

Semi or fully automatic tooth segmentation in CBCT images: a review

Qianhan Zheng¹, Yu Gao¹, Mengqi Zhou¹, Huimin Li¹, Jiaqi Lin¹, Weifang Zhang^{1,2} and Xuepeng Chen^{1,3}

¹ Stomatology Hospital, Zhejiang University School of Medicine, Hangzhou, China

² Social Medicine & Health Affairs Administration, Zhejiang University, Hangzhou, China

³ Clinical Research Center for Oral Diseases of Zhejiang Province, Cancer Center of Zhejiang University, Hangzhou, China

ABSTRACT

Cone beam computed tomography (CBCT) is widely employed in modern dentistry, and tooth segmentation constitutes an integral part of the digital workflow based on these imaging data. Previous methodologies rely heavily on manual segmentation and are time-consuming and labor-intensive in clinical practice. Recently, with advancements in computer vision technology, scholars have conducted in-depth research, proposing various fast and accurate tooth segmentation methods. In this review, we review 55 articles in this field and discuss the effectiveness, advantages, and disadvantages of each approach. In addition to simple classification and discussion, this review aims to reveal how tooth segmentation methods can be improved by the application and refinement of existing image segmentation algorithms to solve problems such as irregular morphology and fuzzy boundaries of teeth. It is assumed that with the optimization of these methods, manual operation will be reduced, and greater accuracy and robustness in tooth segmentation will be achieved. Finally, we highlight the challenges that still exist in this field and provide prospects for future directions.

Subjects Computational Biology, Artificial Intelligence, Computer Vision, Neural Networks

Keywords CBCT, Tooth segmentation, Level set, Deep learning, UNet

Submitted 18 September 2023

Accepted 27 March 2024

Published 19 April 2024

Corresponding authors

Weifang Zhang, chzwf@zju.edu.cn

Xuepeng Chen, cxp1979@zju.edu.cn

Academic editor

Paulo Coelho

Additional Information and
Declarations can be found on
page 26

DOI [10.7717/peerj-cs.1994](https://doi.org/10.7717/peerj-cs.1994)

© Copyright

2024 Zheng et al.

Distributed under

Creative Commons CC-BY 4.0

OPEN ACCESS

INTRODUCTION

Radiographic images are indispensable for diagnosing and designing treatments in the medical field. X-ray, the most commonly used radiological examination method, has difficulty detecting lesions due to the overlapping tissues caused by the absorption of certain organs in the human body (*Oprea et al., 2008*). Computed tomography (CT) was developed to overcome these limitations; this method involves conducting cross-sectional scans around specific body parts one by one to obtain tomographic images (*Goldman, 2007*), compensating for the drawbacks of X-rays. Starting in the second half of the 1990s (*Loubele et al., 2006*), specialized dental cone beam computed tomography (CBCT) was widely used in various fields, including orthodontics (*Kapila & Nervina, 2015*), periodontology (*Woelber et al., 2018*), implant dentistry (*Jacobs et al., 2018*), temporomandibular joint diseases (*Larheim et al., 2015*), and endodontics (*Patel et al., 2019*). As documented by *Pauwels et al. (2000)*, it provides high-resolution images with less radiation exposure than traditional CT.

The progress and widespread utilization of CBCT have profoundly advanced the digitization of dental imaging, constituting a crucial component of the contemporary workflow of digital dentistry (Eaton, 2022). Tooth segmentation *via* CBCT primarily involves discerning distinctive density features (Pal & Pal, 1993) in imaging to distinguish teeth from the periodontal ligament and alveolar bones. This process ensures the precise extraction of each tooth from the CBCT image, allowing for a meticulous analysis of the morphology, position, and relationships of the tooth with surrounding structures. It enables the simulation and evaluation of diverse treatment modalities in a virtual environment, thereby significantly influencing how dental professionals approach diagnostics, treatment planning, and patient care. In orthodontic analysis, three-dimensional (3D) tooth segmentation allows for a more detailed evaluation of the volume and position of tooth eruption and root resorption (Tanna, AlMuzaini & Mupparapu, 2021). Tooth segmentation also serves as a valuable tool in guiding orthognathic surgeries by providing precise anatomical information (Elnagar, Aronovich & Kusnoto, 2019; de Waard *et al.*, 2022). Additionally, in dental implantology (Orentlicher, Horowitz & Abboud, 2012) and transplantation (Wu *et al.*, 2019), tooth segmentation contributes to accurate positioning and navigation during procedures, ultimately improving the overall success rates of these interventions.

The conventional approach to 3D tooth segmentation is manual and involves generating handcrafted outlines of teeth slice by slice. However, this method requires doctor experience and is highly subjective and time-consuming. Due to its drawbacks, the manual method is often used as a parameter for other segmentation methods (Sercan, Bari, s & Aysun, 2021). With the evolution of computer vision technology, various semi or fully automatic image segmentation methods have been introduced. However, when applied to CBCT images, these algorithms need to be further improved to accommodate inherent image complexity. The challenges faced by these segmentation algorithms primarily include diverse tissue types on CBCT images and the subtle boundaries and structural intricacies of teeth (Singh *et al.*, 2020; Anwar *et al.*, 2018). Furthermore, variations in image quality and artifacts, such as noise and distortions, introduce uncertainties in segmentation tasks (Singh *et al.*, 2020; Anwar *et al.*, 2018). Moreover, real-time or near-real-time processing in clinical settings requires segmentation algorithms to be computationally efficient without compromising accuracy.

Scholars and researchers have responded to these challenges by proposing pragmatic segmentation algorithms aimed at elevating the quality of automatic 3D medical image segmentation. In tooth segmentation, semiautomatic methods are used for the initial epoch. These methods are primarily based on predefined mathematical models to assist operators in completing semisegmentation tasks, significantly reducing the time required for tooth segmentation. In recent years, propelled by the introduction of deep learning (DL), fully automatic methods have experienced notable advancements. These models, which are proficient at learning intricate image features, diminish the need for manual intervention and achieve efficient tooth segmentation.

Although numerous models have been proposed, there is a paucity of review articles on tooth segmentation algorithms. To address this gap, a recent study by [Polizzi et al. \(2023\)](#) extensively examined 23 research studies, scrutinizing the technological advancements in automatic tooth segmentation methods employing CBCT. Notably, the study predominantly concentrated on the application of DL methods for achieving fully automatic segmentation, providing limited coverage of semiautomatic approaches and overlooking insights from pivotal conference reports in the field. This article aims to fill this void by conducting a comprehensive review of semiautomatic or fully automatic tooth segmentation methods. Our objective is to carefully categorize and discuss the included articles, elucidating how they automate the process of tooth segmentation by applying and refining existing image segmentation algorithms. Moreover, we analyzed the results of these methods, demonstrating their continuous progress in reducing manual intervention and enhancing the accuracy and robustness of the models. This provides in-depth insights into the intricate advancements in this field.

SURVEY METHODOLOGY

The literature search for this study was conducted in December 2023 utilizing electronic databases, namely, PubMed, Web of Science (WOS), and IEEE Xplore (IEEE). The search applied the following query parameters: (((tooth) OR (teeth)) AND ((segmentation) OR (segment))) AND (((dental) AND (CT)) OR (CBCT)). Except for the patent articles we found on the WOS, all the studies were included.

The primary stages of the literature search are illustrated in [Fig. 1](#). After applying the queries to the aforementioned databases, we obtained a total of 1,140 articles, 419 from PubMed, 649 from WOS and 72 from IEEE. After eliminating duplicate articles by title comparison, 667 records were retained.

Then, we manually examined the titles and abstracts to remove articles that did not address tooth segmentation in CBCT images. After this phase, a total of 61 articles were retained. Finally, we conducted a thorough examination of the full text, excluding two review articles and four articles that were not focused on automating the segmentation process. Our study included a total of 55 articles on tooth segmentation.

RESULTS

Classification and yearly distribution of included articles

The articles included in our review primarily arose from the rapidly evolving landscape of fully automatic and semiautomatic image segmentation algorithms developed in recent years. The methods employed for tooth segmentation can be broadly categorized into two groups: knowledge-based methods and deep learning-based methods.

As illustrated in [Fig. 2](#), studies conducted before 2019 predominantly focus on knowledge-based methodologies, whereas in the post-2019 era, most studies highlight deep learning-based approaches. Notably, the year 2022 had the highest number of relevant articles, underscoring the ongoing exploration on tooth segmentation.

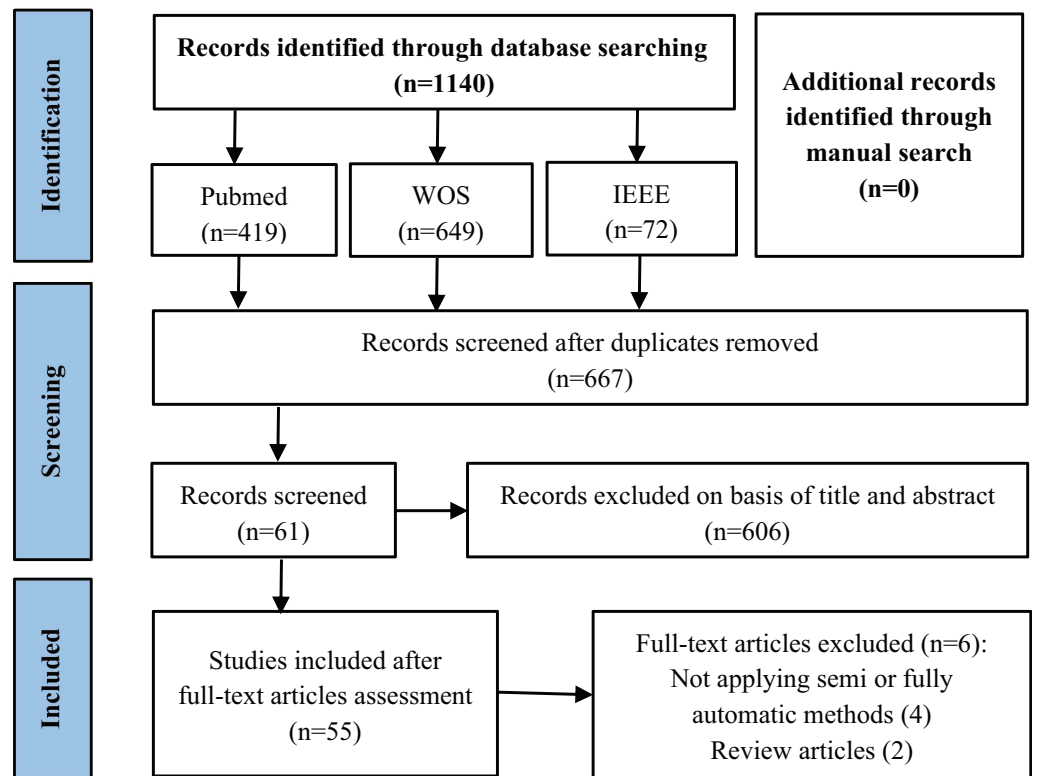


Figure 1 Flowchart of the literature review article selection process.

Full-size DOI: 10.7717/peerj-cs.1994/fig-1

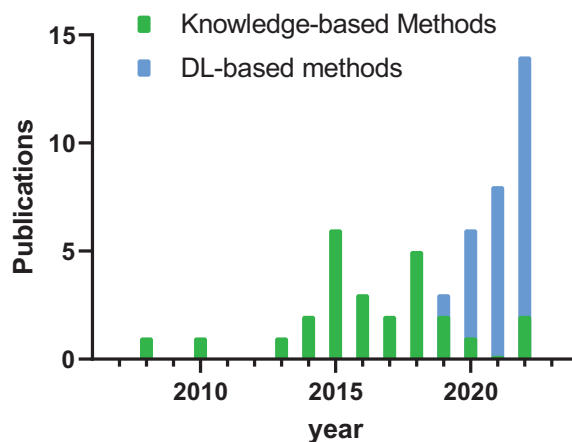


Figure 2 Histogram of publication year of literature. Full-size DOI: 10.7717/peerj-cs.1994/fig-2

Evaluation metrics

Segmentation outcomes in various studies are evaluated by diverse assessment metrics, broadly categorized into three groups: overlap-based metrics, distance-based metrics, and volume-based metrics. The calculation methods for these metrics are detailed in [Table 1](#).

Overlap-based metrics gauge the dissimilarity in the overlap between automated segmentation outcomes and manual segmentation outcomes, leveraging the confusion

Table 1 Evaluation metrics employed to perform measurement of automatic tooth segmentation.

