Application of super-resolution convolutional neural network technique to improve the quality of soft-tissue window cone-beam CT images

Motoki Fukuda (✉ halpop@dpc.agu.ac.jp)

Aichi Gakuin University

Yoshiko Arijii

Aichi Gakuin University

Munetaka Nitoh

Aichi Gakuin University

Michihito Nozawa

Aichi Gakuin University

Chiaki Kuwada

Aichi Gakuin University

Masako Nishiyama

Aichi Gakuin University

Takuma Funakoshi

Aichi Gakuin University

Hiroshi Fujita

Gifu University

Akitoshi Katsumata

Asahi University

Eiichiro Arijii

Aichi Gakuin University

Research Article

Keywords: dental cone-beam computed tomography, soft-tissue, deep learning, U-Net, image conversion

Posted Date: February 26th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-242807/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Objectives

To assess the feasibility of using a super-resolution convolutional neural network to improve the quality of cone-beam computed tomography (CBCT) images for visualizing soft-tissue structures.

Methods

Multidetector computed tomography (CT) images of 200 subjects who were assessed for the status of an impacted third molar were collected as training datasets. CBCT images of 10 subjects who were also examined with CT were collected as testing datasets. The training process used a modified U-Net and bone and soft-tissue window CT images. After creating a model to convert bone images to soft-tissue images, CBCT images were provided as input and the model outputted estimated CBCT images. These estimated CBCT images were then compared with soft-tissue window CBCT and CT images, using slices through approximately the same anatomical regions. Image evaluation was performed with subjective observations and histogram descriptions.

Results

The visibility of soft-tissue structures was improved by the technique, with high visibility being attained in the submandibular region, although visibility remained a little obscured in the maxillary region.

Conclusions

The feasibility of a deep learning-based super resolution technique to improve the visibility of soft-tissue structures on estimated CBCT images was verified.

Introduction

Cone-beam computed tomography for dental use (CBCT) was developed by Arai et al.¹ and Mozzo et al.² in the 1990s, and there has been much interest in how to use this innovative imaging modality to reveal three-dimensional structures of oral and maxillofacial anatomy.³⁻⁷ In addition to improving diagnostic performance, CBCT examination imparts a lower radiation dose than multidetector computed tomography (CT).⁸ However, CBCT has a critical disadvantage in comparison with CT, namely its poor differentiation of soft-tissue structures. CBCT sacrifices soft-tissue image quality and accurate knowledge of radiodensity (CT number) to achieve a low patient dose by visualizing only teeth and bone structures.⁹ Therefore, even if CBCT images are manipulated to visualize soft tissue structures by

changing the window level and width, they may still be insufficient for effective diagnosis. Consequently, for the evaluation of soft-tissue diseases, CT is more likely to be chosen than CBCT.³

Many groups have recently developed computer-aided diagnostic systems using convolutional neural network (CNNs), some of which have been applied to radiological studies of oral and maxillofacial structures.^{10–14} The various CNN types reported include those designed to improve image resolution, the so-called “super-resolution techniques”.¹⁵ These methods can be implemented during image conversion procedures to create images resembling high-dose CT images from low-dose images,^{16–26} and images resembling 7-T magnetic resonance images from 3-T images.²⁷ If CBCT images displayed using window levels and widths suitable for soft tissue structures could be improved by such a technique, they could contribute to clinical CBCT-based diagnosis.

Although there are some differences in the raw data between CT and CBCT, the reconstructed images provide almost the same image quality for bone and hard tissue structures.²⁸ Therefore, if a CNN could learn to identify differences between bone and soft-tissue in reconstructed CT images, the learning model may also be applied to CBCT images.

The present study aimed to assess the feasibility of a super-resolution technique to improve the quality of CBCT images for visualization of soft-tissue structures, doing so through the use of CT data as training datasets to create a trained model.

Results

The training process took 6 days 14 hours and 11 minutes to complete, while the testing process required 3 minutes. Examples of estimated CBCT images are shown in Figs. 3–4, together with the corresponding the soft-tissue window CBCT (swCBCT) and the 256 × 256-pixel 8-bit jpeg image compressed from the original soft-tissue window CT (compressed swCT) images.

