Very deep super-resolution for efficient cone-beam computed tomographic image restoration

Jae Joon Hwang [ID][1], Yun-Hoa Jung [ID][1], Bong-Hae Cho [ID][1], Min-Suk Heo [ID][2,*]

[1]*Department of Oral and Maxillofacial Radiology, School of Dentistry, Pusan National University, Yangsan, Korea*
[2]*Department of Oral and Maxillofacial Radiology and Dental Research Institute, School of Dentistry, Seoul National University, Seoul, Korea*

ABSTRACT

Purpose: As cone-beam computed tomography (CBCT) has become the most widely used 3-dimensional (3D) imaging modality in the dental field, storage space and costs for large-capacity data have become an important issue. Therefore, if 3D data can be stored at a clinically acceptable compression rate, the burden in terms of storage space and cost can be reduced and data can be managed more efficiently. In this study, a deep learning network for super-resolution was tested to restore compressed virtual CBCT images.

Materials and Methods: Virtual CBCT image data were created with a publicly available online dataset (CQ500) of multidetector computed tomography images using CBCT reconstruction software (TIGRE). A very deep super-resolution (VDSR) network was trained to restore high-resolution virtual CBCT images from the low-resolution virtual CBCT images.

Results: The images reconstructed by VDSR showed better image quality than bicubic interpolation in restored images at various scale ratios. The highest scale ratio with clinically acceptable reconstruction accuracy using VDSR was 2.1.

Conclusion: VDSR showed promising restoration accuracy in this study. In the future, it will be necessary to experiment with new deep learning algorithms and large-scale data for clinical application of this technology. *(Imaging Sci Dent 2020; 50: 331-7)*

KEY WORDS: Cone-Beam Computed Tomography; Data Compression; Radiographic Image Enhancement

Introduction

Cone-beam computed tomography (CBCT) is used in many dental hospitals and clinics for the purpose of diagnosis using 3-dimensional (3D) images. However, the storage and management of 3D data is enormously expensive in terms of time and cost because of the large file size of each scan, which may range from 500 to 1,200 MB depending on the size of the field of view. In addition, the loss of detailed information by the lossy compression of CBCT images is irreversible. Therefore, additional storage devices and technical support are required to implement effective backup strategies, incurring additional expenses and consuming additional time. Furthermore, data backup and management are not securely carried out in smaller dental clinics to the same extent as in dental hospitals.

Therefore, if CBCT data can be compressed and later restored to a clinically acceptable image quality, dental and medical institutions can reduce their amount of data storage, depending on the compression rate. Expanding this technique to the entire hospital community could save astronomical amounts of time and money.

However, there has not been much progress in research on efficient compression and reconstruction of CBCT images. Furthermore, no commercial algorithm is available that can reconstruct a compressed 3D CBCT image that is close to the original image quality in a highly reliable manner.
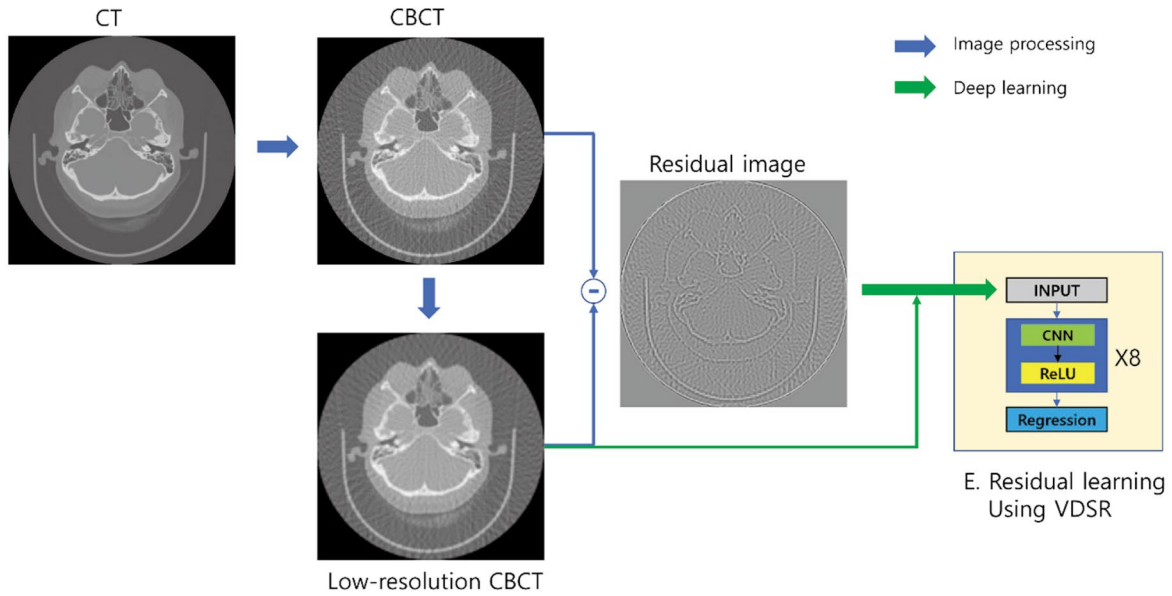
Fig. 1. Schematic view of the method used in this study. After reconstructing virtual cone-beam computed tomography (CBCT) data, the very deep super-resolution (VDSR) network learns to predict the residual image from the low-resolution image.