Metric	Abbreviation	Definition
Overlap-based metrics		
Accuracy	Acc	$Acc = \frac{ Y_r \cap Y_p + (1 - Y_r) \cap (1 - Y_p) }{ U } = \frac{TP + TN}{TP + TN + FP + FN}$
Precision	Pre	$Pre = \frac{ Y_r \cap Y_p }{ Y_p } = \frac{TP}{TP + FP}$
Positive predictive value	PPV	$PPV = Pre$
Recall	Rec	$Rec = \frac{ Y_r \cap Y_p }{ Y_r } = \frac{TP}{TP + FN}$
Sensitivity	Sen	$Sen = Rec$
Specificity	Spe	$Spe = \frac{ (1 - Y_r) \cap (1 - Y_p) }{ 1 - Y_r } = \frac{TN}{TN + FP}$
Dice similarity coefficient	Dice	$Dice = \frac{2 Y_r \cap Y_p }{ Y_r + Y_p } = \frac{2TP}{2TP + FP + FN}$
Jaccard similarity coefficient	Jac	$Jac = \frac{ Y_r \cap Y_p }{ Y_r \cup Y_p } = \frac{TP}{TP + FP + FN}$
Intersection over union	IoU	$IoU = Jac$
Distance-based metrics		
Average symmetric surface distance	ASSD	$ASSD(A, B) = \frac{d(A, B) + d(B, A)}{2}$ where $d(A, B) = \frac{1}{N} \sum_{a \in A} \min_{b \in B} a - b$
Maximum symmetric surface distance	MSSD	$MSSD(A, B) = \max(h(A, B), h(B, A))$ where $h(A, B) = \max_{a \in A} \min_{b \in B} a - b$
95% Hausdorff distance	95HD	$95HD = \max(h^{95\%}(A, B), h^{95\%}(B, A))$ where $h^{95\%}(A, B) = \max_{a \in A} \min_{b \in B^{95\%}} a - b$
Volume-based metrics		
Volume overlap error	VOE	$VOE = 1 - \frac{ Y_r \cap Y_p }{ Y_r \cup Y_p }$
Relative volume difference	RVD	$RVD = \frac{ Y_r - Y_p }{ Y_r } \text{ or } \frac{ Y_p - Y_r }{ Y_r }$

Note: mm, millimeters; %, percents. Indicates the pixels in the reference standard (ground truth), and is the pixels in the automatic segmentation. $| \cdot |$ represents the number of voxels. $\| \cdot \|$ represents the L2 norm. a and b are corresponding points on the boundary of A and B.

matrix. This matrix, presented in tabular format, provides a comprehensive overview of a model's predictions compared to the actual outcomes. The model is divided into four quadrants, defining four distinct outcomes: true positive (TP), false positive (FP), false negative (FN), and true negative (TN). TP denotes the accurate segmentation of foreground pixels, FP indicates the count of pixels erroneously classified as foreground, FN represents the total count of pixels incorrectly classified as background, and TN signifies the correct classification of background pixels (Qiu et al., 2021). These elements facilitate the computation of various secondary indicators, such as accuracy (Ac), precision (Pre),

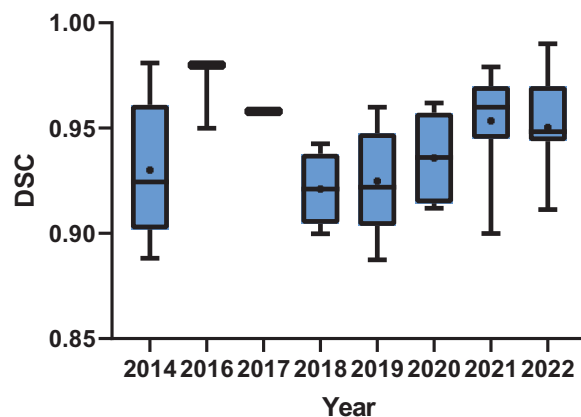


Figure 3 Box plot of the dice score of the publications by year.

Full-size DOI: 10.7717/peerj-cs.1994/fig-3

recall (Rec), specificity (Spe), Dice similarity coefficient (Dice), and Jaccard similarity coefficient (Jac), or the intersection over union (IoU).

Distance-based metrics assess the geometric disparities between the segmentation result and the ground truth. In the context of tooth segmentation, commonly used distance-based metrics include the average symmetric surface distance (ASSD), maximum symmetric surface distance (MSSD), and 95% Hausdorff distance (95 HD).

Volume-based metrics concentrate on the volume and shape attributes of the segmented region, encompassing the volume overlap error (VOE) and relative volume difference (RVD). These three metric categories collectively provide an exhaustive evaluation, quantifying the performance of segmentation algorithms from diverse perspectives. These algorithms empower researchers to gain a comprehensive understanding of their accuracy and reliability in medical image analysis.

Result analysis

In the realm of results analysis, Dice, ASSD, and MSSD emerged as the most frequently employed metrics in knowledge-based methodological research. Conversely, in deep learning-based methods, Dice coefficient and IoU/Jac are the most commonly utilized metrics. Nevertheless, due to the diverse range of evaluation metrics employed across various studies for assessing segmentation outcomes, it is challenging to quantitatively evaluate the results of all publications under a unified standard. This article conducts a brief statistical analysis of publications that adopt Dice coefficient as one of the metrics for assessing result accuracy. As depicted in Fig. 3, there was an upward trend in the Dice scores of publications over time, indicating the ongoing advancement in tooth segmentation methods.

The subsequent sections of this article provide a distinct summary and discussion of the included literature on various image segmentation methods. We believe that by elucidating the principles of different methods and presenting their application instances in the specific task of tooth segmentation, readers will gain a thorough understanding of the current status and developmental trajectories of this field.

Table 2 Summary of the knowledge based methods.

Authors	Method	Performance						Strengths	Weaknesses			
		Dice	ASSD (mm)	MSSD (mm)	RVD	PPV	Sen			Spe	Ac	IoU/Ja
Threshold based segmentation methods												
<i>Martin et al. (2015)</i>	Adaptive thresholds	/	/	/	/	/	/	/	/	/	Automatic delimitation for touching teeth	Limited capacity for pathology
<i>Indraswari et al. (2018)</i>	Histogram threshold, Region split and merge	/	/	/	/	/	0.8022	0.9831	0.9775	/	High accuracy and recall	Sensitivity improvement needed
Edge based segmentation methods												
<i>Pavaiou et al. (2015a, 2015b)</i>	Canny operators	/	/	/	/	/	/	/	/	/	Reducing computational load	Potential challenges in fully automation
Region based segmentation methods												
<i>Kang et al. (2015)</i>	Region growing	/	/	/	0.0229 ± 0.0056	/	/	/	/	/	Fast and accurate	Requires user-defined seed points
<i>Jiang et al. (2019)</i>	Region growing	0.9352 ± 0.0057	0.29 ± 0.04	2.17 ± 0.63	/	/	/	/	/	/	Effective in multiple root tooth	Sharp edges in segmented curves
<i>Naumovich, Naumovich & Goncharenko (2015)</i>	Watershed algorithm	/	/	/	/	/	/	/	/	/	User-friendly software	Limited in CBCT spatial resolution
<i>Galibourg et al. (2018)</i>	Watershed algorithm	/	/	/	/	/	/	/	/	/	Clear geometric visualization	Manual refinement needed
<i>Kakehbaraei, Seyedarabi & Zenouz (2018)</i>	Watershed algorithm	/	/	/	/	0.8681	0.9414	0.9994	0.9993	/	Superior in sensitivity, specificity, and accuracy	Susceptibility to artifacts
Contour based segmentation methods												
<i>Barone, Paoli & Razonale (2016)</i>	B-spline	/	/	/	/	/	/	/	/	/	Enhanced clearness of tooth contours	Limited to anterior and premolar teeth with mono-radicular anatomies
<i>Zhang et al. (2016)</i>	Deformable surface	/	/	/	/	/	/	/	/	/	Effective in complex structures	Manual data integration

(Continued)

Table 2 (continued)

Authors	Method	Performance										Strengths	Weaknesses
		Dice	ASSD (mm)	MSSD (mm)	RVD	PPV	Sen	Spe	Ac	IoU/Ja	95% HD (mm)		
<i>Harrison et al. (2019)</i>	Deformable surface	0.9219 ± 0.0231	0.24 ± 0.05	/	/	/	/	/	/	0.8559 ± 0.0389	1.54 ± 0.52	Efficient surface evolution, high accuracy	Limited parameter details
<i>Gao & Chae, 2010</i>	Variational level set	/	/	/	/	/	/	/	/	UCI: 0.987 2nd LM: 0.983 UM: 0.965	/	Successful crown and root separation	Sensitivity to noise
<i>Ji, Ong & Foong (2014)</i>	Variational level set	0.981 ± 0.008	/	1.13 ± 0.57	/	/	/	/	/	0.964 ± 0.011	/	Enhanced segmentation accuracy	Manual shape prior input
<i>Wang et al. (2018)</i>	Variational level set	/	/	/	0.1141	/	/	/	/	/	/	Improved RVD over existing methods	Manual selection of initial slice and center point
<i>Gan et al. (2015)</i>	Hybrid level set	Incisor: 0.89 ± 0.02 Canine: 0.92 ± 0.08	Incisor: 0.29 ± 0.03 Canine: 0.27 ± 0.02	Incisor: 1.25 ± 0.58 Canine: 1.06 ± 0.40	Incisor: 0.38 ± 0.13 Canine: 0.49 ± 0.09	/	/	/	/	/	/	Significant accuracy improvement	Open bite requirement, impacted teeth challenges
<i>Gan et al. (2017)</i>	Hybrid level set	Premolar: 0.92 ± 0.02 Molar: 0.94 ± 0.01	Premolar: 0.29 ± 0.05 Molar: 0.3 ± 0.08	Premolar: 1.28 ± 0.72 Molar: 1.52 ± 0.75	Premolar: 0.38 ± 0.10 Molar: 0.52 ± 0.18	/	/	/	/	/	/	Overcomes limitations of existing methods for angled teeth	Complexity in VOI estimation, reliance on accurate 3D axis

Table 2 (continued)

Authors	Method	Performance										Strengths			Weaknesses
		Dice	ASSD (mm)	MSSD (mm)	RVD	PPV	Sen	Spe	Ac	IoU/Ja	95% HD (mm)				
Gan et al. (2018)	Hybrid level set	Incisor:	Incisor:	Incisor:	/	/	/	/	/	/	/	/	/	Low ASSD and MSSD	Sensitive to metal artifacts, struggle with angled teeth coherence
		0.90 ± 0.03	0.28 ± 0.03	1.21 ± 0.63											
		Canine:	Canine:	Canine:											
		0.92 ± 0.01	0.27 ± 0.01	0.92 ± 0.32											
Qian et al. (2021)	Hybrid level set	Premolar:	Premolar:	Premolar:										Multimodal precision, automatic segmentation	Dependency on scanned data
		0.92 ± 0.02	0.28 ± 0.02	1.23 ± 0.54											
		Molar:	Molar:	Molar:											
		0.94 ± 0.01	0.3 ± 0.06	1.41 ± 0.66											
Jiang et al. (2022)	Hybrid level set	/	0.1789 ± 0.0552	0.8561 ± 0.1436	/	/	/	/	/	/	/	/	Excels in impacted or malposed teeth	Requires manual user initiation	
		0.9407 ± 0.0218	0.16 ± 0.08	/	/	/	/	/	/	/	/	/			
Methods based on specific theories															
Mortahab, Rezaeian & Soltanian-Zadeh (2013)	Mean shift algorithm	/	/	/	/	0.8167	0.8536	0.9733	/	/	/	/	/	Outperforms other methods	Challenges in low-contrast root slice segmentation
		/	/	/	/	0.7277	0.8324	0.9835	/	/	/	/	Robust metal artifact reduction		
		Anterior: 0.98	Anterior: 0.11	Anterior: 0.97	/	/	/	/	Anterior: 0.97	/	/	/			
Mortahab & Rezaeian (2016)	Mean shift algorithm	Premolar: 0.98	Premolar: 0.12	Premolar: 0.96	/	/	/	/	/	/	/	/	High IoU	Relies on exemplars	
		Molar: 0.95	Molar: 0.26	Molar: 0.89	/	/	/	/	/	/	/	/			
		0.958 ± 0.023	/	/	/	0.972 ± 0.023	0.999 ± 0.001	0.998 ± 0.001	0.998 ± 0.001	/	/	/			Overcomes poor image quality, achieves precise results with reasonable computation time
Eivain et al. (2017)	Graph cut	/	/	/	/	/	/	/	/	/	/	/	Artifacts reduction	Manual adjustment of image artifacts	
		Distance transform	/	/	/	/	/	/	/	/	/	/			/

Knowledge-based methods

Primarily, knowledge-based approaches leverage existing image processing methods, mathematical modeling, and domain-specific expertise to recognize grayscale features in CBCT images, thereby achieving the automatic extraction of tooth structures from the background. These methods can be divided into distinct classes: threshold-based segmentation methods, edge-based segmentation methods, region-based segmentation methods, contour-based segmentation methods and other methods based on specific theories. [Table 2](#) provides a summary of these knowledge-based approaches.

