Deep Learning-Based Fully Automatic Segmentation of the Maxillary Sinus on Cone-Beam Computed Tomographic Images

Han Seung Choi

Yonsei University College of Dentistry

Kug Jin Jeon

Yonsei University College of Dentistry

Young Hyun Kim

Yonsei University College of Dentistry

Eun-Gyu Ha

Yonsei University College of Dentistry

Chena Lee

Yonsei University College of Dentistry

Sang-Sun Han (✉ sshan@yuhs.ac)

Yonsei University College of Dentistry

Research Article

Keywords: Artificial intelligence, Cone-beam computed tomography (CBCT), Maxillary sinus, Segmentation

Posted Date: January 19th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1259065/v1>

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

[Read Full License](#)

Abstract

This study proposes a deep learning-based method for fully automatic segmentation of the maxillary sinus, including clear or hazy states, on cone-beam computed tomographic (CBCT) images.

The model, which consisted of a U-Net architecture with post-processing, was developed using 19,350 CBCT images from 90 maxillary sinuses (34 clear sinuses, 56 hazy sinuses). Post-processing was performed to improve the accuracy of the prediction results of U-Net. Model performance was evaluated compared to manual labeling by an oral radiologist.

The artificial intelligence (AI) prediction results before post-processing were DSC of 0.9071 ± 0.1842 , HD of 2.6207 ± 4.9279 in clear sinuses and DSC of 0.9107 ± 0.1988 , HD of 2.7812 ± 4.2835 in hazy sinuses. After post-processing, the results all improved (DSC: 0.9072 ± 0.1842 , HD: 2.4057 ± 2.4496 in clear sinuses and DSC: 0.9123 ± 0.1977 , HD: 1.9158 ± 2.0887 in hazy sinuses).

The proposed deep learning model showed good performance for clear and hazy maxillary sinus segmentation with minimal time and effort, and the post-processing method improved the accuracy of automatic segmentation.

Introduction

The maxillary sinus is the largest air-filled space around the nasal cavity and occupies most of the maxilla. It acts as a ventilation passage for incoming air through the nose¹. The maxillary sinus begins to form as an invagination from the lateral wall of the middle meatus during fetal development and grows laterally and downward to the alveolar process at the age of 15-18 years. The inferior wall comprises the alveolar process, and the roots of premolars and molars are adjacent to or protrude into the maxillary sinus. The maxillary sinus is highly relevant in the field of dentistry, and its evaluation and analysis are very important. Periapical lesions can cause sinusitis and an oroantral fistula may develop during tooth extraction procedures^{1,2}. A sinus lift procedure may be required due to the lack of alveolar bone at the time of implant placement^{3,4}. Conditions such as mucosal thickening of the maxillary sinus significantly affect the success rate and complications of implant surgery⁵. It is difficult to diagnose the disease early because the maxilla has multiple overlapping structures, so it is often discovered after disease is exacerbated⁶. Odontogenic cysts or tumors cause changes such as elevation of the sinus floor, and malignant tumors destroy the sinus floor. A close assessment of the maxillary sinus is helpful for making an early diagnosis.

Accurate segmentation (an anatomical annotation that delineates the outlines of important structures), is required as a cornerstone for automated diagnosis and treatment planning⁷. Panoramic radiography is the most common imaging modality in dentistry⁸, but it has limitations in the accurate diagnoses of the maxillary sinus due to the presence of overlapping anatomical structures, such as the alveolar bone and zygomatic bone, and the geometric distortion of images. Cone-beam computed tomography (CBCT) has

recently been used for the three-dimensional (3D) evaluation and diagnosis of oral and maxillofacial disease due to its advantages, which include a lower radiation dose, lower cost, and more convenient access compared with computed tomography (CT)⁹. CBCT provides information through a multi-directional reconstruction process, but this process requires special training for data reconstruction and evaluation. In clinical practice, manual segmentation by experts and semi-automatic segmentation are performed^{10–12}. For an accurate analysis, multi-directional images must be considered, but doing so is laborious and time-consuming. Moreover, the results of the analysis may be highly dependent on experts' knowledge and experience. For these reasons, deep learning-based studies have been conducted with the goal of achieving equivalent accuracy in various cases, and several studies have made attempts to perform automatic segmentation of the maxillary sinus^{13–15}. However, automatic segmentation of the maxillary sinus is challenging because it is connected to adjacent anatomical structures such as the frontal, ethmoid, and sphenoid sinuses, as well as nasal structures. Furthermore, CBCT images are poorer-quality than CT images due to image noise and poor soft tissue contrast^{16,17}. For this reason, automatic segmentation in CBCT images is difficult.

