

The Application of Artificial Intelligence for Tooth Segmentation in CBCT Images: A Systematic Review

Mihai Tarce ^{1,*} , You Zhou ¹, Alessandro Antonelli ^{2,3}  and Kathrin Becker ³ 

¹ Periodontology and Implant Dentistry Division, Faculty of Dentistry, The University of Hong Kong, Pokfulam Road, Hong Kong SAR, China; ottoyz@connect.hku.hk

² School of Dentistry, Health Sciences Department, Magna Graecia University of Catanzaro, 88100 Catanzaro, Italy; alessandro.antonelli@unicz.it

³ Department of Orthodontics and Dentofacial Orthopaedics, Charité-Universitätsmedizin Berlin, Aßmannshäuser Str. 4-6, 14197 Berlin, Germany; kathrin.becker@charite.de

* Correspondence: mihai@hku.hk

Abstract: Objective: To conduct a comprehensive and systematic review of the application of existing artificial intelligence for tooth segmentation in CBCT images. Materials and Methods: A literature search of the MEDLINE, Web of Science, and Scopus databases to find publications from inception through 21 August 2023, non-English publications excluded. The risk of bias and applicability of each article was assessed using QUADAS-2, and data on segmentation category, research model, sample size and groupings, and evaluation metrics were extracted from the articles. Results: A total of 34 articles were included. Artificial intelligence methods mainly involve deep learning-based techniques, including Convolutional Neural Networks (CNNs), Fully Convolutional Networks (FCNs), and CNN-based network structures, such as U-Net and V-Net. They utilize multi-stage strategies and combine other mechanisms and algorithms to further improve the semantic or instance segmentation performance of CBCT images, and most of the models have a Dice similarity coefficient greater than 90% and accuracy ranging from 83% to 99%. Conclusions: Artificial intelligence methods have shown excellent performance in tooth segmentation of CBCT images, but still face problems, such as the small size of training data and non-uniformity of evaluation metrics, which still need to be further improved and explored for their application and evaluation in clinical applications.

Keywords: artificial intelligence; image processing, computer-assisted; cone-beam computed tomography



Citation: Tarce, M.; Zhou, Y.; Antonelli, A.; Becker, K. The Application of Artificial Intelligence for Tooth Segmentation in CBCT Images: A Systematic Review. *Appl. Sci.* **2024**, *14*, 6298. <https://doi.org/10.3390/app14146298>

Academic Editor: Bruno Chrcanovic

Received: 21 June 2024

Revised: 12 July 2024

Accepted: 12 July 2024

Published: 19 July 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Cone Beam Computed Tomography (CBCT) has been widely used in the diagnosis of oral disease and treatment. This includes applications such as studying orthodontic plans, planning orthognathic surgery, determining the position of dental implants, and diagnosing maxillofacial diseases [1–3]. Compared to multi-slice CT, CBCT offers a higher spatial resolution and improved contrast-to-noise ratio, while significantly reducing radiation dose [4]. CBCT images offer high-quality diagnostic images of oral soft and hard tissues, allowing medical professionals to obtain more comprehensive information, such as precise 3D information about teeth and bones [5,6]. This enhanced spatial characteristic allows for more accurate localization and evaluation, ultimately improving treatment plans.

Correctly identifying and segmenting teeth from CBCT images plays an important role in supporting clinicians with diagnosis and treatment. The segmentation results can be further used for orthodontic treatment planning, 3D-guided implant surgery, or auto-transplantation of teeth in children, thereby improving the accuracy and success rate of the procedures. Nevertheless, achieving precise segmentation or extracting regions of interest (ROI) is a challenging task. In maxillofacial CBCT images, the pixel proportion of teeth is relatively low, especially in the apical regions, and challenges, such as noise, low contrast, and uneven exposure, are often encountered [7]. The roots also have similar densities as the

surrounding alveolar bone, making it difficult to precisely distinguish their grayscale values in the images and capture the details of roots [8]. Moreover, there are some other reasons that cause the segmentation regions to overlap or have indistinct boundaries, including the occlusion between upper and lower dental arches, crowded teeth, impacted teeth, and artifacts resulting from dental restorative materials, such as fillings, crown restorations, and implants [9]. Long scan time also increases the possibility of image blurring due to patient movement. Hence, precise tooth detection and segmentation from CBCT images has become a significant research concern.

Current tooth segmentation methods primarily rely on traditional manual segmentation with human–computer interaction, including threshold-based, region-based, and edge-based segmentation. The threshold-based segmentation is a traditional digital image processing algorithm that separates foreground from background by setting an appropriate grayscale level [10]. However, this segmentation method is highly influenced by the definition of foreground targets. When there are significant changes in the gray value of the target, it is challenging to accurately determine the segmentation results [11]. The region-based segmentation methods split the image into several regions based on discontinuities in intensity levels of pixels, and then merge the regions based on consistency [12]. However, the performance of this segmentation method is affected by similarity measurements and sensitivity to image noise [13]. It is also more computationally required due to the necessary continual region comparison and classification. Currently, the most widely used method for tooth segmentation is the Level Set Method (LSM) through geometric operations to detect topological changes in contours [14]. LSM offers more accurate localization and faster segmentation speed. Gan et al. [15] initially applied the LSM globally to obtain high-density alveolar bone regions and then applied this locally to identify individual teeth within the alveolar bone. Jiang et al. used two different level sets in alternate evolution, further enhancing tooth segmentation accuracy [16]. Although LSM addresses some challenges in tooth segmentation, its accuracy remains limited for uneven regions and boundaries of missing teeth, and it lacks robustness for complex tooth occlusion relationships.

Artificial intelligence (AI) is emerging as a new field in dentistry, aiming to enhance dental care by making it more seamless, efficient, time-saving, and cost-effective for practitioners [17,18]. With the progress of AI, machine learning and deep learning methods are also starting to become popular in segmenting medical images. Conventional manual segmentation approaches often require the establishment of intricate rules, while data-driven AI models demonstrate superior accuracy and generalization capabilities [19]. Fernandez and Chang successfully used artificial neural networks to analyze palatal view photographs of the maxilla, effectively distinguishing between the teeth and soft tissues [20]. Deleat-Besson et al. proposed a two-stage machine learning-based tooth segmentation system [21]. They extracted and verified meaningful root canal parameters as well as dental crown features by evaluating and testing datasets, enabling automatic segmentation of upper and lower root canals. The root was then separated from the crown and integrated into classification and labeling. Deep learning methods, such as convolutional neural networks (CNNs), have further improved model structures and fortified the capacity of computation while improving results' precision through multi-level learning optimization [22]. Miki et al. enhanced the classical CNN model, achieving precise classification of teeth for seven different types, including central incisors, lateral incisors, canines, first and second premolars, and first and second molars [23]. Ronneberger et al. introduced the U-Net image semantic segmentation network, which was subsequently successfully applied to automatic tooth segmentation in X-ray images by other researchers [24,25]. Additionally, He et al. proposed Mask R-CNN, which was built on a CNN with region proposal networks, enabling classification, localization and segmentation of each detected object, for instance, tooth segmentation [26].

Due to the introduction of many artificial intelligence models, the performance of 2D X-ray image segmentation has been significantly improved. However, application in CBCT images of teeth is limited. On one hand, three-dimensional segmentation imposes

higher requirements on computer performance as well as the design of neural network model architecture. On the other hand, the establishment of CBCT training datasets is more difficult, with no typical way of guiding it, and leading to great variations in segmentation results. Thus, the objective of this study is to perform a systematic review on the application progress of artificial intelligence in tooth segmentation from the CBCT image, so as to lay a comprehensive theoretical foundation for novel methods towards diagnosing and treating oral diseases in the future.

2. Methods

2.1. Research Questions

The primary research question was “What artificial intelligence/machine learning-based models are described in the literature for segmenting permanent human teeth in CBCT images?”. The secondary research questions were: “What is the size of the dataset these models are built from, including the training set and the testing set?” and “How is the accuracy of tooth segmentation assessed?”.

2.2. Search Strategy

A literature search was conducted of the MEDLINE, Web of Science, and Scopus databases to find publications published from the inception through 21 August 2023, without language restrictions, using the following keywords in Table 1. The retrieval process was conducted independently by Y.Z. and M.T.

Table 1. Searching strategy for database.

Database	Query	Results
MEDLINE	("artificial intelligence" OR "deep learning" OR "machine learning" OR "neural networks" OR "automatic" OR "automated") AND ("cone-beam computed tomography" OR "CBCT" OR "3D") AND ("tooth segment*" OR "teeth segment*")	154
Web of Science	TS = (((artificial intelligence) OR (deep learning) OR (machine learning) OR (neural networks) OR (automatic) OR (automated)) AND ((cone-beam computed tomography) OR (CBCT) OR (3D)) AND ((tooth segment*) OR (teeth segment*)))	276
Scopus	TITLE-ABS-KEY (((artificial AND intelligence) OR (deep AND learning) OR (machine AND learning) OR (neural AND networks) OR (automatic) OR (automated)) AND ((cone-beam AND computed AND tomography) OR (CBCT) OR (3d)) AND ((tooth AND segment*) OR (teeth AND segment*)))	299
Total:		729

2.3. Exclusion Criteria

In accordance with the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) 2020 guidelines shown in Figure 1, duplicate articles, reviews, case reports, as well as non-English publications, were also excluded [27]. Titles and abstracts of articles were reviewed, and eligible full text was retrieved for further evaluation. Articles in which the scope of the study was not dentistry, the subjects of the study were not teeth and CBCT, and the techniques employed were not related to artificial intelligence were also excluded.

2.4. Data Extraction

Two authors (Y.Z. and A.A.) independently screened and extract data. Any disagreement was resolved by discussion until consensus is reached or by consulting a third author (M.T.). An initial screening was conducted based on the title and abstract of the articles, followed by a further full-text screening of all articles that met the inclusion criteria. Data extraction was performed from the articles that finally met all the criteria, including the title, author, publication year of the article, segmentation category, research model, sample size and grouping, and evaluation metrics. (Table 2).

Table 2. Research details of all included articles.

