differences in the image quality metrics between
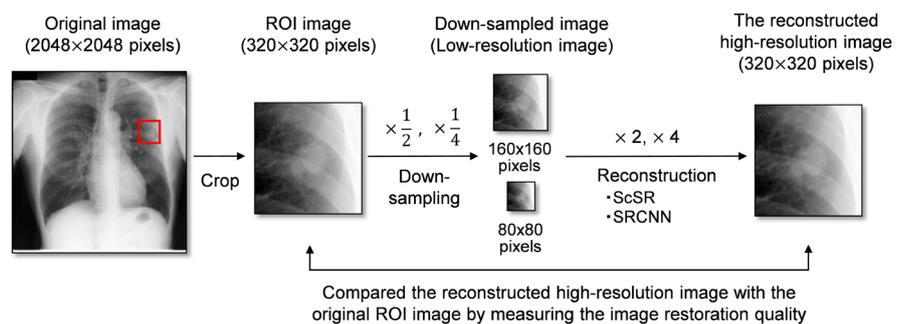

Figure 3. Overview of the evaluation scheme. Abbreviations: ROI, region of interest; ScSR, sparse-coding super-resolution; SRCNN, super-resolution convolutional neural network.

linear interpolation and super-resolution methods was analyzed by one-way analysis of variance (ANOVA) and Tukey's post-hoc test. The statistical significance of the differences in computation times was tested by Student's *t*-test. A *p*-value less than 0.05 was considered statistically significant. All statistical analyses were conducted using IBM SPSS Statistics version 22.0 (IBM Corp., Armonk, NY). Data are presented as mean ± standard deviation (SD).

3. Results

3.1. Comparison of Image Quality

Figure 4 shows the PSNRs and the SSIMs of the four schemes for 2X magnification. The means ± SDs of the nearest neighbor, bilinear, ScSR, and SRCNN methods were 39.87 ± 2.24 dB, 40.39 ± 2.32 dB, 41.56 ± 2.37 dB, and 41.79 ± 2.49 dB, respectively, of the PSNRs (**Figure 4(a)**); and 0.924 ± 0.033, 0.928 ± 0.035, 0.945 ± 0.028, and 0.947 ± 0.029, respectively, of the SSIMs (**Figure 4(b)**). **Table 1** and **Table 2** show the statistical results of the PSNR and SSIM, respectively, for
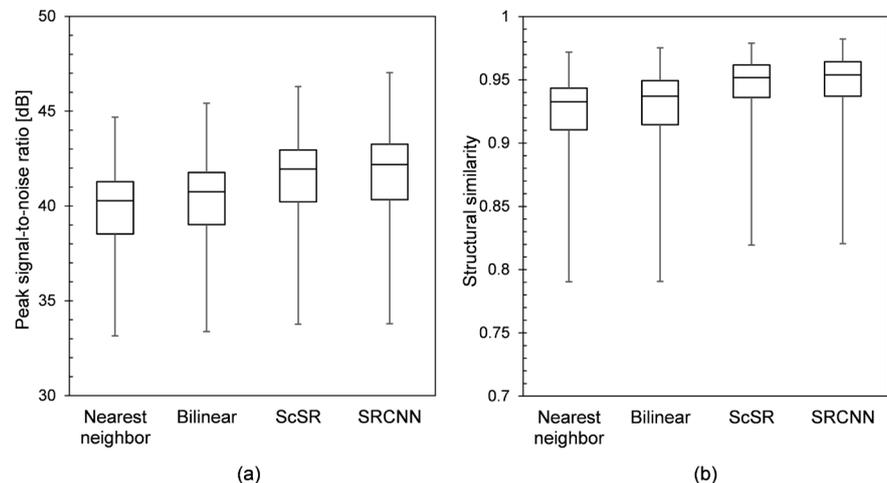

Figure 4. Comparison of the image quality of each method for 2X magnification: (a) peak signal-to-noise ratio (PSNR), (b) structural similarity (SSIM). Abbreviations: ScSR, sparse-coding super-resolution; SRCNN, super-resolution convolutional neural network.

Table 1. Comparisons of the peak signal-to-noise ratio (PSNR) in each method for 2X magnification.

(I) Method	(J) Method	Mean difference (I-J)	95% CI		*p*-value
			Lower limit	Upper limit	
Bilinear	Nearest neighbor	0.53	−0.17	1.22	0.204
ScSR	Nearest neighbor	1.69	1.00	2.39	<0.001
ScSR	Bilinear	1.17	0.47	1.86	<0.001
SRCNN	Nearest neighbor	1.92	1.23	2.62	<0.001
SRCNN	Bilinear	1.40	0.70	2.09	<0.001
SRCNN	ScSR	0.23	−0.46	0.92	0.826

Abbreviations: ScSR, sparse-coding super-resolution; SRCNN, super-resolution convolutional neural network; CI, confidence interval.

