

# PROCEEDINGS OF SPIE

SPIEDigitalLibrary.org/conference-proceedings-of-spie

## Super-resolution convolutional neural network for the improvement of the image quality of magnified images in chest radiographs

Kensuke Umehara, Junko Ota, Naoki Ishimaru, Shunsuke Ohno, Kentaro Okamoto, et al.

Kensuke Umehara, Junko Ota, Naoki Ishimaru, Shunsuke Ohno, Kentaro Okamoto, Takanori Suzuki, Naoki Shirai, Takayuki Ishida, "Super-resolution convolutional neural network for the improvement of the image quality of magnified images in chest radiographs," Proc. SPIE 10133, Medical Imaging 2017: Image Processing, 101331P (24 February 2017); doi: 10.1117/12.2249969

**SPIE.**

Event: SPIE Medical Imaging, 2017, Orlando, Florida, United States

# Super-resolution convolutional neural network for the improvement of the image quality of magnified images in chest radiographs

Kensuke Umehara<sup>a</sup>, Junko Ota<sup>a</sup>, Naoki Ishimaru<sup>b</sup>, Shunsuke Ohno<sup>b</sup>, Kentaro Okamoto<sup>b</sup>, Takanori Suzuki<sup>b</sup>, Naoki Shirai<sup>b</sup>, and Takayuki Ishida<sup>a\*</sup>

<sup>a</sup>Department of Medical Physics and Engineering, Graduate School of Medicine, Osaka University, 1-7 Yamadaoka, Suita 565-0871, JAPAN;

<sup>b</sup>Course of Medical Physics and Engineering, School of Allied Health Sciences, Osaka University, 1-7 Yamadaoka, Suita 565-0871, JAPAN

## ABSTRACT

Single image super-resolution (SR) method can generate a high-resolution (HR) image from a low-resolution (LR) image by enhancing image resolution. In medical imaging, HR images are expected to have a potential to provide a more accurate diagnosis with the practical application of HR displays. In recent years, the super-resolution convolutional neural network (SRCNN), which is one of the state-of-the-art deep learning based SR methods, has proposed in computer vision. In this study, we applied and evaluated the SRCNN scheme to improve the image quality of magnified images in chest radiographs. For evaluation, a total of 247 chest X-rays were sampled from the JSRT database. The 247 chest X-rays were divided into 93 training cases with non-nodules and 152 test cases with lung nodules. The SRCNN was trained using the training dataset. With the trained SRCNN, the HR image was reconstructed from the LR one. We compared the image quality of the SRCNN and conventional image interpolation methods, nearest neighbor, bilinear and bicubic interpolations. For quantitative evaluation, we measured two image quality metrics, peak signal-to-noise ratio (PSNR) and structural similarity (SSIM). In the SRCNN scheme, PSNR and SSIM were significantly higher than those of three interpolation methods ( $p < 0.001$ ). Visual assessment confirmed that the SRCNN produced much sharper edge than conventional interpolation methods without any obvious artifacts. These preliminary results indicate that the SRCNN scheme significantly outperforms conventional interpolation algorithms for enhancing image resolution and that the use of the SRCNN can yield substantial improvement of the image quality of magnified images in chest radiographs.

**Keywords:** Super-resolution, deep learning, deep convolutional neural network, super-resolution convolutional neural network, image restoration, high-resolution, computer vision, chest radiographs

## 1. INTRODUCTION

Chest radiography is the most commonly performed diagnostic imaging technique for identifying chest diseases such as lung nodule, pneumonia, and pneumoconiosis. In recent years, the high-resolution (HR) displays are available and HR images are desired and often required. For example, HR medical images are useful for radiologists to make a correct diagnosis.<sup>1</sup> However, HR images are not often available without expensive dedicated devices. Image interpolation methods are commonly used for improving the resolution of a low-resolution (LR) image to obtain a HR one. However, conventional image interpolation methods tend to generate over-smoothed images with ringing and jagged artifacts. Therefore, it is required to reduce these artifacts by an image processing approach.