In recent studies, deep-learning algorithms have shown good performance in automated detection, segmentation, and diagnosis based on 3D medical imaging^{13–15}. Various networks have been used for segmentation, and many studies have used U-Net models^{18–20}. Those studies have mostly utilized CT images, whereas very few studies have analyzed CBCT images using deep learning¹⁵.

This study proposes a deep learning-based model for fully automated segmentation and presents an analysis of a modification of the model by adding a post-processing method for accurate segmentation of the maxillary sinus using CBCT images, including some that were clear and others with various levels of haziness.

Methods

Datasets and image preparation

This study was approved by the Institutional Review Board (IRB) of Yonsei University Dental Hospital (NO. 2-2021-0059) and was conducted in accordance with relevant guidelines and ethical regulations. As a retrospective study, the requirement of patient informed consent was waived from IRB of Yonsei University Dental Hospital. The data were anonymized to avoid the identification of the patients.

In total, 19,350 CBCT images were acquired from 90 maxillary sinuses of 45 patients (26 females and 19 males) who visited Yonsei University Dental Hospital from June 2020 to March 2021. There were 430 images per person, and the number of images including sinus varied from about 156 to 215. All CBCT data were acquired with a 16 × 10 cm field of view using a RAYSCAN Alpha Plus device (Ray Co. Ltd, Hwaseong-si, Korea). Cases with abnormalities or a history of surgery in the maxillary sinus were excluded, as were poor-quality images with artifacts. The 90 maxillary sinuses were clear or had various states of haziness including mucosal thickening, mucous retention cysts, fluid, etc. (Fig. 1). The datasets

for training, validation, and testing were split into 6:2:2 ratios^{21,22}. The characteristics of the subjects are summarized in Table 1.

Table 1
Number of maxillary sinuses in the datasets.

Sinus	Training	Validation	Testing
Clear	17	8	9
Hazy	37	10	9
Total	54	18	18

CBCT data were used in the Digital Imaging and Communications in Medicine (DICOM) format with a matrix size of 678 (width) $\times$ 678 (height) pixels. To generate a label mask for the ground truth, an oral radiologist with over 20 years of experience conducted manual segmentation of the maxillary sinus in each axial image using 3D Slicer (free open-source software for biomedical image analysis)²³. The label images were exported in the DICOM format and used as the ground truth for network training. Input images of 678 $\times$ 678 pixels were normalized by using two-dimensional min/max normalization and resized to 256 $\times$ 256 pixels. Data augmentation was performed with rotation (-5° to 5°), translation shift (0 to 30%), and zoom (0 to 30%). Figure 2 shows the overall process.

Network architecture

The segmentation model based on U-Net was designed for maxillary sinus segmentation. U-Net models have demonstrated powerful performance in the segmentation of medical images^{18–20,24}. This architecture is shown in Figure 3. It consisted of a fully convolutional encoder and decoder, including 9 convolutional blocks, 4 max-pooling layers, and 4 transposed convolution layers. The convolutional block consisted of a 3 $\times$ 3 convolution with stride 1, batch normalization, and the ReLU activation function. At the end of the network, the softmax activation function in the output layer was used for multi-class segmentation of the background, right maxillary sinus, and left maxillary sinus. In the encoding block, the U-Net extracted the maxillary sinus features and passed them on to the next block. In the decoding block, the spatial information features were extracted, and then the size of feature maps was recovered to be the same as the input. The deep learning networks were implemented in Python 3 with the Keras library and trained on an NVIDIA Titan RTX (24GB).