Table 2. Comparisons of the structural similarity (SSIM) between each method for 2X magnification.

(I) Method	(J) Method	Mean difference (I-J)	95% CI		p-value
			Lower limit	Upper limit	
Bilinear	Nearest neighbor	0.004	−0.005	0.013	0.709
ScSR	Nearest neighbor	0.021	0.012	0.030	<0.001
ScSR	Bilinear	0.017	0.008	0.026	<0.001
SRCNN	Nearest neighbor	0.023	0.014	0.032	<0.001
SRCNN	Bilinear	0.019	0.010	0.029	<0.001
SRCNN	ScSR	0.002	−0.007	0.011	0.937

Abbreviations: ScSR, sparse-coding super-resolution; SRCNN, super-resolution convolutional network; CI, confidence interval.

2X magnification. Briefly, the PSNR was significantly higher for super-resolution methods than for linear interpolation methods ($p < 0.001$), whereas it was not significantly different between super-resolution methods ($p = 0.826$) (Table 1). The same pattern was found for the SSIM results: Super-resolution methods were significantly better than linear interpolation methods ($p < 0.001$), but not significantly different from each other ($p = 0.937$) (Table 2).

Figure 5 shows the PSNRs and the SSIMs as above, but for 4X magnification. The PSNRs for the nearest neighbor, bilinear, ScSR, and SRCNN methods were 36.49 ± 2.11 dB, 37.78 ± 2.25 dB, 38.59 ± 2.22 dB, and 38.66 ± 2.28 dB, respectively (Figure 5(a)); the SSIMs were 0.850 ± 0.055, 0.880 ± 0.051, 0.894 ± 0.045, and 0.895 ± 0.046, respectively (Figure 5(b)). Table 3 and Table 4 present the statistical results of the image quality tests for 4X magnification. The results for 4X magnification were similar to those for 2X magnification: For both PSNR and SSIM measures, super-resolution methods were significantly better than linear interpolation methods (PSNR, $p < 0.01$; SSIM, $p < 0.05$), but not significantly different from each other (PSNR, $p = 0.992$; SSIM, $p = 0.998$).

3.2. Comparison of Computation Time

SRCNN required 1.87 ± 0.04 s and 1.85 ± 0.04 s to process 2X and 4X magnification images, respectively; ScSR required 55.83 ± 0.84 s and 53.33 ± 0.79 s, respectively. For both magnifications, SRCNN was significantly faster ($p < 0.001$).

3.3. Visual Examples

Figure 6 and Figure 7 present representative images of the resulting high-resolution images focused on the lung nodule generated by all four schemes for 2X and 4X magnifications, respectively. The super-resolution methods produced visibly sharper (higher quality) edges in comparison with the linear interpolation methods, especially for 4X magnification (Figure 7).

4. Discussion

In this study, we used two types of super-resolution schemes to improve the image
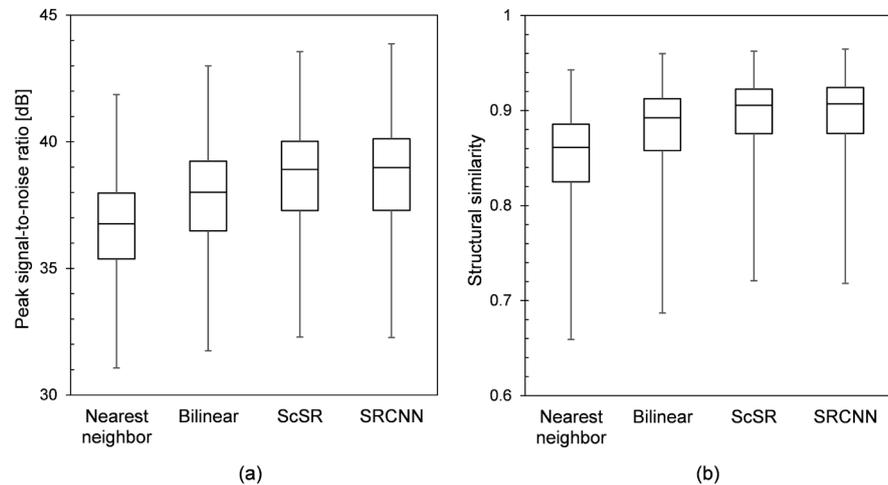
Figure 5. Comparison of the image quality of each method for 4X magnification: (a) peak signal-to-noise ratio (PSNR), (b) structural similarity (SSIM). Abbreviations: ScSR, sparse-coding super-resolution; SRCNN, super-resolution convolutional neural network.

