

Development and Clinical Validation of an Enhanced Deep Learning Model for Automated Segmentation and Implant Planning in Maxillary Edentulous Regions Using CBCT Images

Mehdi Abrishami

Islamic Azad University

Hasan Momeni

Islamic Azad University

Mehdi Rahmani

Islamic Azad University

Elham Sheykhi Some

elhamsheykhii60@gmail.com

Iran University of Medical Sciences

Research Article

Keywords: Artificial intelligence, Deep learning, CBCT, Dental implants, Segmentation, Swin-UNETR, Digital implant planning

Posted Date: July 6th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10084556/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Objectives

This study aimed to develop and clinically validate an advanced deep learning framework for automated segmentation of maxillary edentulous regions on cone-beam computed tomography (CBCT) images and to evaluate its role in supporting digital implant planning workflows.

Materials and Methods

A total of 450 CBCT scans were retrospectively screened, and 320 scans with partial maxillary edentulism were included. Manual segmentation was performed using 3D Slicer by experienced clinicians with consensus annotation. Multiple deep learning architectures, including U-Net, Attention U-Net, nnU-Net, and Swin-UNETR, were developed and trained using the MONAI framework. Dataset division followed an 80:10:10 split for training, validation, and testing. Model performance was assessed using Dice Similarity Coefficient (DSC), Intersection over Union (IoU), and Hausdorff Distance. Additionally, a clinical validation study was conducted in which implant planning was performed with and without AI assistance.

Results

All models successfully segmented maxillary edentulous regions, with Swin-UNETR achieving the highest performance (mean DSC: 92.1%). AI-assisted segmentation significantly reduced planning time (43.5% reduction) and improved interobserver agreement compared with conventional workflows. High usability scores indicated strong clinical acceptance. Discrepancies between manual and automated segmentation were mainly attributed to variability in manual annotation and anatomical complexity.

Conclusion

The proposed deep learning framework provides accurate and reliable segmentation of maxillary edentulous regions and significantly enhances efficiency and reproducibility in digital implant planning. Transformer-based architectures demonstrated superior performance, supporting their potential for integration into clinical implant workflows. Further multicenter validation studies are required to confirm generalizability and facilitate clinical translation.

Introduction

Dental implants have become a predictable and widely accepted treatment modality for the rehabilitation of partially and completely edentulous patients, with long-term survival rates exceeding 90%. Successful implant therapy depends heavily on accurate preoperative assessment of bone quantity, bone quality,

and the spatial relationship of anatomical structures, making prosthetically driven implant planning an essential component of contemporary implant dentistry(1). Cone-beam computed tomography (CBCT) has emerged as the imaging modality of choice for implant treatment planning because it provides three-dimensional visualization of alveolar bone morphology and critical anatomical landmarks while exposing patients to lower radiation doses compared with conventional computed tomography (CT)(2).

Despite the advantages of CBCT imaging, digital implant planning remains a time-consuming and operator-dependent process. Clinicians are required to manually identify edentulous regions, delineate alveolar boundaries, evaluate bone dimensions, and determine optimal implant positions while avoiding adjacent anatomical structures such as the maxillary sinus and nasal cavity. This workflow is susceptible to interobserver variability, fatigue-related errors, and inconsistencies arising from differences in clinical experience(3). Consequently, there is an increasing demand for automated and standardized approaches capable of improving efficiency and reproducibility in implant treatment planning.

Recent advances in artificial intelligence (AI), particularly deep learning (DL), have revolutionized medical image analysis. Convolutional neural networks (CNNs) and encoder-decoder architectures such as U-Net have demonstrated remarkable performance in segmentation tasks involving both medical and dental imaging. Automated segmentation techniques have been successfully applied to tooth identification, mandibular canal localization, maxillary sinus segmentation, and craniofacial structure analysis, significantly reducing processing time while maintaining high levels of accuracy(4).