Threshold-based segmentation methods

Threshold-based methods execute image segmentation by categorizing each pixel based on its gray value compared to a predefined threshold. These encompass fixed, histogram, iterative, and adaptive threshold models.

[Marin et al. \(2015\)](#) employ an adaptive threshold model to filter images and identify potential tooth edges. [Indraswari et al. \(2018\)](#) applied a histogram threshold model for grayscale binarization. However, in [Indraswari et al. \(2018\)](#), due to the concentrated gray levels between teeth and alveolar bones, the histogram threshold model was employed in conjunction with other segmentation approaches to achieve a high accuracy of 0.9775.

Edge-based segmentation methods

Edge-based segmentation methods identify disjointed regions in an image by discerning variations in grayscale, color, and texture among different objects. Detecting target image edges is central to accomplishing segmentation tasks. Prominent differential operators for edge detection include the Sobel, Roberts, Prewitt, Canny, and Laplacian operators.

The approach advocated by [Pavaloiu et al. \(2015a, 2015b\)](#) employs Canny operators for edge detection, leveraging anatomical knowledge of teeth to alleviate computational burdens and enhance outcomes. Therefore, as indicated by [Pavaloiu et al. \(2015b\)](#), this approach has achieved an approximately sixfold increase in speed and an approximately 10% improvement in accuracy compared to their previous work.

Region-based segmentation methods

Region-based methods are applied to segment teeth in CBCT images by identifying regions of homogeneity or similarity. These algorithms include seeded region growing, region split and merge and watershed algorithms. Compared to edge-based methods, these approaches place greater emphasis on the overall contextual information within the segmented regions, enabling them to adapt to gradual variations or blurred tooth boundaries. To validate the performance of these methods in different patients, several scholars ([Kang et al., 2015](#); [Jiang et al., 2019](#); [Kakehbaraei, Seyedarabi & Zenouz, 2018](#)) started collecting clinical CBCT images (7–10 patients) and conducted quantitative statistical analyses on the segmentation results.

Region growing algorithm

The region growing algorithm starts from seed points in the tooth region, iteratively evaluating whether each pixel meets the predefined grayscale value. The process involves

incrementally incorporating neighboring pixels into the segmented region until no additional pixels meet the similarity conditions, thereby effectively demarcating the entirety of the tooth region.

Kang et al. (2015) propose a segmentation model based on seeded region growth combined with iterative methods, achieving a 2.4-fold speed improvement using multicore processors. *Jiang et al. (2019)* used an improved per-pixel region growing model combined with an edge detection model and achieved an average Dice coefficient of 0.9352 for multiple root teeth. However, the precision and reliability methods in *Kang et al. (2015)* and *Jiang et al. (2019)* hinge on the meticulous selection of seed points and the optimization of criteria.

Region split and merge algorithm

The region split and merge algorithm involves recursively partitioning images based on homogeneity criteria and merging regions with analogous attributes. Segmentation is sequentially refined by considering pixel values and facilitating the identification of individual teeth.

Indraswari et al. (2018) applied the region merging algorithm before histogram thresholding was performed to distribute the grayscale intensity inside the teeth and alleviate oversegmentation and undersegmentation.

Watershed algorithm

The watershed algorithm views CBCT images as topographic landscapes, where the gradients of grayscale values of different pixels are analogous to variations in terrain elevation. This method combines gradient information from the image with points that potentially represent tooth locations, creating a series of watershed lines. These lines serve to extract regions representing teeth from the background.

To compensate for the shortcomings of existing software products for low-contrast CBCT images, *Naumovich, Naumovich & Goncharenko (2015)* developed software and algorithms based on watershed transformation. The method proposed by *Kakehbaraei, Seyedarabi & Zenouz (2018)* utilizes the marker-controlled watershed (MCW) algorithm and achieves outstanding results in terms of sensitivity (0.9414), specificity (0.9994), and accuracy (0.9993). *Galibourg et al. (2018)* employed a watershed-based method to conduct automatic tooth segmentation *via* CBCT and compared its performance with that of a validated semiautomatic segmentation method. The results indicate a difference in volume between them, and the difference increases with increasing voxel size on CBCT.

Overall, the aforementioned methods have achieved precise delineation of tooth edges. However, since they all depend on the low-level visual properties of the image (such as grayscale and texture), continuous fine-tuning of parameters and manual intervention are needed. Consequently, more flexible framework models are needed to seamlessly integrate the intrinsic low-level visual attributes of CBCT images with higher-level attributes, such as prior knowledge of dental morphology, for more robust and automated segmentation.

Contour-based segmentation methods

Contour-based methods for tooth segmentation refer to image processing techniques that primarily rely on identifying and delineating the boundaries or contours of teeth in dental images. These methods utilize mathematical curves, such as B-spline curves or active contours, to capture the intricacies and variable shapes of teeth.

B-spline-based model

B-spline serves as a mathematical representation for creating smooth curves. The core principles of B-spline-based methods include defining a B-spline curve that closely aligns with the tooth contours, strategically positioning control points to shape the curve, and utilizing optimization to improve segmentation accuracy. The approach in [Barone, Paoli & Razonale \(2016\)](#) initially extracts 2D tooth contours in multiple planes and then automatically reconstructs the overall 3D tooth shape through B-spline algorithms. Since these planes are automatically defined based on specific anatomical morphologies, their method is similar to the approaches mentioned earlier ([Pavaloiu et al., 2015a, 2015b](#)) and are discussed later ([Ji, Ong & Foong, 2014](#); [Wang et al., 2018](#)), as they are influenced by anatomical considerations.

Deformable surface model

The deformable surface model sets a method of tooth segmentation with shape transformations and adjustments. [Harrison et al. \(2019\)](#) and [Zhang et al. \(2016\)](#) proposed segmentation methods that apply deformable surface models. [Zhang et al. \(2016\)](#), after testing on a collected set of 510 clinical CBCT images, verified that the proposed method has an average processing speed of approximately 3–6 s and a higher volume variation ratio. This indicates that the proposed method can converge more quickly than the other methods and requires fewer iterations to achieve better performance. [Harrison et al. \(2019\)](#) reported an average Dice score of 0.921 for teeth with irregular shapes. However, due to some constraints, the evolution of the surface toward distant parts of the tooth may be somewhat restricted, leading to occasional segmentation errors at the root tip (apex) or at incisal edges and cusp locations ([Harrison et al., 2019](#)).

Active contour model

The active contour model (ACM) is also a type of curve evolution algorithm. The model first defines an initial contour of teeth in CBCT images and then adjusts the contour's shape by minimizing the energy function. This process aims to better adapt the contour to the boundaries of the teeth. As summarized in [Xu, Yezzi & Prince \(2000\)](#), these methodologies are typically categorized into parametric and geometric ACMs. The parametric ACM, rooted in the Lagrange framework, constrains its energy function solely to selecting curve parameters, thereby imposing limitations on its broader applicability. In contrast, the geometric ACM relies on geometric measurements rather than curve expression parameters and is employed to address issues found in the parametric ACM.

The integration of the level set method with curve evolution theory has significantly facilitated the development of the geometric ACM and thereby expanded the applicability

of the ACM ([Chang, Li & Xu, 2013](#)). Consequently, most ACM-based methods are also referred to as level set-based methods.

Single level set model

The single-level set model (SLSM) is a fundamental approach based on the framework of level set methods. [Gao & Chae \(2008\)](#) applied the SLSM for tooth root segmentation, addressing challenges posed by complex image conditions and the issue of root branching. However, when dealing with tooth crown segmentation, especially for adjacent tooth crowns ([Jiang et al., 2022](#)), the performance of SLSM may be limited by certain factors.

Variational level set model

Compared to the SLSM, the variational level set model achieves greater flexibility by minimizing the comprehensive energy function. To avoid missing boundaries between the crowns of two adjacent teeth, in [Gao & Chae \(2008, 2010\)](#), a coupled variational level set method was proposed to address the conflation between adjacent tooth regions. This method separates touching teeth by generating a virtual common boundary.

In contrast, prior shape information is introduced to guide the image segmentation algorithm with greater automation. This allows for automatic layer-by-layer segmentation after specifying a rough initial contour in a single slice, contributing to increased algorithmic efficiency ([Gao & Chae, 2010](#)) and addressing issues of leakage and shrinkage. [Ji, Ong & Foong \(2014\)](#) enhanced the variational level set framework and achieved an average Dice score of 0.981 by utilizing the thickness of the tooth dentine wall as a robust shape prior. Similarly, the variational level set model employed by [Wang et al. \(2018\)](#) leverages the initial intensity and shape information from the preceding slice to enhance and shape-constrain the subsequent slice.

Hybrid level set model

Additionally, [Gan et al. \(2015\)](#) conducted extensive research and proposed a hybrid level set model. They employed a global convex level set model to extract the connected region of teeth and alveolar bones from CBCT images ([Gan et al., 2015, 2018](#)). Subsequently, individual teeth and alveolar bone are separated from the connected region. Moreover, to achieve a more precise segmentation of angled teeth, [Gan et al. \(2017\)](#) presented a protocol that initially extracted the region of interest (ROI) and the corresponding axis of the target tooth from CBCT images. Subsequently, local images of the target tooth are rotated to align its axis perpendicular to the transverse section. Ultimately, the hybrid level set model is employed to iteratively extract the tooth contours from the rotated images.

Furthermore, [Jiang et al. \(2022\)](#) alternately integrated the edge-based hybrid level set method with the geodesic active contour (GAC) method, controlling the evolution of these two models *via* a switch. Compared to several of the previously mentioned methods ([Gan et al., 2015, 2018](#); [Harrison et al., 2019](#); [Jiang et al., 2019](#)), this approach yields the smallest ASSD values (0.16 ± 0.08 mm).

In recent years, with advancements in dental image acquisition techniques, segmentation outcomes from level set methods can be combined with image data from other modalities to increase precision. [Qian et al. \(2021\)](#) used tooth crown meshes

obtained from laser scans to guide the tooth segmentation process automatically for corresponding CBCT data. They also replace the tooth crowns of the reconstructed meshes from CBCT data with those obtained from laser scans. The results obtained using multimodal data demonstrated a more accurate representation of teeth than models segmented from CBCT images.

Other methods based on specific theories

In addition to the aforementioned methodologies, certain studies explore the application of specific machine learning or image segmentation algorithms to automate the tooth segmentation process.

Mean shift algorithm

Derived from a clustering algorithm, the mean shift algorithm was applied to tooth segmentation by [Mortaeheb, Rezaeian & Soltanian-Zadeh \(2013\)](#). [Mortaeheb & Rezaeian \(2016\)](#) introduce a multistep method that employs a least squares support vector machine (LSM) to mitigate metal artifacts and a mean shift algorithm for automatic segmentation. This method not only demonstrated high sensitivity and accuracy at that time but also achieved initial robust metal artifact reduction.

Random walk model

The random walk model, a direct identification method utilizing random numbers to determine the search direction, is also applied in tooth segmentation. [Pei et al. \(2016\)](#) proposed a method that obtains the initial segmentation of teeth through a pure random walk approach. As an iterative refinement, they employ regularization through 3D exemplar registration and label propagation *via* random walks with soft constraints. This integrated approach, which combines semisupervised label propagation and regularization through 3D exemplar registration, allows for a more comprehensive utilization of information in the dataset.

Graph cut

The minimum cut-max flow algorithm is used to partition the image into foreground and background parts. The semiautomatic framework proposed by [Evain et al. \(2017\)](#) acquires shape priors through a statistical shape model and graph cut optimization, perfecting the outcome through the application of morphological opening and the watershed algorithm.

Distance transform model

The primary concept behind binary image distance transformation is to convert a binary image into a grayscale image by assessing the spatial distance between points (target and background points). [Kakehbaraei et al. \(2023\)](#) employed a distance transform model and succeeded in reducing artifacts through histogram adjustment and morphology filtering in axial slices.

Summary

Typically, the knowledge-based methodologies explored in this study leverage prior knowledge and heuristic principles, demonstrating a notable level of interpretability in

addressing distinctive attributes of dental structures and segmentation outcomes. Nevertheless, since the effectiveness of these methods is heavily contingent upon the accuracy and comprehensiveness of prior knowledge, they are potentially less robust when confronted with complex variations in tooth structures or specific scenarios. Additionally, due to the significant degree of manual intervention, these methods may be susceptible to subjective influences.