Loss function

For training, the Adam optimizer²⁵ was applied, and the learning rate was initially set to 10^{-3} , which was reduced by one-fifth up to less than 20^{-4} each time when the validation loss was saturated for 5 epochs. While monitoring the loss of the validation set, if the validation loss did not decrease for 20 epochs, the training was stopped using an early stopper. The loss function of a deep learning network is calculated for validating the quantitative difference between the network output and ground truth. The

network parameters are iteratively adjusted in the training phase to minimize loss. In order to train the proposed model, we employed the Dice coefficient–based multi-label loss function (DL)²². The DL is defined as:

$$DL(P, G) = 1 - \frac{1}{K} \sum_{c=1}^K \frac{2 \sum_i^N p_i^c g_i^c}{\sum_i^N (p_i^c + g_i^c)}$$

where p_i and g_i represent the pixel values of prediction and ground truth, respectively, N runs for all pixels, and K is the number of classes (in our model, background vs. right maxillary sinus vs. left maxillary sinus).

Post-processing

Post-processing was performed to reduce the prediction errors of the developed U-Net model. If the maxillary sinus is adjacent to an air area such as the ethmoid sinus or connected to the airway, it tends to be incorrectly predicted. An intrinsic limitation of neural networks can lead to pixel-level prediction errors²⁶, especially at maxillary sinus boundaries. To solve these problems, we conducted post-processing to eliminate false positives in the maxillary sinus segmentation results of the U-Net. A 3D volume of the maxillary sinus can be obtained by stacking 2D slice images, and volume information can be acquired by using a connection component of a voxel. Background voxels are labeled with 0, and voxels of the maxillary sinus region are labeled with 1. Labels of 1 for small volumes not connected to the maxillary sinus were removed. Then, the 3D volumetric image was converted to 2D images (Fig. 4). This method was implemented using MATLAB 2021a (MathWorks, Natick, MA, USA).

Evaluation metrics

The artificial intelligence (AI) prediction results and the post-processing results were evaluated using the dice similarity coefficient (DSC), precision, recall, and hausdorff distance (HD)²⁷, as shown in equations (1), (2), (3), (4):

$DSC = \frac{2*TP}{2*TP + FP + FN}$	(1)
$Precision = \frac{TP}{TP + FP}$	(2)
$Recall = \frac{TP}{TP + FN}$	(3)

where TP, FP, and FN are true positive, false positive, and false negative, respectively. TP represents the number of pixels for which the maxillary sinus area was accurately predicted. FP represents the number of pixels that were not maxillary sinus areas but was predicted to be maxillary sinuses. FN represents the number of unpredicted pixels in the maxillary sinus area.

The HD evaluation method evaluates the degree of difference between the two results by considering the spatial distance between the segmented objects as an index, and the smaller the value, the more accurate²⁸. The HD measures the maximum distance between two point sets, X and Y.

$$HD(X, Y) = \max(\max_{x \in X} \min_{y \in Y} d(x, y), \max_{y \in Y} \min_{x \in X} d(x, y)) \quad (4)$$

where x and y denote two points between the ground truth and prediction result of the maxillary sinuses, and $d(x, y)$ is the distance between the two points.

Results

Table 2 shows the AI prediction results according to the condition of the sinus. The AI model showed a DSC 0.9071 ± 0.1842 for clear maxillary sinuses and a DSC of 0.9107 ± 0.1988 for hazy maxillary sinuses. Figure 5 and 6 show the best and worst results of maxillary sinus segmentation. Table 3 shows prediction results after post-processing. The prediction results all improved, with a DSC of 0.9072 ± 0.1842 for clear maxillary sinuses and a DSC of 0.9123 ± 0.1977 for hazy maxillary sinuses. The 3D reconstruction results are shown in Figure 7. False positives were removed in post-processing. The average segmentation time per one maxillary sinus was 46.2 seconds for the AI method whereas it took 48.7 minutes for the manual method, and the post-processing process took an average of 16 seconds.

Table 2
Artificial intelligence prediction results according to sinus condition (mean $\pm$ SD).