In implant dentistry, AI-assisted systems have recently attracted considerable attention as promising tools for facilitating CBCT interpretation and implant planning. Bayrakdar et al.(2021) developed a deep learning approach for automated implant planning and demonstrated encouraging accuracy in identifying edentulous regions and estimating implant positions. Furthermore, recent systematic reviews and meta-analyses have shown that AI-based methods achieve high diagnostic accuracy and may substantially enhance the efficiency of digital workflows(5). However, most currently available studies have focused primarily on tooth segmentation or mandibular structures, whereas relatively few investigations have addressed segmentation of maxillary edentulous ridges, which represent a particularly challenging anatomical region because of their complex morphology, variable resorption patterns, and proximity to the maxillary sinus(6).

Moreover, previous studies have been limited by relatively small datasets, underrepresentation of anterior maxillary defects, heterogeneous annotation protocols, and a lack of clinical validation involving experienced clinicians(7). Variability in manual segmentation and the absence of standardized reference labels remain important obstacles to the development of robust and clinically applicable AI systems. Recent evidence suggests that larger and more diverse datasets, advanced network architectures such as Attention U-Net and nnU-Net, and clinician-centered validation frameworks may substantially improve model performance and facilitate translation into routine clinical practice(8).

Therefore, the aim of the present study was to develop and clinically validate an enhanced deep learning model for automated segmentation of maxillary edentulous regions using CBCT images and to investigate its potential application in digital implant planning. In addition, the performance of the proposed model was evaluated using quantitative segmentation metrics and clinician-based assessments to determine its accuracy, usability, and potential impact on treatment planning workflows.

Materials and Methods

Study Design and Ethical Approval

This diagnostic accuracy and clinical validation study was conducted following approval by the Institutional Review Board and Research Ethics Committee of Islamic Azad University Of Isfahan, Iran. All CBCT scans were retrospectively collected from the radiographic archives of the Department of Oral and Maxillofacial Radiology. Patient confidentiality was maintained throughout the study, and all data were anonymized before analysis.

Data Collection and Selection Criteria

A total of 450 cone-beam computed tomography (CBCT) scans acquired between 2019 and 2025 were retrieved using Romexis software (Planmeca Oy, Helsinki, Finland)(1, 9). Imaging parameters ranged from 90 to 120 kVp, 5 to 60 mA, field of view (FOV) from 5 × 5 cm to 16 × 16 cm, and isotropic voxel sizes between 150 and 300 µm.

Following the initial screening, 320 scans fulfilled the inclusion criteria and were included in the study. To minimize anatomical bias and improve model generalizability, cases were selected to ensure balanced representation of anterior, posterior, and combined edentulous regions of the maxilla.

Inclusion criteria comprised:

- Partial edentulism in the maxillary arch;
- Availability of CBCT images with adequate diagnostic quality;
- Residual ridge width ≥ 4 mm and bone height ≥ 2 mm;
- Presence of adjacent teeth allowing prosthetically driven implant planning.

Exclusion criteria included:

- Complete maxillary edentulism;
- Unhealed extraction sockets;
- Severe metal artifacts or motion artifacts;
- Maxillofacial pathology, cysts, tumors, or fractures affecting the region of interest;
- Presence of impacted teeth or tooth buds within the edentulous area;

- Previous bone grafting procedures or sinus augmentation.

Manual Segmentation Protocol

All images were imported into 3D Slicer software (version 5.6.0, Harvard University, USA)(10). Manual segmentation was independently performed by three board-certified oral and maxillofacial surgeons with more than five years of experience in implant dentistry.

Edentulous ridges were delineated in axial, sagittal, and coronal planes. The superior boundary was defined by the floor of the maxillary sinus or nasal cavity, whereas the inferior boundary corresponded to the alveolar crest. Mesiodistal limits were determined according to the adjacent teeth and the extent of the edentulous span.

Initial labeling was performed using the “Grow from Seeds” algorithm, followed by refinement with “Scissors,” “Islands,” and “Joint Smoothing” tools. To improve anatomical fidelity and reduce variability, smoothing parameters were standardized across all cases.