Super-resolution is the process of restoring high-resolution images from low-resolution images. Super-resolution is challenging because high-frequency image content typically cannot be recovered from low-resolution images. Without high-frequency information, the quality of the high-resolution image is limited. Until now, super-resolution has been attempted using traditional image processing methods, but it did not demonstrate high accuracy. Deep learning, when applied to this field, has begun to exceed the accuracy of traditional image processing methods.[1] Recently, with the use of deep network architecture and residual blocks, accuracy has rapidly increased.[2]

Very deep super-resolution (VDSR) is a convolutional neural network (CNN) designed to reconstruct the original resolution of a compressed image.[3] The VDSR network learns the mapping between the low-resolution and the high-resolution images. This mapping is possible because low-resolution and high-resolution images have similar image content and differ primarily in their high-frequency details.

For VDSR network training, a residual image is created by subtracting a downsampled low-resolution image from the original image. Next, residual learning occurs, wherein the CNN learns to estimate a residual image containing information about the high-frequency details of the residual image. No previous reports have been published on the possibility of applying the VDSR technique to the dental field, although it has shown high performance for images in general.

Using the VDSR algorithm, which has a deep CNN architecture, this study aimed to explore the possibility of using a deep learning algorithm in clinical practice to restore compressed virtual CBCT images. Through this study of an efficient and accurate image restoration algorithm, it may be possible to contribute to the efficiency and cost reduction of image storage in the dental field.

Materials and Methods

Figure 1 shows the overview of the method used in this study. The simulation was carried using MATLAB 2020a (MathWorks, Natick, MA, USA) on an i7-8700K processor.

In order to experiment with verifiable images, virtual CBCT image data were created using TIGRE software[4] (the University of Bath, Bat, UK and CERN, Geneva, Switzerland) with multidetector computed tomography (MDCT) images from the CQ500 dataset.[5] The CQ500 comprises 491 head scans with 193,317 slices that are publicly available. The TIGRE software created virtual CBCT images by initially adding noise to the projection image generated from the MDCT image, which it then back-projected using a Feldkamp-type algorithm (Fig. 1).[6] The intensity value of the created image was normalized between 0 and 1 using the equation below:

$$\text{Nomalized image} = \frac{\text{Created image} + m}{m + M}$$

where m and M refer to the minimum and maximum pixel intensity, respectively.

In order to improve the accuracy of deep learning, the size of the dataset was increased using geometric and scale augmentation. Geometric augmentation was performed by first randomly rotating images by 90°, 180°, and 270°, and next, flipping them horizontally.

By using multiple scale factors, the dataset was also augmented by scale augmentation. As the scale factor increases, restoring the original image becomes more difficult because the lower-resolution image loses more detailed image information. Scale augmentation improves the image quality at a larger compression rate because the CNN network can take advantage of the detailed information from images with a smaller compression rate. In this study, scale augmentations with factors of 2, 3, and 4 were used.