Table 3. Comparisons of the peak signal-to-noise ratio (PSNR) in each method for 4X magnification.

(I) Method	(J) Method	Mean difference (I-J)	95% CI		*p*-value
			Lower limit	Upper limit	
Bilinear	Nearest neighbor	1.29	0.64	1.94	<0.001
ScSR	Nearest neighbor	2.10	1.45	2.75	<0.001
ScSR	Bilinear	0.81	0.16	1.46	0.007
SRCNN	Nearest neighbor	2.17	1.52	2.82	<0.001
SRCNN	Bilinear	0.88	0.23	1.53	0.003
SRCNN	ScSR	0.07	−0.58	0.72	0.992

Abbreviations: ScSR, sparse-coding super-resolution; SRCNN, super-resolution convolutional network; CI, confidence interval.

Table 4. Comparisons of the structural similarity (SSIM) between each method for 4X magnification.

(I) Method	(J) Method	Mean difference (I-J)	95% CI		*p*-value
			Lower limit	Upper limit	
Bilinear	Nearest neighbor	0.030	0.015	0.044	<0.001
ScSR	Nearest neighbor	0.044	0.030	0.059	<0.001
ScSR	Bilinear	0.015	0.0001	0.029	0.048
SRCNN	Nearest neighbor	0.045	0.031	0.060	<0.001
SRCNN	Bilinear	0.016	0.001	0.030	0.031
SRCNN	ScSR	0.0009	−0.014	0.015	0.998

Abbreviations: ScSR, sparse-coding super-resolution; SRCNN, super-resolution convolutional network; CI, confidence interval.

quality of magnified images of chest radiographs, and compared them to the commonly-used linear interpolation methods. The super-resolution schemes
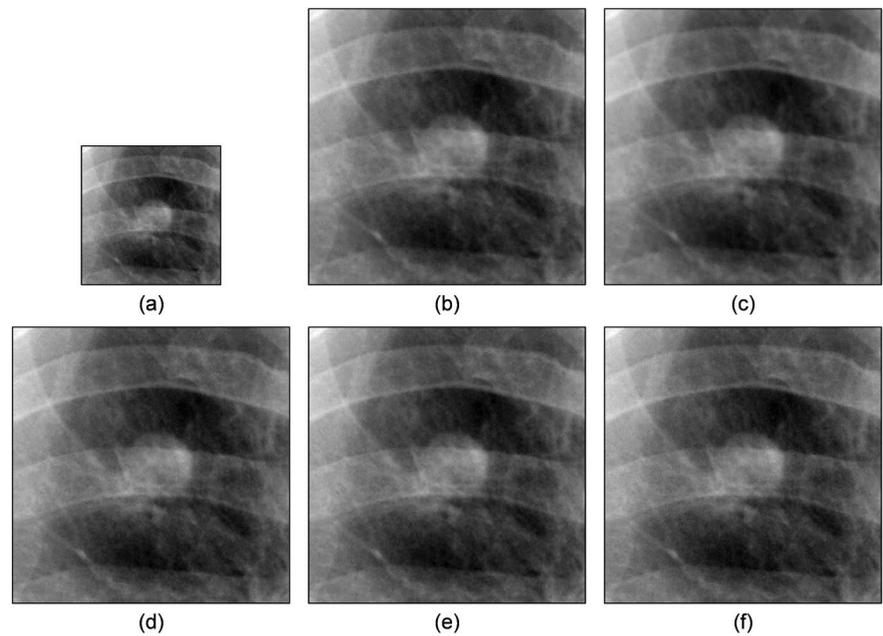
Figure 6. Representative reconstructed high-resolution images for 2X magnification: (a) down-sampled low-resolution image (matrix size: 160 × 160 pixels), (b) nearest neighbor, (c) bilinear, (d) sparse-coding super-resolution, (e) super-resolution convolutional neural network, and (f) original region of interest image (the ground-truth image, matrix size: 320 × 320 pixels).
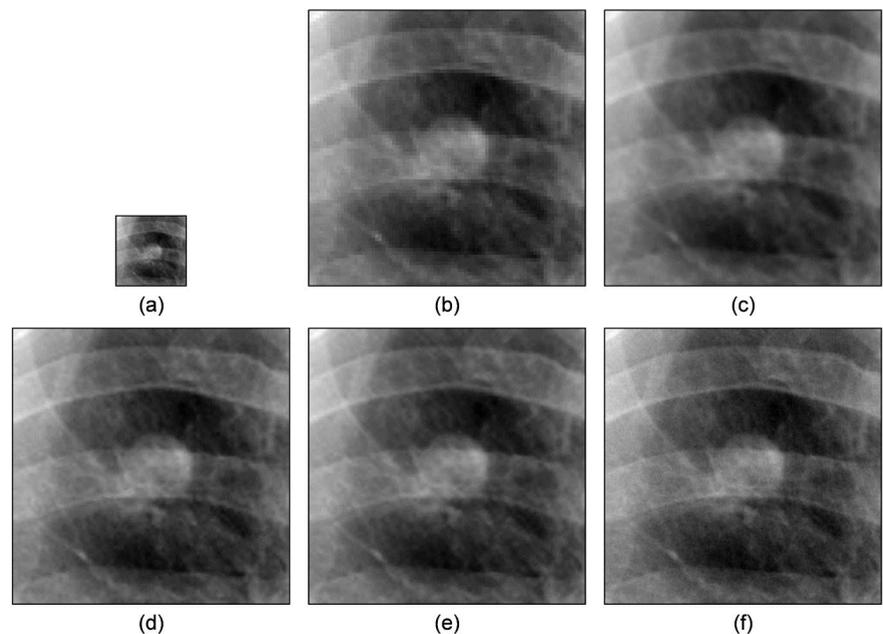