Author	Category	Framework	Capture Method	Number of Samples	Evaluation Metrics	Model Reproducibility
Yin et al. 2023 [28]	Semantic	A context-transformed TransUNet++ (CoT-UNet++) architecture utilized a hybrid encoder to obtain contextual information between adjacent keys and global context, decode, and then fuse at multiple scales through dense concatenation to obtain more accurate location information for tooth segmentation.	CBCT	20 groups of 300 images -Number of training and testing are not available	-Dice Similarity Coefficient (DSC): 0.9206 -Mean Intersection over Union (mIoU): 0.8605 -Mean Pixel Accuracy (MPA): 95.91% -True Positive Rate (TPR): 93.85% -95% Hausdorff Distance (HD): 1.06 mm -Average System Surface Distance (ASSD): 0.48 mm	Not available
Ayidh Alqahtani et al. 2023 [29]	Instance	A multi-class deep CNN based tool for segmentation and classification of teeth with brackets. The CNN model was proposed by previous research (Shaheen et al. 2021).	CBCT	215 scans -Training: 140 -Validation: 35 -Test: 40	-Dice Similarity Coefficient (DSC): 0.99 -Intersection Over Union (IoU): 0.99 -Precision: 99% -Recall: 99% -Accuracy: 99% -95% Hausdorff Distance (HD): 0.12 mm -Segmentation time: 43.56 s	Request by contacting; Virtual Patient Creator (https://creator.relu.eu , accessed on 21 August 2023)
Xie et al. 2023 [30]	Semantic	A deep learning method (FCOS) to detect location and size of each tooth and generate prior ellipses to constrain the evolution of level set by distance, and find out joint point using curvature direction, and then segment tooth.	CBCT	10 scans (453 slices) -Training: 7 -Testing: 3	-Dice coefficient: 0.9480 -Jaccard coefficient (JS): 0.9023 -Precision (PN): 94.84% -Boundary F1 (BF) score: 0.9795	Code available (https://github.com/ruicx/Individual-Tooth-Segmentation-with-Rectangle-Labels , accessed on 21 August 2023)
Wang et al. 2023 [31]	Instance	A 3D ERFNet base bone neural network adopted three branches to learn the spatial embedding, seed map, and identification simultaneously for tooth instance segmentation and further obtained root canal segmentation.	CBCT	201 volumes -Split into three folds for cross-validation	-Symmetric best dice (SBD): 0.9584 -Average instance dice (AID): 0.9425 -Identification accuracy (FA): 97.97% -Average symmetric surface distance (ASSD): 0.12 mm	Confidential
Chen et al. 2023 [32]	Semantic	A CNN-Transformer Architecture UNet (CTA-UNet) network, which combined the advantages of CNNs and Transformers through a parallel architecture, integrated local features extracted by CNNs and global representations obtained by self-attention modules (MSAB) to enhance the segmentation performance.	CBCT	45 volumes -Training: 27 -Validation: 9 -Testing: 9	-Dice similarity coefficient (DSC): 0.8650 -Intersection over union (IoU): 0.7812 -95% Hausdorff Distance (HD95): 0.64 mm -Average Symmetric Surface Distance (ASSD): 0.21 mm	Request by contacting
Yang et al. 2022 [33]	Semantic	A U-Net model was first pre-trained using five labeled classes of images and then combined with a watershed approach to effectively segment the teeth, pulp cavity, and cortical bone.	CBCT	5 photos -Number of training and testing are not available	-Dice score (DSC): 0.9859	Not available
Xie et al. 2022 [34]	Instance	A novel segmentation approach based on multi-task CNN and watershed transform (MCW). Multi-task CNN based on U-Net segmented the tooth foreground and landmark from 2D CBCT slices, 3D marked controlled watershed transform method separated the overlapping 3D tooth objects, and the post-processing method based on prior knowledge merged the individual tooth with the detached tooth root.	CBCT	78 scans (38,082 slides) -Training: 39 (19,416 slides) -Validation: 14 (6820 slides) -Testing: 25 (11,846 slides)	-Dice Similarity Coefficient (DC): 0.88 -Precision (P): 98% -Recall (R): 93% -Average Symmetric Surface Distance (ASSD): 0.53 mm	Not available

Table 2. Cont.

Author	Category	Framework	Capture Method	Number of Samples	Evaluation Metrics	Model Reproducibility
Lee et al. 2022 [35]	Instance	A two-stage point-based detection network using the FCN layers followed by an encoding–decoding structure to extract feature maps, and 3D U-Net architecture for individual tooth segmentation. The adjacent teeth were detected by introducing a novel GD loss function within heatmap regression.	CBCT	120 scans -Training: 80 -Validation: 20 -Testing: 20	-Intersection over Union (IOU): 0.704 -Precision: 93.2% -Average Precision 50 (AP50): 90.91% -Recall: 91.9% -Object Include Ratio (OIR): 96.6%	Not available
Khan et al 2022 [36]	Semantic	A novel deep learning model consists of 38 layers having 11 blocks of 3D convolutional layers followed by batch normalization layers and Relu layers.	CBCT	70 volumes (Dataset augmentation by flipping to 140 volumes) -Training: 84 -Validation: 28 -Testing: 14	-Layers: 38 -Mean Dice score: 0.90 -Mean intersection over union (IoU): 0.60 -Validation accuracy: 95.54% -Training time: 23 h -Model size: 4.3 MB	Not available
Cui et al. a, 2022 [37]	Instance	A deep learning-based AI system with a hierarchical morphology-guided network to segment individual teeth and a filter-enhanced network to extract alveolar bony structures. Images were preprocessed by V-Net for ROI detection and two-stage tooth segmentation, which detected each tooth and represented it by the predicted skeleton. Then multi-task learning network predicted each tooth's volumetric mask by simultaneously regressing the corresponding tooth apices and boundaries.	CBCT	Internal dataset 4938 scans (4215 patients) -Training and Validation: 3457 (3172 patients) -Testing: 1481 (1043 patients) External dataset 407 scans (404 patients)	Internal: -Average Dice score: 0.941 -Average sensitivity: 93.9% -Average ASD error: 0.17 mm External: -Average Dice: 0.9254 -Sensitivity: 92.1% -ASD error: 0.21 mm	Partial CBCT data available (https://pan.baidu.com/s/1LdyUA2QZvmU6ncXK1_bDTw , password:1234, accessed on 21 August 2023); Code available (https://pan.baidu.com/s/194DfSPbgi2vTIVsRa6fbmA , password:1234, accessed on 21 August 2023)
Hsu et al. 2022 [38]	Semantic	A 3.5D U-Net was generated via majority voting for the predictions of 2D U-Nets from three orthogonal slices, 2.5D U-Net, and 3D U-Net at different combination strategies.	CBCT	24 patients -Divided into 4 groups, 6 patients per group for cross validation	DSC: 0.911 Accuracy: 99.9% Sensitivity: 88.8% Sp: 1.00 PPV: 97.0% NPV: 99.9%	Request by contacting
Gerhardt et al. 2022 [39]	Instance	A two-stage 3D U-Net architecture to assess the accuracy of automated detection of teeth and small edentulous regions, which was proposed by previous research (Shaheen et al. 2021).	CBCT	175 scans -Training: 140 -Testing: 35 -Validation: 46 from extra	For fully dentate patients: -Intersection over union (IoU): 0.96 -Accuracy: 99.7% -Recall: 99.7% -Precision: 100% -95% Hausdorff Distance (95HD): 0.33 mm For patients presenting small edentulous areas: -Intersection over union (IoU): 0.97 -Accuracy: 99% -Recall: 100% -Precision: 98.7% -95% Hausdorff Distance (95HD): 0.15 mm Time needed for the human versus machine detection -dental specialist-median time: 98 s to perform the analysis -the AI-median time: 1.5 s to do the same task	Virtual Patient Creator (https://creator.relu.eu , accessed on 21 August 2023)

Table 2. Cont.

Author	Category	Framework	Capture Method	Number of Samples	Evaluation Metrics	Model Reproducibility
Fontenele et al. 2022 [40]	Instance	A two-stage 3D U-Net architecture to assess the influence of dental fillings on performance for tooth segmentation, which was proposed by previous research (Shaheen et al. 2021).	CBCT	175 scans -Training: 140 -Validation: 35 -Test: 74	-Dice similarity coefficient (DSC): 0.96 -Intersection over union (IoU): 0.92 -Accuracy: 100% -Recall: 96% -Precision: 95% -95% Hausdorff Distance (95HD): 0.27 mm	Virtual Patient Creator (https://creator.relu.eu , accessed on 21 August 2023)
Fang et al. 2022 [41]	Instance	A novel curvature enhanced implicit function network for high-quality tooth model generation, which combines the CNN-based segmentation network (HMG-Net) with an implicit function network to generate 3D tooth models with fine-grained geometric details.	CBCT Intra-oral scanning	50 scans -Training: 20 -Validation: 10 -Testing: 20	-Intersection over Union (IoU): 0.8303 -Chamfer-L2: 3.00×10^{-4} -Normal Consistency (Normals): 96.25% -Occupancy accuracy (OccAcc): 79.7%	Not available
Dou et al. 2022 [42]	Instance	A new two-stage deep learning network (TSDNet) from tooth centroid localization to tooth instance segmentation. The first stage used a centroid prediction network (V-Net framework +density-based fast search clustering algorithm) to predict the tooth centroid to achieve accurate spatial localization of individual teeth. Then a tooth instance segmentation network (self-attention mechanism-based guidance module for tooth geometry structure information and tooth feature integration module based on multi-scale fusion of dilated convolutions) was used to obtain instance-level tooth information of individual teeth. The second stage achieves robust and accurate tooth segmentation from CBCT data.	CBCT	40 CBCT scans -training: 30 -validation: 5 -testing: 5	-Dice: 0.952 -Jaccard: 90.2% -Detection accuracy (DA): 99.6% -Average surface distance (ASD): 0.15 mm -Hausdorff distance (HD): 2.12 mm	Not available
Jang et al. 2022 [43]	Instance	A hierarchical multi-step deep learning model by reconstruction panoramic Image from 3D CBCT Images, identification and 2D segmentation of individual teeth in the panoramic images, extraction of loose and tight 3D tooth ROIs using the detected bounding boxes and segmented tooth regions, and finally 3D segmentation for individual teeth from the 3D tooth ROIs.	CBCT	97 scans: -Training: 66 -Testing: 31 For the 3D segmentation -Training: 7 -Testing: 4	-Dice similarity coefficient (DSC): 0.9479 -Precision: 95.97% -Recall: 93.71% -Hausdorff distance (HD): 1.66 mm -Average symmetric surface distance (ASSD): 0.14 mm	Not available
Cui et al. b, 2022 [44]	Semantic	Established a fully annotated CBCT dataset CTooth with tooth gold standard, which contained 22 volumes (7363 slices) with fine tooth labels. An attention-based segmentation framework based on U-Net with an attention branch at the bottleneck position was proposed.	CBCT	CTooth Database: -5803 slices (4243 contain tooth annotations) -5504 annotated images from 22 patients	-Dice similarity coefficient (DSC): 0.8804 -Intersection over union (IoU): 0.7871 -Weighted dice similarity coefficient (WDSC): 95.14% -Sensitivity (SEN): 94.71% -Predictive value (PPV): 82.3%	CTooth (https://github.com/liangjiubujiu/CTooth , accessed on 21 August 2023)

Table 2. Cont.