Single image super-resolution (SR) method is one of the representative approaches to suppress these artifacts. Most of the SR methods are learning-based methods<sup>2-6</sup> that learn a mapping between the LR and HR image spaces. SR method can generate a HR image from a LR image using the relationship between LR and HR images. The sparse-coding based super-resolution (ScSR) method is a well-known single SR method.<sup>7,8</sup> Previous studies demonstrated that the ScSR is useful for enhancing image resolution in medical imaging.<sup>9</sup> However, in running time, the ScSR method still remains challenging for clinical use.

\* Send correspondence to T.I.: tishida@sahs.med.osaka-u.ac.jp

Deep convolutional neural network (DCNNs) have attracted much attention in computer vision, and have successfully applied to many classification tasks. Moreover, DCNNs are applied to image restoration such as image denoising and inpainting<sup>10</sup> and deblurring.<sup>11</sup> In recent years, the super-resolution convolutional neural network (SRCNN)<sup>12</sup>, which is one of the state-of-the-art deep learning based SR methods<sup>13-17</sup>, has proposed in computer vision. The SRCNN directly learns to an end-to-end mapping between LR and HR images. Recent studies demonstrated that the SRCNN yields superior performance to the previous SR methods in processing speed and restoration quality.<sup>18</sup> In this study, we applied and evaluated the SRCNN method to improve the image quality of magnified images in chest radiographs.

## 2. METHODS

### 2.1 Materials

A total of 247 chest X-rays were sampled from the Standard Digital Image Database created by the Japanese Society of Radiological Technology.<sup>19</sup> The database contained 154 cases with lung nodules and 93 cases with non-nodules. The 247 cases divided into a training and a test datasets. The training dataset included 93 cases with non-nodules and the test dataset included 154 cases with lung-nodules.

### 2.2 Super-Resolution Convolutional Neural Network (SRCNN)

Figure 1 shows the architecture of the SRCNN. The SRCNN method can be divided into three parts, patch extraction and representation, non-linear mapping, and reconstruction. The patch extraction and representation part refers to the first layer, which extracts patches from the low-resolution input image. The operation of first layer is as follows:

$$F_1(\mathbf{Y}) = \max(0, W_1 * \mathbf{Y} + B_1) \quad (2.2.1)$$

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

The non-linear mapping part refers to the middle layer, which maps the feature vectors non-linearly to another set of feature vectors, or namely HR features. The operation of middle layer is as follows:

$$F_2(\mathbf{Y}) = \max(0, W_2 * F_1(\mathbf{Y}) + B_2) \quad (2.2.2)$$

The last reconstruction part aggregates these HR features to generate the final HR image. The operation of the last layer is as follows:

$$F(\mathbf{Y}) = W_3 * F_2(\mathbf{Y}) + B_3 \quad (2.2.3)$$

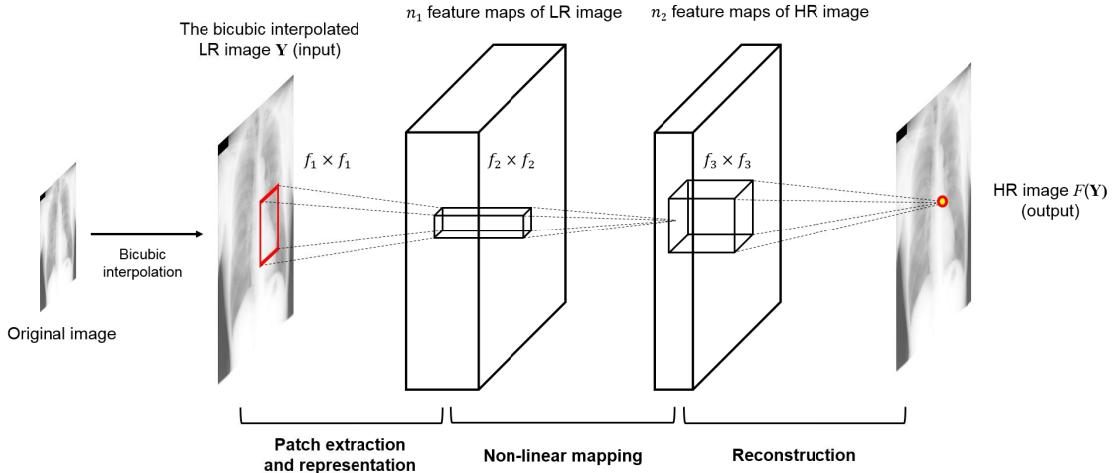


Figure 1. Overview of the SRCNN scheme. Dong *et al.* proposed that a typical and basic network setting is  $f_1 = 9$ ,  $f_2 = 1$ ,  $f_3 = 5$ ,  $n_1 = 64$ ,  $n_2 = 32$ .

## 2.3 Experiments

Figure 2 shows an overview of the evaluation scheme. In training phase, the SRCNN was trained using the training dataset, and extracted the trained filters and biases. In test phase, a total of 154 ROIs (matrix size: 320×320 pixels) centered on the nodules were cropped from each original test image. Next, the LR images were generated by down-sampling using bicubic interpolation. Finally, down-sampled images were up-sampled using conventional interpolation methods as the reference methods or the trained SRCNN. In the experiments, we compared the image quality of the trained SRCNN and three interpolation methods, nearest neighbor, bilinear, and bicubic interpolations.

For exploration of the better network settings in image quality and running time, we applied three types of network settings: (a) 9-1-5 ( $f_1 = 9, f_2 = 1, f_3 = 5, n_1 = 64, n_2 = 32$ ), (b) 9-3-5 ( $f_1 = 9, f_2 = 3, f_3 = 5, n_1 = 64, n_2 = 32$ ), (c) 9-5-5 ( $f_1 = 9, f_2 = 5, f_3 = 5, n_1 = 64, n_2 = 32$ ).

For quantitative evaluation of the HR images, we measured two image quality metrics, peak signal-to-noise ratio (PSNR) and structural similarity (SSIM). These metrics are widely used to measure the image quality objectively. PSNR measures the image quality based on the pixel difference between two images. SSIM measures the similarity between two images for assessing perceptual image quality.<sup>20</sup> The statistical significance of the differences of two image quality metrics between the SRCNN and three image interpolation methods were tested by use of one-way ANOVA and Tukey's post-hoc test.