Consensus segmentation masks were generated after independent review by the three examiners. In cases of disagreement, a fourth senior oral and maxillofacial surgeon performed the final adjudication. All segmentation masks were exported in Neuroimaging Informatics Technology Initiative (NIfTI) format (.nii.gz).

Quality Assurance and Interobserver Agreement

Before annotation, all examiners underwent calibration sessions using twenty representative CBCT scans. Interobserver agreement was assessed using the Dice Similarity Coefficient (DSC)(11) and intraclass correlation coefficient (ICC)(12). Cases showing substantial disagreement were reevaluated until consensus was achieved.

Deep Learning Model Development

Four deep learning architectures were evaluated:

1. Conventional 3D U-Net(13);
2. Attention U-Net(14);
3. nnU-Net(15);
4. Swin-UNETR transformer-based architecture(16).

All models were implemented using the Medical Open Network for Artificial Intelligence (MONAI) framework based on PyTorch(17).

Input volumes underwent preprocessing including intensity normalization, resampling to isotropic voxel spacing, cropping around the region of interest, and data augmentation techniques including random rotations, flipping, scaling, elastic deformation, and Gaussian noise addition to improve robustness and reduce overfitting(18).

The dataset was randomly divided into training (80%), validation (10%), and testing (10%) subsets(15). Five-fold cross-validation was employed to optimize hyperparameters and assess model stability.

Skip connections and multi-scale feature extraction mechanisms were utilized to preserve fine anatomical structures and improve segmentation accuracy. The models were trained using a combination of Dice loss and cross-entropy loss functions(19). Early stopping and learning-rate scheduling were applied to prevent overfitting.

Performance Evaluation

Segmentation performance was quantitatively assessed using:

- Dice Similarity Coefficient (DSC);
- Intersection over Union (IoU);
- Hausdorff Distance (HD95)(20);
- Precision;
- Recall;
- F1-score.

Cases with DSC values below 0.70 underwent qualitative analysis to identify potential causes of discrepancy, including anatomical variations, sinus morphology, ridge irregularities, and annotation-related inconsistencies.

Clinical Validation

To evaluate the clinical applicability of the proposed system, twenty experienced clinicians (oral and maxillofacial surgeons and prosthodontists) participated in the validation phase.

Each clinician performed implant planning on twenty randomly selected CBCT scans under two conditions:

1. Conventional planning without AI assistance;
2. AI-assisted planning using the proposed segmentation model.

The following parameters were recorded:

- Planning time (minutes);
- Implant position and angulation;
- Implant length and diameter selection;
- Number of manual corrections required;
- User satisfaction assessed using a five-point Likert scale(21);
- System usability assessed according to the System Usability Scale (SUS)(22).

Statistical Analysis

Statistical analysis was performed using SPSS version XX (IBM Corp., Armonk, NY, USA).

Continuous variables were expressed as mean $\pm$ standard deviation. Differences between AI-assisted and conventional planning were analyzed using paired t-tests or Wilcoxon signed-rank tests depending on data distribution. Interobserver agreement was evaluated using intraclass correlation coefficients (ICC)(12). Comparisons among deep learning architectures were performed using one-way analysis of variance (ANOVA) followed by Bonferroni post hoc tests.

Statistical significance was set at $p < 0.05$.

Results

Dataset Summary

A total of 450 CBCT scans were initially screened for eligibility. Following the application of predefined inclusion and exclusion criteria, 320 scans with partial maxillary edentulism were included in the final analysis, while 130 scans were excluded because of complete edentulism, poor image quality, severe artefacts, or the presence of pathological lesions.

The final dataset consisted of 178 male and 142 female patients with a mean age of 54.8 ± 12.3 years. According to the location of the edentulous spaces, 108 cases (33.8%) involved anterior regions, 126 (39.4%) involved posterior regions, and 86 cases (26.8%) exhibited combined anterior-posterior defects. In terms of laterality, 94 cases were right-sided, 78 were left-sided, and 148 demonstrated bilateral edentulous areas.

The dataset was randomly divided into training (n = 256), validation (n = 32), and testing (n = 32) subsets.