Figure 7. Representative reconstructed high-resolution images for 4X magnification: (a) down-sampled low-resolution image (matrix size: 80 × 80 pixels), (b) nearest neighbor, (c) bilinear, (d) sparse-coding super-resolution, (e) super-resolution convolutional neural network, and (f) original region of interest image (the ground-truth image, matrix size: 320 × 320 pixels).

yielded substantially higher image quality than linear interpolation methods for

both 2X and 4X magnifications for two different test metrics. However, processing (computation) speed is also important in a clinical setting, so we compared the computation times of both super-resolution schemes. We found that SRCNN, at less than 2 seconds per image, required much less computation time than ScSR.

We did compare the ScSR and SRCNN schemes in terms of image quality of the magnified images. Our experimental results on chest radiographs suggested that the SRCNN scheme yields higher image quality than the ScSR scheme, however, we saw no significant differences. Previous studies using non-medical images showed that the same pattern was found, however, statistical analysis was not performed because they used a small number of test images [10] [11]. Our experimental results herein indicate that there is little difference between the ScSR and SRCNN schemes in terms of the image quality metrics tested. It should be noted that we quantitatively evaluated image quality with our test metrics. Identifying whether the difference between these results is due to using objective instead of subjective tests, or to using chest radiographs instead of non-medical images, or a different factor altogether, will require further study.

To identify the preferred super-resolution scheme in a clinical setting, we compared the computation time between the ScSR and SRCNN schemes. Our experimental results clearly indicated that the SRCNN scheme maintains the high image quality of the super-resolution schemes, but with significantly faster processing speeds than ScSR. Thus, the SRCNN scheme provides an effective approach for the clinical application of super-resolution processing, whereas the ScSR scheme could produce delays resulting from its longer processing time. In this study, though, we measured the CPU-based run-time using a standard personal computer. If parallel processing by a GPU (graphics processing unit) can be utilized to accelerate processing speed, SRCNN is effectively capable of real-time processing, and could thus be applied not only to radiographs, but to real-time X-ray imaging as well. Further study is needed to optimize the processing speed if the potential value of SRCNN in real-time X-ray fluoroscopy is to be realized.

This study had a few limitations. In non-medical images, previous studies revealed that changing the number of layers does not result in high image quality [11]. Therefore, we used the basic and typical SRCNN settings in this study. However, to explore the optimal structure of the SRCNN scheme for use in radiographs, further study will be needed to identify the optimal network setting when using the deeper structure.

Additionally, the number of training images was relatively small. In general, deep-learning benefits from training on larger datasets. The SRCNN scheme can also deal relatively well with a larger training dataset. Therefore, the results of this study need to be confirmed in a larger dataset.

5. Conclusion

In this study, we applied and evaluated the ScSR and the SRCNN super-resolution schemes for the improvement of the image quality of magnified images in chest

radiographs. Our experimental results indicated that the super-resolution methods significantly outperformed the linear interpolation methods currently used for enhancing image resolution in chest radiographs. Our results also revealed that the SRCNN scheme provides an effective approach for clinical application of super-resolution processing to medical images due to its combination of high image quality and near-real-time processing speed.

Conflicts of Interest

The authors have no conflicts of interest directly relevant to the content of this article.