The results of the subjective evaluation are summarized in Table 1. For all anatomical structures, the compressed swCT images showed scores of 4.0, which were equivalent to those of the original CT images. Both the swCBCT and estimated CBCT images showed lower scores than the compressed swCT images for all anatomical structures. In the comparison of the swCBCT images with the estimated CBCT images, all structures showed higher scores on the estimated CBCT images than on the swCBCT images, and the differences were statistically significant in the parapharyngeal and submandibular spaces and lymph nodes. The structures situated inferior to the mandible, such as the digastric muscle, submandibular space, and lymph nodes, showed relatively higher scores, probably because of a small amount of X-ray attenuation by surrounding bony structures.

Table 1
Mean evaluation scores for the three images

		estimated CBCT image	swCBCT image	compressed swCT image
Muscle	Medial pterygoid muscle	1.11 ± 1.08	0.96 ± 0.62	4.00
	Digastric muscle	2.56 ± 1.14	2.17 ± 0.70	4.00
Fascial space	Parapharyngeal space	0.81 ± 1.21*	0.64 ± 0.50	4.00
	Submandibular space	2.37 ± 1.15*	2.16 ± 0.82	4.00
Salivary gland	Submandibular gland	2.44 ± 1.27	2.37 ± 0.92	4.00
Lymph node	Submental or submandibular lymph node	2.37 ± 1.02*	2.13 ± 0.84	4.00
CT: computed tomography, CBCT: cone-beam CT				
swCT: soft-tissue window CT, swCBCT: soft-tissue window CBCT				
Boldfaced characters denote significant differences from the compressed swCT image				
*A significant difference was found in the scores of the estimated CBCT images relative to those of the swCBCT images (Steel-Dwass test, $p < 0.01$)				

Histograms of the voxel values for fat tissue in the submandibular space and digastric muscle in all 10 slices are shown in Figs. 6a–c. The compressed swCT images showed the highest voxel values in the digastric muscle, with the areas of fat tissue and muscle signal showing very little overlap. Comparison of the swCBCT images with the estimated CBCT images reveals that although the means of the voxel values differed significantly between the muscle and fat tissues in both images, the range and standard deviations were wide in the swCBCT images, resulting in considerable overlap in the areas of the two tissues. This would contribute to a difference in visibility between the two images.

Discussion

CBCT images are generally obtained with a lower radiation dose than multidetector CT images, and therefore their soft-tissue differentiation is poor.²⁸ Additionally, radiation is significantly attenuated by the teeth and jawbone, resulting in insufficient x-rays to create clear images. Therefore, it is usually difficult to identify soft-tissue structures situated within the mandibular arch on CBCT images.

In the present testing data (CBCT bony images), the brightness and contrast were adjusted before images were input into the testing process to create the learning model, imitating the contrast on typical bone window CT images. However, the adjusted images did not reveal the same density profiles as the CT images. Another procedure is therefore required to obtain CBCT images more approximating the density profiles of CT images.

According to the results of the subjective evaluation, the estimated CBCT images could be usable in a clinical setting, except for those of the medial pterygoid muscle and parapharyngeal space. The modified U-Net could create images with sufficient visible contrast for soft-tissue diagnosis from CBCT images, but it could not reflect the different densities of muscle and fat enclosed in the maxillary dental arch and cervical vertebrae. To the contrary, in the inferior parts of the coverage, where there was little X-ray attenuation from bony structures, soft-tissue density information was already sufficient in the original CBCT images, and the super resolution CNN technique converted them into images with more visible contrast, especially between the muscle and fat in the fascial space.

A common disadvantage of super-resolution CNNs is that the training process requires a high-capacity graphics processing unit to handle the large amount of image data. Whole CT images could not be analyzed during the training process because of the relatively low capacity of our computer system. To solve this problem, images were compressed to a lower resolution, but future improvements in processor speed could negate the requirement for such a procedure.