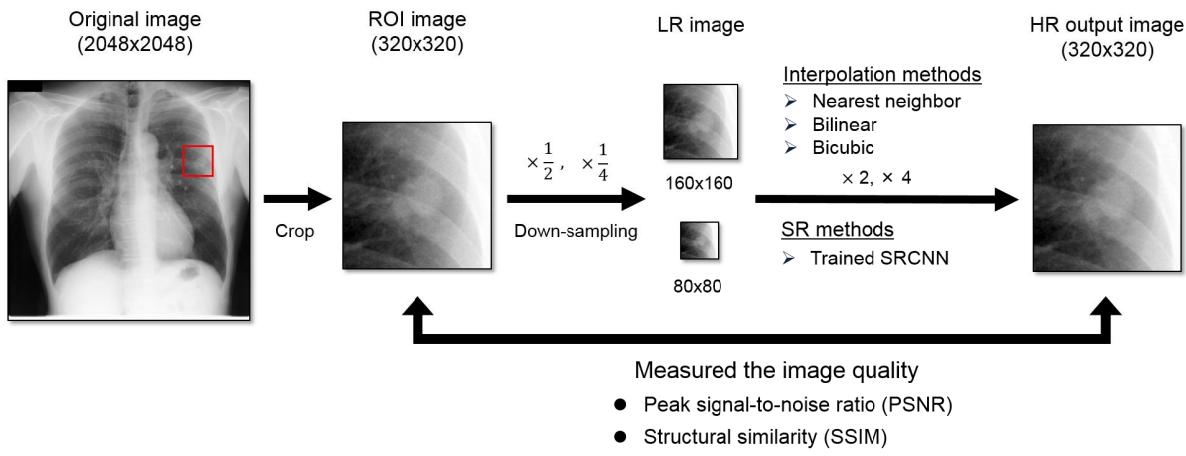


Figure 2. Overview of the evaluation scheme.

## 3. RESULTS

Figure 3 shows the comparisons of the image quality by each method for a magnification of 2. In the SRCNN scheme, the means  $\pm$  SDs of PSNR and SSIM were  $50.429 \pm 1.875$  dB,  $0.9897 \pm 0.0051$ , respectively, which were higher than those of nearest neighbor interpolation ( $48.545 \pm 1.600$  dB,  $0.9850 \pm 0.0057$ , respectively), bilinear interpolation ( $48.770 \pm 1.642$  dB,  $0.9863 \pm 0.0057$ , respectively), and bicubic interpolation ( $49.082 \pm 1.578$  dB,  $0.9867 \pm 0.0053$ , respectively). The differences in the image quality between three interpolation methods and the SRCNN scheme were statistically significant ( $p < 0.001$ ).

Figure 4 shows the comparisons of the image quality by each method for a magnification of 4. In the SRCNN scheme, the means  $\pm$  SDs of PSNR and SSIM were  $47.825 \pm 1.722$  dB,  $0.9822 \pm 0.0072$ , respectively, which were also higher than those of nearest neighbor interpolation ( $45.809 \pm 1.728$  dB,  $0.9735 \pm 0.0098$ , respectively), bilinear interpolation ( $46.867 \pm 1.698$  dB,  $0.9797 \pm 0.0080$ , respectively), and bicubic interpolation ( $47.126 \pm 1.656$  dB,  $0.9803 \pm 0.0077$ , respectively). The differences in PSNR between three interpolation methods and the SRCNN scheme were statistically significant also for a magnification of 4. The differences in SSIM between nearest neighbor, bilinear and the SRCNN were statistically significant. However, there was no statistically significant differences between bicubic and the SRCNN ( $p=0.155$ ).

Figure 5 shows an example of the magnified images by three interpolation methods and the SRCNN scheme for a magnification of 2. Figure 6 shows an example of the magnified images by three interpolation methods and the SRCNN scheme for a magnification of 4. As can be seen from these figures, the SRCNN produced much sharper edge than conventional interpolation methods without any obvious artifacts.