The residual image is formed from the difference between the high-resolution and downsampled low-resolution images. The VDSR network learns to predict the residual image from the low-resolution image. After the VDSR network learns to estimate the residual image, high-resolution images can be reconstructed by adding the estimated residual image to the downsampled low-resolution image.

The dataset of 12,000 virtual CBCT images was divided into training and test sets at a ratio of 8 : 2. An NVIDIA Titan RTX GPU with cuDNN version 5.1 acceleration was used for network training. The models were trained for 100 epochs with a stochastic gradient descent with a momentum optimizer (initial learning rate: 0.1, momentum: 0.9, learning rate drop period: 10).[7]

In this experiment, the peak signal-to-noise ratio (PSNR), structural similarity index (SSIM), and naturalness image quality evaluator (NIQE)[8] of at scale ratios of 2, 3, and 4 were calculated. Furthermore, the loss of image information incurred via the use of VDSR was analyzed according to various compression ratios. Larger PSNR and SSIM values indicate closer image quality relative to the reference image, whereas smaller NIQE scores indicate a perceptually better image. The definitions of PSRN, SSIM, and NIQE, which are widely used in the literature for objective comparisons of image quality, are below:

$$\text{PSNR} = 10 \log_{10}(\text{peakval}^2/\text{MSE})$$

$$\text{SSIM}(x,y) = [l(x,y)]^{\alpha} \cdot [c(x,y)]^{\beta} \cdot [s(x,y)]^{\gamma}$$

NIQE: the distance of the Gaussian model between the NSS-based features of a test image and the training dataset

Where, x and y are image coordinates, MSE is mean squared error, l is luminance, c is contrast, s is structure, and NSS refers to natural scene statistics.

In order to obtain a scale ratio similar to the clinically acceptable PNSR (36.36) of low-dose CBCT,[9] 0.1-unit intervals between compression ratios of 2 and 3 were evaluated.
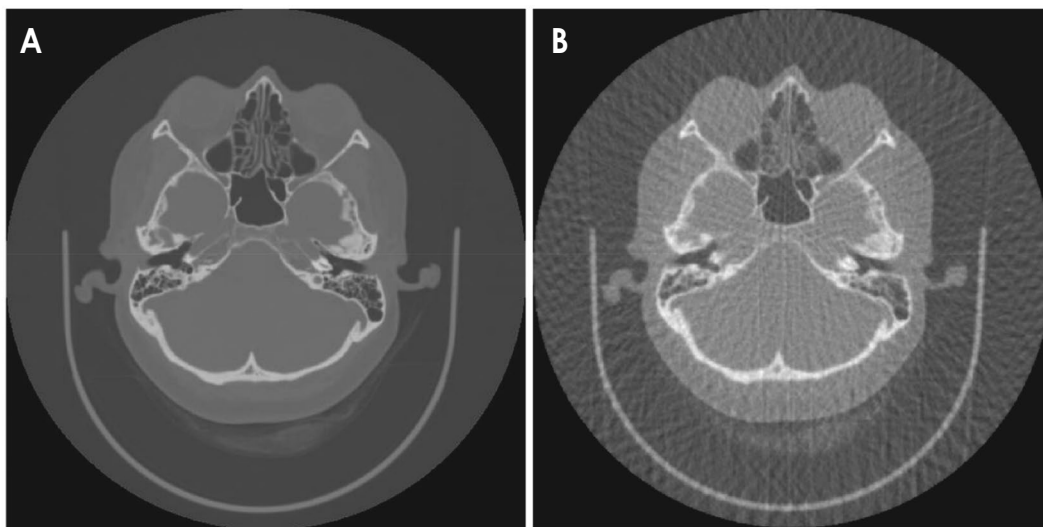