References

[1] Siu, W.C. and Hung, K.W. (2012) Review of Image Interpolation and Super-Resolution. *Proceedings of the* 2012 *Asia Pacific Signal and Information Processing Association Annual Summit and Conference*, California, 3-6 December 2012, 1-10.

[2] Yang, J., Wright, J., Huang, T. and Ma, Y. (2008) Image Super-Resolution as Sparse Representation of Raw Image Patches. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Anchorage, 23-28 June 2008, 1-8. https://doi.org/10.1109/CVPR.2008.4587647

[3] Yang, J., Wright, J., Huang, T.S. and Ma, Y. (2010) Image Super-Resolution via Sparse Representation. *IEEE Transactions on Image Processing*, **19**, 2861-2873. https://doi.org/10.1109/TIP.2010.2050625

[4] Trinh, D.H., Luong, M., Dibos, F., Rocchisani, J.M., Pham, C.D. and Nguyen, T.Q. (2014) Novel Example-Based Method for Super-Resolution and Denoising of Medical Images. *IEEE Transactions on Image Processing*, **23**, 1882-1895. https://doi.org/10.1109/TIP.2014.2308422

[5] Ota, J., Umehara, K., Ishimaru, N., Ohno, S., Okamoto, K., Suzuki, T., Shirai, N. and Ishida, T. (2017) Evaluation of the Sparse Coding Super-Resolution Method for Improving Image Quality of Up-Sampled Images in Computed Tomography. *Proceedings of SPIE Medical Imaging* 2017: *Image Processing*, **10133**, Orlando, 11-16 February, 101331S-1-101331S-9. https://doi.org/10.1117/12.2253582

[6] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V. and Rabinovich, A. (2015) Going Deeper with Convolutions. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Boston, 7-12 June 2015, 1-9. https://doi.org/10.1109/CVPR.2015.7298594

[7] Krizhevsky, A., Sutskever, I. and Hinton, G.E. (2017) ImageNet Classification with Deep Convolutional Neural Networks. *Communications of the ACM*, **60**, 84-90. https://doi.org/10.1145/3065386

[8] Xie, J., Xu, L. and Chen, E. (2012) Image Denoising and Inpainting with Deep Neural Networks. *Advances in Neural Information Processing Systems*, **25**, 341-349.

[9] Xu, L., Ren, J.S., Liu, C. and Jia, J. (2014) Deep Convolutional Neural Network for Image Deconvolution. *Advances in Neural Information Processing Systems*, **27**, 1790-1798.

[10] Dong, C., Loy, C.C., He, K. and Tang, X. (2014) Learning a Deep Convolutional Network for Image Super-Resolution. *Proceedings of European Conference on*

Computer Vision, Zurich, 6-12 September, 184-199.
https://doi.org/10.1007/978-3-319-10593-2_13

[11] Dong, C., Loy, C.C., He, K. and Tang, X. (2016) Image Super-Resolution Using Deep Convolutional Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **38**, 295-307. https://doi.org/10.1109/TPAMI.2015.2439281

[12] Umehara, K., Ota, J., Ishimaru, N., Ohno, S., Okamoto, K., Suzuki, T., Shirai, N. and Ishida, T. (2017) Super-Resolution Convolutional Neural Network for the Improvement of the Image Quality of Magnified Images in Chest Radiographs. *Proceedings of SPIE Medical Imaging* 2017: *Image Processing*, **10133**, Orlando, 11-16 February, 101331P-1-101331P-7. https://doi.org/10.1117/12.2249969

[13] Shiraishi, J., Katsuragawa, S., Ikezoe, J., Matsumoto, T., Kobayashi, T., Komatsu, K., Matsui, M., Fujita, H., Kodera, Y. and Doi, K. (2000) Development of a Digital Image Database for Chest Radiographs with and without a Lung Nodule: Receiver Operating Characteristic Analysis of Radiologists' Detection of Pulmonary Nodules. *American Journal of Roentgenology*, **174**, 71-74.
https://doi.org/10.2214/ajr.174.1.1740071

[14] Huynh-Thu, Q. and Ghanbari, M. (2008) Scope of Validity of PSNR in Image/Video Quality Assessment. *Electronics Letters*, **44**, 800-801.
http://doi.org/10.1049/el:20080522

[15] Wang, Z., Bovik, A.C., Sheikh, H.R. and Simoncelli, E.P. (2004) Image Quality Assessment: From Error Visibility to Structural Similarity. *IEEE Transactions on Image Processing*, **13**, 600-612. https://doi.org/10.1109/TIP.2003.819861