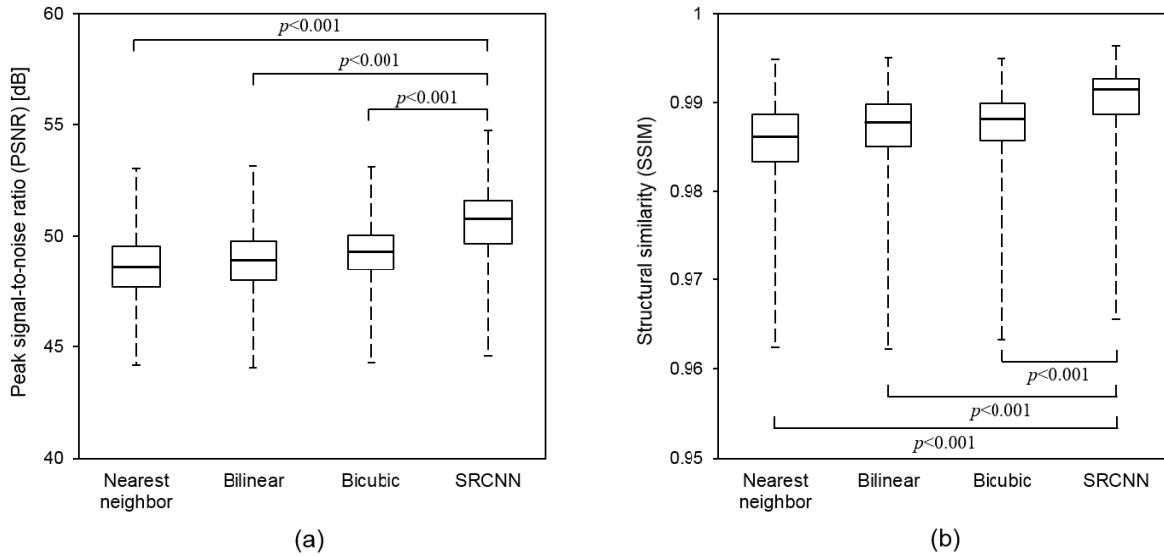


Figure 3. Comparisons of the image quality of each method for a magnification of 2. (a) Peak signal-to-noise ratio (PSNR). (b) Structural similarity (SSIM).

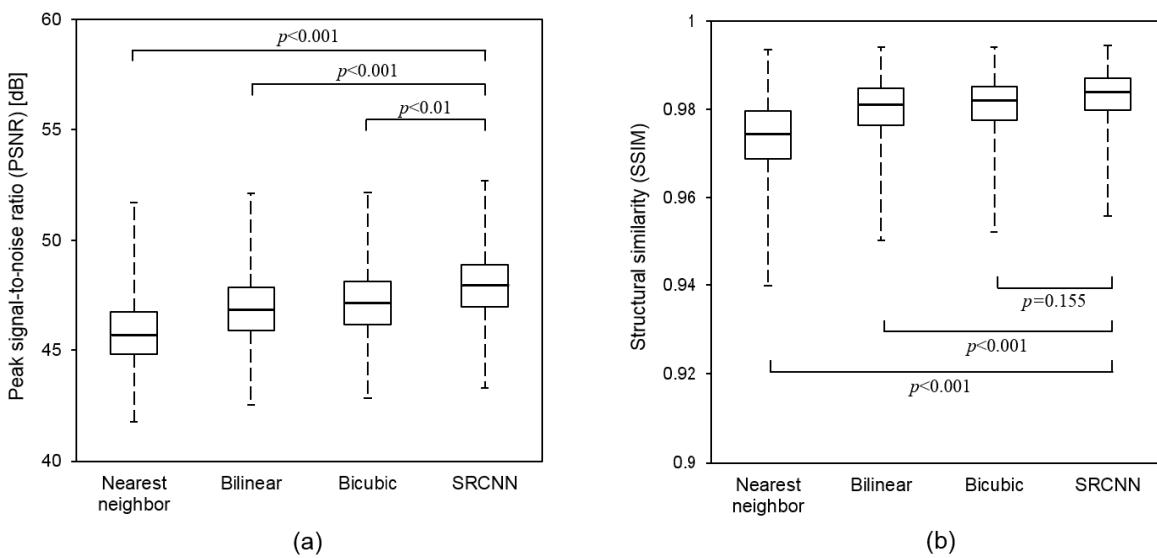


Figure 4. Comparisons of the image quality of each method for a magnification of 4. (a) Peak signal-to-noise ratio (PSNR). (b) Structural similarity (SSIM).

