International Orthodontics 2026; 24: 101081

Websites:
www.em-consulte.com
www.sciencedirect.com

Systematic review

Effectiveness of automated segmentation of maxillofacial structures in cone-beam computed tomography images using artificial intelligence: A systematic review

Alin M. Iacob [1], Alessio Verdecchia [1,2], Yolanda García-Mesa [3,4], Enrico Spinas [2], Teresa Cobo [1]

Available online: 8 October 2025

1. Orthodontics Division, Asturian Dentistry Institute, University of Oviedo, Street Catedrático Jose Serrano 10, 33006 Oviedo, Spain
2. Department of Surgical Sciences, Postgraduate School in Orthodontics, University of Cagliari, 09124 Cagliari, Italy
3. Grupo SINPOS, Department of Cell Biology and Morphology, University of Oviedo, 33006 Oviedo, Spain
4. Instituto de Investigación Sanitaria del Principado de Asturias (ISPA), Oviedo, Spain

Correspondence:
Alin M. Iacob, Orthodontics Division, Asturian Dentistry Institute, University of Oviedo, Street Catedrático Jose Serrano 10, 33006 Oviedo, Spain.
alini63@yahoo.com

Keywords
Artificial intelligence
Convolutional neural network
Cone-beam computed Tomography
Image processing
Jawbone
Segmentation
Orthodontics
Dentistry

Summary

Background > The automated segmentation of maxillary and mandibular bones in cone-beam computed tomography (CBCT) using artificial intelligence (AI) is redefining the standards of digital dentistry and orthodontics, with applications in mini-implant placement, dental implantology, orthognathic surgery, and bone graft planning.
Objective > To systematically assess the performance of AI models – particularly U-Net-based convolutional neural networks (CNNs) – for automated segmentation of maxillary bone structures in CBCT, following the PICOS model (Population – CBCT scans of human maxillae; Intervention – AI-based segmentation; Comparator – manual segmentation; Outcome – accuracy; Study design – diagnostic accuracy studies).
Material and methods > This systematic review adhered to PRISMA 2020 guidelines and was registered in PROSPERO (CRD42024592182). Eligibility criteria included studies applying AI to maxillary bone segmentation in CBCT and reporting quantitative accuracy metrics. Risk of bias was evaluated using the QUADAS-2 tool. The GRADE tool for formulating and grading recommendations in clinical practice was also employed. Data collected comprised number of CBCT scans, AI model architecture, evaluation metrics, and reported clinical applications.
Results > Thirty-one studies, analysing 11,432 CBCT scans, met the inclusion criteria. AI models consistently achieved high segmentation accuracy, with Dice similarity coefficients frequently exceeding 0.98, while substantially reducing processing time compared to manual segmentation.

10.1016/j.ortho.2025.101081

Applications ranged from implant planning and orthognathic surgery to digital orthodontics. Persistent challenges included anatomical variability, imaging artifacts, and the limited availability of high-quality annotated datasets.

Conclusions > AI-based segmentation of maxillary and mandibular bones in CBCT demonstrates promising accuracy and efficiency compared with manual techniques. Nevertheless, the certainty of evidence is limited by retrospective designs and small, heterogeneous samples. Large-scale, prospective multicentre studies with standardized evaluation are needed before these methods can be reliably adopted in routine clinical practice.

Background

In orthodontics, precision is not optional – it is essential. Achieving optimal treatment outcomes requires an in-depth assessment of the dentoskeletal structures that define each patient's anatomy. The clinical value of segmentation lies in the reliability of the 3D models it generates. For implant planning, precise measurements of bone height, width, and angulation determine the choice of implant diameter and length, as well as insertion trajectory, directly influencing long-term survival [1]. In orthognathic surgery, accurate 3D reconstructions allow objective quantification of mandibular asymmetries, enabling individualized osteotomy planning and more predictable postoperative stability [2]. In orthodontics, segmentation combined with grey-value analysis provides information on cortical bone thickness and density, which are decisive for the primary stability of temporary anchorage devices [3]. Similarly, assessment of alveolar bone levels supports safe tooth movement and reduces the risk of iatrogenic bone dehiscence or fenestrations. Thus, robust segmentation is not merely a technical step but a prerequisite for safe interventions and improved treatment outcomes.

Cone-beam computed tomography (CBCT) has revolutionized orthodontic diagnostics by providing high-resolution, three-dimensional (3D) reconstructions of the jaws and surrounding anatomical regions. Compared with conventional two-dimensional imaging, CBCT delivers a richer and more accurate depiction of dental and skeletal relationships, enabling more informed clinical decisions and personalized treatment planning [4–6].

However, fully harnessing this potential still relies on manual segmentation of the jaw bones in CBCT images – a process that is time-consuming, demands substantial technical expertise, and remains prone to human error [7]. In response, there is a growing interest in the development of automated segmentation techniques driven by artificial intelligence (AI) to enhance the efficiency and accuracy of this process [4–6].

To address these limitations, growing interest has focused on AI-driven automated segmentation, with deep learning – particularly convolutional neural networks (CNNs) – showing high precision in medical image segmentation [4,5]. The application of these techniques in orthodontics has the potential to significantly optimize clinical workflows by reducing image analysis time and increasing the reliability of results [4,5,8].

The effectiveness of these automated methods is evaluated using quantitative metrics such as the Dice-Sørensen coefficient (DSC), which is widely used to quantify the overlap between automated segmentation and a manual reference [9,10]. High DSC values indicate strong agreement, which is essential to ensure the clinical applicability of these AI-based methods [10,11].

In orthodontics, achieving high precision in jawbones segmentation is critical for accurate treatment planning, as it directly influences the outcomes of key procedures such as the placement of temporary anchorage devices (TADs) like miniscrews, dental implants, and orthognathic surgery [12–14].

Despite advances in automated segmentation techniques, several challenges remain. A major obstacle is the inherent variability of CBCT images due to differences in patient anatomy, imaging protocols, and the presence of artifacts that can obscure bony structures [6,15].

In addition, the segmentation of trabecular bone – crucial for understanding the mechanical properties of the jaws – poses a unique challenge owing to its complex internal architecture [6].

To overcome these limitations, researchers are exploring approaches such as the use of multimodal data, hybrid models that combine traditional methods with deep learning, and more robust datasets encompassing greater anatomical variability [4–6,8]. Emerging technologies, including generative adversarial networks (GANs) and transfer learning, are also being incorporated to enhance the accuracy and adaptability of algorithms [4,5,8]. Ultimately, the successful implementation of automated segmentation tools in orthodontics could lead to improved patient outcomes, more efficient treatment planning, and a reduction in the overall cost of care.

Therefore, the aim of this systematic review is to comprehensively synthesize and analyse the existing evidence on automated segmentation techniques for jawbones in CBCT images using AI, with a focus on their effectiveness, accuracy, and clinical applications in dentistry.

Material and methods

This systematic review was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines proposed by Page et al. [16] in 2020 and was registered in PROSPERO (International Prospective Register of Systematic Reviews) under the registration number CRD42024592182.

Search strategy

The database search was conducted on September 19th, 2024, and updated on March 10th, 2025. The search was limited to studies involving humans, with no language restrictions. The literature search encompassed all studies published up to March 2025, with no restrictions on the lower bound of publication date. Keywords were used in all possible combinations through the Boolean operators "or" and "and" along with the wildcard "*" for truncation, in the databases PubMed, Elsevier Scopus®, Web of Science, Embase, and Cochrane (*table I*). OpenGrey was also searched to identify grey literature.