Therefore, many researchers have introduced artificial intelligence (AI)-based techniques to achieve fully automatic tooth segmentation. Compared to manual operation and semiautomatic segmentation, these methods intelligently learn features of teeth in CBCT images through end-to-end training on large-scale data. Therefore, they can adapt to more diverse datasets and varied tooth morphologies.

Deep learning-based methods

Machine learning (ML) is a subset of AI which addresses the question of how to build computers that improve automatically through experience (*Jordan & Mitchell, 2015*). As an advanced type of ML, deep learning (DL) employs artificial neural networks with multiple layers (deep neural networks) to model and process complex data representations. It discovers intricate structure in large data sets by using the backpropagation algorithm to indicate how a machine should change its internal parameters that are used to compute the representation in each layer from the representation in the previous layer (*LeCun, Bengio & Hinton, 2015*). Medical image segmentation primarily involves deep neural network architectures, including convolutional neural networks (CNNs), recurrent neural networks (RNNs), and encoder-decoder structures.

CNNs are among the most successful and widely utilized architectures in DL, particularly in computer vision tasks (*Minaee et al., 2022*). They employ convolutional, nonlinear, and pooling layers to extract features efficiently. Through locally connected units and shared weights within receptive fields, CNNs reduce parameter count while enabling hierarchical feature learning across multiple resolutions (*Aloysius & Geetha, 2017*).

RNNs are specifically designed for processing sequential data. However, they sometimes struggle with long sequences due to their inability to capture long-term dependencies and susceptibility to gradient vanishing or exploding problems (*Minaee et al., 2022*). LSTM, a specific type of RNN, addresses these issues by incorporating gates (input gate, output gate, and forget gate) to regulate information flow into and out of memory cells, allowing storage of values over arbitrary time intervals (*Byeon et al., 2015*).

Encoder-decoder structures are popular in sequence-to-sequence modeling for image-to-image translation, where the output can be an enhanced version of an image or a segmentation map (*Cho et al., 2014*).

These architectures constitute the fundamental elements of numerous state-of-the-art automatic image segmentation models based on DL. Notably, these models have gained extensive traction in the domain of dentistry, effectively mitigating limitations in knowledge-based tooth segmentation methodologies. Categorically, tooth segmentation

models based on DL can be broadly classified into semantic segmentation models and instance segmentation models.

Semantic segmentation models

Semantic segmentation entails partitioning an image into distinct regions based on features such as grayscale, color, spatial texture, and geometric shape (Mo et al., 2022). Table 3 provides a summary of these semantic segmentation models. Consequently, these models specifically achieve pixel-level segmentation, extracting all teeth from both the upper and lower jaws as a unified entity from the background.

CNN-based models

Semantic segmentation models for teeth are primarily based on CNNs. Occasionally, these models integrate other knowledge-based algorithms or DL networks to improve segmentation outcomes.

Ma & Yang (2019) integrated a lightweight CNN with a classical level set method known as the distance regularized GAC model, achieving state-of-the-art segmentation results with a Dice score of 0.8874 and a PPV of 0.9982. Wang et al. (2021) introduce an innovative mixed-scale dense (MS-D) CNN that incorporates scales within each layer and establishes dense connections among all feature maps. This design enhances feature extraction and transmission capabilities for tooth segmentation, resulting in an improvement of 0.945 in Dice score compared to that of Ma & Yang (2019). Employ a CNN-based efficient neural network (Enet) model to reduce network parameters while maintaining model accuracy. The model was tested on the public tooth CT image dataset in West China, demonstrating its fast speed in tooth segmentation (one second per CBCT scan).

In general, these models emphasize the simplification and optimization of CNN architectures and parameters, resulting in enhanced network efficiency and accurate tooth segmentation (Wang et al., 2021; Ma & Yang, 2019).

U-Net-based models

Proposed by Ronneberger, Fischer & Brox (2015), U-Net has emerged as the predominant DL model for medical/biomedical image semantic segmentation due to its ability to extract contextual information efficiently with shorter training times and fewer sample data (Siddique et al., 2021). The U-Net model, primarily structured as an encoder-decoder model, draws inspiration from the full convolutional network (FCN, a CNN-based model characterized by exclusively incorporating convolutional layers) proposed by Long, Shelhamer & Darrell (2015). It comprises a contraction path for context capture and a symmetric expansion path for precise positioning. The expansion path is more or less symmetrical with respect to the contraction path, resulting in a U-shaped architecture (Ronneberger, Fischer & Brox, 2015). Hsu et al. (2022) demonstrated that the performance of U-Nets for automatic tooth segmentation in CBCT images varies among different training strategies, such as 2D, 2.5D and 3D U-Nets. The studies included in this review were primarily based on models that use 2D and 3D U-Nets.

Table 3 Summary of the semantic segmentation models.

Authors	No. of cases	Performance						Strengths	Weaknesses			
		Dice	ASSD (mm)	RVD	IoU/Ja	95%HD (mm)	PPV			Sen	Spe	F1
CNN-based models												
<i>Ma & Yang (2019)</i>	100	0.8874	/	/	/	1.987	0.9982	/	/	/	State-of-the-art results in a challenge dataset	Limitations in touching teeth, highlight artifacts and small teeth
<i>Wang et al. (2021)</i>	30	0.945 ± 0.021	0.204 ± 0.061	/	/	/	/	/	/	/	Achieving high Dice score	Potential issues with thin structures and image quality
2D U-Net based models												
<i>Li et al. (2020)</i>	24	0.9526	0.1445	/	0.9142	/	/	0.9549	0.9577	/	High segmentation accuracy	Challenges with deep tooth roots and partial loss.
<i>Lee et al. (2020)</i>	102	0.912	/	/	/	/	/	0.881	0.945	/	Increased average precision, effective adjacent tooth distinction	Requiring realignment for varying poses
<i>Rao et al. (2020)</i>	110	0.9166	0.25	/	/	/	/	/	/	/	Performance enhancement	Not mentioned
<i>Yang et al. (2021)</i>	10	0.9791 ± 0.0145	/	/	0.9525 ± 0.0271	/	0.9733 ± 0.0159	/	/	0.9824 ± 0.0324	Superior tooth segmentation performance	Complex reliance on mathematical procedures
<i>Wu et al. (2022)</i>	100	/	/	/	0.8637	/	/	/	/	/	Robust, precise and effective	Limited dataset, challenges in complex dental occlusion
<i>Tao & Wang (2022)</i>	500	/	/	/	0.8373	/	0.8591	/	/	/	Better segmentation performance and efficiency	Further improvement of segmentation performance needed
3D U-Net based models												
<i>Liu et al. (2021)</i>	170	/	/	/	/	/	/	/	/	/	Swift, precise dental segmentation	Limited tooth landmark detection in the coarse stage
<i>Deng et al. (2023)</i>	61	Upper: 0.97 Lower: 0.96	Upper: 0.1 ± 0.1 Lower: 0.01 ± 0.01	/	/	/	/	/	/	/	Fast, efficient automatic tooth segmentation with high precision and recall	Limited assessment of segmentation in the presence of artifacts
<i>Dot et al. (2022)</i>	453	Upper: 0.95 ± 0.02 Lower: 0.94 ± 0.02	/	/	/	/	/	/	/	/	Performance surpassing 2DU-Net	Minor noise and shadow artifacts
V-Net based models												
<i>Chen et al. (2020)</i>	25	0.936 ± 0.012	0.363 ± 0.145	0.088 ± 0.031	0.881 ± 0.019	/	/	/	/	/	Suitable for diverse teeth, non-open bite positions	Potential blurring of tooth boundaries
<i>Dou et al. (2022)</i>	40	0.952	0.15	/	0.902	2.12	0.996	/	/	/	Emphasizing tooth geometry, detail, and multiscale features	Room for improvement in detail preservation

2D U-Net

2D U-Net utilizes a 2D image as the fundamental unit for input data, while 2Da, 2Dc, and 2D U-Net employ axial, coronal, and sagittal slices, respectively, as input data units (Hsu et al., 2022).

Rao et al. (2020) and Lee et al. (2020) substituted specific convolutional layers within the network and achieved Dice scores exceeding 0.915. Rao et al. (2020) substituted regular convolutional layers in 2D U-Net with three specialized deep bottleneck architectures (DBAs). This effectively deepens the network, optimizes feature extraction and enhances accuracy in localizing the tooth. Lee et al. (2020) replaced certain convolutional layers with dense blocks and introduced spatial dropout layers between contraction and extraction paths. The proposed UDS-Net (U-Net + dense block + spatial dropout) achieves improved segmentation performance with a reduced parameter count compared to the original U-Net.

The expanded modules were introduced by Wu et al. (2022) and Tao & Wang (2022) into the 2D U-Net. Wu et al. (2022) integrate a local feature enhancement module (LE) into the decoder network to fully leverage accurate semantic and location context information across the input image. Tao & Wang (2022) introduce an attention module into the 2D U-Net network to amplify the importance of critical information. These models (Wu et al., 2022; Tao & Wang, 2022) optimize the 2D U-Net-based model, and their IoU scores are approximately 0.84.

Moreover, the 2D U-Net can be synergistically combined with other DL-based or knowledge-based models to mitigate some of its inherent limitations. Li et al. (2020) integrate a 2D attention U-Net with a bidirectional convolution-LSTM (BDC-LSTM). This approach compensates for the 2D U-Net's sole extraction of intraslice contexts by enabling the BDC-LSTM to extract interlayer information from the tooth root sequence. Yang et al. (2021) amalgamate the 2D U-Net model with an enhanced ACM, addressing the challenges that DL methods may face in automatic topology changes. Compared to other 2D U-Net-based models (Rao et al., 2020; Lee et al., 2020; Wu et al., 2022; Tao & Wang, 2022), these two approaches can achieve higher Dice/IoU scores with a smaller dataset (10 in Yang et al. (2021) and 24 in Li et al. (2020)).

3D U-Net

The structure of the 3D U-Net closely resembles that of the 2D U-Net, with the distinction that 2D operations are substituted with 3D operations. Employing a cuboid as the unit for input data, the 3D U-Net can process the entire 3D image instead of individual slices during training (Çiçek et al., 2016). Liu et al. (2021) introduced a coarse-to-fine 3D U-Net-based framework, named SkullEngine, which was designed for high-resolution segmentation and large-scale landmark detection in the skull. Deng et al. (2023) studied 61 patients from the digital archive of the Department of Oral and Maxillofacial Surgery at Houston Methodist Hospital between January 2021 and December 2021 and independently validated the segmentation accuracy of SkullEngine for teeth (Dice score for upper jaw teeth: 0.97; Dice score for lower jaw teeth: 0.96).

Additionally, [Dot et al. \(2022\)](#) assessed the performance of another segmentation model named nnU-Net, an out-of-the-box tool proposed by [Isensee et al. \(2021\)](#). nnU-Net generates three U-Net configurations: a 2D U-Net, a full-resolution 3D U-Net, and a cascaded 3D U-Net ([Isensee et al., 2021](#)). In comparison with those of [Deng et al. \(2023\)](#), the patients included in their study by [Dot et al. \(2022\)](#) presented diverse anatomical deformities and had undergone orthognathic surgery. [Dot et al. \(2022\)](#) demonstrated that nnU-Net performs well on CBCT images of such patients (Dice score for upper jaw teeth: 0.95; Dice score for lower jaw teeth: 0.94).

V-Net-based models

A volumetric FCN model named V-Net is derived from 3D U-Net, attaining enhanced segmentation efficiency through the incorporation of residual architectures in each convolutional stage ([Milletari, Navab & Ahmadi, 2016](#)). [Chen et al. \(2020\)](#) applied a modified V-Net architecture to handle tooth regions and tooth surface prediction simultaneously. [Dou et al. \(2022\)](#) integrated an attention mechanism and a self-regulatory mechanism into the V-Net network structure, achieving higher Dice (0.952) and IoU (0.902) scores and lower ASSD (0.15 mm) values than did the methods proposed by [Chen et al. \(2020\)](#).

However, the semantic segmentation approaches mentioned above lack the ability to perform more detailed analysis and manipulation for each individual tooth. Therefore, many studies introduce instance segmentation models for individual tooth segmentation.

Instance segmentation models

Instance segmentation entails discriminating and outlining individual teeth in CBCT images as a distinct instance ([Tian et al., 2022](#)). This approach surpasses semantic segmentation by offering a nuanced comprehension of the spatial boundaries and type of each tooth. [Table 4](#) shows a summary of these instance segmentation models. These models can be broadly classified into two groups: two-stage models and one-stage models.