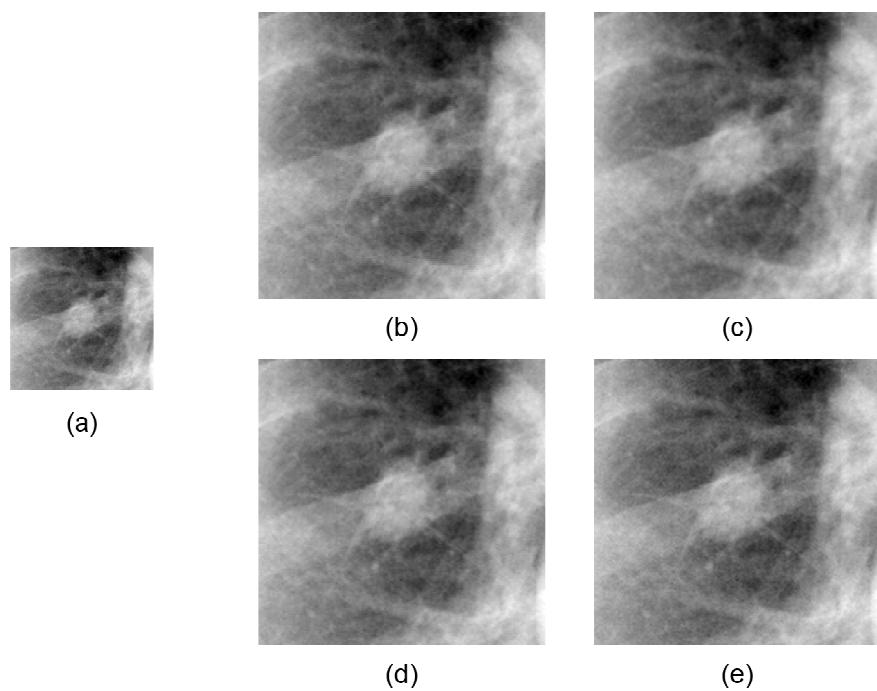


Figure 5. An example of the magnified images for a magnification of 2. (a) Input LR image. (b) Nearest neighbor. (c) Bilinear. (d) Bicubic. (e) SRCNN.

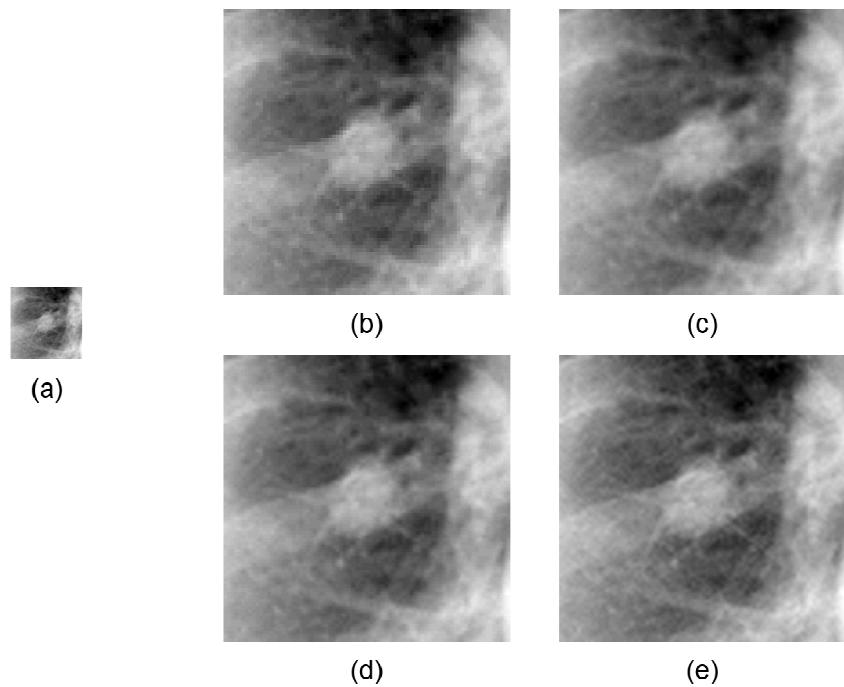


Figure 6. An example of the magnified images for a magnification of 4. (a) Input LR image. (b) Nearest neighbor. (c) Bilinear. (d) Bicubic. (e) SRCNN.