Research question

Given the growing use of AI in jawbone segmentation in CBCT, this review focuses on answering the following question: "What is the accuracy of jawbone segmentation in CBCT using artificial intelligence compared with manual segmentation?".
To structure the search and selection of the studies included in this review, the following PICOS criteria were defined:

- population (P): cone-beam computed tomography (CBCT) images of the dental and maxillofacial region in human subjects;
- intervention (I): AI-based jawbone segmentation model;
- comparison (C): expert manual segmentation performed by clinicians (dentists, radiologists, or maxillofacial surgeons) or validated algorithms used as the reference standard;
- outcomes (O): segmentation performance of the AI model, evaluated through metrics such as DSC, accuracy, sensitivity, specificity, processing time, among others;
- study design (S): evidence studying the effectiveness of deep learning methods for segmenting maxillary and mandibular bones in CBCT images; non-experimental cross-sectional studies.

Table I
Search strategy

Databases	Search strategy
PubMed	(Artificial intelligence [MeSH] or AI or "machine learning" or "deep learning") and (CBCT or "cone-beam CT" or "cone-beam computed tomography" or CT OR "computed tomography") and (segmentation OR segment* or "image analysis" or "image processing") and (alveolar OR alveol* or jaw or dental or maxillofacial) and bone
Web of Science	TS = ("artificial intelligence" or "AI" or "machine learning" or "deep learning") and TS = ("CBCT" or "cone-beam CT" or "cone-beam computed tomography" or "computed tomography" or "CT") and TS = ("segmentation" or "segment*" or "image analysis" or "image processing") and TS = ("alveolar" or "alveol*" or "jaw" or "dental" or "maxillofacial") and TS = ("bone")
Elsevier Scopus	TITLE-ABS-KEY ("artificial intelligence" or "AI" or "machine learning" or "deep learning") and TITLE-ABS-KEY ("CBCT" or "cone-beam CT" or "cone-beam computed tomography" or "computed tomography" or "CT") and TITLE-ABS-KEY ("segmentation" or "segment*" or "image analysis" or "image processing") and TITLE-ABS-KEY ("alveolar" or "alveol*" or "jaw" or "dental" or "maxillofacial") and TITLE-ABS-KEY ("bone")
Embase	("Artificial intelligence"/exp or "artificial intelligence" or "AI" or "machine learning" or "deep learning") and ("CBCT" or "cone-beam CT" or "cone-beam computed tomography" or "CT" or "computed tomography") and ("segmentation" or "segment*" or "image analysis" or "image processing") and ("alveolar" or "alveol*" or "jaw" or "dental" or "maxillofacial") and "bone"
Cochrane	("Artificial intelligence" or "AI" or "machine learning" or "deep learning") and ("CBCT" or "cone-beam CT" or "cone-beam computed tomography" or "CT" or "computed tomography") and ("segmentation" or "segment*" or "image analysis" or "image processing") and ("alveolar" or "alveol*" or "jaw" or "dental" or "maxillofacial") and "bone"
Opengrey	"Artificial intelligence" or "machine learning" or "deep learning" and ("CBCT" or "cone-beam CT" or "cone-beam computed tomography" or "computed tomography") and (segmentation or "image analysis" or "image processing") and (alveolar or jaw or dental or maxillofacial) and bone

Inclusion criteria

Studies published from 2018 onwards, with no language restrictions, using CBCT of the maxillofacial region in human subjects. Studies applied AI-based jawbone segmentation models – preferably deep learning – and compared them with a reference standard (expert manual segmentation or validated algorithms). Performance had to be reported through quantitative metrics (e.g., DSC, accuracy, sensitivity, specificity, processing time).

Exclusion criteria

Studies not using CBCT as the primary imaging modality, not involving oral or maxillofacial structures, or lacking sufficient methodological details (dataset, imaging modality, model description) were excluded. Also excluded were review articles, conference abstracts without full text, non-human studies, works without quantitative performance metrics, or those using synthetic data without validation on real clinical datasets.

Study selection and data extraction

The electronic search was independently performed by two reviewers (blinded for review), who screened the titles and abstracts of the studies identified through the various database search strategies in parallel. Duplicates were removed beforehand using the Zotero reference manager. Potentially relevant studies were selected for full-text review, applying the same predefined inclusion and exclusion criteria at this stage. Excluded articles and exclusion reason were recorded in an *table S1*, online resource. Discrepancies between the two reviewers were resolved by a third author (blinded for review). In all cases, the reasons for exclusion were documented.

Data extraction was performed by (blinded for review) using a standardized table. The extracted variables included: author and year of publication, study design, type of CNN used, number of scans and their distribution into training, validation, and test sets, voxel size, DSC score with standard deviation (SD), and the reference method (ground truth) used for comparison.

Risk of bias assessment

Based on the JBI recommendations [17] and the review by Ma et al. [18], the Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) tool was applied to evaluate the methodological quality of the included studies and identify potential sources of bias. This tool examines four domains: patient selection, index test (AI technique or diagnostic method), reference standard, and study flow/timing. It also assesses applicability in terms of patient selection and reference standard.

Before evaluation, reviewers agreed on specific criteria for study inclusion/exclusion and applied them consistently. In this review, studies were classified into two categories: "low risk'', "high risk'' and "unclear'', the latter grouping studies with potential bias and those lacking sufficient information for a clear judgment.

Two reviewers independently conducted the assessment, resolving disagreements by consensus. Inter-rater reliability was calculated using Cohen's kappa (κ) for each QUADAS-2 domain, with interpretation according to Landis and Koch, where κ values range from slight agreement (< 0.20) to almost perfect agreement (> 0.80).

The GRADE tool was applied following established methodological guidance to assess the certainty of evidence [19,20].

Results

The search strategy yielded a total of 529 results, the filtering and selection process of which is depicted in the flow diagram shown in *figure 1*. All included articles were published between 2019 and 2025, in English and Hungarian. The 31 eligible studies were diagnostic accuracy studies [4,5,7,21–48].

Of these references, 131 were retrieved from PubMed, 55 from Elsevier SCOPUS, 209 from Web of Science, 130 from Embase, and 4 from Cochrane. No grey literature was identified. After manual filtering by title and author, 157 duplicate articles were removed.

Screening of titles and abstracts led to the exclusion of 278 articles that were outside the scope of this systematic review, leaving 94 potentially eligible articles for full-text assessment.

Subsequently, 63 articles were excluded based on one or more predefined criteria. The most frequent reasons were the absence of maxillary bone segmentation (n = 32), insufficient description of the deep learning methodology employed (n = 9), use of imaging modalities other than CBCT (n = 15), and study objectives not aligned with the established inclusion/exclusion criteria (n = 3). Additional reasons included the omission of key quantitative performance metrics such as DSC or accuracy (n = 2), absence of artificial intelligence application (n = 1), and studies limited to case series (n = 1). The results of the included studies were compiled in a standardized table (*table II*).