Fig. 2. Examples of a multidetector computed tomographic (MDCT) image (A) and a virtual cone-beam computed tomographic (CBCT) image (B). Virtual CBCT image made from MDCT (A) using CBCT reconstruction software (TIGRE).
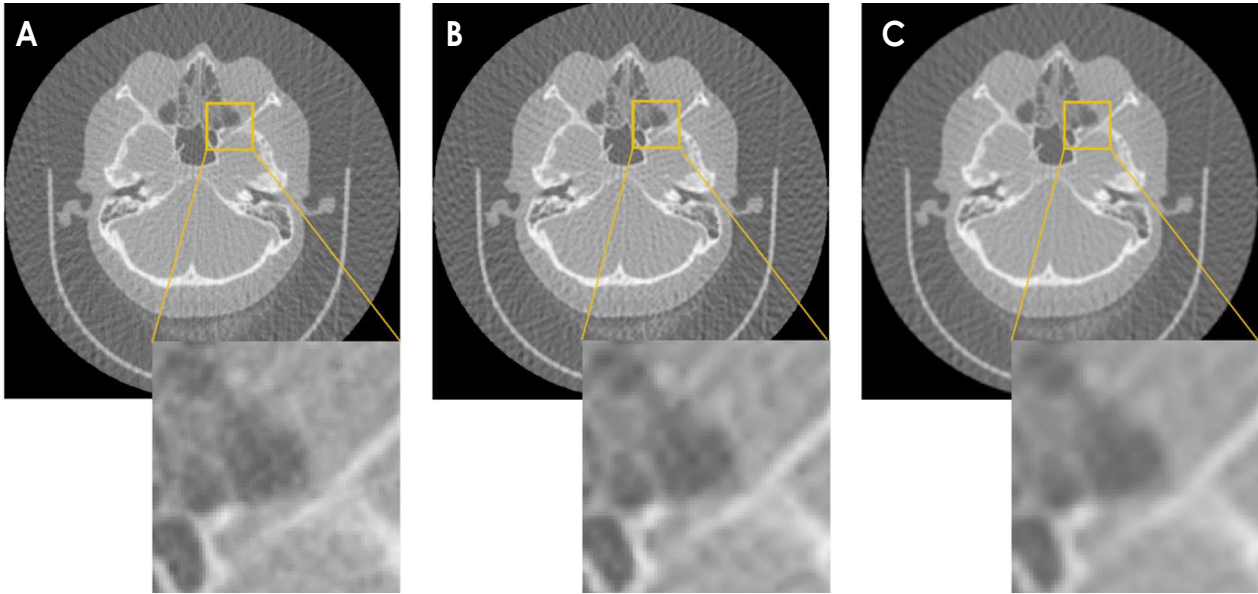
Fig. 3. High-resolution image of reference, bicubic interpolation, and very deep super-resolution (VDSR) at a scale factor of 3. A. High-resolution reference image. B. Restored high-resolution image using VDSR. C. Restored high-resolution image using bicubic interpolation. The VDSR image shows finer details and sharper edges than the image upsampled using bicubic interpolation.
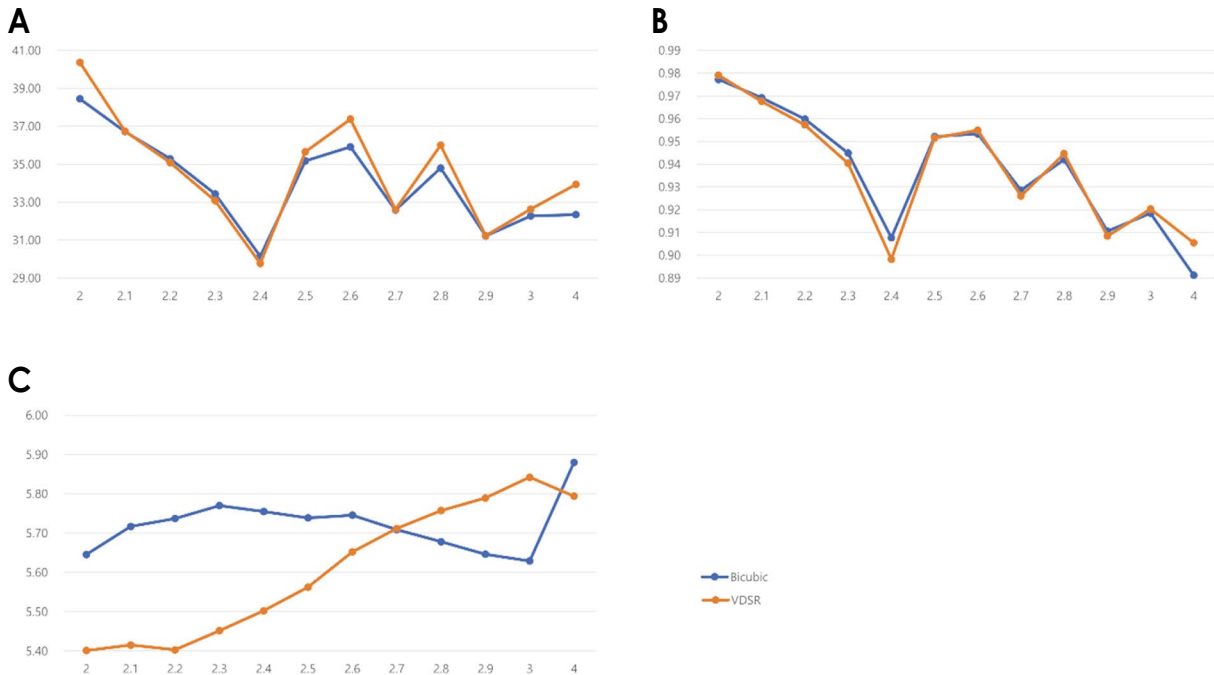