	Clear sinus	Hazy sinus	Total
DSC	0.9071 ± 0.1842	0.9107 ± 0.1988	0.9090 ± 0.1921
Precision	0.8851 ± 0.1784	0.9089 ± 0.1944	0.8977 ± 0.1874
Recall	0.9404 ± 0.1935	0.9173 ± 0.2060	0.9282 ± 0.2005
HD (mm)	2.6207 ± 4.9279	2.7812 ± 4.2835	2.7013 ± 4.6154
* DSC: Dice similarity coefficient, HD: Hausdorff distance, SD: Standard deviation.			

Table 3
Prediction results after post-processing of AI prediction (mean $\pm$ SD).

	Clear sinus	Hazy sinus	Total
DSC	0.9072 $\pm$ 0.1842	0.9123 $\pm$ 0.1977	0.9099 $\pm$ 0.1914
Precision	0.8852 $\pm$ 0.1784	0.9131 $\pm$ 0.1917	0.8999 $\pm$ 0.1860
Recall	0.9404 $\pm$ 0.1935	0.9173 $\pm$ 0.2060	0.9282 $\pm$ 0.2005
HD (mm)	2.4057 $\pm$ 2.4496	1.9158 $\pm$ 2.0887	2.1470 $\pm$ 2.2790
* AI: Artificial intelligence, DSC: Dice similarity coefficient, HD: Hausdorff distance, SD: Standard deviation.			

Discussion

The inferior wall of the maxillary sinus is part of the alveolar process, making it highly relevant for dentistry (e.g., tooth extraction, implants, and apical lesions). When a disease such as a cyst or a tumor occurs in the maxilla, the change in the maxillary sinus is an important aspect of the diagnosis. Therefore, automatic maxillary sinus segmentation will play a starting point in clinical diagnoses. The maxillary sinus is connected to adjacent anatomical structures, which are very difficult to separate accurately. Inflammation of the sinus appears in various forms, such as thickening of the mucous membrane. If there is mucosal thickening, the border is not clearly defined, which may make automatic segmentation more difficult. A few studies have explored the application of automatic maxillary sinus segmentation using CT images^{13,14}. Xu et al.¹³ reported a DSC of 0.94 for clear maxillary sinuses and stated that the results for hazy maxillary sinuses were not satisfactory. In another study, Iwamoto et al.¹⁴ reported a DSC of 0.83 in both clear and hazy maxillary sinuses.

CBCT is widely used in dental clinics because it has several advantages over traditional CT and panoramic radiography. It provides 3D information, unlike panoramic radiography, and it has a significantly lower radiation dose and substantially lower costs than CT. CBCT is often used to diagnose the condition of the maxillary sinus and for planning extraction and implant placement in the maxilla. To the best of our knowledge, only one study has performed automatic maxillary sinus segmentation using CBCT images¹⁵. Jung et al.¹⁵ divided the air and the lesion area of maxillary sinus and the results of entire maxillary sinus segmentation are not shown. The results of segmentation of the air and lesion area were DSC of 0.93 and DSC of 0.76, respectively. It seems more reasonable to first accurately segment the maxillary sinus and then segment the air and the lesion area. The severity of hazy sinuses is important because there is a higher likelihood of complications after tooth extraction or implantation, but the performance of segmentation is worse than that of clear sinuses. Therefore, we added more data for hazy sinuses. The performance of our model was similar for clear and hazy sinuses, with DSC values of 0.9 or higher.

The non-maxillary sinus area was incorrectly predicted as sinus area in some cases. We investigated the use of post-processing to reduce false positive errors. After post-processing, the segmentation results were a DSC of 0.9099 ± 0.1914 , precision of 0.8999 ± 0.1860 , and HD of 2.1470 ± 2.2790 , reflecting improvements compared to before post-processing. In the clear sinus, since the size of the false positive is very small, although the difference in the DSC value was small, the improvement of the HD value using the distance value was larger. Our U-Net and post-processing methods showed similar or higher performance in comparison with previous CT studies^{13,14}.

We proposed a deep learning-based method for automatic segmentation of the maxillary sinus using a U-Net model with post-processing. There is no need for an expert to manually select the maxillary sinus from all images for segmentation. We automatically and accurately segmented the maxillary sinus from CBCT images in less than 1 minute, which was much more efficient than manual segmentation, which took more than 30 minutes. The developed model makes it possible to analyze the CBCT images of many patients in a short time.