The present study had several limitations. First, the subjective evaluation could not exclude any observer bias. An option to overcome this problem may be to assess whether the reliability of judgments on diseases affecting soft tissue conditions improved following application of the super-resolution technique. Second, the evaluated images were compressed to 8-bit gray values from 10 bits. This might have resulted in a loss of some valuable attenuation information, but all seven radiologists showed no significant differences between the original soft-tissue window CT images and the compressed swCT images in their evaluations. To verify with more accurate CT numbers, it should be improved using some normalization processes. Third, the resultant images were difficult to reconstruct in three-dimensions because of the png format that was used. A possible solution would be to create three-dimensional images in DICOM format before inputting them into the super-resolution model. Lastly, the images in the training dataset were compressed to a smaller size to reduce the time cost in the training process, but this reduced the image resolution. To solve this problem, the deep learning machine performance should be improved.

Materials And Methods

Human rights:

The requirement for informed patient consent for inclusion of data was waived by the ethics committee because of the retrospective nature of the data use and this study obtained ethical approval from Aichi-Gakuin University ethics committee (No.

496). All methods were carried out in accordance with relevant guidelines and regulations.

Preparation of datasets

A total of 60 904 original CT images including 30 452 bone window images and their corresponding 30 452 soft-tissue window images acquired from 200 patients, between June 2019 and December 2019,

were selected from our hospital image database for the model learning process. All CT examinations were performed to evaluate the status of an impacted third molar. Cases with severe inflammation were excluded. When these images were downloaded, the window level and width were set at 900 and 4500 HU, respectively, for the bony-window images, and 60 and 300 HU, respectively, for the soft-tissue-window images. The images were downloaded in Joint Photographic Experts Group (jpeg) format with a resolution of 900×900 pixels, and the image patches for the learning process were created by compressing them to 256×256 pixels with 8-bit gray values. Of the 60 904 patches obtained, 54 740 patches (from 180 patients) were assigned as training data, and 6164 patches (from 20 patients) were assigned as validation data (Fig. 1). CT images were acquired using an Aquilion PRIME scanner (Canon Medical Systems, Otawara, Japan) using the following parameters: tube voltage, 120 kV; tube current-time product, 100 mAs; slice thickness, 0.5 mm; field of view, 20 cm.

CBCT images of 10 patients (4272 images) who were also examined with CT were selected from the same image database and prepared for the testing process (Fig. 1). In all 10 patients, the CBCT examinations were performed to clarify the relationship between the mandibular third molar and canal before extraction, whereas the CT was acquired to evaluate post-extraction mandibular nerve damage and inflammation. The two examinations were carried out within 2 years of each other. No severe inflammation, which can affect soft-tissue findings, was observed in the CT images. Each downloaded CBCT image (1039×1264 pixels with 8-bit gray level values in jpeg format) was compressed to 256×256 pixels for use as an image patch. To imitate bone window CT images, the brightness and contrast of all the CBCT images were manually adjusted before they were input into the testing process. The CBCT scans were acquired using an Alphard Vega scanner (Asahi Roentgen Ind. Co. Ltd., Kyoto, Japan) with a field of view of 102×102 mm and a voxel size of 0.2 mm^3 .

All image processing was performed using IrfanView version 4.44 (<http://www.Irfanview.com>).

Learning architecture and processes

Training and testing processes were performed using Neural Network Console (Sony Corporation, Tokyo, Japan) with a Geforce 1080 Ti graphics processing unit (Nvidia, Santa Clara, CA). The learning method used a modification of the U-Net CNN reported by Ronneberger et al.²⁹ The network consisted of a convolutional layer, rectified linear unit (ReLU) activation function layer, and pooling layer (Fig. 2). The training parameters were: learning epochs, 300; initial learning rate, 0.001; solver type, Adam. The trained model was then used to convert the testing CBCT image datasets to soft-tissue quality CBCT images in the Portable Network Graphics (png) format. These images are referred to as estimated CBCT images.