Fig. 4. Peak signal-to-noise ratio (PNSR, A), structural similarity index (SSIM, B), and naturalness image quality evaluator (NIQE, C) values of very deep super-resolution (VDSR) and bicubic interpolation at various scale factors. At most compression ratios, the VDSR of PNSR and SSIM was higher than that of bicubic interpolation. Conversely, the VDSR of NIQE tended to be lower than that of bicubic interpolation.

Results

An example of an input virtual CBCT image is shown in Figure 2. Figure 3 shows the high-resolution reference image and restoration results using VDSL and bicubic interpolation with a scale factor of 3.

At most compression ratios, the PNSR and SSIM of the images restored using VDSR were higher than the PNSR

Table 1. Image quality comparison using peak signal-to-noise ratio (PNSR), structural similarity index (SSIM), and naturalness image quality evaluator (NIQE) at various magnifications

Compression ratio	Bicubic			Very deep super-resolution		
	PSNR	SSIM	NIQE	PSNR	SSIM	NIQE
2	38.5	1.0	5.7	40.4	1.0	5.4
2.1	36.7	1.0	5.7	36.7	1.0	5.4
2.2	35.3	1.0	5.7	35.1	1.0	5.4
2.3	33.4	0.9	5.8	33.1	0.9	5.5
2.4	30.2	0.9	5.8	29.8	0.9	5.5
2.5	35.2	1.0	5.7	35.7	1.0	5.6
2.6	35.9	1.0	5.8	37.4	1.0	5.7
2.7	32.6	0.9	5.7	32.6	0.9	5.7
2.8	34.8	0.9	5.7	36.0	0.9	5.8
2.9	31.2	0.9	5.7	31.2	0.9	5.8
3	32.3	0.9	5.6	32.6	0.9	5.8
4	32.3	0.9	5.9	33.9	0.9	5.8
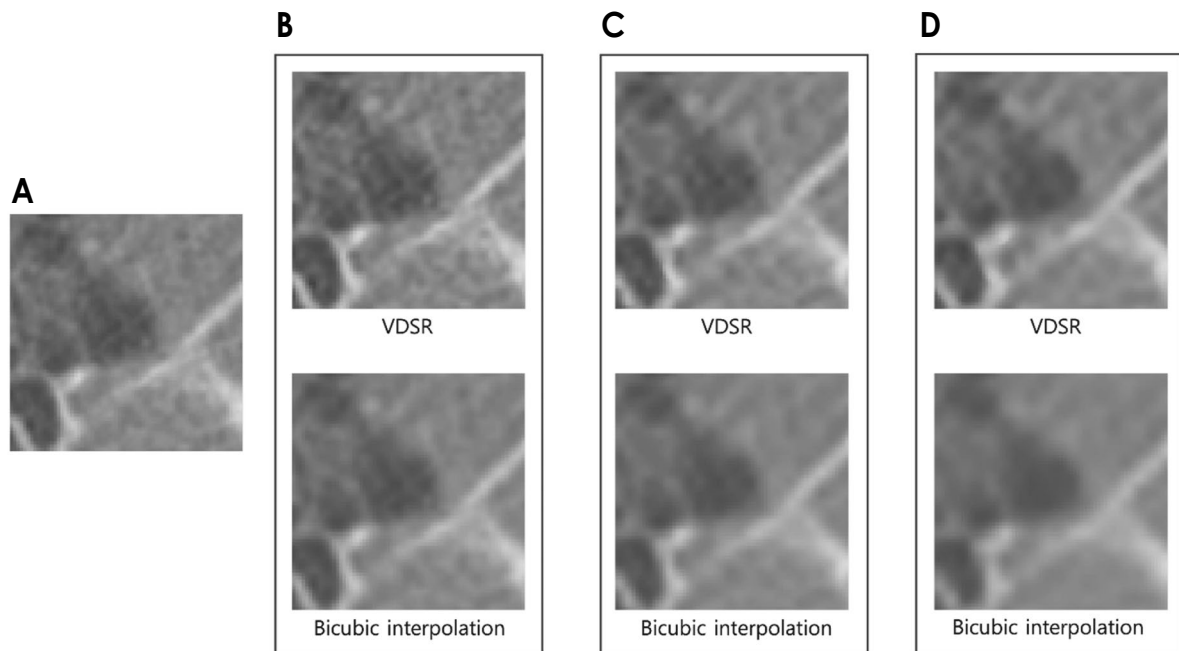
Fig. 5. High resolution image of bicubic interpolation and very deep super-resolution (VDSR) at scale factors of 2.1, 3, and 4. A. High-resolution reference image. B. 2.1× compression. C. 3× compression. D. 4× compression.

and SSIM of the images restored using bicubic interpolation. Conversely, the NIQE of the images restored using VDSR tended to be lower than the NIQE of the images restored using bicubic interpolation.

After upscaling low-resolution images with a scale factor of 2, the PNSR and SSIM of the images restored using VDSR exceeded 40 and 25, respectively. However, at a scale factor of 3, the PNSR and SSIM of the images restored using VDSR dropped to 32 and 22, respectively. This tendency was also found for the PNSR and SSIM of the images restored using bicubic interpolation. The PNSR difference of the images restored using VDSR between scale factors of 3 and 4 was insignificant compared to the PNSR difference of the images restored using VDSR between scale factors of 2 and 3.