#### 4. DISCUSSION

Table 1 shows the comparisons of the image quality and the running time in different network settings for a magnification of 2, and Table 2 shows the comparisons of the image quality and the running time in different network settings for a magnification of 4. There were no statistically significant differences in PSNR and SSIM between three types of network settings, both for a magnification of 2 and for a magnification of 4.

In natural non-medical images, previous studies demonstrated that the 9-5-5 network achieves the high image quality.<sup>18</sup> In contrast, Table 1 and 2 shows that the 9-1-5 and 9-3-5 networks achieve slightly high performance over the 9-5-5 network in chest radiographs. It is possible that the network setting of the SRCNN has not a large influence on the image quality of HR images in chest radiographs.

In this study, we measured the running time using same machine (CPU: Intel Core i7-4770S 3.1 GHz, RAM: 8 GB). Table 1 and 2 shows that the 9-1-5 network is the fastest network. Previous studies in non-medical images demonstrated that the 9-1-5 network is the fastest, while the image quality achieves relatively low performance.<sup>18</sup> It is quite possible that the use of the SRCNN in chest radiographs is not always a trade-off relationship between image quality and running time.

This pilot study had a few limitations. In non-medical images, previous studies revealed that the number of layers does not result in high performance.<sup>18</sup> Therefore, we did not use the deeper structure of the SRCNN in this study. However, for exploration of the optimal structure of the SRCNN in medical imaging, further study is needed to identify the optimal network setting using the deeper structure.

Also the number of training images was relatively small. In general, deep learning benefits from the training of big data. The SRCNN scheme can also deal with a large training dataset. Therefore, it is evident that the results of this study need to be confirmed in larger population. This will be a topic for a future study.

Table 1. Comparisons of the performance in different network settings for a magnification of 2.

| SRCNN network setting | PSNR [dB] (Mean ± SD) | SSIM (Mean ± SD)       | Running time per image on average [s] |
|-----------------------|-----------------------|------------------------|---------------------------------------|
| 9-1-5                 | 50.428 ± 1.867        | <b>0.9898 ± 0.0005</b> | <b>2.324</b>                          |
| 9-3-5                 | <b>50.429 ± 1.875</b> | 0.9897 ± 0.0051        | 2.396                                 |
| 9-5-5                 | 50.404 ± 1.875        | 0.9897 ± 0.0051        | 2.588                                 |

Table 2. Comparisons of the performance in different network settings for a magnification of 4.

| SRCNN network setting | PSNR [dB] (Mean ± SD) | SSIM (Mean ± SD)       | Running time per image on average [s] |
|-----------------------|-----------------------|------------------------|---------------------------------------|
| 9-1-5                 | 47.816 ± 1.724        | 0.9822 ± 0.0072        | <b>2.329</b>                          |
| 9-3-5                 | <b>47.825 ± 1.722</b> | <b>0.9822 ± 0.0072</b> | 2.363                                 |
| 9-5-5                 | 47.823 ± 1.725        | 0.9822 ± 0.0073        | 2.457                                 |

#### 5. CONCLUSIONS

We applied and evaluated the SRCNN scheme to improve the image quality of magnified images in chest radiographs. These preliminary results indicate that the SRCNN scheme significantly outperforms conventional image interpolation methods for enhancing image resolution in chest radiographs. The use of the SRCNN scheme can yield substantial improvement of the image quality and that it has potential to provide an effective and a robust approach for clinical application of super-resolution technique in chest radiographs.