Author	Category	Framework	Capture Method	Number of Samples	Evaluation Metrics	Model Reproducibility
Cui et al. c, 2022 [45]	Semantic	Established a 3D dental CBCT dataset CTooth+, with 22 fully annotated volumes and 146 unlabeled volumes, and further evaluate several tooth segmentation strategies based on fully supervised learning, semi-supervised learning and active learning, with definition of the performance principles.	CBCT	CTooth+ Database: -5504 annotated CBCT images of 22 patients -25,876 unlabeled images of 146 patients 31,380 scans -Training: 80% for the fully supervised (with labelled images) and semi-supervised methods (with labelled and unlabeled images). -Evaluation: 20% image volumes	<p>1. Compared 8 fully-supervised segmentation methods: (3D SkipDenseNet, DenseVoxelNet, 3D Unet, VNet, Voxresnet, nnUnet, Dense Unet, Attention Unet) -Dice similarity coefficient (DSC): Attention UNet: 0.866 -Intersection-over-union (IoU): Attention UNet: 0.7645 -Sensitivity (SEN): Dense UNet: 90.80% -Positive predictive value (ppv): Attention UNet: 87.79% -Hausdorff distance (HD): nnUNet: 1.29 mm -Average symmetric surface distance (ASSD): Attention UNet and nnUNet: 0.27 mm -Surface overlap (SO): Dense UNet: 95.98% -Surface dice (SD): Dense UNet: 95.91%</p> <p>2. Compared 4 semi-supervised methods (trained by 9 labelled volumes and 8 unlabeled volumes) (MT, CPS, DCT, CTCT) -Dice similarity coefficient (DSC): CTCT: 0.8532 -Intersection-over-union (IoU): CTCT: 0.746 -Sensitivity (SEN): CTCT: 87.55% -Positive predictive value (ppv): CTCT: 84.22% -Hausdorff distance (HD): MT: 2.76 mm -Average symmetric surface distance (ASSD): CTCT: 0.43 mm</p> <p>3. Compared 3 active learning-based methods (trained by 9 labelled volumes and 8 unlabeled volumes) (ENT, MAR, CEAL) -Dice similarity coefficient (DSC): FSL 82: 0.866 -Intersection-over-union (IoU): FSL 82: 0.7645 -Sensitivity (SEN): CEAL: 87.85% -Positive predictive value (ppv): FSL 82: 87.79% -Hausdorff distance (HD): MT: 2.76 mm -Average symmetric surface distance (ASSD): CEAL: 1.05 mm -Surface overlap (SO): CEAL: 95.92% -Surface dice (SD): CEAL: 0.9589</p>	CTooth+ (https://github.com/liangjiubujiu/CTooth , accessed on 21 August 2023)
Al-Sarem et al. 2022 [46]	Semantic	A pre-trained deep learning model (DenseNet169) based on U-Net for detecting and classifying tooth regions.	CBCT	500 scans -Training: 70% -validation: 20% -Testing: 10%	-Accuracy: 90.81% -Precision: 96% -Recall: 97% -F1-score: 0.97	Request by contacting;

Table 2. Cont.

Author	Category	Framework	Capture Method	Number of Samples	Evaluation Metrics	Model Reproducibility
Yang et al. 2021 [47]	Instance	A two-stage tooth segmentation model with deep convolutional neural networks and level set method. First to detect the center point, direction and length of the tooth by deep convolutional neural networks (segment dental pulp by U-Net to locate the tooth) and use a series of mathematical methods to fit an ellipse curve as the shape prior information, which is used to define the prior constraint term. Then, to combine the image data term, the length term, the regularization term and the prior constraint term to define the level set formulation of the energy functional and propose an accurate tooth segmentation model.	CBCT	10 patients (512 scanning slices for each) For training U-Net for pulp segmentation: -Training: 2 patients (1024 slices) -Validation: 1 patient (512 slices)	-Dice coefficient: 0.9791, -Jaccard coefficient: 0.9595 -Detection accuracy: 97.33% -Mean boundary F1 score: 0.9824	Not available
Shaheen et al. 2021 [48]	Instance	A CNN-based system for segmentation of each individual tooth and classification to a particular tooth class, which uses a 3D U-Net to segment tooth with bounding box.	CBCT from two machines	186 CBCT scans -Training: 140 (teeth: 400) -Validation: 35 (teeth: 100) -Testing: 11 (teeth: 332)	For segmentation -Dice similarity coefficient (DSC): 0.90 -Intersection over union (IoU): 0.82 -Recall: 83% -Precision: 98% -95% Hausdorff Distance (HD): 0.56 mm -Time: 13.7 ± 1.2 s For classification -Recall: 98.5% -Precision: 97.9% -Accuracy: 96.6%	Virtual Patient Creator (https://creator.relu.eu , accessed on 21 August 2023)
Lin et al. 2021 [49]	Semantic	A novel data pipeline based on micro-CT data to train the 2D U-Net for an accurate pulp cavity and tooth segmentation on CBCT images. The 2D U-Net containing region proposal network (RPN) with a feature pyramid network (FPN) structure was proposed in previous research to locate the extracted tooth and segmentation (Duan et al. 2021).	CBCT Micro CT	30 Teeth -Training: 25 groups (3200 sagittal slices and 6400 axial slices) -Testing: 5 groups	-Dice similarity coefficient (DSC): 0.962 -precision rate (PR): 97.31% -recall rate (RR): 95.11% -average symmetric surface distance (ASSD): 0.09 mm -Hausdorff distance (HD): 1.54 mm	Not available
Cui et al. 2021 [50]	Instance	A hierarchical morphological guide model with 3D V-Net as backbone located tooth centroids and predicted skeletons first, and then predicted the detailed geometric features (tooth volume, boundary, and root landmarks) with a multi-task learning mechanism under guidance.	CBCT	100 CBCT -Training set: 50 -Validation set: 10 -Testing: 40	-Dice: 0.948 -Jaccard: 0.891 -Average surface distance (ASD): 0.18 mm -Hausdorff distance (HD): 1.52 ± 0.28 mm	Not available
Lahoud et al. 2021 [51]	Instance	A CNN based AI-driven tooth segmentation, automated detection and segmentation of tooth structures.	CBCT	2924 slices, -Training: 2095 -Optimization: 501 -Validation: 328	-DSC: 0.93 -IoU: 0.87 -Segmentation volumes: 536 mm ³ -Average median surface deviation: 7.85 mm -Time: 0.5 min	Not available

Table 2. Cont.

Author	Category	Framework	Capture Method	Number of Samples	Evaluation Metrics	Model Reproducibility
Duan et al. 2021 [52]	Instance	A two-phase deep learning solution for tooth and pulp segmentation using U-Net in CBCT images. First, the single tooth bounding box is extracted by using the Region Proposal Network (RPN) with the Feature Pyramid Network (FPN) method from the perspective of panorama. Second, U-Net model is iteratively performed for refined tooth and pulp segmentation.	CBCT	20 sets Ten-fold cross-validation	Single root tooth/Multi-root tooth -Dice similarity coefficient (DSC): 0.957/0.962 -Average symmetric surface distance (ASD): 0.104/0.137 mm -Relative volume difference (RVD): 0.049/0.053	Not available
Wang et al. 2021 [53]	Semantic	A novel CNN architecture, mixed-scale dense (MS-D) CNN, for multiclass segmentation of the jaw, the teeth, and the background in CBCT scans.	CBCT	30 scans (9507 slices) -Divided into 4 subsets, 7 scans each (4-fold cross-validation scheme) -Training: 3 -Testing: 1	-Dice similarity coefficient (DSC): 0.945 -MAD: 0.204 ± 0.061 mm	Not available
Wu et al. 2020 [54]	Instance	A two-level hierarchical deep neural network for tooth segmentation. First embed center-sensitive mechanism with global stage heatmap and a deep supervised 3D-Unet, to ensure accurate tooth centers and guide the localization of tooth instances. Then, in the local stage, DenseASPP-UNet is proposed for fine segmentation and classification and a boundary-aware dice loss is proposed to gain accurate tooth boundaries.	CBCT	20 scans -Training: 12 (324 teeth) -Testing: 8 (219 teeth)	-Dice similarity coefficient (DSC): 0.962 -Detection accuracy (DA): 99.1% -Identification accuracy (FA): 99.5% -Average Symmetric Surface Distance (ASD): 0.122 mm	Not available
Rao et al. 2020 [55]	Semantic	A symmetric fully convolutional residual network, with dense conditional random fields (DCRF) to refine the posterior probability map. It used a novel Deep Bottleneck Architecture (DBA) to replace the general convolutional layer in U-Net and introduce a skip connection structure to enhance the propagation and reuse of the features. The DCRF was applied for overall structured prediction to get rid of the noises in the images, which can locate the tooth contour precisely.	CBCT	-Training: 86 images (conventional/unconventional = 51/35) -Testing: 24 images	-Volume Difference (VD): 18.86 -Dice Similarity Coefficient (DSC): 0.9166 -Average Symmetric Surface Distance (ASSD): 0.25 mm -Maximum Symmetric Surface Distance (MSSD): 1.18 mm	Not available
Lee et al. 2020 [56]	Semantic	A CNN based method (UDS-Net) with multi-phase training and preprocessing based on the U-Net structure. For multi-phase training, sub-volumes of different sizes were defined and used to produce stable and fast convergence. Then, a histogram-based method was used as a preprocessing step to estimate the average gray density level of the bone and tooth regions. Finally, a posterior probability function was developed and regularized the CNN models with spatial dropout layers and replaced the convolutional layers with dense convolution blocks, further improving the segmentation performance.	CBCT	102 datasets -Training: 69 (1066 images) -Validation: 1 (400 images) -Testing: 32 (151 images)	For validation -Dice: 0.938 -Recall: 95.2% -Precision: 92.4% For testing -Dice: 0.918 -Recall: 93.2% -Precision: 90.4%	Not available

Table 2. Cont.

