

Review

Automatic Segmentation of Mandible from Conventional Methods to Deep Learning—A Review

Bingjiang Qiu ^{1,2,3}, Hylke van der Wel ^{1,4}, Joep Kraeima ^{1,4}, Haye Hendrik Glas ^{1,4}, Jiapan Guo ^{2,3,*},
Ronald J. H. Borra ⁵, Max Johannes Hendrikus Witjes ^{1,4} and Peter M. A. van Ooijen ^{2,3} 

¹ 3D Lab, University Medical Center Groningen, University of Groningen, Hanzeplein 1, 9713 GZ Groningen, The Netherlands; b.qiu@umcg.nl (B.Q.); h.van.der.wel@umcg.nl (H.v.d.W.); j.kraeima@umcg.nl (J.K.); h.h.glas@umcg.nl (H.H.G.); m.j.h.witjes@umcg.nl (M.J.H.W.)

² Department of Radiation Oncology, University Medical Center Groningen, University of Groningen, Hanzeplein 1, 9713 GZ Groningen, The Netherlands; p.m.a.van.ooijen@umcg.nl

³ Data Science Center in Health (DASH), University Medical Center Groningen, University of Groningen, Hanzeplein 1, 9713 GZ Groningen, The Netherlands

⁴ Department of Oral and Maxillofacial Surgery, University Medical Center Groningen, University of Groningen, Hanzeplein 1, 9713 GZ Groningen, The Netherlands

⁵ Medical Imaging Center (MIC), University Medical Center Groningen, University of Groningen, Hanzeplein 1, 9713 GZ Groningen, The Netherlands; r.j.h.borra@umcg.nl

* Correspondence: j.guo@umcg.nl



Citation: Qiu, B.; van der Wel, H.; Kraeima, J.; Glas, H.H.; Guo, J.; Borra, R.J.H.; Witjes, M.J.H.; van Ooijen, P.M.A. Automatic Segmentation of Mandible from Conventional Methods to Deep Learning—A Review. *J. Pers. Med.* **2021**, *11*, 629. <https://doi.org/10.3390/jpm11070629>

Academic Editor: Jorge Luis Espinoza

Received: 28 May 2021

Accepted: 28 June 2021

Published: 1 July 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 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/>).

Abstract: Medical imaging techniques, such as (cone beam) computed tomography and magnetic resonance imaging, have proven to be a valuable component for oral and maxillofacial surgery (OMFS). Accurate segmentation of the mandible from head and neck (H&N) scans is an important step in order to build a personalized 3D digital mandible model for 3D printing and treatment planning of OMFS. Segmented mandible structures are used to effectively visualize the mandible volumes and to evaluate particular mandible properties quantitatively. However, mandible segmentation is always challenging for both clinicians and researchers, due to complex structures and higher attenuation materials, such as teeth (filling) or metal implants that easily lead to high noise and strong artifacts during scanning. Moreover, the size and shape of the mandible vary to a large extent between individuals. Therefore, mandible segmentation is a tedious and time-consuming task and requires adequate training to be performed properly. With the advancement of computer vision approaches, researchers have developed several algorithms to automatically segment the mandible during the last two decades. The objective of this review was to present the available fully (semi)automatic segmentation methods of the mandible published in different scientific articles. This review provides a vivid description of the scientific advancements to clinicians and researchers in this field to help develop novel automatic methods for clinical applications.

Keywords: mandible segmentation; 3D virtual surgical planning; convolutional neural networks; machine learning

1. Introduction

Three-dimensional (3D) medical imaging techniques have a fundamental role in the field of oral and maxillofacial surgery (OMFS) [1,2]. 3D images are used to guide diagnosis, assess the severity of disease and for pre-operative planning and per-operative guidance using 3D images and virtual surgical planning (VSP) [3]. In the field of oral cancer, where surgical resection requiring the partial removal of the mandible is a common treatment, resection surgery is often based on 3D VSP to accurately design a resection plan around tumor margins [4]. In orthognathic surgery and dental implant surgery, 3D VSP is also extensively used to precisely guide mandibular surgery [3]. Image segmentation from the radiography images of the head and neck (H&N), which is a process to create a 3D volume of the target tissue, is a useful tool to visualize the mandible and quantify geometric

parameters [5]. Studies have shown that 3D VSP requires accurate segmentation of the mandible, which is currently performed by medical technicians [4,6]. Mandible segmentation, especially for 3D VSP, was usually done manually, which is a time-consuming and poorly reproducible process [7,8].