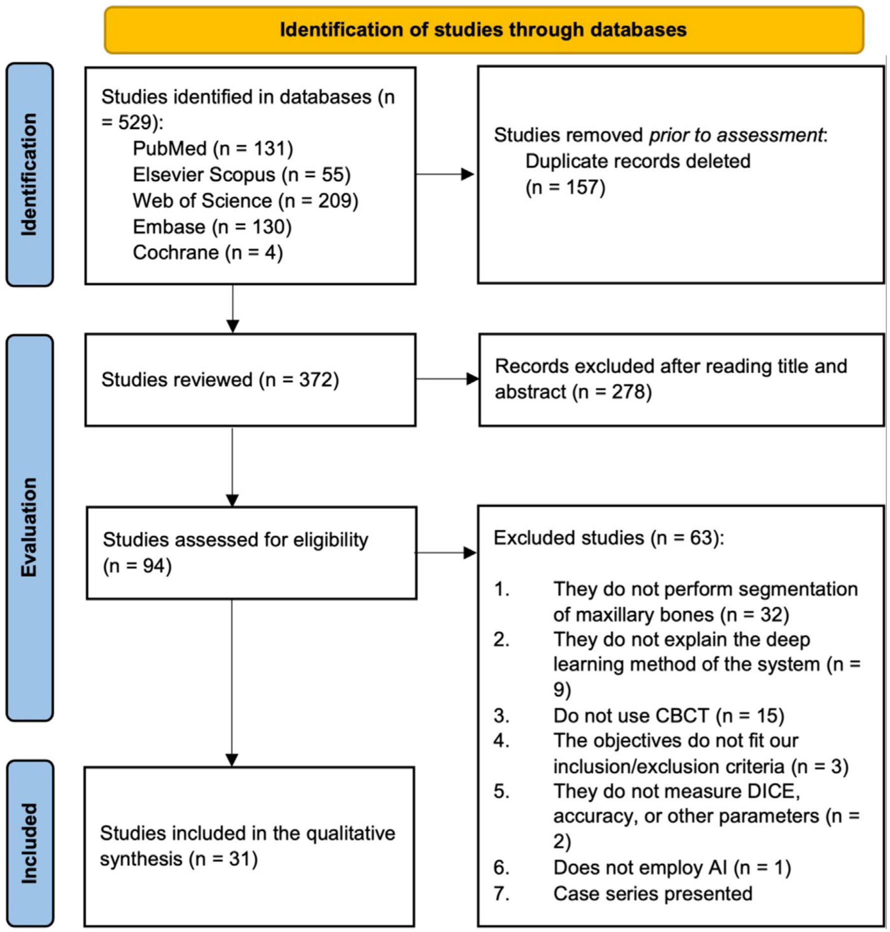

FIGURE 1

PRISMA Flowchart of the study selection process

Study characteristics

In this study, a total of 31 diagnostic accuracy studies (*table II*) investigating the use of artificial intelligence for maxillary bone segmentation in CBCT images were reviewed. Collectively, these studies analysed 11,432 CBCT scans, with sample sizes ranging from 13 [34] to 4.938 [7] cases. The study populations were heterogeneous, comprising healthy individuals, orthodontic patients, individuals with maxillofacial deformities, patients undergoing orthognathic surgery, and edentulous patients undergoing implant planning. While some investigations focused on specific groups, others aimed to enhance the clinical applicability of the evaluated models by including more diverse populations. Regarding age distribution, there was substantial variability, with studies including paediatric patients as young as 11 years and older adults over 90 years of age. Sex distribution, reported in some studies, showed a relatively balanced proportion of male and female participants [5,43].

Table II
Articles included in the qualitative synthesis

Author and year	Type of study	Anatomical structure segmented	CNN used	Total of scans; training/ validation scans; testing scans	Voxel size	DICE score (standard deviation)	Comparison
Wu et al., 2025 [21]	Non-experimental retrospective cross-sectional study	Mandibular condyle	Modified U-Net 3D	Total of scans: 490 Training/Validation: not specified Testing: not specified	N/S	Cortical: 0.9010 (0.0090) [internal], 0.8870 (0.0110) [external] Trabecular: 0.9690 (0.0080) [internal], 0.9770 (0.0120) [external] Complete condyle: 0.9820 (0.0080) [internal], 0.9840 (0.0100) [external]	Manual threshold adjustment in 3D Slicer by an expert, reviewed by two experienced radiologists
Hegyi et al., 2024 [22]	Non-experimental retrospective cross-sectional study	Mandible	Three pre-trained CNN models	Total of scans: 70 Training/Validation: 5713 Testing: 15	Original: 0.15–0.3 mm After Resampling: 0.2 mm	0.95 (0.01)	Manual or semi-automatic segmentation by a maxillofacial radiology expert
Chen et al., 2024 [23]	Non-experimental retrospective cross-sectional study	Maxilla, mandible, teeth	SegResNet	Total of scans: 1417 Training/Validation: not specified Testing: not specified	0.125–0.3 mm	0.9570 (0.0250) [internal], 0.9440 [external]	Manual segmentation by a dentist
Chen et al., 2024 [24]	Non-experimental retrospective cross-sectional study	Maxilla, mandible, teeth, periapical lesions	nnUNet, Dynamic Graph CNN, Minkowski CNN, PointNet++	Total of scans: 138 Training/Validation: 103, Testing: 35	0.125 mm	0.9178 (no SD)	Manual segmentation by a dentist
Liu et al., 2024 [25]	Non-experimental retrospective cross-sectional study	Teeth, maxillary and mandibular alveolar bone	U-Net Architecture	Total of scans: 506 Training/Validation: 451 Testing: 55	0.18–1.0 mm	Mandible: 0.9830 (0.0180) Maxilla: 0.9240 (0.0310)	Manual segmentation by a dentist
Nogueira-Reis et al., 2024 [26]	Non-experimental retrospective cross-sectional study	Maxilla, mandible, mandibular canal, maxillary sinus, teeth, pharyngeal airway space	3D U-Net	Total of scans: 30 Training/Validation: not specified Testing: not specified	0.2–0.3 mm	0.996 (0.057)	Manual segmentation by a dentist
van Nistelrooij et al., 2024 [27]	Non-experimental retrospective cross-sectional study	Mandible segmentation	Six integrated U-Net 3D CNNs to operate simultaneously as a single unit	Total of scans: 335 Training/Validation: not specified Testing: not specified	Unspecified	0.904 (no SD)	Manual segmentation by a clinician using ITK-SNAP

Table II **(Continued).**

Author and year	Type of study	Anatomical structure segmented	CNN used	Total of scans; training/validation scans; testing scans	Voxel size	DICE score (standard deviation)	Comparison
Dot et al., 2024 [28]	Non-experimental retrospective cross-sectional study	Maxilla, mandible, teeth, mandibular canal	Full-resolution U-Net 3D	Total of scans: 273 Training/Validation: not specified Testing: not specified	0.16 mm	Internal test set: 0.9220 (0.0630) External test set: 0.9420 (0.0740)	Manual segmentation by a dentist
Tian et al., 2024 [29]	Non-experimental retrospective cross-sectional study	Maxilla, mandible	Based on the 3D UNesT network	Total of scans: 155 Training/Validation: 115 Testing: 40	0.16–0.45 mm	Mandibular cortical: 0.9369 (0.0139) Mandibular trabecular: 0.9683 (0.0069) Maxillary cortical: 0.8614 (0.0402) Maxillary trabecular: 0.9557 (0.0182)	Manual annotation made by two dentists and validated by a dental expert using ITK-SNAP software
Liu et al., 2023 [30]	Non-experimental retrospective cross-sectional study	Mandible (anterior alveolar bone)	Tooth Swin Transformer Network (TSTNet) – CBCT segmentation	Total of scans: 503 Training/Validation: not specified Testing: not specified	0.125–0.5 mm	0.9399 (0.0120)	Manual segmentation by a dentist
Moufti et al., 2023 [31]	Non-experimental retrospective cross-sectional study	Mandible (edentulous mandibular bone)	IOSNet – IOS segmentation	Total of scans: 43 Training/Validation: 33 Testing: 10	Unspecified	Training: 0.89, test: 0.78 No SD	Manual segmentation by researchers using ITK-SNAP
Tao et al., 2023 [32]	Non-experimental retrospective cross-sectional study	Maxilla (palatal bone)	U-Net	Total of scans: 70 Training/Validation: not specified Testing: not specified	0.25 mm	0.958 (0.025)	Manual segmentation by a dentist
Vinayahalingam et al., 2023 [33]	Non-experimental retrospective cross-sectional study	Mandiblular condyle, glenois fossae	U-Net 3D	Total of scans: 162 Training/Validation: not specified, Testing: not specified	0.4 mm	Condyle: 0.976 (0.01)	Manual segmentation by dentists
Xu et al., 2023 [34]	Non-experimental retrospective cross-sectional study	Maxilla, mandible, teeth, periodontal ligament	CNN mixed scale	Total of scans: 13 Training/Validation: 5 Testing: 4	0.15 mm	0.9514 (0.0121)	Manual segmentation by a dentist
Nogueira-Reis et al., 2023 [35]	Non-experimental retrospective cross-sectional study	Maxilla, maxillary teeth, maxillary sinus	Modified Multiplanar Net	Total of scans: 40 Training/Validation: not specified Testing: not specified	0.2–0.3 mm	0.998 (no OD)	Manual segmentation by a dentist

Table II (Continued).