Two-stage models

Within the domain of two-stage instance segmentation models, research has focused primarily on two paths: bottom-up semantic segment-based models and top-down detection-based models.

Bottom-up semantic segment-based models

Bottom-up semantic segment-based models generate instance masks by initially classifying pixels through semantic segmentation and subsequently distinguishing different instances of the same kind through clustering or other metric learning methods.

[Cui, Li & Wang \(2019\)](#) introduced a two-stage model named ToothNet for automatic tooth instance segmentation in CBCT images. They initiate the process by extracting the tooth edge with the help of a network that comprises a single encoder, nine convolutional layers and three decoder branches. Subsequently, a network constructed from the 3D region proposal network (RPN) is applied to generate region proposals for further tooth

Table 4 Summary of the instance segmentation models.

Authors	No. of cases	Performance								Strengths	Weaknesses
		Dice	ASSD (mm)	RVD	IoU/Ja	95% HD (mm)	PPV	Sen	Spe		
Bottom-up two-stage models											
<i>Cui, Li & Wang (2019)</i>	20	0.9198	/	/	/	/	0.9775	/	/	Superior results with efficient training	Sensitive to extreme gray values
Anchor-based top-down two-stage models											
<i>Chung et al. (2020)</i>	175	/	0.15 ± 0.04	/	/	0.86 ± 0.44	/	0.93 ± 0.04	0.93 ± 0.07	High sensitivity and recall	Computational concerns
<i>Lahoud et al. (2021)</i>	314	/	0.10 ± 0.06	/	0.87 ± 0.03	/	/	/	/	AI matches manual accuracy	Limited to single/double-rooted teeth, potential sensitivity to artifacts
<i>Lee et al. (2022)</i>	120	/	/	/	0.704	/	/	0.932	0.919	Surpassing existing approaches, especially in challenging scenarios	Not mentioned
Point-based top-down two-stage models											
<i>Wu et al. (2020)</i>	20	0.962	0.122	/	/	/	/	/	/	Effective metal artifact handling	Limited by dataset scale
<i>Duan et al. (2021)</i>	20	ST: 0.96 ± 0.01 MT: 0.96 ± 0.00	ST: 0.10 ± 0.02 MT: 0.14 ± 0.02	ST: 0.05 ± 0.02 MT: 0.05 ± 0.01	/	/	/	/	/	Addressing complex tooth morphology	Challenges in full-mouth CBCT
<i>Cui et al. (2021)</i>	100	94.8 ± 0.4	0.18 ± 0.02	/	/	1.52 ± 0.28 9	/	/	/	Achieving state-of-the-art results	Not mentioned
<i>Cui et al. (2022)</i>	4938	0.9254	0.21	/	/	/	/	/	0.921	High accuracy, efficiency improvement	Limited tooth crown surface detail
One-stage models											
<i>Chung et al. (2020)</i>	175	/	0.15 ± 0.04	/	/	0.86 ± 0.44	/	0.93 ± 0.04	0.93 ± 0.07	Effective overcoming of metal artifacts	Not mentioned
<i>Shaheen et al. (2021)</i>	186	0.90 ± 0.03	/	/	0.82 ± 0.05	0.56 ± 0.38	/	0.98 ± 0.02	0.83 ± 0.05	High precision, recall; fast processing	Potential segmentation quality degradation
<i>Fontenele et al. (2022)</i>	175	Anterior: 0.95 ± 0.03 Premolars: 0.97 ± 0.03 Molars: 0.97 ± 0.03	/	/	Anterior: 0.91 ± 0.05 Premolars: 0.94 ± 0.06 Molars: 0.95 ± 0.04	Anterior: 0.25 ± 0.34 Premolars: 0.17 ± 0.38 Molars: 0.19 ± 0.43	Anterior: 0.99 ± 0.01 Premolars: 0.99 ± 0.01 Molars: 0.99 ± 0.01	Anterior: 1.00 Premolars: 1.00 Molars: 1.00	Anterior: 0.91 ± 0.05 Premolars: 0.94 ± 0.05 Molars: 0.94 ± 0.04	Robust, accurate, time-efficient, even with dental fillings	Observing minor impact of fillings

Table 4 (continued)

Authors	No. of cases	Performance								Strengths	Weaknesses
		Dice	ASSD (mm)	RVD	IoU/Ja	95% HD (mm)	PPV	Sen	Spe		
<i>Jang et al. (2022)</i>	97	0.9441 ± 0.0310	0.27 ± 0.12	/	/	1.34 ± 0.86	/	0.9618 ± 0.0467	0.9317 ± 0.0583	Robust to metal artifacts	Limited to single/double-rooted teeth, dependence on panoramic images
<i>Ayidh Alqahtani et al. (2023)</i>	215	0.99 ± 0.06	/	/	0.99 ± 0.02	0.12 ± 0.15	0.99 ± 0.01	0.99 ± 0.02	0.99 ± 0.01	Outperforms existing methods	Limited to specific CBCT device
<i>Li et al. (2022)</i>	350	0.9113 ± 0.0045	/	/	0.8480 ± 0.0057	1.00 ± 0.27	/	0.9213 ± 0.0029	0.9123 ± 0.0089	Superior performance, especially in challenging scenarios	Not mentioned
<i>Xie, Yang & Chen (2023)</i>	10	0.9480 ± 0.0257	/	/	0.9023 ± 0.0437	/	/	0.9484 ± 0.0287	/	Superior accuracy and stability	Occasional interference-related failures
<i>Gerhardt et al. (2022)</i>	170	/	/	/	0.97	0.15	/	/	/	Achieving high accuracy	Potential false positives

identification. This model is the first DL solution for tooth instance segmentation *via* CBCT.

However, the bottom-up approach relies primarily on unextracted low-level features during the first step of segmentation. Therefore, in *Cui, Li & Wang (2019)*, the Dice score for segmentation on the test set was approximately 0.91, leaving room for further improvement.

Top-down detection-based methods

The top-down detection methods initiate instance region detection in the image, followed by pixel-level segmentation of the candidate area. Initially, the location or boundaries of the instance are determined through target detection; these methods acquire the ROI of each individual tooth and subsequently conduct semantic segmentation in the ROIs.

In algorithms for ROI detection, a bounding box serves as a rectangular representation delineating the spatial coordinates of a detected object in the image (*Padilla, Netto & Da Silva, 2020*). The anchor-based object detection method depends on the generation of predefined anchors, which are subsequently refined by predicted bounding box offsets to localize the tooth ROIs precisely in the CBCT images.

Chung et al. (2020) implemented a faster R-CNN (*Ren et al., 2017*) framework based on the RPN for initial ROI detection. *Lahoud et al. (2021)* utilized the RPN with a feature pyramid network (FPN) (*Lin et al., 2017*) model for precise single-tooth localization and bounding box determination. *Lee et al. (2022)* employed a 3D FCN layer followed by an encoding-decoding structure based on a 3D hourglass network (*Xu & Takano, 2021*) to

extract the bounding box. These methods enhance the ROI detection accuracy, with average precision at an IoU of 50 (AP50) of approximately 90% and object inclusion ratio (OIR) values exceeding 99.9%.

Instead of bounding boxes, point-based object detection models apply key points for the precise localization of ROIs. Its distinct advantage over anchor-based detection lies in the absence of anchor boxes, which simplifies the complex hyperparameter choices in the network required to generate a large number of proposed regions (Duan et al., 2021).

Wu et al. (2020) introduced a two-level method featuring a center-sensitive heatmap mask strategy at the global stage for precise tooth center determination and a DenseASPP-U-Net at the local stage for detailed segmentation of individual teeth. The approach proposed by Duan et al. (2021) is also grounded in the differentiation of individual teeth through heatmaps and box regressions.

The method proposed by Cui et al. (2021) and Cui et al. (2022) is designed to extract the centroid and skeleton of each tooth as coarse-level morphological representations. Its robustness and generalizability were evaluated and validated on the largest dataset (a total of 4,938 CBCT scans) to date (Cui et al., 2022).

In general, the majority of two-stage instance segmentation models use top-down approaches. These methods exhibit lower dependence on computing power and rely extensively on precise target detection.

One-stage models

Inspired by single-stage target detection, which omits the independent and explicit extraction of candidate regions, current one-stage segmentation methods are based on the FCN framework and demonstrate enhanced segmentation accuracy and computational efficiency. For instance, Xie, Yang & Chen (2023) employ a fully convolutional one-stage object detector (FCOS) model proposed by Tian et al. (2019) to identify each individual tooth. The average Dice score of the model is 0.9480, with a standard deviation of 0.0257, indicating the outstanding and stable performance of the proposed model.

Essentially, instance segmentation can be conceptualized as region-level, instance location-aware semantic segmentation, where the same pixel may carry different semantics in distinct regions. Jang et al. (2022) proposed a three-step method involving the generation of 2D panoramic images from 3D CT images and the identification and segmentation of individual teeth in 2D panoramic images. Additionally, Shaheen et al. (2021) and Gerhardt et al. (2022) both employed 3D U-Net-based models, leveraging rough segmentation results for ROI identification. Furthermore, Li et al. (2022) initially applied a 3D U-Net-based model to segment four quadrants of teeth from input CBCT data as ROIs before employing a semantic graph attention mechanism (SGANet) based on graph convolutional networks (GCNs) proposed by Defferrard, Bresson & Vandergheynst (2016) for tooth segmentation in each quadrant. SGANet explicitly incorporates anatomical topology information to learn fine-grained discriminative features, reducing confusion in delineating boundaries between adjacent teeth (Li et al., 2022).

To enhance the robustness and accuracy of segmentation models across diverse CBCT datasets, various methodologies have been devised to mitigate the impact of metal artifacts,

dental filling materials, and orthodontic brackets. The method proposed by [Fontenele et al. \(2022\)](#) comprises multiple configured 3D U-Nets, which progress step by step from rough detection to refined segmentation of teeth with high-density dental filling material. Similarly, [Ayidh Alqahtani et al. \(2023\)](#) validated a segmentation pipeline configured with multiple 2D U-Nets. These methods demonstrate high classification and segmentation accuracy (over 0.99) for teeth filled with high-density dental filling material ([Fontenele et al., 2022](#)) or with brackets ([Ayidh Alqahtani et al., 2023](#)).

Summary

The outlined deep learning-based approaches overcome the limitations of traditional methods and achieve fully automatic and robust tooth segmentation. The significant advantages of DL in tooth segmentation include precise segmentation, flexibility in handling various imaging challenges, and continuous advancements in model architectures. Collectively, these aspects establish DL as a pivotal tool in the evolution of dental image segmentation methodologies.

Finally, when different DL models are used for automated tooth segmentation tasks, dentists should take into consideration patient needs, possible side effects during dental practice, and the characteristics of the radiographic image dataset.

DISCUSSION

Challenges and advances

Automating tooth segmentation encounters numerous challenges, such as the intricacy of dental and bone tissues, issues arising from low-quality images, and difficulties in procuring sufficient training datasets.

Research indicates that segmentation inaccuracies predominantly manifest as fine anatomical structures such as cervical margins, pits and fissures ([Galibourg et al., 2018](#)) or apical lesions in the context of multiple-rooted teeth ([Qian et al., 2021](#)). The presence of surrounding periodontal tissues often obscures the true boundaries of teeth ([Mortaeheb, Rezaeian & Soltanian-Zadeh, 2013](#); [Ji, Ong & Foong, 2014](#); [Li et al., 2020](#); [Xie, Yang & Chen, 2023](#)), thereby intensifying segmentation complexities. Segmenting extremely small teeth ([Ma & Yang, 2019](#)), irregularly arranged teeth ([Gan et al., 2017](#)), and impacted third molars ([Qian et al., 2021](#)) with diverse eruption positions and shapes ([Wu et al., 2020](#)) is challenging. Additionally, the proximity of adjacent teeth ([Mortaeheb, Rezaeian & Soltanian-Zadeh, 2013](#); [Kakehbaraei, Seyedarabi & Zenouz, 2018](#); [Kakehbaraei, Seyedarabi & Zenouz, 2018](#); [Ma & Yang, 2019](#)) or closely positioned teeth during occlusion ([Gan et al., 2017](#); [Kakehbaraei, Seyedarabi & Zenouz, 2018](#); [Qian et al., 2021](#); [Wu et al., 2020](#)) may induce mutual interference, leading to segmentation failure in delineating joint contours.