The mandible is located in the lower part of the facial skeleton and consists of the body, rami, angle, and condyles. The different anatomical regions of the mandible have varying densities, which is especially true if the teeth and surrounding soft tissue are also considered, which makes accurate manual segmentation of the mandible time-consuming and challenging [9]. Because a 3D model of the mandible is necessary for 3D VSP in OMFS, and time is often a limiting factor in the clinical workflow, fast mandible segmentation has become a frequent topic of research in recent decades [7]. Researchers have attempted to automate the segmentation process of the mandible, as well as all kinds of tissues from medical images, to reduce the processing time and also the inter-observer variability [10]. Recent advances in image segmentation [11–15] have enabled their applications to medical image segmentation, which stimulate progress in automating mandible segmentation. Although significant progress has been made in mandible segmentation, state-of-the-art methods still experience unsatisfactory outcomes, with several challenges to be solved. For oncology segmentation, the segmentation has to be made for an affected bone (with the tumor), which is a different task from segmenting a bone in implantology or for orthognathic purposes. In implantology, many patients have incomplete teeth, which can accordingly lead to resorption of the bone causing shape variation in the mandible among patients. In orthognathic surgery, the population considered varies greatly in shape in addition to asymmetric, which is also a possible challenge for (automatic) segmentation. Figure 1 shows research challenges in automatic segmentation algorithms for the mandible. The challenges associated with mandibular bone segmentation can be categorized as follows:

- The presence of anatomically complex bony structures in the scans. As the examples show in Figure 1a, a normal H&N scan includes other bony structures, with a complex anatomy and a similar density. Determining the correct boundaries and separating the mandibular bone from the other bones may be challenging.
- Artifacts. When X-rays pass through high-density structures or materials, including teeth, postoperative metal implants, etc., the signal on the detectors will change, which will lead to attenuation calculation errors in the (cone beam) computed tomography (CBCT/CT) reconstruction process and consequently cause high noise and strong artifacts in the visual impression of the scans [16]. The mandible boundaries nearby teeth tend to be blurred and hard to detect. In particular, the boundaries of mandible rami are difficult to be identified when dental braces and metal implants badly affect the image quality [17], as shown in Figure 1b. Furthermore, the fact that the superior and the inferior teeth are at the same slice and even overlapping that makes segmentation methods challenging, as shown in Figure 1c.
- Low contrast. Due to the tomography process and the thinness of condyles, the slices in condyles commonly have low contrast [18], as shown in Figure 1d. Especially in cone-beam CT (CBCT) scans, the condyles in the scans are more blurred than in CT because of its intensity inhomogeneity [19,20].
- Variation in the appearance. As the examples show in Figure 1e, mandibular shapes vary in shape and size between individuals [18]. Additionally, the shape of the mandibular condyles and body is extremely variable among patients and different age groups [21].
- Annotation Bias. The manual mandible segmentation often leads to inter-observer variability (Dice score of 94.09% between two clinical experts) [22], which directly influences the quality of treatment planning.

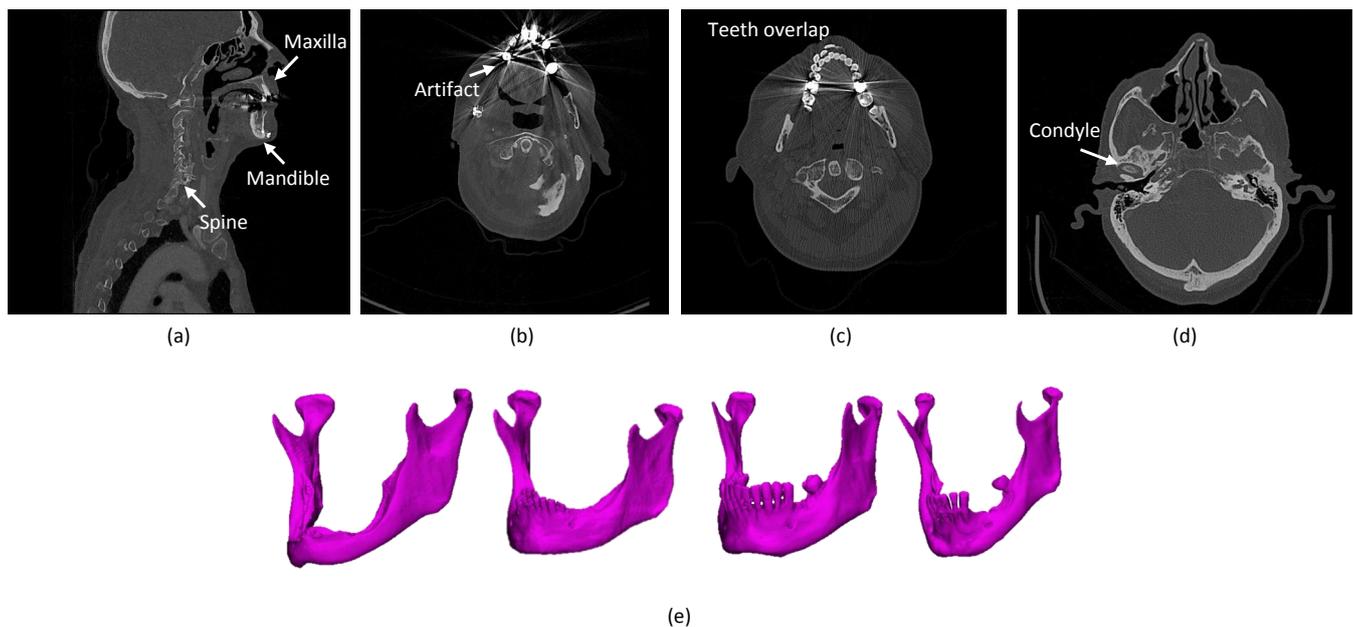


Figure 1. Examples of typical cases that challenge accurate mandible segmentation. (a) Various bone-structured organs in the H&N scans. (b) Metal artifacts. (c) Presence of inferior and superior teeth in the same slice. (d) Lower intensity in the condyles. (e) Large variation in mandibles between patients.