Author and year	Type of study	Anatomical structure segmented	CNN used	Total of scans; training/validation scans; testing scans	Voxel size	DICE score (standard deviation)	Comparison
Fontenele et al., 2023 [4]	Non-experimental retrospective cross-sectional study	Maxilla (alveolar bone)	3D U-Net	Total of scans: 141 Training/Validation: 99/12 Testing: 30	0.1–0.3 mm	0.970 (0.020)	Manual segmentation by clinical experts
Ilesan et al., 2023 [36]	Non-experimental retrospective cross-sectional study	Mandible, mandibular canal, mandibular teeth	FCN1, FCN2	Total of scans: 180 Training/Validation: 160 Testing: 20	0.88–1.37 mm	0.916 (0.033)	Manual segmentation by a dentist
Cui et al., 2022 [7]	Non-experimental retrospective cross-sectional study	Maxilla, mandible, teeth	V-Net	Total of scans: 4938 Training/Validation: not specified Testing: not specified	0.4 mm (normalized)	0.945 (internal data) 0.938 (external data)	Combining manual segmentation and other segmentation methods
Fan et al., 2022 [37]	Non-experimental retrospective cross-sectional study	Mandible	U-Net 3D Architecture	Total of scans: 120 Training/Validation: not specified Testing: not specified	0.5 mm	0.9696 (0.005)	Manual segmentation by dentists using ITK-SNAP
Gillot et al., 2022 [38]	Non-experimental retrospective cross-sectional study	Maxilla, mandible, cranial base, cervical vertebrae, skin	UNETR (U-Net Transformers)	Total of scans: 578 Training/Validation: not specified Testing: not specified	0.4 mm (normalized, resampling)	Mandible: 0.962 (0.020) Maxilla: 0.853 (0.064)	Manual segmentation by 3 clinical experts
Jha et al., 2022 [39]	Non-experimental retrospective cross-sectional study	Mandibular condyle	U-Net 3D, U-Net 3D Cascading	Total of scans: 234 Training/Validation: not specified Testing: not specified	0.3 mm	0.932 ± 0.023	Manual segmentation by clinical experts
Preda et al., 2022 [40]	Non-experimental retrospective cross-sectional study	Maxillary bone, palatine bone, zygomatic bone, nasal bone, lacrimal bone	3D U-Net	Total of scans: 144 Training/Validation: not specified Testing: not specified	Raw segmentation: 1.5 mm/voxel, refined segmentation: 0.23 mm/voxel	0.926 (0.010)	Manual segmentation by clinical experts
Kim et al., 2021 [41]	Non-experimental retrospective cross-sectional study	Mandibular condyle	U-Net	Total of scans: 18 Training/Validation: 5 Testing: 2	0.3 mm	Trabecular: 0.870 (0.023) Cortical: 0.734 (0.032)	Manual segmentation by a dentist
Le et al., 2021 [42]	Non-experimental retrospective cross-sectional study	Mandible (ramus and condyle)	U-Net	Total of scans: 109 Training/Validation: 90 Testing: 19	0.2–0.4 mm	0.91 (0.03)	Manual segmentation by dentists
Lo Giudice et al., 2021 [43]	Non-experimental retrospective cross-sectional study	Mandible	Tiramisu Model	Total of scans: 40 Training/Validation: 20 Testing: 20	0.47 × 0.47 × 0.3 mm	0.972, no SD	Manual segmentation by dentists

TABLE II **(Continued).**

Author and year	Type of study	Anatomical structure segmented	CNN used	Total of scans; training/ validation scans; testing scans	Voxel size	DICE score (standard deviation)	Comparison
Verhelst et al., 2021 [44]	Non-experimental retrospective cross-sectional study	Mandible	Two CNNs using a 3D U-Net architecture	Total of scans: 190 Training/Validation: 160 Testing: 30	0.3 mm	0.9722 (0.0062)	Semi-automatic segmentation or AI-refined segmentation of the user
Wang et al., 2021 [45]	Non-experimental retrospective cross-sectional study	Maxilla	3D U-Net	Total of scans: 60 Training/Validation: 30 Testing: not specified	0.25 mm	0.92 (0.01) (maxilla)	Manual segmentation by dentists
Wang et al., 2021 [5]	Non-experimental retrospective cross-sectional study	Maxilla, mandible	Three consecutive 3D U-Nets	Total of scans: 32 Training/Validation: 28/4 Testing: 2	Unspecified	Jaw bones: 0.934 (0.019) Teeth: 0.945 (0.021)	Manual segmentation by a dentist
Chen et al., 2020 [46]	Non-experimental retrospective cross-sectional study	Maxilla	Tiramisu, U-Net, LSTM	Total of scans: 96 Training/Validation: 36 Testing: 60	0.3 mm	0.800 (0.029)	Manual segmentation by dentists using ITK-SNAP software
Zhang et al., 2020 [47]	Non-experimental retrospective cross-sectional study	Maxilla (midface), mandible	U-Net (Multiple), Encoder, GAP, Fully Connected Layers	Total of scans: 107 Training/Validation: not specified Testing: not specified	0.3–0.4 mm	Midface: 0.9319 (0.0089) Mandible: 0.9327 (0.0097)	Both manual segmentation and other automatic segmentation methods
Torosdagli et al., 2019 [48]	Non-experimental retrospective cross-sectional study	Mandible	Tiramisu, U-Net, LSTM	Total of scans: 250 Training/Validation: not specified Testing: not specified	$0.754 \times 0.754 \times 0.377$ mm or $0.584 \times 0.584 \times 0.292$ mm	0.9386 (no SD)	Manual segmentation by three expert interpreters

CNN: convolutional neural network; N/S: not specified.

AI architectures employed

All studies included in this review employed artificial intelligence-based models for bone segmentation, with a clear predominance of convolutional neural networks (CNNs). The most frequently used architecture was the 3D U-Net and its variants, reported in at least 15 studies [4,5,21,25,26,28–35,39,40,44,48]. Other approaches included models such as V-Net [7], SegResNet [22], and transformer-based architectures like Swin-UNETR [24]. Some studies explored hybrid strategies, integrating multiple neural networks to segment different anatomical structures [35] or combining CBCT data with intraoral scans to enhance segmentation accuracy [23,30].

Voxel size

Voxel size, a critical parameter in automated segmentation, varied considerably across studies, ranging from 0.10 to 1.5 mm. Most employed sizes between 0.2 and 0.4 mm, balancing computational efficiency with segmentation accuracy [26,29,34,38,41,42,44]. Some studies used finer resolutions (< 0.15 mm) when segmenting small anatomical structures requiring greater detail, such as teeth and the periodontal ligament [4,25,26]. In contrast, studies using voxel sizes greater than 0.4 mm were scarce and generally focused on evaluating overall bone morphology rather than the detailed segmentation of specific structures [7,37,43,48]. In some cases, variability in voxel resolution was observed within the same study, with resampling techniques applied to standardize input data and improve model robustness [25].

Performance metrics used

The variables used to assess model accuracy were diverse, with the DSC being the most widely applied. Across the included studies, DSC values ranged from 0.734 [41] to 0.996 [26] with the highest scores generally associated with convolutional neural network models based on the 3D U-Net architecture [5,21,25,29,30,38,40,43]. Another frequently reported metric was the Hausdorff Distance (HD), examined in 20 studies, with values ranging from 0.045 [35] to 8.538 mm [34]. In general, lower HD values were observed for bone structure segmentation, whereas larger discrepancies occurred in the segmentation of more complex structures such as the periodontal ligament. Additional metrics, including the Intersection over Union (IoU) coefficient and the Average Surface Distance (ASD), were also investigated, supporting the reliability of segmentation in structures with well-defined boundaries [4,23,37].