The PNSR and SSIM of the images restored using

VDSR showed a tendency to decrease with an increasing scale factor, but did not decrease proportionally within the 2-3 interval. NIQE showed a tendency to increase with an increasing scale factor in VDSR, but there was no clear tendency for bicubic interpolation.

In this experiment, the scale factor that showed the closest PNSR of the images restored using VDSR to the clinically acceptable PNSR (36.36) was 2.1 (Table 1, Fig. 4).

The image quality difference between the images restored using bicubic interpolation or VDSR at the scale factors of 2.1, 3, and 4 is compared in Figure 5. The images restored using VDSR showed finer details and sharper edges than the images upsampled using bicubic interpolation.

Discussion

Image compression is widely used for digital photography in everyday life, but it has not been used effectively in dental and medical imaging. This is because lossy compression of the original image diminishes image quality, and there is a concern that this degradation may cause errors in interpretation.[10]

However, as 3D images become more popular in the dental field, the cost of storing and managing CBCT images, each of which exceeds 500 MB in size, is increasing. As 3D CBCT has gained popularity in the dental field, the storage and management of 3D images have become an issue. At some general hospitals, old data that exceed the capacity of the data server may be stored on magnetic tape. However, it takes time to retrieve the images stored on tape, which delays the comparison and interpretation time when trying to make a diagnosis. Therefore, if Digital Imaging and Communications in Medicine images can be compressed, stored, and restored with sufficient accuracy that there is no difference in interpretation results, the efforts and costs of storing and managing 3D images can be greatly reduced.