Author	Category	Framework	Capture Method	Number of Samples	Evaluation Metrics	Model Reproducibility
Chung et al. 2020 [57]	Instance	A neural network, pose-aware TRCNN, for pixel-wise labeling to exploit an instance segmentation framework that is robust to metal artifacts. First, the alignment information of the patient was extracted by pose regression neural networks to attain a volume-of-interest (VOI) region and realign the input image, which reduces the inter overlapping area between tooth bounding boxes. Then VOI region was realigned based on the pose. Finally, a 3D U-Net was performed for individual tooth segmentation by converting the pixel-wise labeling task to a distance regression task.	CBCT	-Training: 100 images -Training: 50 -Testing: 25	-F1 score: 0.93 -Aggregated Jaccard index (AJI): 0.86 -Precision: 93% -Sensitivity: 93% -Hausdorff distance (HD): 1.59 mm -Average symmetric surface distance (ASSD): 0.20 mm	Not available
Chen et al. 2020 [58]	Instance	A multi-task 3D fully convolutional network (FCN) and marker-controlled watershed transform (MWT) to segment individual tooth. The multi-task FCN learns to simultaneously predict the probability of tooth region and the probability of tooth surface. Through the combination of the tooth probability gradient map and the surface probability map as the input image, MWT is used to automatically separate and segment individual tooth.	CBCT	25 patient -Training: 20 -Testing: 5	-Jaccard similarity coefficient (Omega): 0.936 -Dice similarity coefficient (DSC): 0.881 -Relative volume difference (RVD): 0.072 -Average symmetric surface distance (ASSD): 0.363 mm	Not available
Ezhov et al. 2019 [59]	Semantic	A V-Net based fully convolutional network for both coarse and fine segmentation. First, the model was trained to predict coarse segmentation using a large weakly labeled dataset, and then finetuned on a smaller, precisely labeled dataset while still predicting coarse masks.	CBCT	-Training: 93% -Testing: 7%	-Intersection over union (IoU): 0.94 -Average surface distance (ASD): 0.17 mm	Not available
Cui et al. 2019 [60]	Instance	A two-stage deep supervised neural network using 3D Mask R-CNN as the base network for segmentation and identification. First, the edge map was extracted from CBCT images enhance image contrast along shape boundaries. Then, the 3D region proposal network (RPN) was built with a novel learned similarity matrix to help efficiently remove redundant proposals, speed up training and save GPU memory.	CBCT	20 scans -Training: 12 -Testing: 8	-Dice similarity coefficient (DSC): 0.9237 -Detection accuracy (DA): 99.55% -Identification accuracy (FA): 96.85%	Not available
Gou et al. 2019 [61]	Semantic	A novel tooth-based approach that integrated U-Net with a level set model. Level set method was used to build the mask for CT images. U-Net structure was changed for the feasibility to images of any size.	CBCT	400 images -Training: 300 -Validation: 100	-Accuracy: 66.7% -Time: 10 s	Not available

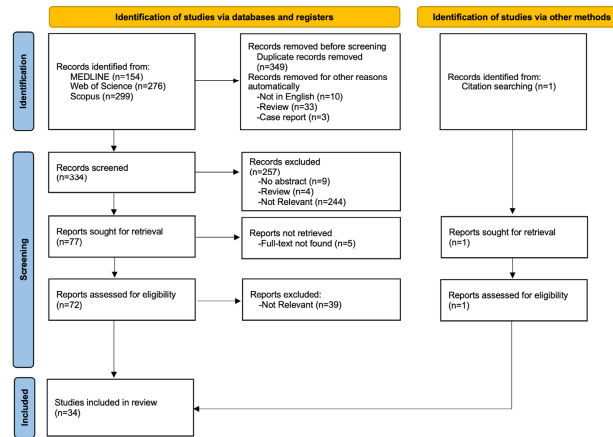


Figure 1. PRISMA 2020 Flowchart review process and identification of the number of relevant studies.

2.5. Risk of Bias Assessment

Risk of bias assessments were analyzed using the Quality Assessment Tool for Diagnostic Accuracy Studies-2 (QUADAS-2). Study bias and applicability were assessed through four key domains: patient selection, index test, reference standard, and flow and timing. The risk of bias and concerns regarding applicability assessment results are summarized in Figure 2 and shown in charts (Figure 3).

	Risk of Bias				Applicability Concerns		
	Patient Selection	Index Test	Reference Standard	Flow and Timing	Patient Selection	Index Test	Reference Standard
Al-Sarem et al., 2022	●	●	●	●	●	●	●
Ayidh Alqahtani et al., 2023	●	●	●	●	●	●	●
Chen et al., 2020	●	●	●	●	●	●	●
Chen et al., 2023	●	●	●	●	●	●	●
Chung et al., 2020	●	●	●	●	●	●	●
Cui et al., 2019	●	●	●	●	●	●	●
Cui et al., 2021	●	●	●	●	●	●	●
Cui et al. a, 2022	●	●	●	●	●	●	●
Cui et al. b, 2022	●	●	●	●	●	●	●
Cui et al. c, 2022	●	●	●	●	●	●	●
Dou et al., 2022	●	●	●	●	●	●	●
Duan et al., 2021	●	●	●	●	●	●	●
Ezhov et al., 2019	●	●	●	●	●	●	●
Fang et al., 2022	●	●	●	●	●	●	●
Fontenele et al., 2022	●	●	●	●	●	●	●
Gerhardt et al., 2022	●	●	●	●	●	●	●
Gou et al. 2019	●	●	●	●	●	●	●
Hsu et al., 2022	●	●	●	●	●	●	●
Jang et al., 2022	●	●	●	●	●	●	●
khan et al., 2022	●	●	●	●	●	●	●
Lahoud et al., 2021	●	●	●	●	●	●	●
Lee et al., 2020	●	●	●	●	●	●	●
Lee et al., 2022	●	●	●	●	●	●	●
Lin et al., 2021	●	●	●	●	●	●	●
Rao et al., 2020	●	●	●	●	●	●	●
Shaheen et al., 2021	●	●	●	●	●	●	●
Wang et al., 2021	●	●	●	●	●	●	●
Wang et al., 2023	●	●	●	●	●	●	●
Wu et al., 2020	●	●	●	●	●	●	●
Xie et al., 2022	●	●	●	●	●	●	●
Xie et al., 2023	●	●	●	●	●	●	●
Yang et al., 2021	●	●	●	●	●	●	●
Yang et al., 2022	●	●	●	●	●	●	●
Yin et al., 2023	●	●	●	●	●	●	●

● High ● Unclear ● Low

Figure 2. Summary for risk of bias and concerns regarding applicability of selected articles under QUADAS-2.

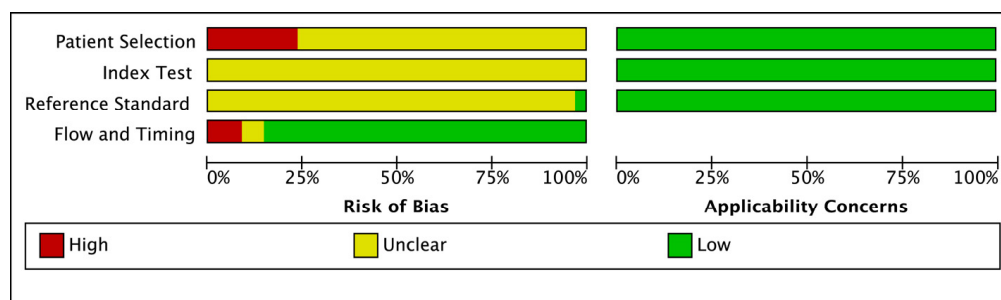


Figure 3. Percentage for risk of bias and application concerns of the selected articles under QUADAS-2.

3. Results

3.1. Search Results and Study Selection

A total of 729 records were obtained from searching various databases using specified keywords: 154 from MEDLINE, 276 from Web of Science, and 299 from Scopus. After automatically removing duplicates, the titles and abstracts of the remaining 334 articles were initially screened to determine their alignment with the inclusion criteria. Following excluding reviews, case reports, and articles not written in English, 77 articles were identified for full-text retrieval. Then, 44 of them were excluded for reasons including the full text being unavailable, the lack of artificial intelligence in the research methodology, or the focus being unrelated to tooth segmentation. Additionally, one article that met the criteria was obtained through manual search. A total of 34 studies was finally included in this systematic review. The PRISMA diagram illustrating the search results is depicted in Figure 1.

3.2. Risk of Bias and Applicability Concerns

The risk of bias and applicability concerns for each study were assessed following the QUADAS-2 guidelines and the results are illustrated in Figure 2. It was found that 23.53% of the studies ignored patients with restorations, such as fillings or crown restorations, when selecting raw CBCT images. Nearly all studies using machine learning or deep learning models examined and evaluated tooth segmentation methods based on established judgment criteria. It is not clear whether researchers can ignore the criteria in favor of fair data analysis. A total of 85.29% of studies utilized the entire collected image dataset for the learning, testing, and validation of the artificial intelligence model. While every study showed some uncertainty or a serious risk of bias, there were no applicability concerns. Figure 3 presents the risk of bias and the percentage of applicability concerns for the selected articles based on QUADAS-2.

3.3. Study Characteristics

The research details of all included articles are summarized in Table 2. Most studies applied artificial intelligence methods based on deep learning techniques, such as CNNs [30,32,34,39,50,52,55,56], fully convolutional networks (FCNs) [35,61], and network structures based on CNNs. For example, U-Net [28,29,33,36,43–47,49,52,55] and V-Net [31,37–40,42,48,50,53,54,58,59,61] are applied to segment two-dimensional and three-dimensional images, respectively. Some studies adopt multi-stage or multi-task strategies based on CNNs to divide the tooth segmentation problem into several sub-problems [34,36,48,50,52,54,58]. Other approaches used priori knowledge or morphological information to determine the center, shape, or border of teeth [48–50,52,55,58,59]. Still, some researchers applied the attention mechanism to capture the dependencies and contextual information among the teeth [32,48]. Moreover, some other models integrate machine learning techniques, such as LSM [28,35,49], conditional random field (CRF) [55], region proposal network (RPN) [29,52], and feature pyramid network (FPN) [29,51,52] to extract and utilize feature information from CBCT images for tooth detection, segmentation, and identification.