Model Performance

All four deep learning architectures successfully detected and segmented maxillary edentulous regions in the test dataset.

Among the evaluated models, Swin-UNETR achieved the highest segmentation performance, followed by nnU-Net, Attention U-Net, and conventional U-Net.

The average Dice Similarity Coefficient (DSC), Intersection over Union (IoU), and Hausdorff Distance (HD95) values for each model are summarized in Table 1.

Table 1
Quantitative performance of different deep learning models.

Model	DSC (%)	IoU (%)	Precision (%)	Recall (%)	HD95 (mm)
U-Net	84.3	76.8	85.1	83.7	3.6
Attention U-Net	87.8	80.9	88.4	87.2	2.9
nnU-Net	90.4	84.7	91.3	89.8	2.3
Swin-UNETR	92.1	87.6	93.2	91.4	1.8

Compared with the conventional U-Net model, Swin-UNETR demonstrated significantly higher DSC values ($p < 0.001$).

The DSC values of the test cases ranged from 0.72 to 0.97. Only nine cases (2.8%) exhibited DSC values below 0.80, and no case recorded a DSC lower than 0.70.

Interobserver Agreement

Interobserver agreement among the three examiners was excellent. The mean intraclass correlation coefficient (ICC) for volumetric measurements was 0.94 (95% CI: 0.91–0.97), while the average Dice Similarity Coefficient between observers was 0.91 ± 0.03 .

Discrepancy Analysis

Detailed evaluation of low-performing cases revealed several sources of disagreement between manual and automated segmentation are summarized in Table 2.

The most frequent causes included:

- Irregular sinus floor morphology (n = 5);
- Thin residual bone (< 4 mm) identified by the model but excluded manually (n = 8);

- Partial volume effects and low-density bone regions adjacent to extraction sockets (n = 4);
- Mild metal artefacts causing false-positive segmentation (n = 3);
- Variations introduced by manual smoothing procedures (n = 11);
- Minor inconsistencies among annotators regarding the mesiodistal boundaries of edentulous spaces (n = 6).

Interestingly, in several cases the AI-generated masks more accurately followed the anatomical contours of the maxillary sinus floor and alveolar crest compared with the consensus manual segmentations.

Table 2	
Main causes of discrepancy between manual and automated segmentation.	
Cause of discrepancy	Number of affected cases
Smoothing-induced under-segmentation	11
Thin bone omitted during manual segmentation	8
Irregular sinus floor morphology	5
Low-density bone adjacent to extraction sites	4
Mesiodistal boundary variation	6
Metal artefacts	3

Clinical Validation

Twenty clinicians participated in the clinical validation phase. The overall clinical validation outcomes are summarized in Table 3. The mean planning time using the conventional workflow was 14.7 ± 3.8 minutes per case, whereas AI-assisted planning significantly reduced this time to 8.3 ± 2.1 minutes ($p < 0.001$), corresponding to a mean reduction of 43.5%.

The average number of manual corrections required after AI segmentation was 1.2 ± 0.8 per case.

Agreement among clinicians regarding implant position and angulation improved significantly when AI assistance was used (ICC = 0.93) compared with conventional planning (ICC = 0.84).

Furthermore, the mean System Usability Scale (SUS) score was 88.6 ± 5.4 , indicating excellent usability. Overall user satisfaction measured by a five-point Likert scale was 4.7 ± 0.4 .

Table 3
Clinical validation outcomes of conventional and AI-assisted implant planning.

Parameter	Conventional Planning	AI-assisted Planning	P-value
Planning time (min)	14.7 ± 3.8	8.3 ± 2.1	< 0.001
Manual corrections (per case)	3.6 ± 1.4	1.2 ± 0.8	< 0.001
Interobserver agreement (ICC)	0.84	0.93	< 0.01
System Usability Scale (SUS)	69.8 ± 7.5	88.6 ± 5.4	< 0.001
User satisfaction (Likert 1–5)	3.8 ± 0.7	4.7 ± 0.4	< 0.001

Comparison of Implant Planning Parameters

A detailed comparison of implant planning parameters between conventional and AI-assisted workflows is presented in Table 4. No statistically significant differences were observed between AI-assisted and conventional workflows regarding implant diameter or implant length selection ($p > 0.05$). However, AI-assisted planning showed significantly lower interoperator variability in implant angulation and mesiodistal positioning ($p < 0.01$).