In the present article, a review has been conducted, where the literature available in PubMed and Web of Science databases relating to these studies was considered. In the literature, a variety of segmentation methods, based on statistical shape model, atlas and machine learning, are proposed for head and neck scanning. The selected publications were divided into several categories depending on the type of method used and discussed extensively.

2. Method for Literature Selection

In February 2021, a search was conducted on the Web of Science (<https://apps.whoofknowledge.com> (accessed on 23 February 2021)) and PubMed (<https://www.ncbi.nlm.nih.gov/pubmed/> (accessed on 23 February 2021)), with the topic keywords and search builder: (automat* AND segment* AND mandible AND 3D). While searching the literature, no specific time line was set. The literature search was performed by the two observers.

In this article, the lists of literature acquired by the above-mentioned database search were imported into Rayyan QCRI [23]. Using the web application, duplicates were searched and removed accordingly. Subsequently, another search was performed to exclude the articles that did not include “segment” or “mandible” keywords in their title and abstracts. Consequently, the irrelevant studies were removed by manually screening the titles and abstracts of journal articles. The obtained relevant publications were further supplemented with selected publications found in their list of references. A detailed analysis of the resulting publications was then conducted from the perspective of image modality, image database, evaluation metrics, and methodology.

3. Results

In the field of automatic 3D mandible segmentation, the search on the Web of Science and PubMed databases yielded 33 and 45 results, respectively. After detecting the duplicates in Rayyan, 23 duplicates were removed. After reviewing their abstracts by two observers, 12 were considered to be relevant and were further supplemented with selected publications from their list of references. In total, we collected 77 publications focused on mandible segmentation or organs at risk (OARs) (including mandible) segmentation in the H&N region. Here, we emphasize that we collected 41 publications from the reference of [24]. An overview of the results of the literature search is presented in Figure 2.

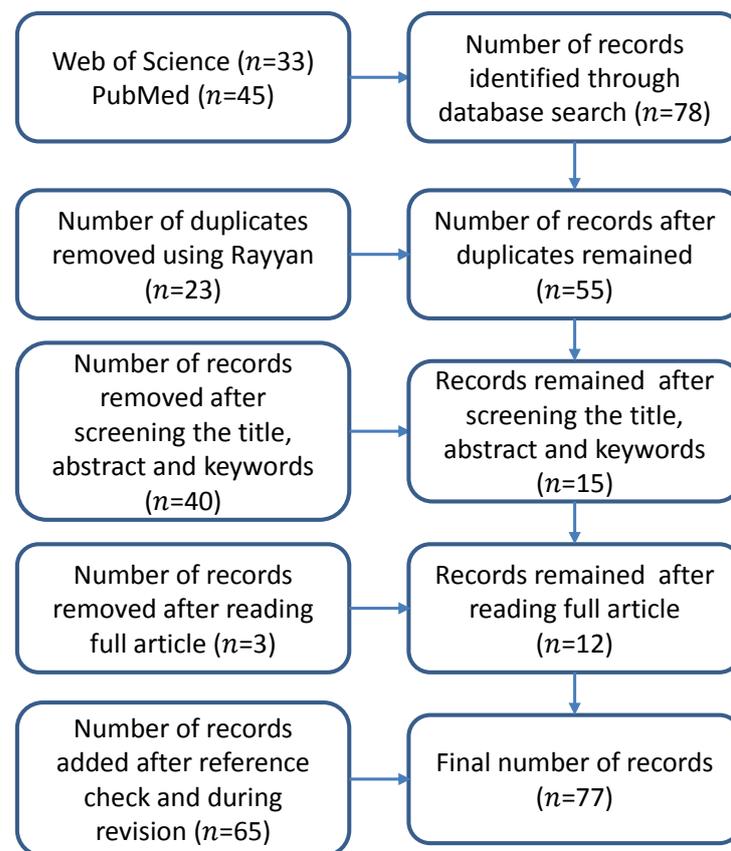


Figure 2. Flowchart of the literature review article selection process.

3.1. Image Modality

CT imaging is primarily performed for mandibular bone segmentation in their works since CT images provide a promising visibility of the bone structures. There are 70, 7 and 4 publications using CT, CBCT and magnetic resonance imaging (MRI), respectively, to visualize H&N (note: some articles applied the scans from multiple image modalities). CBCT images are applied in H&N because of its lower radiation dose, lower cost and faster scanning time than conventional CT [25]. In particular, CBCT is often used for validating patient settings or adjusting treatment plans to adapt to anatomical changes. While CT images provide a good visibility of the bony anatomy, the contrast differences between various soft tissues are much lower than those in MRI. In the consensus on manual delineation of the mandible based on CT, it is also necessary to use MR images, in addition to CT, to help delineate surrounding tumors and soft tissues (note: PET(/CT) imaging technique is also helpful for the tumor delineation [26–28]). Furthermore, considering the ionizing radiation of CT/CBCT scanning and the development of computer vision techniques, a strategy of MRI-only treatment planning has become increasingly valuable and feasible in the field of 3D VSP [3].