The number of CBCT images included in the studies ranged from 5 to 4938 [36,38]. Two of these studies used other data besides CBCT to improve the performance of the tooth segmentation models. One study used Micro-CT data of the teeth to enhance the outcomes of the U-Net model for tooth and pulp cavity segmentation in CBCT images [29], while another study used intra-oral scanning data to replace the surface morphology of the crowns in CBCT images and generate high-resolution tooth models [41]. Besides tooth segmentation, some studies have also explored the segmentation of the pulp cavity, jaw-bone, or small edentulous areas [29,37,39]. In this paper, we only consider the applications of tooth segmentation from CBCT images. To train and evaluate deep learning models, the location and extent of teeth need to be extracted and labeled from the patient's maxillofacial 3D images which are obtained directly in the clinic. Manual annotation can provide accurate ground truth (GT) but often requires significant time and effort. Therefore, it is common to augment the dataset by cropping, rotating, and mirroring after the data has been normalized. Two of these studies have created publicly available datasets for use by other researchers [45,46]. The homogenized CBCT data were divided into training and test sets for AI model training and testing, and the remaining data were used for measuring the performance of tooth segmentation methods.

3.4. Evaluation Metrics

Different metrics were used to compare the segmented images with the ground truth and various neural networks or machine learning algorithms that exist in order to evaluate AI segmentation models' performance. Evaluation metrics can be divided into three categories: overlap and similarity metrics, distance metrics, and volumetric metrics. Overlap and similarity metrics define how well the segmented regions agree with a known reference standard. They include accuracy (detection accuracy (DA), identification accuracy (FA), pixel accuracy (PA)), sensitivity or recall, precision, Boundary F1 (BF) score, Jaccard coefficient (JS), intersection over union (IoU), and Dice index or Dice similarity coefficient (DSC), mean Dice and symmetric best Dice (SBD)) [28,30,33,47,48,60,61]. Distance metrics describe the difference in distance between the segmented graphic contours, or surface pixels, and the reference standard. They include average system surface distance (ASSD), Hausdorff distance (HD), also known as maximum symmetric surface distance (MSSD), average median surface deviation, and mean absolute deviations (MADs) [51,52,57]. Volumetric metrics evaluate the performance of the tooth segmentation model from a volumetric perspective. They include volume differences (VD) and relative volume differences (RVD) [52,59]. Other metrics include object include ratio (OIR), Chamfer-L2, normal consistency (Normals), and occupancy accuracy (OccAcc), also applied to measure the coherence and accuracy of the segmentation algorithms on surfaces and space vectors [41,61]. In addition, some studies have reported the segmentation time [35,39,43,51,54] and the data size of different models [36] as additional evaluation criteria.

3.5. Performance of AI Models

The most commonly used metric is the Dice index, which is used by 25 articles. The Dice index assesses the similarity between segmentation results and ground truth by calculating pixels in the intersection and comparing this to the average size of two areas, visually evaluating the accuracy of AI segmentation models [33]. Similar metrics, including IoU and Jaccard coefficient, are calculated by dividing the area of intersection by the area of the union [61]. A total of 11 and 6 articles, respectively, used them to measure the accuracy and similarity. Compared to IoU, the Jaccard coefficient is not only used for image segmentation but can also be applied to text. In summary, the existing AI tooth segmentation models can achieve a Dice index of 0.935 ± 0.035 (Mean \pm SD) and an IoU or Jaccard coefficient of 0.877 ± 0.075 . The CNN model developed by Ayidh Alqahtani et al. showed the highest DSC of 0.99, as well as the highest IoU, precision, and recall score of 0.99, and the lowest 95% HD of 0.12 mm, demonstrating near-perfect segmentation results [31]. In contrast, CAT-UNet exhibited the lowest segmentation accuracy, with a

DSC of 0.865 and an IoU of 78.12 [32]. The method of Point-based Detection and Gaussian Disentanglement achieved only a minimum average IOU of 0.704 in tooth detection [61]. Moreover, there are four articles used the Boundary F1 Score to assess the identification and segmentation of boundary regions, achieving an accuracy of 0.97 ± 0.02 .

In addition, accuracy, precision, and recall are widely used metrics for evaluating the model performance. Accuracy quantifies the proportion of pixels correctly classified by the model to the total number of pixels. Precision measures the correctness of the segmented area, while recall evaluates the success of the model in segmenting individual teeth [39]. Current AI models can achieve an accuracy of $98.00\% \pm 1.55\%$, a precision of $94.52\% \pm 4.53\%$, and a recall of $94.17\% \pm 4.08\%$ in the tooth segmentation in CBCT images. Gerhardt et al. and Fontenele et al. applied the same segmentation model based on the U-Net to the images [39,42]. Gerhardt et al. focused on images with small edentulous areas, achieving 100% accuracy and 100% precision. Fontenele et al., on the other hand, focused on images containing dental fillings and achieved the highest recall of 99.7%.

Furthermore, 15 papers used ASSD and 11 papers used HD to evaluate the image segmentation performance, respectively. ASSD considers the distance between each segmented area and the ground truth, while HD only considers the maximum distance, typically using the 95% Hausdorff Distance for evaluation. The results indicate that AI segmentation methods achieved an ASSD of 0.22 ± 0.13 mm and a 95% HD of 0.94 ± 0.59 mm. Lin et al. improved the U-Net-based segmentation model proposed by Duan et al. through integrating high precision pulp cavity images from micro-CT with the tooth images from CBCT, achieving the lowest ASSD of 0.09 mm [29,52]. Other evaluation metrics, utilized in only one article, are not described separately in this section.

4. Discussion

Accurate segmentation and identification of teeth in CBCT images is intricate and challenging. Although manual and semi-automatic methods can accomplish segmentation tasks, the accuracy and robustness of segmentation outcomes need improvement. AI-based methods for tooth segmentation in CBCT images represent a promising avenue for future development. Despite this potential, current research methods and results are scattered. This study attempts to provide a complete and systematic review of research published on this topic to date, to summarize current adopted AI tooth segmentation techniques, and to establish a foundation for further development.

Depending on the type of output result, segmentation methods can be classified into semantic segmentation and instance segmentation. In this review, 15 studies performed semantic segmentation and 19 studies performed instance segmentation of CBCT images. Semantic segmentation methods can categorize each pixel or voxel in a CBCT image as a tooth or non-tooth but cannot distinguish between different instances of teeth [62]. Wang et al. utilized the Mixed-Scale Dense (MS-D) CNN for semantic segmentation of the mandible and teeth. The segmentation results achieved a large overlap with the ground truth and exhibited minor surface deviation [30]. Hsu et al. proposed a 3.5D U-Net structure and compared it with five other U-Net structures, showing better semantic segmentation results [40]. On the other hand, instance segmentation methods can simultaneously detect and segment each tooth instance in CBCT images with its location and class information [26]. For example, ToothNet was the first implementation of using a two-stage deep CNN for tooth recognition and instance segmentation of CBCT images. It first extracts edge maps from CBCT images and then utilizes a 3D RPN and a novel learned similarity matrix based on Mask R-CNN to generate candidate regions [60]. Wu et al. also used a two-level hierarchical deep learning method to first determine the tooth centers from thermal maps, and then identify the teeth and segment them into seven different types by DenseASPP-UNet [58]. Comparing different methods, we found that instance segmentation techniques typically utilize two-stage or multi-stage neural networks to accomplish the recognition and segmentation of teeth. Introducing a priori knowledge or morphological information, such as the center, skeleton, boundary, and curvature of the tooth, enhances the robustness and

detailed representation of tooth segmentation. For example, the first stage involves detecting the edges or centers of the teeth to form the region of interest for coarse segmentation. The second stage or the subsequent process uses neural networks to accurately segment and recognize different instances of the teeth for fine segmentation [34,48–50,52,58].

U-Net is a CNN-based architecture designed to classify each pixel in an image. It is a predominant model for deep learning-based segmentation and is extensively employed in segmenting teeth from CBCT images. The structure features symmetric encoders and decoders, enhanced by cross-layer connections that merge feature maps from different resolutions, preserving detailed information [24]. Al-Sarem et al. performed an initial CBCT image segmentation using U-Net and further compared the results of six pre-trained deep learning networks in tooth classification [47]. Building upon U-Net, V-Net introduces residual modules into the encoder and decoder design and uses convolutional layers instead of pooling layers to increase network depth and complexity. This adjustment improves gradient flow during training and allows direct handling of 3D data, marking a significant step forward in volumetric image analysis [63]. Ezhov et al. used V-Net and weakly labeled data to obtain coarse segmentation results, further training the model using accurately labeled datasets, and achieving high-resolution segmentation of individual teeth [53]. Furthermore, by modifying the order of doubling the number of channels and the process of deconvolution in U-Net, 3D U-Net can speed up convergence and avoid bottlenecks in the network structure [64]. Virtual Patient Creator is an online cloud platform that offers 3D U-Net AI models as open-source data on its website, enabling high-resolution tooth segmentation of CBCT images [54]. Several studies have validated its performance by segmenting teeth with different conditions, such as teeth containing different fillings, small edentulous regions, and teeth with brackets, and obtained good segmentation accuracy [31,39,42].