Table 4
Comparison of implant planning parameters between conventional and AI-assisted workflows.

Parameter	Conventional Workflow	AI-assisted Workflow	P-value
Implant diameter selection	Comparable	Comparable	0.05
Implant length selection	Comparable	Comparable	0.05
Implant angulation variability	Higher	Lower	< 0.01
Mesiodistal positioning variability	Higher	Lower	< 0.01

Overall Findings

The proposed AI framework demonstrated excellent segmentation accuracy and clinical applicability. Transformer-based architectures, particularly Swin-UNETR, achieved the highest performance. Integration of AI into the implant planning workflow substantially reduced planning time, improved consistency among clinicians, and provided anatomically reliable segmentation of maxillary edentulous regions.

Discussion

The present study developed and clinically validated an enhanced deep learning framework for automated segmentation of maxillary edentulous regions on CBCT images and investigated its applicability in digital implant planning. The proposed framework demonstrated high segmentation accuracy, excellent reproducibility, and a significant reduction in treatment-planning time. Among the

evaluated architectures, Swin-UNETR achieved the highest performance, with a mean DSC of 92.1%, indicating that transformer-based architectures may offer advantages over conventional convolutional neural networks for complex dentomaxillofacial segmentation tasks.

The findings of the present study are consistent with the growing evidence supporting the use of deep learning in dentomaxillofacial imaging. The U-Net architecture introduced by Weng et al.(2021) revolutionized biomedical image segmentation and has subsequently become the foundation of numerous medical and dental segmentation frameworks (23). More recently, Isensee et al.(2021) demonstrated the remarkable adaptability and robustness of nnU-Net across multiple segmentation tasks, establishing it as one of the current benchmarks in medical image analysis (15).

In implant dentistry, Bayrakdar et al.(2021) evaluated a deep learning system for CBCT-based implant planning and reported that artificial intelligence could accurately identify missing tooth regions and surrounding anatomical structures, emphasizing its potential as a clinical support tool (5).

Likewise, Dot et al. (2024) recently introduced DentalSegmentator, an open-source nnU-Net-based framework trained on more than 700 CT and CBCT scans, demonstrating excellent robustness and high segmentation accuracy across various dentomaxillofacial structures(8) .The superior performance observed in the present study may be explained by the use of a larger and anatomically balanced dataset, consensus annotation by experienced clinicians, and the implementation of transformer-based architectures.

The superiority of Swin-UNETR over conventional U-Net observed in the present investigation agrees with the findings reported by Hatamizadeh et al.(2021), who introduced Swin-UNETR and demonstrated state-of-the-art performance in three-dimensional medical image segmentation (24). Unlike traditional convolutional neural networks, transformers are capable of capturing long-range dependencies and global contextual information, which is particularly important in maxillary edentulous regions characterized by irregular alveolar morphology and variable sinus anatomy.

Interestingly, several discrepancies between AI-generated and manually generated segmentations in the present study appeared to originate from limitations of manual annotation rather than model errors. Similar observations have been reported in medical image analysis, where interobserver variability and imperfections in ground-truth labels constitute major sources of disagreement between AI predictions and manual segmentations (15). In the present study, smoothing procedures and omission of thin residual bone regions represented the most frequent causes of disagreement. In several cases, AI-generated masks followed the anatomical contours of the maxillary sinus floor more accurately than the manually generated reference masks.