Several quality parameters of CBCT images, such as low resolution, noise, contrast deficiency ([Galibourg et al., 2018](#); [Mortaeheb, Rezaeian & Soltanian-Zadeh, 2013](#)), and the presence of metal artifacts ([Kang et al., 2015](#); [Gan et al., 2017](#); [Kakehbaraei, Seyedarabi & Zenouz, 2018](#)), negatively impact segmentation outcomes. In response to these constraints, segmentation pipelines undergo continual refinement. During the preprocessing phase, images typically require homogenization and enhancement before automated software-

driven segmentation to address artifacts and enhance data consistency (*Kakehbaraei, Seyedarabi & Zenouz, 2018; Galibourg et al., 2018; Naumovich, Naumovich & Goncharenko, 2015*). Several studies project 3D images into 2D (*Galibourg et al., 2018*) or regress tooth poses and realign them (*Chung et al., 2020; Lee et al., 2022*) to mitigate the spatial complexities of CBCT images, thereby contributing to improved segmentation model performance.

In DL-based tools, the efficacy of automatic segmentation models is occasionally constrained by the scale and annotation accuracy of the training dataset. *Cui et al. (2022)* emphasized the significance of large-scale, multicenter, and real-clinical data over synthesized small-sized datasets. *Ayidh Alqahtani et al. (2023)* also noted that a large labeled training dataset is essential to avoid overfitting a model, enhance its learning and optimization, and effectively capture the inherent data distributions. Recently, specific data augmentation techniques, such as flipping, rotation, random deformation, and conditional generative models, have been implemented to augment the training dataset (*Cui et al., 2022*).

Clinical application and ethical responsibilities of AI

The ongoing progress in automatic tooth segmentation technology has significantly streamlined the workflow in clinical digital dentistry, providing substantial benefits for patients and dental professionals. In comparison to traditional manual methods, automatic segmentation technology takes approximately 20–30 s to segment a CBCT scan, whereas manual segmentation requires approximately 3–5 h (*Wang et al., 2021*). Moreover, even in the presence of metal or other artifacts in CBCT images, the models of *Fontenele et al. (2022)* and *Ayidh Alqahtani et al. (2023)* can accurately perform segmentation without relying on additional image processing steps.

However, both dentists and computer programmers engaged in AI should recognize the ethical responsibilities associated with the application of AI in dentistry. This approach is particularly pertinent to issues of patient data privacy, compliance, and underlying errors in AI-driven methodologies. *Silva & Soto (2022)* emphasize that in the realm of AI-related research and clinical experiments, direct and candid communication with patients to obtain their consent is established prior to data collection. *Roy (2022)* and *Harvey & Gowda (2021)* stressed the importance of encryption and access restriction protocols to ensure the confidentiality of imaging data from clinical patients. They believe that this relies heavily on stringent cybersecurity measures, comprehensive regulatory frameworks, and regular auditing mechanisms to mitigate the risk of data breaches. Finally, acknowledging the possible errors in AI-driven approaches, *Pesapane et al. (2018)* suggested that dental practitioners must actively identify discrepancies while applying AI technology. Researchers, in turn, should conduct regular assessments and update AI systems based on clinical feedback. In summary, these considerations collectively contribute to ethical responsibility in the application and development of AI in dental practices.

Future directions

First, future research endeavors should focus on the continuous optimization of tooth segmentation algorithms to adapt to increasingly diverse datasets and mitigate the impact of factors such as image artifacts and anatomical complexities. During the algorithm optimization process, researchers may contemplate the incorporation of techniques such as adaptive learning (Kerr, 2016) and transfer learning (Weiss, Khoshgoftaar & Wang, 2016) to enhance the algorithm's performance across various data distributions. Furthermore, for distinct types of dental structures, the use of advanced network architectures such as transformers can be considered (Strudel et al., 2021; Dosovitskiy et al., 2020). Exploring methods that leverage multiscale and contextual information is essential for capturing both detailed anatomical structures and their holistic relationships (Strudel et al., 2021). This integration may entail models with a larger number of parameters and more complex structures, necessitating increased computational resources and training time (Dosovitskiy et al., 2020).

Second, future research endeavors should focus on achieving precise segmentation of each internal structure while segmenting the entire tooth. This includes detailed delineation of the tooth crown and roots and accurate segmentation of structures such as enamel, dentin, dental bone, and root canals. By integrating advanced computer vision and DL technologies, researchers can strive to develop more refined and accurate segmentation models, thereby providing clinicians with more detailed information about dental structures.

Third, for future research on tooth segmentation techniques, segmentation techniques should be integrated with other advanced image processing technologies for a more comprehensive understanding and presentation of dental anatomical structures. For instance, combining 3D imaging (Karatas & Toy, 2014) and virtual reality (Anthes et al., 2016) technologies can offer dental professionals a more innovative 3D display of dental structures, leading to more precise surgical planning and preoperative predictions.

Generally, future research on tooth segmentation should further advance algorithm optimization, detailed structure segmentation, and the integration of multiple technologies to meet increasingly complex clinical demands. This approach will enhance the practicality and accuracy of segmentation techniques in dental practice.

CONCLUSIONS

Fast and precise tooth segmentation in CBCT images constitutes a crucial component of digital dentistry. The automation of this intricate and time-consuming process poses a formidable challenge due to the intricate nature of oral tissues and variations in image quality. This research involved a comprehensive examination of 55 articles and presented an exhaustive analysis of fully automatic or semiautomatic tooth segmentation methods proposed over the past 15 years. Knowledge-based tooth segmentation approaches, encompassing threshold-based, edge-based, region-based, and contour-based methods, utilize existing mathematical models and achieve relatively accurate segmentation in specific test images. However, these methods have certain limitations when dealing with variable segmentation samples and often require varying degrees of manual intervention.

Current studies have shifted toward modern AI-driven approaches. Within the domain of DL-based tooth segmentation, approaches leveraging the U-Net model stand out due to their robust performance in medical image tasks. Instance segmentation models extend the ability to recognize and delineate individual teeth in CBCT images, offering a nuanced understanding of spatial boundaries and tooth types and contributing to more in-depth and detailed analysis. DL-based tooth segmentation significantly contributes to dental diagnostics and treatment planning. Prospective directions underscore the importance of algorithmic refinement, intricate structure segmentation, and the introduction of cutting-edge imaging technologies to effectively address ever-increasing clinical demands.

ACKNOWLEDGEMENTS

A generative AI language model (ChatGPT-3.5; OpenAI, San Francisco, Calif) was used for language polishing and translation of a portion of the content.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

This work was supported by the Zhejiang Provincial Natural Science Foundation of China (LY18H140001), the National Natural Science Foundation of China (81400511), the Key R&D Program of Zhejiang (2023C03072), and the R&D Program of the Stomatology Hospital of Zhejiang University School of Medicine (RD2022DLYB03, RD2022JCEL04). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:

Zhejiang Provincial Natural Science Foundation of China: LY18H140001.

National Natural Science Foundation of China: 81400511.

Key R&D Program of Zhejiang: 2023C03072.

Zhejiang University School of Medicine: RD2022DLYB03, RD2022JCEL04.

Competing Interests

The authors declare that they have no competing interests.

Author Contributions

- Qianhan Zheng conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.
- Yu Gao conceived and designed the experiments, performed the experiments, prepared figures and/or tables, and approved the final draft.
- Mengqi Zhou conceived and designed the experiments, performed the experiments, prepared figures and/or tables, and approved the final draft.
- Huimin Li conceived and designed the experiments, prepared figures and/or tables, and approved the final draft.

- Jiaqi Lin conceived and designed the experiments, prepared figures and/or tables, and approved the final draft.
- Weifang Zhang conceived and designed the experiments, authored or reviewed drafts of the article, and approved the final draft.
- Xuepeng Chen conceived and designed the experiments, authored or reviewed drafts of the article, and approved the final draft.

Data Availability

The following information was supplied regarding data availability:

This is a literature review.

REFERENCES

- Aloysius N, Geetha M. 2017.** A review on deep convolutional neural networks. In: *2017 International Conference on Communication and Signal Processing (ICCSP)*. 588–592 DOI [10.1109/ICCSP.2017.8286426](https://doi.org/10.1109/ICCSP.2017.8286426).
- Anthes C, García-Hernández RJ, Wiedemann M, Kranzlmüller D. 2016.** State of the art of virtual reality technology. In: *2016 IEEE Aerospace Conference*. Piscataway: IEEE, 1–19.
- Anwar SM, Majid M, Qayyum A, Awais M, Alnowami M, Khan MK. 2018.** Medical image analysis using convolutional neural networks: a review. *Journal of Medical Systems* **42**(11):226 DOI [10.1007/s10916-018-1088-1](https://doi.org/10.1007/s10916-018-1088-1).
- Ayidh Alqahtani K, Jacobs R, Smolders A, Van Gerven A, Willems H, Shujaat S, Shaheen E. 2023.** Deep convolutional neural network-based automated segmentation and classification of teeth with orthodontic brackets on cone-beam computed-tomographic images: a validation study. *European Journal of Orthodontics* **45**(2):169–174 DOI [10.1093/ejo/cjac047](https://doi.org/10.1093/ejo/cjac047).
- Barone S, Paoli A, Razionale AV. 2016.** CT segmentation of dental shapes by anatomy-driven reformation imaging and B-spline modelling. *International Journal for Numerical Methods in Biomedical Engineering* **32**(6):75 DOI [10.1002/cnm.2747](https://doi.org/10.1002/cnm.2747).
- Byeon W, Breuel TM, Raue F, Liwicki M. 2015.** Scene labeling with lstm recurrent neural networks. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Piscataway: IEEE, 3547–3555.
- Chang X, Li D, Xu Y. 2013.** A study of image segmentation based on level set method. In: *2013 International Conference on Advanced Computer Science and Electronics Information*. 360–363.
- Chen Y, Du H, Yun Z, Yang S, Dai Z, Zhong L, Yang W. 2020.** Automatic segmentation of individual tooth in dental CBCT images from tooth surface map by a multi-task FCN. *IEEE ACCESS* **8**:97296–97309 DOI [10.1109/ACCESS.2020.2991799](https://doi.org/10.1109/ACCESS.2020.2991799).
- Cho K, Van Merriënboer B, Bahdanau D, Bengio Y. 2014.** On the properties of neural machine translation: Encoder-decoder approaches. ArXiv preprint DOI [10.48550/arXiv.1409.1259](https://doi.org/10.48550/arXiv.1409.1259).
- Chung M, Lee M, Hong J, Park S, Lee J, Lee J, Yang IH, Lee J, Shin YG. 2020.** Pose-aware instance segmentation framework from cone beam CT images for tooth segmentation. *Computers in Biology and Medicine* **120**(11):103720 DOI [10.1016/j.combiomed.2020.103720](https://doi.org/10.1016/j.combiomed.2020.103720).
- Çiçek Ö, Abdulkadir A, Lienkamp SS, Brox T, Ronneberger O. 2016.** 3D U-Net: learning dense volumetric segmentation from sparse annotation. In: *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2016: 19th International Conference, October 17-21, 2016, Proceedings, Part II 19*. Athens, Greece, 424–432.
- Cui Z, Fang Y, Mei L, Zhang B, Yu B, Liu J, Jiang C, Sun Y, Ma L, Huang J, Liu Y, Zhao Y, Lian C, Ding Z, Zhu M, Shen D. 2022.** A fully automatic AI system for tooth and alveolar bone