3.2. Image Database

Among the reviewed publications, there are four databases of H&N images that are publicly available. The Public Domain Database for Computational Anatomy (PDDCA) (<http://www.imagenglab.com/newsite/pddca/> (accessed on 24 January 2019)) [29], an open-access resource of medical images for cancer research, consists of 48 patient CT images from the Radiation Therapy Oncology Group (RTOG) 0522 study of Head-Neck Cetuximab [30] database of The Cancer Imaging Archive (TCIA) [31], together with manual annotation of the mandible, brainstem, etc. [29]. The organizer of PDDCA provides 40 CT scans with mandible annotations. This dataset has been used for the Head

and Neck Auto Segmentation MICCAI Challenge (2015) [29]. According to the protocol of the Challenge, 25 of 40 scans (0522c0001-0522c0328) are employed as the training dataset, and the other 15 scans (0522c0555-0522c0878) are employed as the test dataset [29]. ‘Organ-at-risk segmentation from head & neck CT scans’ is one of four tasks in StructSeg 2019 (<https://structseg2019.grand-challenge.org/Dataset/> (accessed on 23 February 2021)), consisting of 50 nasopharynx cancer patient CT scans with several manually delineated organs at risk including mandibles (left and right). The StructSeg 2019 dataset has been used for the Automatic Structure Segmentation for Radiotherapy Planning Challenge 2019. Each of the annotated CT scans is marked by one experienced oncologist and verified by another experienced oncologist.

Furthermore, a dataset has been augmented or combined into new publicly available databases by Tang et al. [32] in 2019, for example, the manual delineations of 28 OARs in 35 CT scans from the Head-Neck Cetuximab and 105 CT scans from Head-Neck-PET-CT [33] databases (<https://github.com/uci-cbcl/UaNet#Data> (accessed on 23 February 2021)). A subset of 31 CT scans, which was used for the test and validation set in the study from Nikolov et al. [24], was collected from two datasets of The Cancer Genome Atlas Head-Neck Squamous Cell Carcinoma [34] and Head-Neck Cetuximab [30] from TCIA. The subset with the ground truth added is available at <https://github.com/deepmind/tcia-ct-scan-dataset> (accessed on 23 February 2021).

3.3. Evaluation Metrics

In addition to the differences in the used dataset, the results are evaluated and presented in different ways in the reviewed papers. Moreover, there are no standard metrics for the segmentation evaluation, so different evaluation metrics are used to report the segmentation performance. For segmentation, the evaluation metrics are mainly divided into three categories: overlap-based metrics, distance-based metrics, and volume-based metrics. The metrics are summarized in Table 1. The overlap-based metrics indicate the difference in overlap measurement between automatic prediction and manual segmentation, which can be obtained by four indicators: true positive (TP), false positive (FP), false negative (FN) and true negative (TN). The TP value is the number of pixels segmented correctly as foreground. The count of pixels falsely classified as the foreground is given by the FP value. The total count of falsely classified as background pixels is represented by FN. The TN value represents the correctly classified background pixels. The most commonly used overlap-based metrics include the Dice similarity coefficient (Dice), Sensitivity (Sen), false positive volume fraction (FPVF) [35], false negative volume fraction (FNVF) [35], etc. To measure the contour difference between automatic and manual segmentation, the most commonly used metrics are distance-based metrics. In the context of mandibular segmentation, the following distance-based metrics have been frequently used: average symmetric surface distance (ASD), Hausdorff distance (HD), 95th-percentile Hausdorff distance (95HD), mean square error (MSE) [36], and root mean square error (RMSE) [36]. In some medical image segmentation tasks, the volume of the object is also very important for treatment planning, and the metric based on volume is helpful to evaluate the performance of the segmentation method. Volume overlap error and volume error are the common indices to evaluate the results of mandibular segmentation.

3.4. Methodology

In this paper, the mandible segmentation approaches have been broadly divided into six categories, namely, (1) statistical shape model-based (SSM-based), (2) active shape model-based (ASM-based), (3) active appearance model-based (AAM-based), (4) atlas-based, (5) level set-based, (6) classical machine learning-based and (7) deep learning-based approaches, as depicted in Table 2. There are several hybrid segmentation methods proposed in the collected literature. The hybrid methods are introduced in the listed seven categories.

Table 1. Performance metrics applied to performance measurement of automatic segmentation of the mandible and the corresponding references and mathematical definitions. Note: Y_r indicates the pixels in the reference standard (ground truth), and Y_p is the pixels in the automatic segmentation. $|\cdot|$ represents the number of voxels. $\|\cdot\|$ represents the L_2 norm. a and b are corresponding points on the boundary of A and B .