Another important finding was the significant reduction in treatment-planning time achieved with AI assistance. Digital implant planning has traditionally been considered a time-consuming and operator-dependent process requiring precise interpretation of CBCT images and anatomical structures (3). In the present study, AI-assisted planning reduced planning time by approximately 43% while maintaining

excellent agreement regarding implant dimensions. These findings support previous reports suggesting that AI has the potential to streamline digital workflows and improve clinical efficiency (5)

Moreover, the high interobserver agreement and excellent System Usability Scale scores observed in the present study indicate that AI-assisted segmentation may contribute to greater standardization and reproducibility in implant planning. Such findings are particularly relevant because variability in clinical experience and subjective interpretation represent important challenges in contemporary implant dentistry. Gupta (2013) et al. emphasized that accurate assessment of anatomical structures is essential for successful implant treatment and that CBCT interpretation remains highly dependent on clinician expertise (25).

Despite these promising findings, several limitations should be acknowledged. First, the study was retrospective and conducted using data obtained from a single institution. Differences in imaging protocols, scanner characteristics, and patient populations may influence model performance. Second, although the dataset was substantially larger than those used in previous studies, external validation using multicenter datasets is still required. Third, the current model was designed primarily for segmentation and did not automatically determine implant dimensions or final implant positions. Therefore, prospective studies investigating fully automated prosthetically driven implant planning systems are warranted.

Future research should focus on the development of end-to-end frameworks capable of integrating anatomical segmentation, bone quality assessment, automatic implant position recommendation, and surgical guide design. Combining deep learning algorithms with digital workflows and robotic implant placement may represent the next stage in the evolution of implant dentistry. Multicenter datasets and standardized annotation protocols will be essential for ensuring the robustness and generalizability of these systems.

Overall, the findings of the present study suggest that advanced deep learning architectures, particularly transformer-based models, can provide highly accurate and clinically useful segmentation of maxillary edentulous regions. AI-assisted workflows significantly improve efficiency and reproducibility and represent an important step toward standardized and semi-automated digital implant planning.

Conclusion

The present study demonstrated that advanced deep learning models can accurately and reliably segment maxillary edentulous regions on CBCT images with high clinical applicability. Transformer-based architecture (Swin-UNETR) showed superior performance compared with conventional CNN models, indicating improved capability in capturing complex anatomical structures. AI-assisted workflows significantly reduced planning time and improved interobserver agreement, supporting their role as effective decision-support tools in implant dentistry. Overall, these findings highlight the potential of deep learning to enhance efficiency, standardization, and reproducibility in digital implant planning, although further multicenter validation is required before clinical translation.

Declarations

Conflict of Interest

None declared.

Ethics Approval and Consent to Participate

This retrospective study was conducted using fully anonymized cone-beam computed tomography (CBCT) images retrieved from the institutional radiographic archive. No identifiable patient information was accessed, and no intervention or direct patient involvement occurred.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Author Contribution

M.A: investigation, supervision. M.R: investigation, methodology. H.M: writing – original draft, writing – review and editing. E.SH: project administration, writing – review and editing

Declaration of Generative AI Use

The authors declare that generative artificial intelligence tools were used solely for language refinement, editing, and stylistic improvement of the manuscript. These tools were not used for data collection, data analysis, interpretation of results, or drawing scientific conclusions. All scientific content, study design, methodology, and findings were developed, reviewed, and approved by the authors, who take full responsibility for the integrity and accuracy of the work.

References

1. Jacobs R, Salmon B, Codari M, Hassan B, Bornstein MM. Cone beam computed tomography in implant dentistry: recommendations for clinical use. *BMC Oral Health*. 2018;18(1):88.
2. Benavides E, Rios HF, Ganz SD, An C-H, Resnik R, Reardon GT, et al. Use of cone beam computed tomography in implant dentistry: the International Congress of Oral Implantologists consensus report. *Implant Dent*. 2012;21(2):78–86.
3. Greenberg AM. Digital technologies for dental implant treatment planning and guided surgery. *Oral Maxillofacial Surg Clin*. 2015;27(2):319–40.