Subjective evaluation of the quality of the estimated CBCT images

The image quality of the estimated CBCT images (Figs. 3a and 4a) was subjectively evaluated on a personal computer display by seven radiologists, all of whom had more than 3 years of experience in interpretation of CT and CBCT images. The radiologists compared the estimated CBCT images with both

the test CBCT image expressed with an appropriate window level and width for visualizing soft-tissue structures, which we refer to as the swCBCT image (Figs. 3b and 4b), and the compressed swCT image (Figs. 3c and 4c). The quality of these images was scored using a five-point grading system relative to the original soft-tissue window CT images presented on a DICOM display. In the actual evaluations, the observation windows of the displays were manually adjusted to optimize the visualization of six anatomical structures, including the medial pterygoid and digastric muscles, parapharyngeal and submandibular spaces, submandibular gland, and submental or submandibular lymph nodes. For the evaluation of the fascial space, the radiologists paid special attention to the visibility of the included fat tissue. For lymph nodes, the node to be evaluated was indicated beforehand on the images. The subjective scoring was performed according to the following procedure:

Score 0: The anatomical structure was difficult to identify.

Score 1: Between scores 0 and 2.

Score 2: The anatomical structure was sufficiently identifiable for use in a clinical setting, but the quality was inferior to the original soft-tissue window CT image displayed on a DICOM viewer.

Score 3: Between scores 2 and 4.

Score 4: The anatomical structure was clearly identified and the quality was equivalent to the original soft-tissue window CT image displayed on a DICOM viewer.

The means and standard deviations of the subjective scores were calculated for 10 patients, and the differences between image types were assessed using the Steel-Dwass test with statistical significance of $p < 0.01$.

Visibility of the digastric muscle relative to the fat tissue in the submandibular space

The visibility of a soft tissue partially depends on the contrast between the target tissue and adjacent tissues. Therefore, to verify the visibility judgments, the voxel values of the anterior belly of the digastric muscle and the adjacent fat tissue in the submandibular space were measured on a slice of each of the three image types that were subjectively evaluated (estimated CBCT, swCBCT, and compressed swCT images) in the 10 patients. The most appropriate slices showing the maximum area of the muscle were selected by a radiologist (MF), and 160-pixel circular regions of interest (ROIs) were set in the bilateral muscles and adjacent fat tissues (Fig. 5). For the estimated CBCT and compressed swCT images, the window level and width were maintained at the same values used when they were created. The windowing of the swCBCT images was determined by a radiologist (MF), so that sufficiently high contrast was shown between the two tissues. The measured voxel values using ImageJ software^{30–31} for the muscle and fat tissues were totaled for all 10 subjects.

Conclusions

Although the resultant super-resolution CBCT images are not currently of the same quality as multidetector CT images, the feasibility of the super resolution CNN model was verified, it may have a possibility to use in the clinical situation. Our next research directions are to improve the image quality and to apply the super-resolution technique to diseases affecting soft-tissue conditions.

Declarations

Acknowledgements

We thank Karl Embleton, PhD, from Edanz Group (<https://en-author-services.edanzgroup.com/ac>) for editing a draft of this manuscript.

Author Contributions

Motoki Fukuda conducted the full experiment involving the deep learning method, compiled the results and wrote the full manuscript. Yoshiko Arij, Munetaka Nitoh, Michihito Nozawa, Chiaki Kuwada, Masako Nishiyama, and Takuma Funakoshi evaluated estimated CBCT images as the observers. Hiroshi Fujita, Akitoshi Katsumata, and Eiichiro Arij supervised the whole experiment with important instructions and advices. All authors reviewed the manuscript.

Conflict of interest

None of the authors have any conflict of interest associated with this study.