## REFERENCES

- [1] Park, SC., Park, MK., and Kang, MG., "Super-resolution image reconstruction: a technical overview," IEEE signal processing magazine, 20(3), 21-36 (2003).
- [2] Yang, CY., and Yang, MH., "Fast direct super-resolution by simple functions," Proc. of the IEEE International Conference on Computer Vision, 561-568 (2013).
- [3] Timofte, R., Smet, VD., and Gool, LV., "Anchored neighborhood regression for fast example-based super-resolution," Proc. of the IEEE International Conference on Computer Vision, 1920-1927 (2013).
- [4] Timofte, R., Smet, VD., and Gool, LV., "A+: Adjusted anchored neighborhood regression for fast super-resolution," Asian Conference on Computer Vision, 111-126 (2014).
- [5] Cui, Z., Chang, H., Shan, S., Zhong, B., Chen, X., "Deep network cascade for image super-resolution," Proc. of European Conference on Computer Vision, 49-64 (2014).
- [6] Schulter, S., Leistner, C., Bischof, H., "Fast and accurate image upscaling with super-resolution forests," Proc. of the IEEE Conference on Computer Vision and Pattern Recognition, 3791-3799 (2015).
- [7] Yang, J., Wright, J., Huang, T., and Ma, Y., "Image super-resolution as sparse representation of raw image patches," Proc. of the IEEE Conference on Computer Vision and Pattern Recognition, 1-8 (2008).
- [8] Yang, J., Wright, J., Huang, T., and Ma, Y., "Image super-resolution via sparse representation," IEEE transactions on image processing, 19 (11), 2861-2873 (2010).
- [9] Trinh, DH., Luong, M., Dibos, F., Rocchisani, JM., Pham, CD., Nguyen, TQ., "Novel example-based method for super-resolution and denoising of medical images," IEEE transactions on image processing, 23(4), 1882-1895 (2014).
- [10] Xie, J., Xu, L., Chen, E., "Image denoising and inpainting with deep neural networks," Advances in Neural Information Processing Systems, 350-358 (2012).
- [11] Xu, L., Ren, JS., Liu, C., Jia, J., "Deep convolutional neural network for image deconvolution," Advances in Neural Information Processing Systems, 1790-1798 (2014).
- [12] Dong, C., Loy, CC., He, K., and Tang, X., "Learning a deep convolutional network for image super-resolution," Proc. of European Conference on Computer Vision, 184-199 (2014)
- [13] Wang, Z., Liu, D., Yang, J., Han, W., Huang, T., "Deep networks for image super-resolution with sparse prior," Proc. of the IEEE International Conference on Computer Vision, 370-378 (2015).
- [14] Liu, D., Wang, Z., Wen, B., Yang, J., Han, W., Huang, TS., "Robust single image super-resolution via deep networks with sparse prior," IEEE Transactions on Image Processing, 25(7), 3194-3207 (2016).
- [15] Kim, J., Kwon Lee, J., Mu Lee, K., "Accurate image super-resolution using very deep convolutional networks," Proc. of the IEEE Conference on Computer Vision and Pattern Recognition, 1646-1654 (2016).
- [16] Kim, J., Kwon Lee, J., Mu Lee, K., "Deeply-recursive convolutional network for image super-resolution," Proc. of the IEEE Conference on Computer Vision and Pattern Recognition, 1637-1645 (2016).
- [17] Shi, W., Caballero, J., Huszár, F., Totz, J., Aitken, AP., Bishop, R., Wang, Z., "Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network," Proc. of the IEEE Conference on Computer Vision and Pattern Recognition, 1874-1883 (2016).
- [18] Dong, C., Loy, CC., He, K., and Tang, X., "Image super-resolution using deep convolutional networks," IEEE transactions on pattern analysis and machine intelligence, 38(2), 295-307 (2016).
- [19] Shiraishi, J., Katsuragawa, S., Ikezoe, J., Matsumoto, T., Kobayashi, T., Komatsu, K., Matsui, M., Fujita, H., Kodera, Y., and Doi, K., "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," AJR 174, 71-74 (2000).
- [20] Wang, Z., Bovik, AC., Sheikh, HR., and Simoncelli, EP., "Image quality assessment: from error visibility to structural similarity," IEEE transactions on image processing, 13(4), 600-612 (2004).