4. Xiang B, Lu J, Yu J. Evaluating tooth segmentation accuracy and time efficiency in CBCT images using artificial intelligence: a systematic review and meta-analysis. *J Dent.* 2024;146:105064.
5. Kurt Bayrakdar S, Orhan K, Bayrakdar IS, Bilgir E, Ezhov M, Gusarev M, Shumilov E. A deep learning approach for dental implant planning in cone-beam computed tomography images. *BMC Med Imaging.* 2021;21(1):86.
6. Alqutaibi AY, Algabri R, Ibrahim WI, Alhajj MN, Elawady D. Dental implant planning using artificial intelligence: A systematic review and meta-analysis. *J Prosthet Dent.* 2024.
7. Kuckertz S, Papenberg N, Honegger J, Morgas T, Haas B, Heldmann S, editors. Deep-learning-based CT-CBCT image registration for adaptive radio therapy. *Medical Imaging 2020: Image Processing*; 2020: SPIE.
8. Dot G, Chaurasia A, Dubois G, Savoldelli C, Haghighat S, Azimian S, et al. DentalSegmentator: Robust open source deep learning-based CT and CBCT image segmentation. *J Dent.* 2024;147:105130.
9. Tyndall DA, Price JB, Tetradis S, Ganz SD, Hildebolt C, Scarfe WC. Position statement of the American Academy of Oral and Maxillofacial Radiology on selection criteria for the use of radiology in dental implantology with emphasis on cone beam computed tomography. *Oral surgery, oral medicine, oral pathology and oral radiology.* 2012;113(6):817–26.
10. Fedorov A, Beichel R, Kalpathy-Cramer J, Finet J, Fillion-Robin J-C, Pujol S, et al. 3D Slicer as an image computing platform for the Quantitative Imaging Network. *Magn Reson Imaging.* 2012;30(9):1323–41.
11. Dice LR. Measures of the amount of ecologic association between species. *Ecology.* 1945;26(3):297–302.
12. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med.* 2016;15(2):155–63.
13. Çiçek Ö, Abdulkadir A, Lienkamp SS, Brox T, Ronneberger O, editors. 3D U-Net: learning dense volumetric segmentation from sparse annotation. *International conference on medical image computing and computer-assisted intervention*; 2016: Springer.
14. Oktay O, Schlemper J, Folgoc LL, Lee M, Heinrich M, Misawa K et al. Attention u-net: Learning where to look for the pancreas. *arXiv preprint arXiv:180403999.* 2018.
15. Isensee F, Jaeger PF, Kohl SA, Petersen J, Maier-Hein KH. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. *Nat Methods.* 2021;18(2):203–11.
16. Hatamizadeh A, Nath V, Tang Y, Yang D, Roth HR, Xu D, editors. Swin unetr: Swin transformers for semantic segmentation of brain tumors in mri images. *International MICCAI brainlesion workshop.* Springer; 2021.
17. Cardoso MJ, Li W, Brown R, Ma N, Kerfoot E, Wang Y et al. Monai: An open-source framework for deep learning in healthcare. *arXiv preprint arXiv:221102701.* 2022.
18. Shorten C, Khoshgoftaar TM. A survey on image data augmentation for deep learning. *J big data.* 2019;6(1):1–48.

19. Milletari F, Navab N, Ahmadi S-A, editors. V-net: Fully convolutional neural networks for volumetric medical image segmentation. 2016 fourth international conference on 3D vision (3DV); 2016: Ieee.
20. Karimi D, Salcudean SE. Reducing the hausdorff distance in medical image segmentation with convolutional neural networks. IEEE Trans Med Imaging. 2019;39(2):499–513.
21. Likert R. A technique for the measurement of attitudes. Archives Psychol. 1932.
22. Brooke J. A quick and dirty usability scale, sl: Usability evaluation in industry. 1996.
23. Weng W, Zhu X. INet: convolutional networks for biomedical image segmentation. Ieee Access. 2021;9:16591–603.
24. Hatamizadeh A, Nath V, Tang Y, Yang D, Roth H, Xu D. Swin transformers for semantic segmentation of brain tumors in MRI images. Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries.272 – 84.
25. Gupta J, Ali SP. Cone beam computed tomography in oral implants. Natl J maxillofacial Surg. 2013;4(1):2–6.