- segmentation from cone-beam CT images. *Nature Communications* **13**(1):2096
DOI [10.1038/s41467-022-29637-2](https://doi.org/10.1038/s41467-022-29637-2).
- Cui Z, Li C, Wang W. 2019.** ToothNet: automatic tooth instance segmentation and identification from cone beam CT images. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 6368–6377.
- Cui Z, Zhang B, Lian C, Li C, Yang L, Wang W, Zhu M, Shen D. 2021.** Hierarchical morphology-guided tooth instance segmentation from CBCT images. In: *Information Processing in Medical Imaging: 27th International Conference, IPMI 2021, Virtual Event, June 28–June 30, 2021, Proceedings* 27. 150–162.
- de Waard O, Baan F, Bruggink R, Bronkhorst EM, Kuijpers-Jagtman AM, Ongkosuwito EM. 2022.** The prediction accuracy of digital orthodontic setups for the orthodontic phase before orthognathic surgery. *Journal of Clinical Medicine* **11**(20):6141 DOI [10.3390/jcm11206141](https://doi.org/10.3390/jcm11206141).
- Defferrard M, Bresson X, Vandergheynst P. 2016.** Convolutional neural networks on graphs with fast localized spectral filtering. In: *30th Conference on Neural Information Processing Systems (NIPS 2016), Barcelona, Spain*. Available at https://proceedings.neurips.cc/paper_files/paper/2016/file/04df4d434d481c5bb723be1b6df1ee65-Paper.pdf.
- Deng HH, Liu Q, Chen A, Kuang T, Yuan P, Gateno J, Kim D, Barber JC, Xiong KG, Yu P, Gu KJ, Xu X, Yan P, Shen D, Xia JJ. 2023.** Clinical feasibility of deep learning-based automatic head CBCT image segmentation and landmark detection in computer-aided surgical simulation for orthognathic surgery. *International Journal of Oral and Maxillofacial Surgery* **52**(7):793–800 DOI [10.1016/j.ijom.2022.10.010](https://doi.org/10.1016/j.ijom.2022.10.010).
- Dosovitskiy A, Beyer L, Kolesnikov A, Weissenborn D, Zhai X, Unterthiner T, Houshy N. 2020.** An image is worth 16x16 words: transformers for image recognition at scale. ArXiv preprint ArXiv. 2010.11929.
- Dot G, Schouman T, Dubois G, Rouch P, Gajny L. 2022.** Fully automatic segmentation of craniomaxillofacial CT scans for computer-assisted orthognathic surgery planning using the nnU-Net framework. *European Radiology* **32**(6):3639–3648 DOI [10.1007/s00330-021-08455-y](https://doi.org/10.1007/s00330-021-08455-y).
- Dou W, Gao S, Mao D, Dai H, Zhang C, Zhou Y. 2022.** Tooth instance segmentation based on capturing dependencies and receptive field adjustment in cone beam computed tomography. *Computer Animation and Virtual Worlds* **33**(5):e2100 DOI [10.1002/cav.2100](https://doi.org/10.1002/cav.2100).
- Duan W, Chen Y, Zhang Q, Lin X, Yang X. 2021.** Refined tooth and pulp segmentation using U-Net in CBCT image. *Dentomaxillofacial Radiology* **50**(6):20200251 DOI [10.1259/dmfr.20200251](https://doi.org/10.1259/dmfr.20200251).
- Eaton KA. 2022.** The development of digital dentistry in the UK: an overview. *Primary Dental Journal* **11**(4):94–98 DOI [10.1177/20501684221134198](https://doi.org/10.1177/20501684221134198).
- Elnagar MH, Aronovich S, Kusnoto B. 2019.** Digital workflow for combined orthodontics and orthognathic surgery. *Oral and Maxillofacial Surgery Clinics of North America* **32**(1):1–14 DOI [10.1016/j.coms.2019.08.004](https://doi.org/10.1016/j.coms.2019.08.004).
- Evain T, Ripoche X, Atif J, Bloch I. 2017.** Semi-automatic teeth segmentation in cone-beam computed tomography by graph-cut with statistical shape priors. In: *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*.
- Fontenele RC, Gerhardt MDN, Pinto JC, Van Gerven A, Willems H, Jacobs R, Freitas DQ. 2022.** Influence of dental fillings and tooth type on the performance of a novel artificial intelligence-driven tool for automatic tooth segmentation on CBCT images—A validation study. *Journal of Dentistry* **119**:104069 DOI [10.1016/j.jdent.2022.104069](https://doi.org/10.1016/j.jdent.2022.104069).

- Galibourg A, Dumoncel J, Telmon N, Calvet A, Michetti J, Maret D. 2018.** Assessment of automatic segmentation of teeth using a watershed-based method. *Dentomaxillofacial Radiology* **47(1)**:20170220 DOI [10.1259/dmfr.20170220](https://doi.org/10.1259/dmfr.20170220).
- Gan Y, Xia Z, Xiong J, Li G, Zhao Q. 2018.** Tooth and alveolar bone segmentation from dental computed tomography images. *IEEE Journal of Biomedical and Health Informatics* **22(1)**:196–204 DOI [10.1109/JBHI.2017.2709406](https://doi.org/10.1109/JBHI.2017.2709406).
- Gan Y, Xia Z, Xiong J, Zhao Q, Hu Y, Zhang J. 2015.** Toward accurate tooth segmentation from computed tomography images using a hybrid level set model. *Medical Physics* **42(1)**:14–27 DOI [10.1118/1.4901521](https://doi.org/10.1118/1.4901521).
- Gan Y, Xia Z, Xiong J, Zhou X, Li G, Zhao Q. 2017.** Angled tooth segmentation from computerized tomography images. In: *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. 3413–3416.
- Gao H, Chae O. 2008.** Touching tooth segmentation from CT image sequences using coupled level set method. In: *2008 5th International Conference on Visual Information Engineering (VIE 2008)*. 382–387.
- Gao H, Chae O. 2010.** Individual tooth segmentation from CT images using level set method with shape and intensity prior. *Pattern Recognition* **43(7)**:2406–2417 DOI [10.1016/j.patcog.2010.01.010](https://doi.org/10.1016/j.patcog.2010.01.010).
- Gerhardt MDN, Fontenele RC, Leite AF, Lahoud P, Van Gerven A, Willems H, Smolders A, Beznik T, Jacobs R. 2022.** Automated detection and labelling of teeth and small edentulous regions on cone-beam computed tomography using convolutional neural networks. *Journal of Dentistry* **122(2014)**:104139 DOI [10.1016/j.jdent.2022.104139](https://doi.org/10.1016/j.jdent.2022.104139).
- Goldman LW. 2007.** Principles of CT and CT technology. *Journal of Nuclear Medicine Technology* **35(3)**:115–128 DOI [10.2967/jnmt.107.042978](https://doi.org/10.2967/jnmt.107.042978).
- Harrison J, Chantrel S, Schmittbuhl M, de Guise JA. 2019.** Segmentation and 3D-modelling of single-rooted teeth from CBCT data: an automatic strategy based on dental pulp segmentation and surface deformation. In: *World Congress on Medical Physics and Biomedical Engineering*. 201–205.
- Harvey HB, Gowda V. 2021.** Regulatory issues and challenges to artificial intelligence adoption. *Radiologic Clinics* **59(6)**:1075–1083.
- Hsu K, Yuh DY, Lin SC, Lyu PS, Pan GX, Zhuang YC, Chang CC, Peng HH, Lee TY, Juan CH, Juan CE, Liu YJ, Juan CJ. 2022.** Improving performance of deep learning models using 3.5D U-Net via majority voting for tooth segmentation on cone beam computed tomography. *Scientific Reports* **12(1)**:19809 DOI [10.1038/s41598-022-23901-7](https://doi.org/10.1038/s41598-022-23901-7).
- Indraswari R, Kurita T, Arifin AZ, Suciati N, Astuti ER, Navastara DA. 2018.** 3D region merging for segmentation of teeth on cone-beam computed tomography images. In: *2018 Joint 10th International Conference on Soft Computing and Intelligent Systems (SCIS) and 19th International Symposium on Advanced Intelligent Systems (ISIS)*. 341–345.
- Isensee F, Jaeger PF, Kohl SAA, Petersen J, Maier-Hein KH. 2021.** nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. *Nature Methods* **18(2)**:203–211 DOI [10.1038/s41592-020-01008-z](https://doi.org/10.1038/s41592-020-01008-z).
- Jacobs R, Salmon B, Codari M, Hassan B, Bornstein M. 2018.** Cone beam computed tomography in implant dentistry. Recommendations for clinical use. *BMC Oral Health* **18(1)**:88 DOI [10.1186/s12903-018-0523-5](https://doi.org/10.1186/s12903-018-0523-5).
- Jang TJ, Kim KC, Cho HC, Seo JK. 2022.** A fully automated method for 3D Individual tooth identification and segmentation in dental CBCT. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **44(10)**:6562–6568 DOI [10.1109/TPAMI.2021.3086072](https://doi.org/10.1109/TPAMI.2021.3086072).

- Ji DX, Ong SH, Foong KWC. 2014. A level-set based approach for anterior teeth segmentation in cone beam computed tomography images. *Computers in Biology and Medicine* 50:116–128 DOI 10.1016/j.combiomed.2014.04.006.
- Jiang B, Zhang S, Shi M, Liu HL, Shi H. 2022. Alternate level set evolutions with controlled switch for tooth segmentation. *IEEE Access* 10(3):76563–76572 DOI 10.1109/ACCESS.2022.3192411.
- Jiang B, Zhang Y, Tang X, Shi H. 2019. Region growing model with edge restrictions for multiple roots tooth segmentation. In: *Proceedings of the Third International Symposium on Image Computing and Digital Medicine*. 171–174.
- Jordan MI, Mitchell TM. 2015. Machine learning: trends, perspectives, and prospects. *Science* 349(6245):255–260 DOI 10.1126/science.aaa8415.
- Kakehbaraei S, Arvanaghi R, Seyedarabi H, Esmaeili F, Zenouz AT. 2023. 3D tooth segmentation in cone-beam computed tomography images using distance transform. *Biomedical Signal Processing and Control* 79(3):104122 DOI 10.1016/j.bspc.2022.104122.
- Kakehbaraei S, Seyedarabi H, Zenouz AT. 2018. Dental segmentation in conebeam computed tomography images using watershed and morphology operators. *Journal of Medical Signals and Sensors* 8(2):119–124.
- Kang H, Choi C, Shin J, Lee J, Shin Y. 2015. Fast and accurate semiautomatic segmentation of individual teeth from dental CT images. *Computational and Mathematical Methods in Medicine* 2015(1):810796 DOI 10.1155/2015/810796.
- Kapila SD, Nervina JM. 2015. CBCT in orthodontics: assessment of treatment outcomes and indications for its use. *Dentomaxillofacial Radiology* 44(1):20140282 DOI 10.1259/dmfr.20140282.
- Karatas OH, Toy E. 2014. Three-dimensional imaging techniques: a literature review. *European Journal of Dentistry* 8(01):132–140 DOI 10.4103/1305-7456.126269.
- Kerr P. 2016. Adaptive learning. *ELT Journal* 70(1):88–93 DOI 10.1093/elt/ccv055.
- Lahoud P, EzEldeen M, Beznik T, Willems H, Leite A, Van Gerven A, Jacobs R. 2021. Artificial intelligence for fast and accurate 3D tooth segmentation on CBCT. *Journal of Endodontics* 47(5):827–835 DOI 10.1016/j.joen.2020.12.020.
- Larheim TA, Abrahamsson AK, Kristensen M, Arvidsson LZ. 2015. Temporomandibular joint diagnostics using CBCT. *Dentomaxillofacial Radiology* 44(1):20140235 DOI 10.1259/dmfr.20140235.
- LeCun Y, Bengio Y, Hinton G. 2015. Deep learning. *Nature* 521(7553):436–444 DOI 10.1038/nature14539.
- Lee J, Chung M, Lee M, Shin YG. 2022. Tooth instance segmentation from cone-beam CT images through point-based detection and Gaussian disentanglement. *Multimedia Tools and Applications* 81(13):18327–18342 DOI 10.1007/s11042-022-12524-9.
- Lee S, Woo S, Yu J, Seo J, Lee J, Lee C. 2020. Automated CNN-based tooth segmentation in cone-beam CT for dental implant planning. *IEEE Access* 8:50507–50518 DOI 10.1109/ACCESS.2020.2975826.
- Li Q, Chen K, Han L, Zhuang Y, Li J, Lin J. 2020. Automatic tooth roots segmentation of cone beam computed tomography image sequences using U-Net and RNN. *Journal of X-Ray Science and Technology* 28(5):905–922 DOI 10.3233/XST-200678.
- Li P, Liu Y, Cui Z, Yang F, Zhao Y, Lian C, Gao C. 2022. Semantic graph attention with explicit anatomical association modeling for tooth segmentation from CBCT images. *IEEE Transactions on Medical Imaging* 41(11):3116–3127 DOI 10.1109/TMI.2022.3179128.