Despite these advancements, achieving optimal segmentation solely with one deep learning approach remains challenging. U-Net struggles with long-range dependencies and global information extraction, particularly when the segmentation target has low contrast with the background. Furthermore, both V-Net and 3D U-Net are computationally and memory intensive due to their direct processing of 3D images. To solve these problems, researchers have proposed numerous enhancements to these foundational models to refine detection and segmentation outcomes. On the one hand, the multi-task strategy can be used to decompose the tooth segmentation problem into different sub-problems, such as detection and identification, coarse segmentation, fine segmentation, and classification, in order to improve accuracy and efficiency. Wang et al. utilized a framework containing three task branches, including spatial embedding, seed mapping, and identification, while performing deep learning on CBCT images, for instance tooth segmentation and classification [37]. On the other hand, some studies have introduced attention or self-attention mechanisms to capture pixels' correlation and contextual information between or within teeth to improve the accuracy and consistency of tooth segmentation [65]. Dou et al. inserted a 3D self-attention module into the encoder of V-Net, which can obtain the spatial relationship between pixel points in the tooth geometry, thus capturing the complete tooth spatial features and achieving an effective separation of the boundaries from the tooth root and bone [48]. Cui et al. introduced a series of 3D attentional structures in U-Net to reduce the impact of background and noise on segmenting the tooth volume [45]. A recently developed deep learning network, Transformer, can extract more effective image features by attentively capturing global contextual information and long-range dependencies of the target [66]. Based on this idea, Chen et al. proposed a pre-trained CNN Transformer Architecture UNet (CTA UNet), which combines the benefits of CNN and Transformer to segment CBCT images from multiple scales and merge dental spatial features effectively [32]. TransUNet [67] has a similar structure to CTA UNet, and CoTNET [68] is also developed based on it, which can simultaneously extract global and local fine features, combine with self-attention methods, and link context to achieve good target detection and instance segmentation performance. Yin et al. combined CoTNET with U-Net++ (a U-Net

variant) to achieve feature fusion at different levels and multiple scales and obtain finer segmented images [33].

In addition, by combining other machine learning techniques and algorithms, the detection effect and segmentation results can be further optimized. The Watershed Transform (WT) is a segmentation algorithm that enables automatic segmentation and separation of individual teeth by analyzing topological morphology to obtain boundary information [69]. Chen et al. proposed a two-branch V-Net that predicted both the probability of tooth region and the tooth surface, up-sampled the feature map at the decoder path and integrated the Marker-controlled watershed transform to achieve accurate segmentation of individual teeth and minimize the information loss [59]. Yang et al. combined the advantages of semantic and instance segmentation and used a U-Net model for labeling and semantic segmentation of each pixel in the image [36]. Xie et al. used a multi-task CNN with a U-Net structure to segment the foreground and center regions of a tooth first. Then both models followed by WT to segment the teeth into separate units, especially for overlapping teeth [34]. Gaussian distributions based on heatmap responses effectively preserved spatial information in the image through tooth surface pixel distributions, the task of pixel-by-pixel classification is converted to a distance map regression task to perform individual tooth segmentation, especially for overlapping teeth, with better segmentation results [58,61]. Moreover, the Region Proposal Network with Feature Pyramid Network can accurately locate the position of each tooth from the panoramic image and initially obtain the segmentation area, which improves the accuracy and edge smoothness of CBCT image segmentation and allows further effective segmentation of the pulp chamber [29,51,52]. Artificial intelligence models have also been used to assist traditional manual segmentation methods, such as LSM, by automatically locating the center point to control the segmentation range and achieve higher segmentation efficiency and accuracy [28,35,49].

Considering the existing AI methods used for tooth segmentation in CBCT images, the evaluation metrics used in the various studies varied widely and lacked a uniform standard. Most models have DSCs greater than 90% and accuracy ranging from 83% to 99%, whereas a true application of machine learning methods to clinical work requires an accuracy rate of approximately 99% [5]. However, AI-based image segmentation methods are affected by computer literacy, annotation cost and high diversity of medical images. Tooth voxels only account for 1–3% of the overall CBCT data, and many slices do not contain any tooth-related voxels, which makes direct processing of the entire CBCT image challenging. Therefore, most current deep learning models adopt a stage-based supervised learning approach, which consists of roughly detecting the tooth region and then performing fine segmentation. Lee et al. used a pre-trained U-Net with multiple stages on 3 different CBCT sub-volumes, including tooth, tooth-containing slices, and the whole CBCT, but the segmentation results were not significantly improved [56]. Furthermore, supervised neural networks required to be trained with a large number of samples that have been manually annotated, but significant differences between different CBCT images, such as brackets, metal fillings, crown restorations, dental implants, artifacts, and missing tooth areas, increase the cost of labeling and affect the training results. Meanwhile, most studies use a relatively small number of samples, which increases the risk of model overfitting. Therefore, data augmentation is usually used to improve the robustness and generalization ability of the model. Examples include spatially rotating, scaling, and flipping the image [33], using blending modes to perform random linear combinations of images or cropping and recollaging images [39]. In addition, researchers have also improved the model generalization ability to some extent by normalizing the neural network, such as employing the dropout layer, but it still cannot solve the problem fundamentally [55]. It is worth mentioning that the open CBCT annotated datasets CTooth and CTooth++ were established and made public in two studies [45,46]. They provide a basis for subsequently training network weights based on large-scale datasets. At the same time, the use of pre-trained neural networks may also improve the model's generalization and migration ability.

Despite the development of artificial intelligence in the field of tooth segmentation, it is important to note that all current studies have a high risk of uncertainty. On the one hand, CBCT images are not randomly selected, and researchers artificially exclude images containing fillings, crown restorations, or images containing edentulous areas when selecting patients [39,40]. On the other hand, some studies have considered only anterior teeth and premolars when selecting segmentation targets and have intentionally ignored molar teeth containing complex root conditions [51], or have concealed target selection [37,43,47]. All of these factors can lead to an inappropriate increase in segmentation accuracy. In addition, it is unclear whether the researchers were able to maintain objectivity when analyzing the data with a clear understanding of the reference standard. Therefore, the risk of index tests and reference standards cannot be ruled out, and there is subjectivity in the segmentation results obtained by the study, but better research methods and reference standards are lacking. Notably, only two papers have disclosed their model code and part of their datasets. This is common in the academic field, but it highlights the importance of open science and transparent research [35,38]. On one hand, open-source code and datasets not only enhance the reproducibility and verifiability of research but also strengthen academic communication and collaboration. On the other hand, this practice also faces issues related to information and copyright protection. Furthermore, there are currently no standardized formats or secure data storage and access platforms, making the management and acquisition of shared resources complicated.

We believe that this study provides readers with a comprehensive understanding of the emerging applications of artificial intelligence in CBCT tooth segmentation, establishing a theoretical foundation for novel approaches in the diagnosis and treatment of oral diseases. However, it is important to acknowledge that this article still has certain limitations. Only studies written in English were included, and the unavailability of full texts for some relevant articles may have led to potential inaccuracies in the results. Additionally, only studies based on CBCT were included.

5. Conclusions

This systematic review demonstrates that artificial intelligence methods exhibit excellent performance in tooth segmentation in CBCT images. Clinicians can efficiently accomplish tasks, such as disease diagnosis and treatment planning, either directly or indirectly, with the aid of these methods. Despite that, they still faced with problems of small databases and non-uniform evaluation metrics.

The research results strongly advocate for the creation of large-scale standardized and openly available datasets. This initiative is instrumental in elevating the accuracy of artificial intelligence models and enhancing the generalization, transfer learning capabilities, and robustness of segmentation models. Furthermore, it is imperative to promote the adoption of standardized protocols and industry-recognized evaluation metrics. Random sampling and blind data collection methods should be employed to reduce bias. Additionally, the utilization of unsupervised or semi-supervised training approaches can effectively decrease bias resulting from manual interventions during data processing. Only then can we objectively analyze the data, compare the segmentation results, and increase the possibility of clinical translation. In conclusion, novel techniques for automatic segmentation in digital dentistry require further improvement, and their clinical applications require further exploration and evaluation.

Author Contributions: Conceptualization, M.T. and K.B.; Methodology, M.T. and Y.Z.; Validation, M.T., Y.Z. and A.A.; Writing—original draft, Y.Z., Writing—review and editing, Y.Z., M.T., K.B. and A.A.; Supervision, M.T. and K.B.; All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AI	Artificial intelligence
ASSD	Average System Surface Distance
BF Score	Boundary F1 Score
CBCT	Cone Beam Computed Tomography
CNNs	Convolutional Neural Networks
CRF	Conditional Random Field
DA	Detection Accuracy
DSC	Dice Similarity Coefficient
FA	Identification Accuracy
FCNs	Fully Convolutional Networks
FPN	Feature Pyramid Network
GT	Ground Truth
HD	Hausdorff Distance
IoU	Intersection over Union
JS	Jaccard Coefficient
LSM	Level Set Method
MADs	Mean Absolute Deviations
MS-D	Mixed-Scale Dense
MSSD	Maximum Symmetric Surface Distance
OccAcc	Occupancy Accuracy
OIR	Object Include Ratio
PA	Pixel Accuracy
PRISMA	Preferred Reporting Items for Systematic reviews and Meta-Analyses
QUADAS-2	Quality Assessment Tool for Diagnostic Accuracy Studies-2
ROI	Regions of Interest
RPN	Region Proposal Network
RVD	Relative Volume Differences
SBD	Symmetric Best Dice
VD	Volume Differences
WT	Watershed Transform

References

1. Beek, D.-M.; Baan, F.; Liebregts, J.; Nienhuijs, M.; Bergé, S.; Maal, T.; Xi, T. A learning curve in 3D virtual surgical planned orthognathic surgery. *Clin. Oral Investig.* **2023**, *27*, 3907–3915. [[CrossRef](#)] [[PubMed](#)]
2. Zhang, Q.; Gong, Y.; Liu, F.; Wang, J.; Xiong, X.; Liu, Y. Association of temporomandibular joint osteoarthritis with dentoskeletal morphology in males: A cone-beam computed tomography and cephalometric analysis. *Orthod. Craniofac Res.* **2023**, *26*, 458–467. [[CrossRef](#)] [[PubMed](#)]
3. Algahtani, F.N.; Hebbal, M.; Alqarni, M.M.; Alaamer, R.; Alqahtani, A.; Almohareb, R.A.; Barakat, R.; Abdlhafeez, M.M. Prevalence of bone loss surrounding dental implants as detected in cone beam computed tomography: A cross-sectional study. *PeerJ* **2023**, *11*, e15770. [[CrossRef](#)] [[PubMed](#)]
4. Casiraghi, M.; Scarone, P.; Bellesi, L.; Piliero, M.A.; Pupillo, F.; Gaudino, D.; Fumagalli, G.; Del Grande, F.; Presilla, S. Effective dose and image quality for intraoperative imaging with a cone-beam CT and a mobile multi-slice CT in spinal surgery: A phantom study. *Phys. Med.* **2021**, *81*, 9–19. [[CrossRef](#)] [[PubMed](#)]
5. Weese, J.; Lorenz, C. Four challenges in medical image analysis from an industrial perspective. *Med. Image Anal.* **2016**, *33*, 44–49. [[CrossRef](#)] [[PubMed](#)]
6. Barriviera, M.; Duarte, W.R.; Januário, A.L.; Faber, J.; Bezerra, A.C.B. A new method to assess and measure palatal masticatory mucosa by cone-beam computerized tomography. *J. Clin. Periodontol.* **2009**, *36*, 564–568. [[CrossRef](#)] [[PubMed](#)]
7. Rad, A.; Rahim, M.S.M.; Rehman, A.; Altameem, A.; Saba, T. Evaluation of Current Dental Radiographs Segmentation Approaches in Computer-aided Applications. *Iete Tech. Rev.* **2013**, *30*, 210–222. [[CrossRef](#)]
8. Li, S.; Fevens, T.; Krzyżak, A.; Jin, C.; Li, S. Semi-automatic computer aided lesion detection in dental X-rays using variational level set. *Pattern Recognit.* **2007**, *40*, 2861–2873. [[CrossRef](#)]
9. Said, E.; Nassar, D.; Fahmy, G.; Ammar, H. Teeth segmentation in digitized dental X-ray films using mathematical morphology. *IEEE Trans. Inf. Forensics Secur.* **2006**, *1*, 178–189. [[CrossRef](#)]