Importantly, segmentation performance was not uniform across all jawbone structures. Several studies reported higher DSC values for trabecular bone than for cortical bone, with Tian et al. observing mandibular trabecular segmentation at 0.968 compared to 0.936 for cortical regions [25]. Similarly, Gillot et al. found lower accuracy for maxillary cortical bone (0.853) compared with mandibular bone (0.962) [34]. The segmentation of thin cortical plates, alveolar crests, and regions affected by metallic artifacts tended to yield lower scores than larger, well-defined structures such as the mandibular body or condyle.

Reference standard

Manual segmentation performed by experienced clinical experts served as the reference standard in most studies. In those specifying the evaluators' profiles, segmentation was primarily conducted by dentists, radiologists, or oral and maxillofacial surgery specialists with extensive experience [21,29,33,34,38,41,43,48]. Some studies employed semi-automatic methods or AI-refined segmentations as alternatives to purely manual approaches [7,22,44]. Overall, a high level of agreement was observed between automated and manual segmentation, with intraclass correlation coefficients (ICC) reaching values as high as 0.98, indicating excellent reproducibility [24].

CBCT devices used

CBCT images were acquired using a wide range of scanners, with the NewTom® VGi evo, Morita® Veraview, Alphard® 3030, Accuitomo® 3D, and i-CAT® CBCT among the most common. Some studies incorporated data from multiple devices across different clinical centers, enabling more robust model validation [7,25]. Despite promising results, several challenges were identified, including variability in image quality due to differences in voxel resolution and acquisition protocols [31,37,49], as well as the presence of metallic artifacts from dental restorations, implants, and orthodontic appliances, which reduced model accuracy in affected regions [4,35].

Risk of bias

The risk of bias for each study was independently assessed by two reviewers (A.M.I. and A.V.) using the QUADAS-2 tool, with results for each domain summarized in tables. The evaluated domains included patient selection, index test accuracy, reference standard, and study flow and timing. Overall, the QUADAS-2 assessment indicated a low global risk of bias, with some methodological uncertainties related to index test calculation

Table III

Cohen's Kappa (k) Index of inter-rater agreement of the QUADAS-2 tool, according Landis and Koch[1]

	Risk of bias				Applicability criteria	
QUADAS-2 domain	Patient selection	Calculated indices	Reference standard	Time flow	Patient selection	Reference standard
Cohen's Kappa (κ)	0.6517	0.6517	0.6517	0.6517	0.7832	1.0000

[1]Landis J.R., Koch G.G. The measurement of observer agreement for categorical data. Biometrics. 1977;33(1):159–74. PMID: 843571.

and study flow/timing. Inter-rater agreement (*table III*) showed substantial concordance in "patient selection" (κ = 0.8864) and "index test" (κ = 0.8276), and moderate agreement for "reference standard" and "flow/timing" (κ = 0.7727). Regarding applicability, moderate agreement was found for both "patient selection" (κ = 0.7727) and "reference standard" (κ = 0.6429), indicating reasonable consistency between reviewers.

The GRADE evaluation [19,20] showed that the overall certainty of evidence ranged from low to moderate (*table IV*). The primary outcome, segmentation accuracy, was rated as moderate certainty, supported by consistent reporting of high DSC values across most studies but downgraded due to retrospective designs, small and heterogeneous sample sizes, and lack of external validation. Efficiency outcomes (processing time) were judged to have low certainty, as only a limited number of studies reported this metric and with heterogeneous methodologies. The certainty of evidence for generalizability and clinical applicability was considered low to moderate, mainly due to reliance on single-centre datasets and inconsistent reporting of anatomical subregions. No criteria for upgrading were met.

Discussion

The aim of this systematic review was to synthesize and analyse the evidence on the effectiveness of automated segmentation of maxillary and mandibular bones in CBCT images using AI, with a focus on accuracy, efficiency, and clinical applications in dentistry. In particular, architectures based on CNNs were evaluated. Deep learning, especially CNNs, remain the most common AI approach for CBCT-based jawbone segmentation. U-Net architectures and their 2D/3D variants remain the most frequently employed approach and consistently report favourable segmentation outcomes [24,29,33,37,39,40].

However, the robustness of these findings is constrained by methodological weaknesses: all included studies were retrospective, sample sizes were often small and heterogeneous, and no study reported a priori power calculations. Moreover, datasets were largely single-centre, with limited external validation, and performance varied according to anatomical region, with cortical bone and alveolar crest segmentation proving more challenging than trabecular bone. Taken together, these limitations indicate that while AI-based segmentation demonstrates technical promise and faster processing compared to manual methods, the certainty of evidence is low to moderate, and clinical generalizability remains unproven.

Accuracy and efficiency

Segmentation accuracy is typically assessed using metrics such as the DSC, IoU, and HD [4,5,22,25,29,33,43,44]. Reported DSC values for jawbone segmentation range between approximately 80 and 98% [4,5,25,28,29,40,43,44]. For example, Tian et al. achieved mean DSCs of 93.69%, 96.83%, 86.14%, and 95.57% for different mandibular structures [29], whereas Preda

TABLE IV

Summary of findings table according to the Grading of Recommendation, Assessment, Development, and Evaluation (GRADE) approach for the main outcomes of the systematic review on AI-based segmentation of maxillary bones in CBCT images

Outcome	No. of studies (Refs)	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Effect estimate (range DSC)	Certainty of evidence
Overall segmentation accuracy (DSC $\geq$ 0.90)	28 [4,5,7,21–30, 32–40,42–45,47,48]	Non-experimental cross-sectional studies	Low	Low	Low	Low	Undetected	0.90–0.998	High
High accuracy (DSC $\geq$ 0.95)	15 [4,21–23,25,26,29, 32–35,37,38,43,44]	Non-experimental cross-sectional studies	Low	Serious (variation between studies)	Low	Low	Undetected	0.95–0.998	Moderate
External validation datasets	4 [7,21,23,28]	Non-experimental cross-sectional studies	Low	Serious (variation in voxel size and protocols)	Low	Moderate due to small sample sizes	Undetected	0.887–0.984	Moderate

Certainty ratings were determined based on risk of bias (QUADAS-2), inconsistency, indirectness, imprecision, and publication bias. Reference numbers in brackets indicate the studies contributing data to each outcome.

et al. reported 92.6% for maxillofacial bone segmentation [40]. Accuracy may vary according to bone type, CBCT image quality, artifact presence, and dataset characteristics [7,28,29].

These variations have direct clinical implications. Underestimation of cortical bone thickness may compromise assessments of safe sites for TAD insertion or implant stability, while reduced accuracy in the alveolar crest could affect orthodontic movement planning and risk evaluation for dehiscence. Conversely, reliable segmentation of trabecular bone and mandibular condyles supports accurate implant trajectory planning and orthognathic simulations. Therefore, clinicians should be aware that although AI achieves high average accuracy, performance is structure-dependent, and errors in critical regions may translate into clinically significant consequences.

In terms of efficiency, AI-based segmentation methods significantly outperform manual and semi-automatic techniques. Automated segmentation times are usually measured in seconds or minutes per scan, compared to tens or hundreds of minutes for manual segmentation [4,29,33,40,43]. One study reported a mean automated segmentation time of 0.25 minutes versus 12.9–17.7 minutes for manual refinement [29], while another found a 204-fold time reduction, averaging 39.1 seconds per scan [40].

Key elements and optimal performance

The U-Net algorithm is the most widely used and generally accurate for maxillary segmentation [24,29,33,37,39,40]. Two-stage deep learning systems enhance accuracy and manage CBCT variability [29]. The MONAI framework also stands out as a valuable tool for developing medical image segmentation models [22,31,38].

The literature highlights the importance of large, multicentre datasets to train robust and generalizable AI models [7,29]. Variability in image quality and acquisition parameters across centres requires training on diverse CBCT scans [7,29].