Although our study showed good performance for automatic segmentation of the maxillary sinus, some limitations should be addressed in future research. First, if false positive pixels are connected to the maxillary sinus region, they are not removed by post-processing. To overcome these limitations, a process for detecting the edge of the maxillary sinus is required. Using a variety of training data may also help solve this problem. Second, this study used data obtained using a single CBCT device. It is necessary to use additional data from various CBCT devices to improve the performance of the algorithm. Finally, developing a segmentation model of areas of only haziness is needed. If the methodology is further developed, it will be possible to collect quantitative volume information on the segmented maxillary sinuses and automatically calculate the ratio of the volume of the haziness area to the entire maxillary sinus volume.

Conclusion

The proposed deep learning algorithm showed good performance for clear and hazy maxillary sinus segmentation with minimal time and effort, and the post-processing method reduced errors and improved the accuracy of automatic sinus segmentation.

Declarations

Author contributions

S.H. proposed the ideas; H.C., K.J. and Y.K. collected data; H.C. designed the deep learning model; K.J., C.L. analyzed and interpreted data; H.C., K.J., E.H. and S.H. critically reviewed the content; and H.C, K.J, and S.H. drafted and critically revised the article.

Competing interests

The authors declare no conflict of interests.

Funding

This work was funded by a National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (No. 2019R1A2C1007508).

Data availability

The data generated and analyzed during the current study are not publicly available due to privacy laws and policies in Korea but are available from the corresponding author on reasonable request.

References

1. Hauman, C., Chandler, N. & Tong, D. Endodontic implications of the maxillary sinus: a review. *International endodontic journal* **35**, 127–141 (2002).
2. Yoo, B. J. *et al.* Treatment strategy for odontogenic sinusitis. *American Journal of Rhinology & Allergy* **35**, 206–212 (2021).
3. Rosado, L. d. P. L., Barbosa, I. S., de Aquino, S. N., Junqueira, R. B. & Verner, F. S. Dental students' ability to detect maxillary sinus abnormalities: A comparison between panoramic radiography and cone-beam computed tomography. *Imaging science in dentistry* **49**, 191–199 (2019).
4. Kim, J. S. *et al.* What affects postoperative sinusitis and implant failure after dental implant: a meta-analysis. *Otolaryngology–Head and Neck Surgery* **160**, 974–984 (2019).
5. Drumond, J. P. N., Allegro, B. B., Novo, N. F., Miranda, S. L. d. & Sendyk, W. R. Evaluation of the prevalence of maxillary sinuses abnormalities through spiral computed tomography (CT). *International archives of otorhinolaryngology* **21**, 126–133 (2017).
6. Na, J. Y., Kang, J. H., Choi, S.-H., Jeong, H.-G. & Han, S.-S. Squamous cell carcinoma of the maxillary sinus mimicking periodontitis. *Journal of the Korean Dental Association (대한치과의사협회지)* **55**, 276–283 (2017).
7. Sharma, N. & Aggarwal, L. M. Automated medical image segmentation techniques. *Journal of medical physics/Association of Medical Physicists of India* **35**, 3 (2010).
8. Malina-Altzinger, J., Damerau, G., Grätz, K. W. & Stadlinger, P. B. Evaluation of the maxillary sinus in panoramic radiography—a comparative study. *International journal of implant dentistry* **1**, 1–7 (2015).
9. Macleod, I. & Heath, N. Cone-beam computed tomography (CBCT) in dental practice. *Dental update* **35**, 590–598 (2008).
10. Demir, U. L., Akca, M., Ozpar, R., Albayrak, C. & Hakyemez, B. Anatomical correlation between existence of concha bullosa and maxillary sinus volume. *Surgical and Radiologic Anatomy* **37**, 1093–1098 (2015).