10. Barandiaran, I.; Macía, I.; Berckmann, E.; Wald, D.; Dupillier, M.P.; Paloc, C.; Grana, M. An Automatic Segmentation and Reconstruction of Mandibular Structures from CT-Data. In Proceedings of the 10th International Conference on Intelligent Data Engineering and Automated Learning (IDEAL 2009), Burgos, Spain, 23–26 September 2009.
11. Evain, T.; Ripoché, X.; Atif, J.; Bloch, I. Semi-Automatic Teeth Segmentation in Cone-Beam Computed Tomography by Graph-Cut with Statistical Shape Priors. In Proceedings of the IEEE 14th International Symposium on Biomedical Imaging (ISBI)—From Nano to Macro, Melbourne, Australia, 18–21 April 2017.
12. Modi, C.K.; Desai, N.P. A Simple and Novel Algorithm for Automatic Selection of Roi for Dental Radiograph Segmentation. In Proceedings of the 24th Canadian Conference on Electrical and Computer Engineering (CCECE), Niagara Falls, Canada, 8–11 May 2011.
13. Indraswari, R.; Kurita, T.; Arifin, A.Z.; Suciati, N.; Astuti, E.R.; Navastara, D.A. 3D Region Merging for Segmentation of Teeth on Cone-Beam Computed Tomography Images. In Proceedings of the Joint 10th International Conference on Soft Computing and Intelligent Systems (SCIS)/19th International Symposium on Advanced Intelligent Systems (ISIS), Toyama, Japan, 5–8 December 2018.
14. Rad, A.E.; Rahim, M.S.M.; Norouzi, A. Digital dental X-ray Image Segmentation and Feature Extraction. *Telkomnika Indones. J. Electr. Eng.* **2013**, *11*, 3109–3114.
15. Gan, Y.; Xia, Z.; Xiong, J.; Zhao, Q.; Hu, Y.; Zhang, J. Toward accurate tooth segmentation from computed tomography images using a hybrid level set model. *Med. Phys.* **2015**, *42*, 14–27. [[CrossRef](#)]
16. Jiang, B.; Zhang, S.; Shi, M.; Liu, H.-L.; Shi, H. Alternate Level Set Evolutions With Controlled Switch for Tooth Segmentation. *IEEE Access* **2022**, *10*, 76563–76572. [[CrossRef](#)]
17. Ahmed, N.; Abbasi, M.S.; Zuberi, F.; Qamar, W.; Halim MS, B.; Maqsood, A.; Alam, M.K. Artificial Intelligence Techniques: Analysis, Application, and Outcome in Dentistry—A Systematic Review. *BioMed Res. Int.* **2021**, *2021*, 9751564. [[CrossRef](#)] [[PubMed](#)]
18. Karobari, M.I.; Adil, A.H.; Basheer, S.N.; Murugesan, S.; Savadamoorthi, K.S.; Mustafa, M.; Abdulwahed, A.; Almokhatieb, A.A. Evaluation of the Diagnostic and Prognostic Accuracy of Artificial Intelligence in Endodontic Dentistry: A Comprehensive Review of Literature. *Comput. Math. Methods Med.* **2023**, *2023*, 7049360. [[CrossRef](#)] [[PubMed](#)]
19. Garcia-Garcia, A.; Orts-Escolano, S.; Oprea, S.; Villena-Martinez, V.; Martinez-Gonzalez, P.; Garcia-Rodriguez, J. A survey on deep learning techniques for image and video semantic segmentation. *Appl. Soft Comput.* **2018**, *70*, 41–65. [[CrossRef](#)]
20. Fernandez, K.; Chang, C. Teeth/palate and interdental segmentation using artificial neural networks. In Proceedings of the Artificial Neural Networks in Pattern Recognition: 5th INNS IAPR TC 3 GIRPR Workshop, ANNPR 2012, Trento, Italy, 17–19 September 2012; Proceedings 5. Springer: Berlin/Heidelberg, Germany.
21. Deleat-Besson, R.; Le, C.; Al Turkestani, N.; Zhang, W.; Dumont, M.; Brosset, S.; Soroushmehr, R. Automatic Segmentation of Dental Root Canal and Merging with Crown Shape. In Proceedings of the 43rd Annual International Conference of the IEEE-Engineering-in-Medicine-and-Biology-Society (IEEE EMBC), Virtual Event, 1–5 November 2021.
22. Shelhamer, E.; Long, J.; Darrell, T. Fully Convolutional Networks for Semantic Segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 640–651. [[CrossRef](#)] [[PubMed](#)]
23. Miki, Y.; Muramatsu, C.; Hayashi, T.; Zhou, X.; Hara, T.; Katsumata, A.; Fujita, H. Classification of teeth in cone-beam CT using deep convolutional neural network. *Comput. Biol. Med.* **2017**, *80*, 24–29. [[CrossRef](#)] [[PubMed](#)]
24. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In Proceedings of the Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015: 18th International Conference, Munich, Germany, 5–9 October 2015; Proceedings, Part III 18. Springer: Berlin/Heidelberg, Germany.
25. Ronneberger, O.; Fischer, P.; Brox, T. Dental X-ray Image Segmentation Using a U-Shaped Deep Convolutional Network. In Proceedings of the International Symposium on Biomedical Imaging, Brooklyn, NY, USA, 16–19 April 2015.
26. He, K.; Gkioxari, G.; Dollár, P.; Girshick, R. Mask r-cnn. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017.
27. Page, M.J.; McKenzie, J.E.; Bossuyt, P.M.; Boutron, I.; Hoffmann, T.C.; Mulrow, C.D.; Shamseer, L.; Tetzlaff, J.M.; Akl, E.A.; Brennan, S.E.; et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *Syst. Rev.* **2021**, *10*, 89. [[CrossRef](#)] [[PubMed](#)]
28. Yin, Y.; Xu, W.; Chen, L.; Wu, H. CoT-UNet++: A medical image segmentation method based on contextual transformer and dense connection. *Math. Biosci. Eng.* **2023**, *20*, 8320–8336. [[CrossRef](#)]
29. Alqahtani, K.A.; Jacobs, R.; Smolders, A.; Van Gerven, A.; Willems, H.; Shujaat, S.; Shaheen, E. Deep convolutional neural network-based automated segmentation and classification of teeth with orthodontic brackets on cone-beam computed-tomographic images: A validation study. *Eur. J. Orthod.* **2023**, *45*, 169–174. [[CrossRef](#)]
30. Xie, R.; Yang, Y.; Chen, Z. WITS: Weakly-supervised individual tooth segmentation model trained on box-level labels. *Pattern Recognit.* **2023**, *133*, 108974. [[CrossRef](#)]
31. Wang, Y.; Xia, W.; Yan, Z.; Zhao, L.; Bian, X.; Liu, C.; Qi, Z.; Zhang, S.; Tang, Z. Root canal treatment planning by automatic tooth and root canal segmentation in dental CBCT with deep multi-task feature learning. *Med. Image Anal.* **2023**, *85*, 102750. [[CrossRef](#)] [[PubMed](#)]
32. Chen, Z.; Chen, S.; Hu, F. CTA-UNet: CNN-transformer architecture UNet for dental CBCT images segmentation. *Phys. Med. Biol.* **2023**, *68*, 175042. [[CrossRef](#)] [[PubMed](#)]