Metric	Abbreviation	Definition
Overlap-based metrics, reported in percents (%)		
Dice similarity index	Dice	$Dice = \frac{2 Y_r \cap Y_p }{ Y_r + Y_p } = \frac{2TP}{2TP + FP + FN}$
Sensitivity	Sen	$Sen = \frac{ Y_r \cap Y_p }{ Y_r } = \frac{TP}{TP + FN}$
Recall	Rec	$Recall = Sen$
Positive Predictive Value	PVV	$PVV = \frac{ Y_r \cap Y_p }{ Y_p } = \frac{TP}{TP + FP}$
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	$IoU = Jac$
Specificity	Spe	$Specificity = \frac{ (1 - Y_p) \cap (1 - Y_r) }{ Y_p } = \frac{TN}{TN + FP}$
False positive volume fraction	FPVF	$FPVF = \frac{ Y_p - Y_r }{ Y_r } = \frac{FP}{TP + FN}$
False negative volume fraction	FNVF	$FNVF = \frac{ Y_r - Y_p }{ Y_r } = \frac{FN}{TP + FN}$
Distance-based metrics, reported in millimeters (mm)		
Average symmetric surface distance	ASD	$ASD(A, B) = \frac{d(A, B) + d(B, A)}{2}$, where $d(A, B) = \frac{1}{N} \sum_{a \in A} \min_{b \in B} \ a - b\ $,
Hausdorff distance	HD	$D_{HD}(A, B) = \max(h(A, B), h(B, A))$, where $h(A, B) = \max_{a \in A} \min_{b \in B} \ a - b\ $
95th-percentile 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\ $
Mean square error	MSE	$MSE = \frac{1}{n} \sum_{i=1}^n \ A_i - B_i\ ^2$, where A_i is the boundary of the i -th OAR and B_i is the boundary of the i -th prediction.
Root mean square error	RMSE	$RMSE = \sqrt{MSE}$
Volume-based metrics, reported in percents (%)		
Volume overlap error	VOE	$1 - \frac{ Y_r \cap Y_p }{ Y_r \cup Y_p }$
Volume error	VE	$VE = \frac{Y_r - Y_p}{Y_r}$ or $\frac{Y_p - Y_r}{Y_r}$

Table 2. Methodology applied to automatic mandible segmentation in the head and neck region, and the corresponding references.

Methodology Categories	Publications	Number of Publications
SSM-based	[18,37–39]	4
ASM-based	[40–42]	3
AAM-based	[43,44]	2
Atlas-based	[45–62]	18
Level set-based	[9,63]	2
Classical machine learning-based	[19,36,64–76]	15
Deep learning-based	[5,24,32,77–106]	33

3.4.1. SSM-Based, ASM-Based and AAM-Based Methods

Statistical shape model-based (SSM-based) segmentation algorithms take advantage of the prior shape information to extract the structures of the objects. A dataset of object shapes is needed for training in this technique. Procrustes alignment is applied to align the landmark point set of the objects, which is placed at key features and/or along the

contour of the object shapes [107]. Principal component analysis (PCA) is employed to build a template shape of the objects [108].

The active shape model (ASM) [109] is the one of the commonly used image segmentation techniques based on the SSM approach. In 2006, Lamecker et al. [40] used a vanilla SSM-based segmentation method to extract the anatomical variability of developed mandible shapes in the training stage and match the statistical mandible shape to a given CT dataset via a deformable model approach in the testing stage. Kainmueller et al. [42] developed a segmentation method based on an SSM and a Dijkstra-based optimization for reconstructing the mandible including the course of the alveolar nerve. Albrecht et al. [41] adopted a multiatlas-based segmentation to obtain the coarse segmentation for the OARs and then employed the active shape model (ASM) to finely segment the OARs based on the results from coarse segmentation.

The active appearance model (AAM) [110] is an extension of SSM and ASM to further statistically model the texture information of the object. Two statistical models of shape and texture are further merged into an appearance model [110]. Babalola et al. [44] applied the AAM approach to automatic mandible and brainstem segmentation. Mannion-Haworth et al. [43] utilized the groupwise image registration method-based minimum description length approach [111] and AAM [110] built from manually segmented objects in CT images.

Moreover, several publications on SSM-based segmentation methods have been used in combination with other strategies to enhance the segmentation outcomes. Hybrid methods have demonstrated good results [18,37–39]. Kainmueller et al. [39] used SSM adaptation for mandible segmentation and then utilized the graph cut algorithm for fine segmentation in the MICCAI 2019 challenge [112]. Gollmer et al. [37] applied the SSM method with optimized correspondence for mandible segmentation. The authors established correspondence by optimizing a model-based cost function. Moreover, the authors introduced a relaxed SSM method for mandible segmentation [38]. Abdolali et al. [18] proposed a framework based on SSM and fast marching for automatic segmentation of the mandibular bone and canal. Moreover, the authors utilized low-rank decomposition for preprocessing. Table 3 presents the segmentation performance using SSM-based, ASM-based and AAM-based approaches.

3.4.2. Atlas-Based Methods

Atlas-based segmentation methods first utilize deformable registration approaches to register a known reference segmentation mask (that form an atlas) to a patient [50]. The registration optimization problem is usually solved by searching the deformation space. Then, this deformation is applied to contours made on the atlas to project the contours back to the patient-space.