References

1. Arai, Y., Tammisalo, E., Iwai, K., Hashimoto, K. & Shinoda, K. Development of a compact computed tomographic apparatus for dental use. *Dentomaxillofac Radiol.***28**, 245–248. (1999).
2. Mozzo, P., Procacci, C., Tacconi, A., Martini, P. T. & Andreis, I, A. A new volumetric CT machine for dental imaging based on the cone-beam technique: preliminary results. *Eur Radiol.* **8**, 1558–1564. (1998).
3. Horner, K., O'Malley, L., Taylor, K. & Glenney, A. M. Guidelines for clinical use of CBCT: a review. *Dentomaxillofac Radiol.* **44**, 20140225. (2015).
4. Jacobs, R., Salmon, B., Codari, M., Hassan, B. & Bornstein, M. M. Cone beam computed tomography in implant dentistry: recommendations for clinical use. *BMC Oral Health.* **18**, 88. (2018).
5. Kapila, S. D. & Nervina, J. M. CBCT in orthodontics: assessment of treatment outcomes and indications for its use. *Dentomaxillofac Radiol.***44**, 20140282. (2015).
6. Matzen, L. H., Schropp, L., Spin-Neto, R. & Wenzel, A. Radiographic signs of pathology determining removal of an impacted mandibular third molar assessed in a panoramic image or CBCT. *Dentomaxillofac Radiol.* **46**, 20160330. (2017).

7. Leonardi, D. K. et al. Diagnostic accuracy of cone-beam computed tomography and conventional radiography on apical periodontitis: a systematic review and meta-analysis. *J Endod.***42**, 356–364. (2016).
8. Elstrøm, U. V., Muren, L. P., Petersen, J. B. & Grau C. Evaluation of image quality for different kV cone-beam CT acquisition and reconstruction methods in the head and neck region. *Acta Oncol.***50**, 908–917. (2011).
9. Katsumata, A. et al. Effects of image artifacts on gray-value density in limited-volume cone-beam computerized tomography. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod.***104**, 829–836. (2007).
10. Murata, M. et al. Deep-learning classification using convolutional neural network for evaluation of maxillary sinusitis on panoramic radiography. *Oral Radiol.***35**, 301–307. (2019).
11. Hiraiwa, T. et al. A deep-learning artificial intelligence system for assessment of root morphology of the mandibular first molar on panoramic radiography. *Dentomaxillofac Radiol.* **48**, 20180218. (2019).
12. Fukuda, M. et al. Evaluation of an artificial intelligence system for detecting vertical root fracture on panoramic radiography. *Oral Radiol.***36**, 337–343. (2020).
13. Arijji, Y. et al. Contrast-enhanced computed tomography image assessment of cervical lymph node metastasis in patients with oral cancer by using a deep learning system of artificial intelligence. *Oral Surg Oral Med Oral Pathol Oral Radiol.***127**, 458–463. (2019).
14. Kise, Y. et al. Preliminary study on the application of deep learning system to diagnosis of Sjögren's syndrome on CT images. *Dentomaxillofac Radiol.***48**, 20190019. (2019).
15. Dong, C., Loy, C. C., He, K. & Tang, X. Image super-resolution using deep convolutional networks. *IEEE Trans Pattern Anal Mach Intell.* **38**, 295–307. (2016).
16. Zhao, T., McNitt-Gray, M. & Ruan, D. A convolutional neural network for ultra-low-dose CT denoising and emphysema screening. *Med Phys.* **46**, 3941–3950. (2019).
17. Chen, H. et al. Low-dose CT with a residual encoder-decoder convolutional neural network. *IEEE Trans Med Imaging.***36**, 2524–2535. (2017).
18. Chen, H. et al. Low-dose CT via convolutional neural network. *Biomed Opt Express.* **8**, 679–694. (2017).
19. Kang, E., Min, J. & Ye, J. C. A deep convolutional neural network using directional wavelets for low-dose X-ray CT reconstruction. *Med Phys.* **44**, e360–375. (2017).
20. Kang, E., Chang, W., Yoo, J. & Ye, J. C. Deep convolutional framelet denoising for low-dose CT via wavelet residual network. *IEEE Trans Med Imaging.* **37**, 1358–1369. (2018).
21. Kim, B., Han, M., Shim, H. & Baek, J. A performance comparison of convolutional neural network-based image denoising methods: The effect of loss functions on low-dose CT images. *Med Phys.* **46**, 3906–3923. (2019).
22. Yang, Q. et al. Low-dose CT image denoising using a generative adversarial network with Wasserstein distance and perceptual loss. *IEEE Trans Med Imaging.* **37**, 1348–1357. (2018).