33. Yang, H.; Wang, X.; Li, G. Tooth and Pulp Chamber Automatic Segmentation with Artificial Intelligence Network and Morphometry Method in Cone-beam CT. *Int. J. Morphol.* **2022**, *40*, 407–413. [[CrossRef](#)]
34. Xie, L.; Liu, B.; Cao, Y.; Yang, C. Automatic Individual Tooth Segmentation in Cone-Beam Computed Tomography Based on Multi-Task CNN and Watershed Transform. In Proceedings of the 24th IEEE International Conference on High Performance Computing and Communications, 8th IEEE International Conference on Data Science and Systems, 20th IEEE International Conference on Smart City and 8th IEEE International Conference on Dependability in Sensor, Cloud and Big Data Systems and Application, HPCC/DSS/SmartCity/DependSys 2022, Hainan, China, 18–20 December 2022.
35. Lee, J.; Chung, M.; Lee, M.; Shin, Y.G. Tooth instance segmentation from cone-beam CT images through point-based detection and Gaussian disentanglement. *Multimed. Tools Appl.* **2022**, *81*, 18327–18342. [[CrossRef](#)]
36. Khan, S.; Mukati, A.; Rizvi, S.S.H.; Yazdanie, N. Tooth Segmentation in 3D Cone-Beam CT Images Using Deep Convolutional Neural Network. *Neural Netw. World* **2022**, *32*, 301–318. [[CrossRef](#)]
37. Cui, Z.; Fang, Y.; Mei, L.; Zhang, B.; Yu, B.; Liu, J.; Jiang, C.; Sun, Y.; Ma, L.; Huang, J.; et al. A fully automatic AI system for tooth and alveolar bone segmentation from cone-beam CT images. *Nat. Commun.* **2022**, *13*, 2096. [[CrossRef](#)] [[PubMed](#)]
38. Hsu, K.; Yuh, D.-Y.; Lin, S.-C.; Lyu, P.-S.; Pan, G.-X.; Zhuang, Y.-C.; Chang, C.-C.; Peng, H.-H.; Lee, T.-Y.; Juan, C.-H.; et al. Improving performance of deep learning models using 3.5D U-Net via majority voting for tooth segmentation on cone beam computed tomography. *Sci. Rep.* **2022**, *12*, 19809. [[CrossRef](#)] [[PubMed](#)]
39. do Nascimento Gerhardt, M.; Fontenele, R.C.; Leite, A.F.; Lahoud, P.; Van Gerven, A.; Willems, H.; Smolders, A.; Beznik, T.; Jacobs, R. Automated detection and labelling of teeth and small edentulous regions on cone-beam computed tomography using convolutional neural networks. *J. Dent.* **2022**, *122*, 104139. [[CrossRef](#)] [[PubMed](#)]
40. Fontenele, R.C.; Gerhardt, M.D.N.; Pinto, J.C.; Van Gerven, A.; Willems, H.; Jacobs, R.; Freitas, D.Q. 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. *J. Dent.* **2022**, *119*, 104069. [[CrossRef](#)]
41. Fang, Y.; Cui, Z.; Ma, L.; Mei, L.; Zhang, B.; Zhao, Y.; Jiang, Z.; Zhan, Y.; Pan, Y.; Zhu, M.; et al. Curvature-Enhanced Implicit Function Network for High-quality Tooth Model Generation from CBCT Images. In Proceedings of the 25th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI), Singapore, 18–22 September 2022.
42. Dou, W.; Gao, S.; Mao, D.; Dai, H.; Zhang, C.; Zhou, Y. Tooth instance segmentation based on capturing dependencies and receptive field adjustment in cone beam computed tomography. *Comput. Animat. Virtual Worlds* **2022**, *33*, e2100. [[CrossRef](#)]
43. Jang, T.J.; Kim, K.C.; Cho, H.C.; Seo, J.K. A Fully Automated Method for 3D Individual Tooth Identification and Segmentation in Dental CBCT. *IEEE Trans. Pattern Anal. Mach. Intell.* **2022**, *44*, 6562–6568. [[CrossRef](#)] [[PubMed](#)]
44. Cui, W.; Wang, Y.; Zhang, Q.; Zhou, H.; Song, D.; Zuo, X.; Jia, G.; Zeng, L. CTooth: A Fully Annotated 3D Dataset and Benchmark for Tooth Volume Segmentation on Cone Beam Computed Tomography Images. In *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*; Springer: Cham, Switzerland, 2022.
45. Cui, W.; Wang, Y.; Li, Y.; Song, D.; Zuo, X.; Wang, J.; Zhang, Y.; Zhou, H.; Chong, B.S.; Zeng, L.; et al. CTooth+: A Large-Scale Dental Cone Beam Computed Tomography Dataset and Benchmark for Tooth Volume Segmentation. In *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*; Springer: Cham, Switzerland, 2022.
46. Al-Sarem, M.; Al-Asali, M.; Alqutaibi, A.Y.; Saeed, F. Enhanced Tooth Region Detection Using Pretrained Deep Learning Models. *Int. J. Environ. Res. Public Health* **2022**, *19*, 15414. [[CrossRef](#)] [[PubMed](#)]
47. Yang, Y.; Xie, R.; Jia, W.; Chen, Z.; Yang, Y.; Xie, L.; Jiang, B. Accurate and automatic tooth image segmentation model with deep convolutional neural networks and level set method. *Neurocomputing* **2021**, *419*, 108–125. [[CrossRef](#)]
48. Shaheen, E.; Leite, A.; Alqahtani, K.A.; Smolders, A.; Van Gerven, A.; Willems, H.; Jacobs, R. A novel deep learning system for multi-class tooth segmentation and classification on cone beam computed tomography. A validation study: Deep learning for teeth segmentation and classification. *J. Dent.* **2021**, *115*, 103865. [[CrossRef](#)] [[PubMed](#)]
49. Lin, X.; Fu, Y.; Ren, G.; Yang, X.; Duan, W.; Chen, Y.; Zhang, Q. Micro-Computed Tomography-Guided Artificial Intelligence for Pulp Cavity and Tooth Segmentation on Cone-beam Computed Tomography. *J. Endod.* **2021**, *47*, 1933–1941. [[CrossRef](#)] [[PubMed](#)]
50. Cui, Z.; Zhang, B.; Lian, C.; Li, C.; Yang, L.; Wang, W.; Zhu, M.; Shen, D. Hierarchical Morphology-Guided Tooth Instance Segmentation from CBCT Images. In Proceedings of the Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), Virtual Event, 28–30 June 2021.
51. Lahoud, P.; EzEldeen, M.; Beznik, T.; Willems, H.; Leite, A.; Van Gerven, A.; Jacobs, R. Artificial Intelligence for Fast and Accurate 3-Dimensional Tooth Segmentation on Cone-beam Computed Tomography. *J. Endod.* **2021**, *47*, 827–835. [[CrossRef](#)] [[PubMed](#)]
52. Duan, W.; Chen, Y.; Zhang, Q.; Lin, X.; Yang, X. Refined tooth and pulp segmentation using U-Net in CBCT image. *Dentomaxillofacial Radiol.* **2021**, *50*, 20200251. [[CrossRef](#)] [[PubMed](#)]
53. Wang, H.; Minnema, J.; Batenburg, K.J.; Forouzanfar, T.; Hu, F.J.; Wu, G. Multiclass CBCT Image Segmentation for Orthodontics with Deep Learning. *J. Dent. Res.* **2021**, *100*, 943–949. [[CrossRef](#)]
54. Wu, X.; Chen, H.; Huang, Y.; Guo, H.; Qiu, T.; Wang, L. Center-Sensitive and Boundary-Aware Tooth Instance Segmentation and Classification from Cone-Beam CT. In Proceedings of the IEEE 17th International Symposium on Biomedical Imaging (ISBI), Iowa, IA, USA, 3–7 April 2020.
55. Rao, Y.; Wang, Y.; Meng, F.; Pu, J.; Sun, J.; Wang, Q. A Symmetric Fully Convolutional Residual Network With DCRF for Accurate Tooth Segmentation. *IEEE Access* **2020**, *8*, 92028–92038. [[CrossRef](#)]

56. Lee, S.; Woo, S.; Yu, J.; Seo, J.; Lee, J.; Lee, C. Automated CNN-Based Tooth Segmentation in Cone-Beam CT for Dental Implant Planning. *IEEE Access* **2020**, *8*, 50507–50518. [[CrossRef](#)]
57. Chung, M.; Lee, M.; Hong, J.; Park, S.; Lee, J.; Lee, J.; Yang, I.-H.; Lee, J.; Shin, Y.G. Pose-aware instance segmentation framework from cone beam CT images for tooth segmentation. *Comput. Biol. Med.* **2020**, *120*, 103720. [[CrossRef](#)] [[PubMed](#)]
58. Chen, Y.; Du, H.; Yun, Z.; Yang, S.; Dai, Z.; Zhong, L.; Feng, Q.; Yang, W. Automatic Segmentation of Individual Tooth in Dental CBCT Images From Tooth Surface Map by a Multi-Task FCN. *IEEE Access* **2020**, *8*, 97296–97309. [[CrossRef](#)]
59. Ezhov, M.; Zakirov, A.; Gusarev, M. Coarse-to-Fine Volumetric Segmentation of Teeth in Cone-Beam CT. In Proceedings of the 16th IEEE International Symposium on Biomedical Imaging (ISBI), Venice, Italy, 8–11 April 2019.
60. Cui, Z.; Li, C.; Wang, W. ToothNet: Automatic Tooth Instance Segmentation and Identification from Cone Beam CT Images. In Proceedings of the 32nd IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 15–20 June 2019.
61. Gou, M.; Rao, Y.; Zhang, M.; Sun, J.; Cheng, K. Automatic image annotation and deep learning for tooth CT image segmentation. In Proceedings of the Image and Graphics: 10th International Conference, ICIIG 2019, Beijing, China, 23–25 August 2019; Proceedings, Part II 10. Springer: Berlin/Heidelberg, Germany.
62. Long, J.; Shelhamer, E.; Darrell, T. Fully Convolutional Networks for Semantic Segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 7–12 June 2015.
63. Milletari, F.; Navab, N.; Ahmadi, S.A. V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation. In Proceedings of the 4th IEEE International Conference on 3D Vision (3DV), Stanford, CA, USA, 25–28 October 2016.
64. Çiçek, Ö.; Abdulkadir, A.; Lienkamp, S.S.; Brox, T.; Ronneberger, O. 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation. In Proceedings of the Medical Image Computing and Computer-Assisted Intervention—MICCAI 2016, Athens, Greece, 17–21 October 2016; Springer International Publishing: Cham, Switzerland.
65. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention Is All You Need. In Proceedings of the 31st Annual Conference on Neural Information Processing Systems (NIPS), Long Beach, CA, USA, 4–9 December 2017.
66. Liu, Z.; Lin, Y.; Cao, Y.; Hu, H.; Wei, Y.; Zhang, Z.; Lin, S.; Guo, B. Swin Transformer: Hierarchical Vision Transformer using Shifted Windows. In Proceedings of the 18th IEEE/CVF International Conference on Computer Vision (ICCV), Virtual Event, 10 March 2021.
67. Lin, A.; Chen, B.; Xu, J.; Zhang, Z.; Lu, G.; Zhang, D. DS-TransUNet: Dual Swin Transformer U-Net for Medical Image Segmentation. *IEEE Trans. Instrum. Meas.* **2022**, *71*, 4005615. [[CrossRef](#)]
68. Li, Y.; Yao, T.; Pan, Y.; Mei, T. Contextual Transformer Networks for Visual Recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **2023**, *45*, 1489–1500. [[CrossRef](#)]
69. Grau, V.; Mewes, A.U.J.; Alcaniz, M.; Kikinis, R.; Warfield, S.K. Improved watershed transform for medical image segmentation using prior information. *IEEE Trans. Med. Imaging* **2004**, *23*, 447–458. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.