The DSC remains the most consistently reported accuracy metric [4,5,22,25,28,29,40,43,44], while efficiency is typically evaluated through segmentation time [4,29,33,40,43].

Comparison with traditional methods

AI-based segmentation offers clear advantages over traditional maxillary segmentation techniques [4,29], which include thresholding [29], region growing [38], edge detection [29], watershed algorithms [5], clustering, and classical machine learning methods [29].

Cui et al. [7] and Tian et al. [29] report key improvements with AI-based segmentation. Traditional methods often struggle with CBCT limitations such as low contrast, artifacts, and image noise [7,29], whereas AI methods achieve high performance, with DSCs often $\geq$ 90% [7,24,25,29,40]. In some cases, AI models even slightly outperform expert radiologists; for example, Cui et al. [7], obtained a mean DSC 0.5% higher for tooth and bone segmentation.

A major advantage of AI-based segmentation is speed [4,7,26,44]. Manual segmentation can take tens or hundreds of minutes per scan [5,31,37,38], and semi-automatic methods still require substantial operator input [36,44]. AI systems can complete segmentation in seconds or minutes [4,7,26,29,44]. For instance, one study reported 0.25 minutes for AI segmentation compared with 12.9–17.7 minutes for manual refinement, while another achieved complete segmentation in 17 seconds compared to 147–160 minutes for radiologists [7].

Traditional non-fully-automated methods are operator-dependent, introducing variability and potential errors [31,36,37,44]. In contrast, AI methods are automatic, consistent, and reproducible [7,25], with studies showing better intra- and inter-operator consistency compared to manual segmentation [44].

AI solutions to traditional challenges

Classical approaches such as thresholding and edge detection struggle with low contrast, noise, and CBCT artifacts [29,38], whereas deep learning models are less affected by these issues [7,49].

Traditional methods often require manual initialization, landmark selection, or post-processing [7], while newer AI systems aim for fully automated workflows [7,25,29].

AI-based methods have enabled segmentation of internal structures such as cortical and cancellous bone and the mandibular canal, rather than only gross maxillary bone boundaries [26,44]. Xu et al. [34] even achieved segmentation of the periodontal ligament, providing valuable anatomical detail for treatment planning and simulation purposes with finite element analysis. Although in some cases manual segmentation may achieve slightly higher accuracy, this is generally at the cost of considerably longer processing times [4]. The efficiency and improving accuracy of AI-based methods make them increasingly valuable for digital workflows [25,29].

Advantages of AI-based segmentation

Effective management of image imperfections

Traditional segmentation struggles with noisy CBCT data, low contrast, and artifacts [4,29,38,41]. Threshold-based segmentation performs poorly in low signal-to-noise and low-contrast images [4,38], while metallic restorations and implants produce artifacts that hinder accurate segmentation [25,29,34,38].

Deep learning-based AI methods better manage noise, artifacts, and intensity variations in CBCT images [24,25,29,49]. These models learn complex features from large datasets, reducing sensitivity to such issues [43,44,49]. AI segmentation has been shown to be less affected by artifacts than thresholding [49], and robust segmentation of teeth and bone has been reported despite blurred boundaries [7].

Network architecture and training strategies further enhance performance. For example, one study incorporated a filtering module into the network to improve bone boundary contrast [7], while Chen et al. embedded anatomical constraints into a U-

Net to increase accuracy [24]. Transformer-based models have also shown promise in handling artifacts from restorative dental materials in CBCT scans [24].

Generalization across populations and protocols

Generalization remains a debated topic [7,25]. Early AI methods and traditional approaches often suffer from poor generalizability due to reliance on small, single-center datasets with specific imaging protocols and demographics [7,38,44]. Recent AI studies emphasize the need for large, multicenter datasets with diverse populations and protocols to improve robustness [7,25]. Traditional methods frequently fail to generalize to data acquired with different scanners or patient groups [29,38,48]. Limited diversity in CBCT devices and anatomies restricts model applicability [31,35]. Studies evaluating AI systems on external datasets report better generalization [7]. For instance, one system maintained consistent performance on data from 15 centres despite image variations [7], and another improved segmentation accuracy across different scanner manufacturers using a smaller training set [25].

To enhance generalization, strategies include two-stage segmentation to account for field-of-view differences and ROI resizing to standardize voxel sizes [29]. Pre-processing techniques such as rotation, flipping, cropping, and intensity variation further strengthen model robustness [7,29,44].

Limitations of AI-based segmentation

A major limitation is the need for large, high-quality annotated datasets, which are often difficult and costly to obtain [4,7,25,26,29,34,35]. Deep learning performance depends heavily on training data, and small datasets can hinder generalization to different imaging protocols and patient populations [25]. Many existing methods have been tested on limited datasets [7,32], and exhaustive annotation for training imposes a significant workload [29].

Another concern is potential bias from non-representative training data [24]. Lack of heterogeneity – such as training only on one CBCT device type or narrow patient demographics – can introduce bias and reduce generalization [24,35]. Inaccurate manual annotations used as ground truth may also compromise model evaluation [24].

Additional challenges include high-density artifacts from implants, motion artifacts, complex anatomical variations such as deformed mandibular condyles, and difficulty segmenting fine cortical bone boundaries [29,41]. Poorly defined defect margins and crowded dentition add further complexity [5,34], potentially leading to under- or over-segmentation that requires refinement [29].

Clinical implications

AI-based automated maxillary segmentation represents a disruptive advance in dental clinical practice. In surgical planning, AI enables precise, patient-specific 3D models and surgical guides for orthognathic surgery and implant placement, improving accuracy and efficiency. AI can also assist in mandibular asymmetry analysis relevant to surgical planning [37].

In surgery and implantology, AI supports assessment of alveolar bone quality and quantity and identification of vital anatomical structures such as the mandibular condyle [42], mandibular canal, and maxillary sinus [36], enhancing surgical safety [22,25,49], and potentially reducing planning time [4,25,26]. Automated segmentation is particularly valuable in edentulous cases, aiding anatomical reconstruction and implant planning [31].

In orthodontics, AI facilitates longitudinal assessment of bone changes during treatment by comparing sequential 3D models, improving therapeutic monitoring [24]. It also aids in identifying anatomical variations such as mandibular asymmetry [37] and planning precise placement of temporary anchorage devices (TADs) by automatically measuring palatal bone thickness [32]. Automated segmentation has proven useful in orthodontic planning for patients with cleft alveolus and palate, where anatomy is complex [5,38,45], and also for the generation of finite element analysis models used for 3D simulations [34].

By enabling rapid generation of patient-specific 3D models from CBCT data, AI fosters personalized, fully digital dentistry [26,28,29,40], supporting individualized treatment planning, detailed clinical simulations, and device fabrication tailored to patient anatomy [29,40,44]. Integration with other data sources, such as intraoral scans, further enhances model accuracy [23,28].

Future perspectives

Future research should focus on improving AI robustness and accuracy [24,27,29,31,32,35,39], including training on larger, more diverse datasets encompassing various anatomies, pathologies, and CBCT devices [24,27,29,31,32,35,39]. Algorithms should better handle challenging cases with artifacts and anatomical variability [24,25,29,31,39], and explore advanced architectures such as transformer-based models [24,38,39]. Further development is needed for accurate substructure segmentation, including cortical and cancellous bone and the mandibular canal [22,27,39,40], as well as differentiation of individual teeth and bony segments [25,27].

Integrating CBCT with other imaging modalities, such as intraoral scans, can improve crown reconstruction accuracy and generate more complete 3D models [7,25,26,35,47].

Demonstrating clinical utility requires rigorous validation studies and clinical trials [25–29,33,35,40,43], including external validation on diverse datasets and CBCT scanners [27,31,32,35,39,40], and comparison with expert manual segmentation using standard metrics such as the DSC [4,22,24,25,29,40,41,43,48]. Consistent reporting of time efficiency versus manual methods is also needed [4,26,28–30,32,33,35,36,38,40,41,43,47]. More clinical trials are required to assess the impact on treatment planning, surgical outcomes, and overall patient care [25,27,28,35].