23. Yi, X. & Babyn, P. Sharpness-aware low-dose CT denoising using conditional generative adversarial network. *J Digit Imaging*. **31**, 655–669. (2018).
24. You, C. et al. Structurally-sensitive multi-scale deep neural network for low-dose CT denoising. *IEEE Access*. **6**, 41839–41855. (2018)
25. Shan, H. et al. 3-D convolutional encoder-decoder network for low-dose CT via transfer learning from a 2-D trained network. *IEEE Trans Med Imaging*. **37**, 1522–1534. (2018)
26. Wu, D., Kim, K. & Li, Q. Computationally efficient deep neural network for computed tomography image reconstruction. *Med Phys*. **46**, 4763–4776. (2019)
27. Bahrami, K. et al. Reconstruction of 7T-like images from 3T MRI. *IEEE Trans Med Imaging*. **35**, 2085–2097. (2016).
28. Lechuga, L. & Weidlich, G. A. Cone beam CT vs. fan beam CT: a comparison of image quality and dose delivered between two differing CT imaging modalities. *Cureus*. **8**, e778. (2016)
29. Ronneberger, O., Fischer, P. & Brox, T. U-net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W. & Frangi, A., editors. Medical image computing and computer-assisted intervention – *MICCAI 2015 Lecture Notes in Computer Science*. **9351**. Cham: Springer; pp. 234–241. (2015).
30. Rasband, W. S. ImageJ, U.S. National Institutes of Health, Bethesda, Maryland, USA, <http://imagej.nih.gov/ij/> (1997-2012).
31. Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nat Methods*. **9**, 671–675. (2012).

Tables

Table 1 Mean evaluation scores for the three images

		estimated CBCT image	swCBCT image	compressed swCT image
Muscle	Medial pterygoid muscle	1.11±1.08	0.96±0.62	4.00
	Digastric muscle	2.56±1.14	2.17±0.70	4.00
Fascial space	Parapharyngeal space	0.81±1.21*	0.64±0.50	4.00
	Submandibular space	2.37±1.15*	2.16±0.82	4.00
Salivary gland	Submandibular gland	2.44±1.27	2.37±0.92	4.00
Lymph node	Submental or submandibular lymph node	2.37±1.02*	2.13±0.84	4.00

CT: computed tomography, CBCT: cone-beam CT

swCT: soft-tissue window CT, swCBCT: soft-tissue window CBCT

Boldfaced characters denote significant differences from the compressed swCT image

*A significant difference was found in the scores of the estimated CBCT images relative to those of the swCBCT images (Steel-Dwass test, $p < 0.01$)

Figures

Figure 1

Learning and testing processes. CNN: convolutional neural network.

Figure 2

The architecture of the modified U-Net. The network is composed of a convolutional layer, batch normalization layer, rectified linear unit layer, and pooling layer. Conv: convolutional layer, Pooling: maximum pooling layer, ReLU: rectified linear unit layer.

Figure 3

Example images at the maxillary dental arch level. a. estimated CBCT created by the trained model. b. soft-tissue window CBCT image. c. compressed soft-tissue window CT image. White arrows show medial pterygoid muscles and white arrow heads show parapharyngeal spaces.

Figure 4

Example images at the submandibular level. a. estimated CBCT created by the trained model. b. soft-tissue window CBCT image. c. compressed soft-tissue-window CT image. White arrows show the digastric muscles and white arrow heads show submental spaces.

Figure 5

Circular ROIs used for measuring the voxel values in the digastric muscle (white line circle) and fat tissue in the submandibular space (dotted line circle). All ROIs have 10-pixel diameters (160-pixel areas).

Figure 6

Histograms of voxel values of the digastric muscle and fat tissue in the submandibular space. estimated CBCT (a), swCBCT (b), and compressed swCT (c) images. Values in parentheses show the means and standard deviation of voxel values with their ranges.