- Lin TY, Dollár P, Girshick R, He K, Hariharan B, Belongie S. 2017.** Feature pyramid networks for object detection. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2117–2125.
- Liu Q, Deng H, Lian C, Chen X, Xiao D, Ma L, Xia JJ. 2021.** SkullEngine: a multi-stage CNN framework for collaborative CBCT image segmentation and landmark detection. In: *Machine Learning in Medical Imaging: 12th International Workshop, MLMI 2021, Held in Conjunction with MICCAI 2021, September 27, 2021, Proceedings 12*. Strasbourg, France, 606–614.
- Long J, Shelhamer E, Darrell T. 2015.** Fully convolutional networks for semantic segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **39(4)**:640–651 DOI [10.1109/TPAMI.2016.2572683](https://doi.org/10.1109/TPAMI.2016.2572683).
- Loubele M, Maes F, Schutyser F, Marchal G, Jacobs R, Suetens P. 2006.** Assessment of bone segmentation quality of cone-beam CT versus multislice spiral CT: a pilot study. *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontology* **102(2)**:225–234 DOI [10.1016/j.tripleo.2005.10.039](https://doi.org/10.1016/j.tripleo.2005.10.039).
- Ma J, Yang X. 2019.** Automatic dental root CBCT image segmentation based on CNN and level set method. In: *Medical Imaging 2019: Image Processing*. Vol. 10949, 668–674.
- Marin I, Păvăloiu IB, Goga N, Vasilăţeanu A, Drăgoi G. 2015.** Automatic contour detection from dental CBCT DICOM data. In: *2015 E-Health and Bioengineering Conference (EHB)*. 1–4.
- Milletari F, Navab N, Ahmadi SA. 2016.** V-net: fully convolutional neural networks for volumetric medical image segmentation. In: *2016 Fourth International Conference on 3D Vision (3DV)*. 565–571 DOI [10.1109/3DV.2016.79](https://doi.org/10.1109/3DV.2016.79).
- Minaee S, Boykov Y, Porikli F, Plaza A, Kehtarnavaz N, Terzopoulos D. 2022.** Image segmentation using deep learning: a survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **44(7)**:3523–3542 DOI [10.1109/TPAMI.2021.3059968](https://doi.org/10.1109/TPAMI.2021.3059968).
- Mo Y, Wu Y, Yang X, Liu F, Liao Y. 2022.** Review the state-of-the-art technologies of semantic segmentation based on deep learning. *Neurocomputing* **493**:626–646 DOI [10.1016/j.neucom.2022.01.005](https://doi.org/10.1016/j.neucom.2022.01.005).
- Mortaheb P, Rezaeian M. 2016.** Metal artifact reduction and segmentation of dental computerized tomography images using least square support vector machine and mean shift algorithm. *Journal of Medical Signals & Sensors* **6(1)**:1–11 DOI [10.4103/2228-7477.175867](https://doi.org/10.4103/2228-7477.175867).
- Mortaheb P, Rezaeian M, Soltanian-Zadeh H. 2013.** Automatic dental CT image segmentation using mean shift algorithm. In: *2013 8th Iranian Conference on Machine Vision and Image Processing (MVIP)*. 121–126.
- Naumovich SS, Naumovich SA, Goncharenko VG. 2015.** Three-dimensional reconstruction of teeth and jaws based on segmentation of CT images using watershed transformation. *Dentomaxillofacial Radiology* **44(4)**:20140313 DOI [10.1259/dmfr.20140313](https://doi.org/10.1259/dmfr.20140313).
- Oprea S, Marinescu C, Lita I, Jurianu M, Visan DA, Cioc IB. 2008.** Image processing techniques used for dental x-ray image analysis. In: *2008 31st International Spring Seminar on Electronics Technology*. 125–129.
- Orentlicher G, Horowitz A, Abboud M. 2012.** Computer-guided implant surgery: indications and guidelines for use. *Compendium of Continuing Education in Dentistry* **33(10)**:720–732.
- Padilla R, Netto SL, Da Silva EAB. 2020.** A survey on performance metrics for object-detection algorithms. In: *2020 International Conference on Systems, Signals and Image Processing (IWSSIP)*. 237–242.
- Pal NR, Pal SK. 1993.** A review on image segmentation techniques. *Pattern Recognition* **26(9)**:1277–1294 DOI [10.1016/0031-3203\(93\)90135-J](https://doi.org/10.1016/0031-3203(93)90135-J).

- Patel S, Brown J, Pimentel T, Kelly RD, Abella F, Durack C. 2019. Cone beam computed tomography in Endodontics—A review of the literature. *International Endodontic Journal* 52(8):1138–1152 DOI 10.1111/iej.13115.
- Pauwels R, Araki K, Siewerdsen JH, Thongvigitmanee SS. 2000. Technical aspects of dental CBCT: state of the art. *Dentomaxillofacial Radiology* 44(1):20140224 DOI 10.1259/dmfr.20140224.
- Pavaloiu IB, Goga N, Marin I, Vasileteanu A. 2015a. Automatic segmentation for 3D dental reconstruction. In: *2015 6th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*. 1–6.
- Pavaloiu IB, Vasileteanu A, Goga N, Marin I, Ioanitescu R, Dorobantu AA, Patrascu I. 2015b. Knowledge based segmentation for fast 3D dental reconstruction from CBCT. In: *2015 20th International Conference on Control Systems and Computer Science*. 397–401 DOI 10.1109/CSCS.2015.102.
- Pei Y, Ai X, Zha H, Xu T, Ma G. 2016. 3D exemplar-based random walks for tooth segmentation from cone-beam computed tomography images. *Medical Physics* 43(9):5040–5050 DOI 10.1118/1.4960364.
- Pesapane F, Volonté C, Codari M, Sardanelli F. 2018. Artificial intelligence as a medical device in radiology: ethical and regulatory issues in Europe and the United States. *Insights into imaging* 9:745–753 DOI 10.1007/s13244-018-0645-y.
- Polizzi A, Quinzi V, Ronsivalle V, Venezia P, Santonocito S, Lo Giudice A, Leonardi R, Isola G. 2023. Tooth automatic segmentation from CBCT images: a systematic review. *Clinical Oral Investigations* 27(7):3363–3378 DOI 10.1007/s00784-023-05048-5.
- Qian J, Lu S, Gao Y, Tao Y, Lin J, Lin H. 2021. An automatic tooth reconstruction method based on multimodal data. *Journal of Visualization* 24(1):205–221 DOI 10.1007/s12650-020-00697-0.
- Qiu B, van der Wel H, Kraeima J, Glas HH, Guo J, Borra RJH, Witjes MJH, van Ooijen PMA. 2021. Automatic segmentation of mandible from conventional methods to deep learning—a review. *Journal of Personalized Medicine* 11(7):629 DOI 10.3390/jpm11070629.
- Rao Y, Wang Y, Meng F, Pu J, Sun J, Wang Q. 2020. A symmetric fully convolutional residual network with DCRF for accurate tooth segmentation. *IEEE Access* 8:92028–92038 DOI 10.1109/ACCESS.2020.2994592.
- Ren S, He K, Girshick R, Sun J. 2017. Faster R-CNN: towards real-time object detection with region proposal networks. *IEEE Transactions on Pattern Analysis & Machine Intelligence* 39(6):1137–1149 DOI 10.1109/TPAMI.2016.2577031.
- Ronneberger O, Fischer P, Brox T. 2015. U-Net: convolutional networks for biomedical image segmentation. In: *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015: 18th International Conference, October 5–9, 2015, Proceedings, Part III* 18. Munich, Germany, 234–241.
- Roy S. 2022. Privacy prevention of health care data using AI. *Journal of Data Acquisition and Processing* 37(3):769 DOI 10.5281/zenodo.7699408.
- Sercan S, Bari, s OG, Aysun B. 2021. Is manual segmentation the real gold standard for tooth segmentation? A preliminary in vivo study using conebeam computed tomography images. *Meandros Medical and Dental Journal* 22(3):263–273 DOI 10.4274/meandros.galenos.2021.14632.
- Shaheen E, Leite A, Alqahtani KA, Smolders A, Van Gerven A, Willems H, Jacobs R. 2021. A novel deep learning system for multi-class tooth segmentation and classification on cone beam computed tomography. A validation study. *Journal of Dentistry* 115:103865 DOI 10.1016/j.jdent.2021.103865.

- Siddique N, Paheding S, Elkin CP, Devabhaktuni V. 2021.** U-Net and its variants for medical image segmentation: A review of theory and applications. *IEEE Access* **9**:82031–82057 DOI [10.1109/ACCESS.2021.3086020](https://doi.org/10.1109/ACCESS.2021.3086020).
- Silva I, Soto M. 2022.** Privacy-preserving data sharing in healthcare: an in-depth analysis of big data solutions and regulatory compliance. *International Journal of Applied Health Care Analytics* **7**(1):14–23.
- Singh SP, Wang L, Gupta S, Goli H, Padmanabhan P, Gulyás B. 2020.** 3D deep learning on medical images: a review. *Sensors (Basel)* **20**(18):5097 DOI [10.3390/s20185097](https://doi.org/10.3390/s20185097).
- Strudel R, Garcia R, Laptev I, Schmid C. 2021.** Segmenter: Transformer for semantic segmentation. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 7262–7272.
- Tanna NK, AlMuzaini AAAY, Mupparapu M. 2021.** Imaging in orthodontics. *Dental Clinics of North America* **65**(3):623–641 DOI [10.1016/j.cden.2021.02.008](https://doi.org/10.1016/j.cden.2021.02.008).
- Tao S, Wang Z. 2022.** Tooth CT image segmentation method based on the U-Net network and attention module. *Computational and Mathematical Methods in Medicine* **2022**(2):1–8 DOI [10.1155/2022/3289663](https://doi.org/10.1155/2022/3289663).
- Tian D, Han Y, Wang B, Guan T, Gu H, Wei W. 2022.** Review of object instance segmentation based on deep learning. *Journal of Electronic Imaging* **31**(4):041205.
- Tian Z, Shen C, Chen H, He T. 2019.** Fcos: fully convolutional one-stage object detection. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 9627–9636 ArXiv DOI [10.48550/arXiv.1904.01355](https://doi.org/10.48550/arXiv.1904.01355).
- Wang Y, Liu S, Wang G, Liu Y. 2018.** Accurate tooth segmentation with improved hybrid active contour model. *Physics in Medicine & Biology* **64**(1):15012 DOI [10.1088/1361-6560/aaf441](https://doi.org/10.1088/1361-6560/aaf441).
- Wang H, Minnema J, Batenburg KJ, Forouzanfar T, Hu FJ, Wu G. 2021.** Multiclass CBCT image segmentation for orthodontics with deep learning. *Journal of Dental Research* **100**(9):943–949 DOI [10.1177/00220345211005338](https://doi.org/10.1177/00220345211005338).
- Weiss K, Khoshgoftaar TM, Wang DD. 2016.** A survey of transfer learning. *Journal of Big data* **3**(1):1–40 DOI [10.1186/s40537-016-0043-6](https://doi.org/10.1186/s40537-016-0043-6).
- Woelber JP, Fleiner J, Rau J, Ratka-Krüger P, Hannig C. 2018.** Accuracy and usefulness of CBCT in periodontology: a systematic review of the literature. *The International Journal of Periodontics & Restorative Dentistry* **38**(2):289–297 DOI [10.11607/prd.2751](https://doi.org/10.11607/prd.2751).
- Wu X, Chen H, Huang Y, Guo H, Qiu T, Wang L. 2020.** Center-sensitive and boundary-aware tooth instance segmentation and classification from cone-beam CT. In: *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*. 939–942.
- Wu Y, Chen JM, Xie FP, Liu HH, Niu G, Lin LS. 2019.** Simulation of postoperative occlusion and direction in autotransplantation of teeth: application of computer-aided design and digital surgical templates. *British Journal of Oral & Maxillofacial Surgery* **57**(7):638–643 DOI [10.1016/j.bjoms.2019.05.011](https://doi.org/10.1016/j.bjoms.2019.05.011).
- Wu J, Zhang M, Yang D, Wei F, Xiao N, Shi L, Liu H, Shang P. 2022.** Clinical tooth segmentation based on local enhancement. *Frontiers in Molecular Biosciences* **9**:932348 DOI [10.3389/fmolb.2022.932348](https://doi.org/10.3389/fmolb.2022.932348).
- Xie R, Yang Y, Chen Z. 2023.** WITS: Weakly-supervised individual tooth segmentation model trained on box-level labels. *Pattern Recognition* **133**(3):108974 DOI [10.1016/j.patcog.2022.108974](https://doi.org/10.1016/j.patcog.2022.108974).
- Xu T, Takano W. 2021.** Graph stacked hourglass networks for 3d human pose estimation. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 16105–16114.

- Xu C, Yezzi A, Prince JL. 2000.** On the relationship between parametric and geometric active contours. In: *Conference Record of the Thirty-Fourth Asilomar Conference on Signals, Systems and Computers*. Vol. 1483–1489.
- Yang Y, Xie R, Jia W, Chen Z, Yang Y, Xie L, Jiang B. 2021.** Accurate and automatic tooth image segmentation model with deep convolutional neural networks and level set method. *Neurocomputing* **419(10)**:108–125 DOI [10.1016/j.neucom.2020.07.110](https://doi.org/10.1016/j.neucom.2020.07.110).
- Zhang X, Chen X, Liu Y, Han B, Zhuang T, Zuo W. 2016.** An effective approach of teeth segmentation within the 3d cone beam computed tomography image based on deformable surface model. *Mathematical Problems in Engineering* **2016**:9505217 DOI [10.1155/2016/9505217](https://doi.org/10.1155/2016/9505217).