Limitations of the present work

Although reported accuracy values were generally high, the evidence base is weakened by several methodological limitations. All included studies were retrospective, most used small and heterogeneous samples, and none reported power calculations. Furthermore, datasets were largely single-centre, limiting external validity. These factors increase the risk of bias and reduce the certainty of evidence. Therefore, the results should be interpreted with caution, as the current evidence demonstrates technical feasibility rather than established clinical reliability.

Conclusions

This systematic review suggests that AI-based segmentation of maxillary and mandibular bones in cone-beam computed tomography (CBCT) shows promising performance, with high accuracy metrics and reduced processing times compared with manual techniques. These findings indicate a potential to facilitate the generation of anatomically precise 3D models for diagnosis, surgical planning, and orthodontic treatment. However, the certainty of the evidence remains limited: all included studies were retrospective, with small and heterogeneous samples, and many lacked external validation. Accordingly, while deep learning architectures such as convolutional neural networks (CNNs) demonstrate strong technical potential, their clinical integration should be approached cautiously. Future research must prioritize large-scale, prospective, multicentre studies with standardized evaluation metrics to confirm reliability and generalizability before AI-based segmentation can be considered for routine use in maxillofacial diagnostics and treatment planning.

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

Contribution : Alin M. Iacob: conceptualization; methodology; formal analysis; visualization; writing – original draft.
Alessio Verdecchia: project administration; supervision; validation; writing – review and editing.
Yolanda García-Mesa: software; resources; writing – review and editing.
Enrico Spinas: validation; visualization.
Teresa Cobo: investigation; data curation; validation; writing – review and editing.

Disclosure of interest : The authors declare that they have no competing interest.

Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ortho.2025.101081.

References

[1] Marques VR, Soh D, Cerqueira G, Orgev A. Restorative artificial intelligence-driven implant dentistry for immediate implant placement with an interim crown: a clinical report. J Prosthet Dent 2025. doi: 10.1016/j.prosdent.2025.07.006. S022-3913 (25) 00570-0.

[2] Lo HY, Leung PH, Su Y, Leung YY, Yeung AWK, Yang W. AI-driven CBCT segmentation and 3D modeling of the anterior surface of maxilla for computer-assisted surgery: a comparison of multiple algorithms. J Craniomaxillofac Surg 2025;53:1683–90. doi: 10.1016/j.jcms.2025.07.007.

[3] Srinivasan S, Tee BC, Wang A, Gohel A, Kim DG, Deguchi T, et al. Reliability and accuracy of assessing temporary anchorage device-tooth root contact with cone-beam computed tomography. Am J Orthod Dentofacial Orthop 2021;159:271–80. doi: 10.1016/j.ajodo.2020.01.020.

[4] Fontenele RC, Gerhardt M, do N, Picoli FF, Van Gerven A, Nomidis S, et al. Convolutional neural network-based automated maxillary alveolar bone segmentation on cone-beam computed tomography images. Clin Oral Implants Res 2023;34:565–74. doi: 10.1111/clr.14063.

[5] Wang H, Minnema J, Batenburg KJ, Forouzanfar T, Hu FJ, Wu G. Multiclass CBCT image segmentation for orthodontics with deep learning. J Dent Res 2021;100:943–9. doi: 10.1177/00220345211005338.

[6] Zhang S, Jiang B, Shi H. Jawbone segmentation with trabecular bone preservation from cone beam CT images. Appl Sci 2022;12:1556. doi: 10.3390/app12031556.

[7] Cui Z, Fang Y, Mei L, Zhang B, Yu B, Liu J, et al. A fully automatic AI system for tooth and alveolar bone segmentation from cone-beam CT images. Nat Commun 2022;13:2096. doi: 10.1038/s41467-022-29637-2.

[8] Bayrakdar SK, Orhan K, Bayrakdar IS, Bilgir E, Ezhov M, Gusarev M, et al. A deep learning approach for dental implant planning on cone-beam computed tomography images. BMC Med Imaging 2021;21(1):86. doi: 10.21203/rs.3.rs-286578/v1.

[9] Almajalid R, Zhang M, Shan J. Fully automatic knee bone detection and segmentation on three-dimensional MRI. Diagnostics 2022;12:123. doi: 10.3390/diagnostics12010123.

[10] Taha AA, Hanbury A. Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool. BMC Med Imaging 2015;15:29. doi: 10.1186/s12880-015-0068-x.

[11] Bouzidi D, Ghozzi F, Fakhfakh A. Ant colony optimization with brainSeg3D protocol for multiple sclerosis lesion detection. In: Aloulou H, Abdulrazak B, De Marassé-Enouf A, Mokhtari M, editors. Particip. Urban Health