In 2007, Zhang et al. [50] used an atlas-based image segmentation method to automatically segment ROIs in H&N CT images. Chuang et al. [45] presented an atlas-based semiautomatic segmentation for the mandible. The authors firstly cropped a minimum enclosing box of the mandible in a raw CT image followed by a user-determined global threshold to remove nonosseous tissue and then used the output from these registrations to transform the 3D mandible template model of the respective template scans into a mandible model that maps to the input test scan. In another paper [62], a semisupervised registration algorithm based on the atlas was developed to accurately segment the OARs with a real ground contour and all other coarse OARs in the atlas. The method concatenates rigid and deformable blocks, takes an atlas image, a group of atlas-space segmentations and a patient image as inputs and outputs the patient-space segmentation of all OARs defined on the atlas. Qazi et al. [51] combined atlas registration and organ-specific model-based segmentation in a common framework to segment various organs at risk in H&N CT images. The authors applied atlas registration and organ-specific model-based segmentation at a global level and then used a probabilistic refinement step to refine at the voxel level. Ayyalusamy et al. [54] applied an atlas-based method for OAR segmentation in H&N and

analyzed the dependence of atlas-based automatic segmentation for different levels on anatomy matching between sample patients and atlas patients.

In the case of mandible segmentation, the atlas-based method has become increasingly popular, so it has been frequently implemented in commercial or open-source tools. There are several articles using an atlas-based tool, for instance, Advanced Medical Imaging Registration Engine (ADMIRE) v1.05 (Elekta Software) used by [56], Smart Probabilistic Image Contouring Engine (SPICE) used by [57,58], and PLASTIMATCH MABS provided by [61]. Moreover, La Macchia et al. [60] quantitatively analyzed three different automatic atlas-based segmentation software offerings for adaptive radiotherapy in a small dataset.

In regard to the fusion strategy, Mencarelli et al. [46] built a knowledge base and thresholding technique for nine high-contrast structures (including the mandible) for an online or offline image-guided radiotherapy (RT) application. Then, they further used a hidden Markov model to identify the structures. Wang et al. [49] developed a mandible and maxilla segmentation approach in an H&N CBCB scan that estimated a patient-specific atlas applying a sparse label fusion strategy from predefined CT atlases and then converted it into a convex optimization problem using maximum a posteriori probability for mandible/maxilla segmentation. Gorthi et al. [52] developed a novel segmentation method that used active contour-joint registration and an atlas selection strategy.

A multiple atlas strategy or its modified version is also commonly used because the propagated label cannot be generalized in conventional atlas-based models. To address this issue, multiatlas approaches have been proposed. Chen et al. [47] presented a multiatlas-based method for multiple structure segmentation in CT scans. The authors registered CT images with the atlases at the global level so that structures of interest are aligned approximately in space. Based on that approach, the multiatlas-based segmentation can be performed at the local level. A method presented by Haq et al. [55] dynamically selected and weighted the proper number of atlases for weighted label fusion and predicted the segmentations and consensus maps, indicating voxel-wise agreement between different atlases that were selected from those exceeding an alignment weight called the dynamic atlas attention index. Alignment weights were computed at the image level (called global weighted voting) or at the structure level (called structure weighted voting). McCarroll et al. [59] studied an automatic segmentation strategy for H&N cancer to better provide a fully automated radiation treatment planning solution for low- and middle-income countries using an atlas-based deformable-image-registration algorithm. Han et al. [48,53] adopted a novel hierarchical multiatlas registration approach for OAR segmentation. Table 4 shows the summary of atlas-based method used.

3.4.3. Level Set-Based Methods

The level set-based image segmentation algorithm is modified from the snake algorithm [113]. It requires an initial contour curve, and then, the curve evolution is performed via minimizing the functional energy. To obtain a better prediction, the studies [9,63] make use of the atlas method to generate an initial contour in their segmentation tasks. Patch-based sparse representation has been applied to estimate the patient-specific atlas and then combined with a level set framework using the rule of maximizing a posteriori probability. With these methods, the mandible segmentation problem can be converted into a convex optimization problem [9]. Table 5 shows a summary of the level set-based method used.

Table 3. Summary of SSM-based, ASM-based and AAM-based methods.

Study	Year	Datasets	No. of Patients	Performance					Image Modalities	Time/Equipment	Category
				Dice (%)	ASD (mm)	HD (mm)	95HD (mm)	VOE (%)			
Abdolali [18]	2017	In-house	120	91.38 ± 2.06	0.71 ± 0.09	—	—	—	CBCT	5 min/CPU	SSM-based
Gollmer [37]	2012	In-house	30	—	0.50 ± 0.00	11.30 ± 3.50	—	14.0 ± 1.9	CT	—	SSM-based
Gollmer [38]	2013	In-house	30 + 6	—	0.40 ± 0.08	10.26 ± 3.16	—	11.0 ± 2.0	CT(train)/ CBCT(test)	—	SSM-based
Kainmueller [39]	2009	MICCAI 2009	18	88.40	—	8.40	—	—	CT	15 min/CPU	SSM-based
Lamecker [40]	2006	In-house	15	—	—	—	—	—	CT	—	ASM-based
Albrecht [41]	2015	PDDCA	40	88.13 ± 5.55	—	—	2.83 ± 1.18	—	CT	5 min/CPU	ASM-based
Kainmueller [42]	2009	In-house	106	—	0.50 ± 0.10	6.20 ± 2.30	—	—	CBCT	—	ASM-based
Mannion-Haworth [43]	2015	PDDCA	48	92.67 ± 1.00	—	—	1.98 ± 0.59	—	CT	30 min/CPU	AAM-based
Babalola [44]	2009	MICCAI 2009	18	76.10 ± 5.10 (exclude the 13th case)	—	—	—	—	CT	17 min/CPU	AAM-based

Table 4. Summary of atlas-based methods. ⁽¹⁾ Values estimated from figures; actual values not reported.