11. Andersen, T. N. *et al.* Accuracy and precision of manual segmentation of the maxillary sinus in MR images—a method study. *The British journal of radiology* **91**, 20170663 (2018).
12. 12.. Gomes, A. F. *et al.* Development and validation of a formula based on maxillary sinus measurements as a tool for sex estimation: a cone beam computed tomography study. *International journal of legal medicine* **133**, 1241–1249 (2019).
13. Xu, J. *et al.* Automatic CT image segmentation of maxillary sinus based on VGG network and improved V-Net. *International Journal of Computer Assisted Radiology and Surgery* **15**, 1457–1465 (2020).
14. Iwamoto, Y. *et al.* in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. 2789-2792 (IEEE).
15. Jung, S.-K., Lim, H.-K., Lee, S., Cho, Y. & Song, I.-S. Deep Active Learning for Automatic Segmentation of Maxillary Sinus Lesions Using a Convolutional Neural Network. *Diagnostics* **11**, 688 (2021).
16. Zhang, Y. *et al.* Reducing metal artifacts in cone-beam CT images by preprocessing projection data. *International Journal of Radiation Oncology* Biology* Physics* **67**, 924–932 (2007).
17. Katsumata, A. *et al.* Image artifact in dental cone-beam CT. *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontology* **101**, 652–657 (2006).
18. Ronneberger, O., Fischer, P. & Brox, T. in *International Conference on Medical image computing and computer-assisted intervention*. 234-241 (Springer).
19. Li, X. *et al.* H-DenseUNet: hybrid densely connected UNet for liver and tumor segmentation from CT volumes. *IEEE transactions on medical imaging* **37**, 2663–2674 (2018).
20. Dong, H., Yang, G., Liu, F., Mo, Y. & Guo, Y. in *annual conference on medical image understanding and analysis*. 506-517 (Springer).
21. Li, H., Li, A. & Wang, M. A novel end-to-end brain tumor segmentation method using improved fully convolutional networks. *Computers in biology and medicine* **108**, 150–160 (2019).
22. Wang, L., Wang, C., Sun, Z. & Chen, S. An improved dice loss for pneumothorax segmentation by mining the information of negative areas. *IEEE Access* **8**, 167939–167949 (2020).
23. Fedorov, A. *et al.* 3D Slicer as an image computing platform for the Quantitative Imaging Network. *Magnetic resonance imaging* **30**, 1323–1341 (2012).
24. Yang, S. *et al.* Deep learning segmentation of major vessels in X-ray coronary angiography. *Scientific reports* **9**, 1–11 (2019).
25. Kingma, D. P. & Ba, J. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980* (2014).
26. Wang, S., Yang, D. M., Rong, R., Zhan, X. & Xiao, G. Pathology image analysis using segmentation deep learning algorithms. *The American journal of pathology* **189**, 1686–1698 (2019).
27. Huttenlocher, D. P., Klanderman, G. A. & Rucklidge, W. J. Comparing images using the Hausdorff distance. *IEEE Transactions on pattern analysis and machine intelligence* **15**, 850–863 (1993).

28. Taha, A. A. & Hanbury, A. Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool. BMC medical imaging **15**, 1–28 (2015).

Figures

Figure 1

Examples of various maxillary sinus images. (a) clear sinuses (both sides), (b) a slightly hazy sinus (left side), (c) a moderately hazy sinus (right side), (d) a severely hazy sinus (left side).

Figure 2

Overall process of our study.

Figure 3

Structure of U-Net based convolutional neural network segmentation.

Figure 4

The post-processing workflow.

Figure 5

Best results of maxillary sinus segmentation. The right side presents a clear maxillary sinus, and the left side shows a hazy maxillary sinus. Red: the edges of the ground truth image; green: the edges of the artificial intelligence (AI) prediction image.

Figure 6

Worst results of maxillary sinus segmentation. The mispredicted part in the posterior part of the maxillary sinus is the ethmoid sinus. Red: the edges of the ground truth image; green: the edges of the artificial

intelligence (AI) prediction image.

Figure 7

Three-dimensional reconstruction results. (a) clear maxillary sinus, (b) hazy maxillary sinus. Red: the ground truth (manual labels); green: the results of artificial intelligence (AI) prediction; yellow: the results after post-processing of AI prediction. The average segmentation time per one maxillary sinus was 48.7 minutes for the manual method and 30 seconds for the AI method, and the post-processing process took an average of 16 seconds.