Healthy Aging Age AI, vol. 13287. Cham, Switzerland.: Springer International Publishing; 2022. doi: 10.1007/978-3-031-09593-1_19 [p. 234–45].
[12] Syed AZ, Sin C, Rios R, Mupparapu M. Incidental occurrence of an unusually large mastoid foramen on cone-beam computed tomography and review of the literature. Imaging Sci Dent 2016;46:39. doi: 10.5624/isd.2016.46.1.39.
[13] Mah JK, Huang JC, Choo H. Practical applications of cone-beam computed tomography in orthodontics. J Am Dent Assoc 2010;141:7S–13S. doi: 10.14219/jada.archive.2010.0361.
[14] Correa LR, Spin-Neto R, Stavropoulos A, Schropp L, Da Silveira HED, Wenzel A. Planning of dental implant size with digital panoramic radiographs, CBCT -generated panoramic images, and CBCT cross-sectional images. Clin Oral Implants Res 2014;25:690–5. doi: 10.1111/clr.12126.
[15] Nguyen KCT, Duong DQ, Almeida FT, Major PW, Kaipatur NR, Pham TT, et al. Alveolar bone segmentation in intraoral ultrasonographs with machine learning. J Dent Res 2020;99:1054–61. doi: 10.1177/0022034520920593.
[16] Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. doi: 10.1136/bmj.n71.
[17] Campbell JM, Klugar M, Ding S, Carmody DP, Hakonsen SJ, Jadotte YT, et al. Diagnostic test accuracy: methods for systematic review and meta-analysis. Int J Evid Based Healthc 2015;13:154–62. doi: 10.1097/XEB.0000000000000061.
[18] Ma LL, Wang YY, Yang ZH, Huang D, Weng H, Zeng XT. Methodological quality (risk of bias) assessment tools for primary and secondary medical studies: what are they and which is better? Mil Med Res 2020;7:7. doi: 10.1186/s40779-020-00238-8.
[19] Brozek JL, Canelo-Aybar C, Akl EA, Bowen JM, Bucher J, Chiu WA, et al. GRADE Guidelines 30: the GRADE approach to assessing the certainty of modeled evidence – an overview in the context of health decision-making. J Clin Epidemiol 2021;129:138–50. doi: 10.1016/j.jclinepi.2020.09.018.
[20] Prasad M. Introduction to the GRADE tool for rating certainty in evidence and recommendations. Clin Epidemiol Glob Health 2024;25:101484. doi: 10.1016/j.cegh.2023.101484.
[21] Wu Q, Feng B, Li W, Zhang W, Wang J, Wang X, et al. Automatic segmentation and visualization of cortical and marrow bone in mandibular condyle on CBCT: a preliminary exploration of clinical application. Oral Radiol 2025;41:88–101. doi: 10.1007/s11282-024-00780-4.
[22] Hegyi A, Somodi K, Pinter C, Molnar B, Windisch P, Garcia-Mato D, et al. Automatic segmentation of dental cone-beam computed tomography scans using a deep learning framework. Orv Hetil 2024;165:1242–51. doi: 10.1556/650.2024.33098.
[23] Chen H, Qu Z, Tian Y, Jiang N, Qin Y, Gao J, et al. A cross-temporal multimodal fusion system based on deep learning for orthodontic monitoring. Comput Biol Med 2024;180:109025. doi: 10.1016/j.compbiomed.2024.109025.
[24] Chen RQ, Lee Y, Yan H, Mupparapu M, Lure F, Li J, et al. Leveraging pretrained transformers for efficient segmentation and lesion detection in cone-beam computed tomography scans. J Endod 2024;50(10):150514.e1. doi: 10.1016/j.joen.2024.07.012.
[25] Liu Y, Xie R, Wang L, Liu H, Liu C, Zhao Y, et al. Fully automatic AI segmentation of oral surgery-related tissues based on cone beam computed tomography images. Int J Oral Sci 2024;16:34. doi: 10.1038/s41368-024-00294-z.
[26] Nogueira-Reis F, Morgan N, Suryani IR, Tabchoury CPM, Jacobs R. Full virtual patient generated by artificial intelligence-driven integrated segmentation of craniomaxillofacial structures from CBCT images. J Dent 2024;141:104829. doi: 10.1016/j.jdent.2023.104829.
[27] van Nistelrooij N, Schitter S, van Lierop P, El Ghoul K, Koenig D, Hanisch M, et al. Detecting mandible fractures in CBCT scans using a 3-stage neural network. J Dent Res 2024;103(13):1384–91. doi: 10.1177/00220345241256618.
[28] Dot G, Chaurasia A, Dubois G, Savoldelli C, Haghighat S, Azimian S, et al. Dental segmentator: robust open source deep learning-based CT and CBCT image segmentation. J Dent 2024;147:105130. doi: 10.1016/j.jdent.2024.105130.
[29] Tian Y, Hao J, Wang M, Zhang Z, Wang G, Kou D, et al. Automatic jawbone structure segmentation on dental CBCT images via deep learning. Clin Oral Investig 2024;28:663. doi: 10.1007/s00784-024-06061-y.
[30] Liu J, Hao J, Lin H, Pan W, Yang J, Feng Y, et al. Deep learning-enabled 3D multimodal fusion of cone-beam CT and intraoral mesh scans for tooth-bone reconstruction. Patterns (NY) 2023;4(9):100825. doi: 10.1016/j.patter.2023.100825.
[31] Moufti MA, Trabulsi N, Ghousheh M, Fattal T, Ashira A, Danishvar S. Developing an artificial intelligence solution to autosegment the edentulous mandibular bone for implant planning. Eur J Dent 2023;17:1330–7. doi: 10.1055/s-0043-1764425.
[32] Tao T, Zou K, Jiang R, He K, He X, Zhang M, et al. Artificial intelligence-assisted determination of available sites for palatal orthodontic mini implants based on palatal thickness through CBCT. Orthod Craniofac Res 2023;26:491–9. doi: 10.1111/ocr.12634.
[33] Vinayahalingam S, Berends B, Baan F, Moin DA, van Luijn R, Bergé S, et al. Deep learning for automated segmentation of the temporomandibular joint. J Dent 2023;132:104475. doi: 10.1016/j.jdent.2023.104475.
[34] Xu P, Gholamalizadeh T, Moshfeghifar F, Darkner S, Erleben K. Deep-learning-based segmentation of individual tooth and bone with periodontal ligament interface details for simulation purposes. IEEE Access 2023;11:102460–470. doi: 10.1109/ACCESS.2023.3317512.
[35] Nogueira-Reis F, Morgan N, Nomidis S, Van Gerven A, Oliveira-Santos N, Jacobs R, et al. Three-dimensional maxillary virtual patient creation by convolutional neural network-based segmentation on cone-beam computed tomography images. Clin Oral Investig 2023;27:1133–41. doi: 10.1007/s00784-022-04708-2.
[36] Ilesan RR, Beyer M, Kunz C, Thieringer FM. Comparison of artificial intelligence-based applications for mandible segmentation: from established platforms to in-house-developed software. Bioengineering (Basel) 2023;10(5):604. doi: 10.3390/bioengineering10050604.
[37] Fan Y, Zhang Y, Chen G, He W, Song G, Matthews H, et al. Automated assessment of mandibular shape asymmetry in 3-dimensions. Am J Orthod Dentofac Orthop 2022;161:698–707. doi: 10.1016/j.ajodo.2021.07.014.
[38] Gillot M, Baquero B, Le C, Deleat-Besson R, Bianchi J, Ruellas A, et al. Automatic multi-anatomical skull structure segmentation of cone-beam computed tomography scans using 3D UNETR. PLoS One 2022;17(10):e0275033. doi: 10.1371/journal.pone.0275033.
[39] Jha N, Kim T, Ham S, Baek SH, Sung SJ, Kim YJ, et al. Fully automated condyle segmentation using 3D convolutional neural networks. Sci Rep 2022;12:20590. doi: 10.1038/s41598-022-24164-y.
[40] Preda F, Morgan N, Van Gerven A, Nogueira-Reis F, Smolders A, Wang X, et al. Deep convolutional neural network-based automated segmentation of the maxillofacial complex from cone-beam computed tomography: a validation study. J Dent 2022;124:104238. doi: 10.1016/j.jdent.2022.104238.
[41] Kim YH, Shin JY, Lee A, Park S, Han SS, Hwang HJ. Automated cortical thickness measurement of the mandibular condyle head on CBCT images using a deep learning method. Sci Rep 2021;11(1):14852. doi: 10.1038/s41598-021-94362-7.
[42] Le C, Deleat-Besson R, Prieto J, Brosset S, Dumont M, Zhang W, et al. Automatic Segmentation of Mandibular Ramus and Condyles. Annu Int Conf IEEE Eng Med Biol Soc 2021;2021:2952–5. doi: 10.1109/EMBC46164.2021.9630727.

[43] Lo Giudice A, Ronsivalle V, Spampinato C, Leonardi R. Fully automatic segmentation of the mandible based on convolutional neural networks (CNNs). Orthod Craniofac Res 2021;24(Suppl. 2):100–7. doi: 10.1111/ocr.12536.

[44] Verhelst P-J, Smolders A, Beznik T, Meewis J, Vandemeulebroucke A, Shaheen E, et al. Layered deep learning for automatic mandibular segmentation in cone-beam computed tomography. J Dent 2021;114:103786. doi: 10.1016/j.jdent.2021.103786.

[45] Wang X, Pastewait M, Wu T-H, Lian C, Tejera B, Lee Y-T, et al. 3D morphometric quantification of maxillae and defects for patients with unilateral cleft palate via deep learning-based CBCT image auto-segmentation. Orthod Craniofac Res 2021;24(Suppl. 2):108–16. doi: 10.1111/ocr.12482.

[46] Chen S, Wang L, Li G, Wu T-H, Diachina S, Tejera B, et al. Machine learning in orthodontics: Introducing a 3D auto-segmentation and auto-landmark finder of CBCT images to assess maxillary constriction in unilateral impacted canine patients. Angle Orthod 2020;90:77–84. doi: 10.2319/012919-59.1.

[47] Zhang J, Liu M, Wang L, Chen S, Yuan P, Li J, et al. Context-guided fully convolutional networks for joint craniomaxillofacial bone segmentation and landmark digitization. Med Image Anal 2020;60:101621. doi: 10.1016/j.media.2019.101621.

[48] Torosdagli N, Liberton DK, Verma P, Sincan M, Lee JS, Bagci U. Deep geodesic learning for segmentation and anatomical landmarking. IEEE Trans Med Imaging 2019;38:919–31. doi: 10.1109/TMI.2018.2875814.

[49] Chen Y, Du H, Yun Z, Yang S, Dai Z, Zhong L, et al. Automatic segmentation of individual tooth in dental CBCT images from tooth surface map by a multi-task FCN. IEEE Access 2020;8:97296–309. doi: 10.1109/ACCESS.2020.2991799.