Study	Year	Datasets	No. of Patients	Performance						Image Modalities	Time/Equipment	Category
				Dice (%)	ASD (mm)	HD (mm)	95HD (mm)	VE (%)	Sen (%) / PPV (%)			
Chuang [45]	2019	In-house	54 + 20	97.60 ± 10.60	—	—	—	—	—	CT	3–8 h/CPU	Atlas-based
Mencarelli [46]	2014	In-house	188	—	—	—	—	—	—	CT	—	Atlas-based
Chen [47]	2015	PDDCA	40	91.70 ± 2.34	—	—	2.49 ± 0.76	—	—	CT	100 min/CPU	Atlas-based
Han [48]	2008	In-house	10	90.00	—	—	—	—	—	CT	1 h/CPU	Atlas-based
Wang [49]	2014	In-house	13 + 30	91.00 ± 2.00	0.61 ± 0.17	0.92 ± 0.47	—	—	—	CBCT + CT	—	Atlas-based
Zhang [50]	2007	In-house	7	80.00	0.85	—	—	—	—	CT	—	Atlas-based
Qazi [51]	2011	In-house	25	93.00	—	2.64	—	—	—	CT	12 min/CPU	Atlas-based
Gorthi [52]	2009	MICCAI 2009	18	77.80 ± 7.40	—	16.87 ± 6.75	—	—	—	CT	—	Atlas-based
Han [53]	2009	MICCAI 2009	18	90.33 ± 1.49	—	8.07 ± 3.12	—	—	—	CT	1 min/GPU	Atlas-based
Ayyalusamy [54]	2019	In-house	40	85.00 ⁽¹⁾	—	9 ⁽¹⁾	—	—	—	CT	—	Atlas-based
Haq [55]	2019	In-house PDDCA	45	85.00 ⁽¹⁾	—	—	3.00 ⁽¹⁾	—5.00 ⁽¹⁾	—	CT	—	Atlas-based
			32	83.00 ⁽¹⁾			3.00 ⁽¹⁾					
Liu [56]	2016	In-house	6	89.50	—	—	—	—	—	CT	10 min/—	Atlas-based
Zhu [57]	2013	In-house	32	89.00 ± 4.00	—	9.80 ± 4.10	—	—	—	CT	11.1 min/—	Atlas-based
Walker [58]	2014	In-house	40	98.00 ± 2.00	—	—	—	—	—	CT	19.7 min/—	Atlas-based
McCarroll [59]	2018	In-house	128	84.00 ± 7.00	1.89 ± 1.55	18.63 ± 14.90	—	—	—	CT	11.5 min/CPU	Atlas-based
La Macchia [60]	2012	In-house	5	89.00 ± 2.00	—	—	—	−4.76 ± 7.12	87.00 ± 5.00/ 92.00 ± 2.00	CT	10.6 min/CPU	Atlas-based
Zaffino [61]	2016	In-house	25	88.00 ± 7.00	—	—	—	—	—	CT	120 min/CPU	Atlas-based
Huang [62]	2019	In-house	500	84.50 ± 1.60	—	—	—	—	—	CT	—/GPU	Atlas-based

Table 5. Summary of level set-based methods.

Study	Year	Datasets	No. of Patients	Performance			Image Modalities	Time/Equipment	Category
				Dice (%)	ASD (mm)	HD (mm)			
Wang [9]	2014	In-house	15	92.00 ± 2.00	0.65 ± 0.19	0.96 ± 0.53	CBCT	5 h/CPU	Level set-based
Zhang [63]	2009	MICCAI 2009	18	87.93 ± 2.06	—	8.70 ± 3.28	CT	—	Level set-based

3.4.4. Classical Machine Learning-Based Methods

In the past few decades, the role of machine learning in medical applications has greatly increased. Classical machine learning (CML) methods include thresholding techniques, linear regression, support vector machines, random forests, etc. Many studies have used CML methods to segment the mandible, as listed in Table 6.