In this study, the CNN architecture, which has shown excellent performance in the dental field[11], was applied to achieve super-resolution of virtual CBCT images. It was demonstrated that by using VDSR, which showed high performance, the CBCT data could be reduced by up to half the original amount without clinically significant image degradation. However, a limitation of this study is that it used virtual CBCT data generated from MDCT data rather than actual CBCT data. It is necessary to verify the appropriate compression rate in large-scale CBCT data in the future. In addition, the criterion for clinically acceptable PNSR was borrowed from results obtained by analyzing low-dose CBCT. Therefore, additional studies are required to attain an acceptable PNSR during CBCT image restoration after compression. It will also be necessary to study indicators that reflect the reader's subjective satisfaction, such as changes in the mean opinion score,[2] which are relevant during the compression and restoration of CBCT data.

A higher compression rate allows a smaller storage capacity, but more time is needed to compress and restore the images. When compressing and storing 3D dental and medical images, both the accuracy of restoration and the time and computing power required for compression and restoration must be considered. In recent years, deep learning networks that can restore images in real time with high accuracy have been developed;[12] therefore, if performance verification is performed on large-scale datasets and radiologists agree that they do not perceive a difference during the interpretation process, VDSR technology can be applied in the clinical setting.

Conflicts of Interest: None

References

1. Hatvani J, Horváth A, Michetti J, Basarab A, Kouamé D, Gyöngy M. Deep learning-based super-resolution applied to dental computed tomography. IEEE Trans Radiat Plasma Med Sci 2019; 3; 120-8.
2. Yang W, Zhang X, Tian Y, Wang W, Xue J, Liao Q. Deep learning for single image super-resolution: a brief review. IEEE Trans Multimedia 2019; 21; 3106-21.
3. Kim J, Lee JK, Lee KM. Accurate image super-resolution using very deep convolutional networks. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); Las Vegas, USA: 2016. p. 1646-54.
4. Biguri A, Dosanjh M, Hancock S, Soleimani M. TIGRE: a MATLAB-GPU toolbox for CBCT image reconstruction. Biomed Phys Eng Express 2016; 2: 055010.
5. Chilamkurthy S, Ghosh R, Tanamala S, Biviji M, Campeau NG, Venugopal VK, et al. Deep learning algorithms for detection of critical findings in head CT scans: a retrospective study. Lancet 2018; 392: 2388-96.
6. Miao H, Zhao H, Gao F, Gong S. Implementation of FDK reconstruction algorithm in cone-beam CT based on the 3D Shepp-Logan model. 2009 2nd International Conference on Biomedical Engineering and Informatics; Tianjin, China: 2009. p. 1-5.
7. Dogo EM, Afolabi OJ, Nwulu NI, Twala B, Aigbavboa CO. A comparative analysis of gradient descent-based optimization algorithms on convolutional neural networks. 2018 International Conference on Computational Techniques, Electronics

and Mechanical Systems (CTEMS); Belgaum, India: 2018. p. 92-9.

8. Zvezdakova AV, Kulikov DL, Zvezdakov SV, Vatolin DS. BSQ-rate: a new approach for video-codec performance comparison and drawbacks of current solutions. Program Comput Soft 2020; 46: 183-94.

9. Song Y, Zhang W, Zhang H, Wang Q, Xiao Q, Li Z, et al. Low-dose cone-beam CT (LD-CBCT) reconstruction for image-guided radiation therapy (IGRT) by three-dimensional dual-dictionary learning. Radiat Oncol 2020; 15: 192.

10. Reddy BV, Reddy PB, Kumar PS, Reddy AS. Lossless compression of medical images for better diagnosis. 2016 IEEE 6th International Conference on Advanced Computing (IACC); Bhimavaram. India: 2016. p. 404-8.

11. Hwang JJ, Jung YH, Cho BH, Heo MS. An overview of deep learning in the field of dentistry. Imaging Sci Dent 2019; 49; 1-7.

12. Dong C, Loy CC, He K, Tang X. Image super-resolution using deep convolutional networks. IEEE Trans Pattern Anal Mach Intell 2016; 38: 295-307.