Normally, this technique is implemented in a hybrid way. A multiple threshold method by Otsu [114] was used to calculate proper thresholding values in the segmentation phase in the article from Barandiaran et al. [67]. After thresholding, region growing was applied so that many bone structures could be segmented, such as parts of the jaw bone or spine. The largest connected component of the segmented image volume can be simply selected for mandible segmentation [67]. Later, Wang et al. [65] made use of random forests to segment the mandible and maxilla in their expanded CBCT database and achieved better results than their previous method [9]. Linares et al. [19] proposed a semiautomatic two-stage segmentation method for CBCT. The authors first performed bone segmentation using a supervoxel technique and then implemented graph clustering with an interactive user-placed seed process that was used for mandible and skull segmentation. In this way, supervoxel methods can reduce the excessive number of voxels of CBCT 3D volume, leading to shorter processing time. Orbes-Arteaga et al. [68] introduced a new patch-based label fusion approach to weight the label votes using a generative probabilistic approach, in which local classifiers from the dictionary of atlas patches were generated for weighting the probability of the target. Wang et al. [69] integrated shape priors with a hierarchical learning model, in which they introduced 3 novel strategies: hierarchical critical model vertex identification, joint learning of shape and appearance, and hierarchical vertex regression. Qazi et al. [70] applied point-based registration and model-based segmentation and introduced a novel voxel classification approach to improve the results from model-based segmentation. Torosdagli et al. [71] adopted random forest regression to detect the mandible in 3D and then utilized a 3D gradient-based fuzzy connectedness image segmentation algorithm for final segmentation. Wu et al. [72] built population fuzzy anatomy models to find the hierarchical relationship of the objects. Then, the authors further utilized the anatomical information in the hierarchical relationship to locate the objects. Finally, the segmentation could be generated according to the location results. Tam et al. [36] presented support vector regression to predict all the boundary points of the organs of interest. Gacha et al. [75] performed histogram equalization and grayscale morphology for mandible segmentation. Then, image binarization was performed after grayscale method. Finally, 3D template matching was performed using cross-correlation to find a point inside the mandible followed by a geodesic dilation using the binary image as mask. Tong et al. [73] introduced new features into the automatic anatomy recognition approach from [66] in which they combined texture and intensity information into the recognition procedure. Wu et al. [74] developed a methodology, called AAR-RT, which extended their previous AAR framework [66] to RT planning in H&N. Spampinato et al. [76] applied a fully automatic mandible segmentation method to support orthodontists in assessing facial asymmetry, in which they used morphological operations for extraction of the mandible, and then, the connected components were computed and considered part of the mandible in each 2D slice.

Table 6. Summary of classical machine learning-based methods.

Study	Year	Datasets	No. of Patients	Performance						Image Modalities	Time/Equipment	Category
				Dice (%)	ASD (mm)	HD (mm)	Jac (%)	FPVF/FNVF (%)	MSE/RMSE (mm)			
Ji [64]	2013	In-house	12	97.90 ± 1.10	0.20 ± 0.13	—	95.80 ± 2.00	—	—	MRI	128 s/CPU	CML
Wang [65]	2015	In-house	30 60	94.00 ± 2.00 95.00 ± 2.00	0.42 ± 0.15 0.33 ± 0.11	0.74 ± 0.25 0.41 ± 0.20	—	—	—	CBCT CT	20 min/—	CML
Udupa [66]	2014	In-house	15	—	—	3.30 ± 0.56	—	1.00 ± 0.00/ 49.00 ± 8.00	—	MRI	54 s/CPU	CML
Linares [19]	2019	In-house	16	92.88	—	—	86.48	—	—	CBCT	5 min/CPU	CML
Barandiaran [67]	2009	In-house	12	—	—	—	—	—	—	CT	10 s/CPU	CML
Orbes-Arteaga [68]	2015	PDDCA	40	93.08 ± 2.36	—	—	—	—	—	CT	—	CML
Wang [69]	2016	PDDCA	48	94.40 ± 1.30	0.43 ± 0.12	—	—	—	—	CT	—/CPU	CML
Qazi [70]	2010	In-house	25	90.19	—	—	—	—	—	CT	3 min/CPU	CML
Torosdagli [71]	2017	PDDCA	40	91.00	—	<1.00	—	—	—	CT	—/CPU	CML
Wu [72]	2018	In-house	216	89.00	—	1.60	—	—	—	CT	—	CML
Tam [36]	2018	In-house	56	85.20 ± 5.30	—	—	—	—	0.10/3.16	CT	1 s/CPU	CML
Tong [73]	2018	In-house	246	—	—	—	—	—	—	CT	—	CML
Wu [74]	2019	In-house	216	89.00	—	—	—	—	—	CT	30 s/CPU	CML
Gacha [75]	2018	PDDCA	30	80.49	—	—	—	—	—	CT	—	CML
Spampinato [76]	2012	In-house	10	—	—	—	—	—	—	CT	—	CML

Along with the quick development of computer and MRI imaging technology, MRI images have also been used in mandible segmentation. Ji et al. [64] presented a two-stage rule-constrained seedless region growing approach for the mandible segmentation in MRI images. A 3D/2D seedless region growing approach is used to detect a trabecular bone and cortical bone of mandible after using a thresholding technique. The mandibular body was finally merged with some morphological processes. Udupa et al. [66] developed an automatic anatomy recognition methodology in H&N MRI scans that is based on fuzzy modeling ideas and tightly integrated fuzzy models with the iterative relative fuzzy connectedness delineation algorithm to locate and delineate organs in different body scans.

3.4.5. Deep Learning-Based Methods

In recent years, deep learning-based segmentation algorithms have become popular among researchers. With the development of deep learning technology, deep learning methods have shown a tremendous performance in the area of image segmentation. Deep learning techniques provide more flexibility and powerful capabilities than the traditional machine learning methods and require less expert analysis, facilitating extension to the other segmentation tasks [115]. For instance, the popular deep learning architectures, including UNet [13], SegNet [12], etc., are widely used for automatic image segmentation. Consequently, the deep learning approach has shown promising applications in the problem of mandible segmentation. Over the years, many studies have provided better solutions using deep learning architectures than other classical image segmentation methods, as shown in Tables 7–11.

Ibragimov et al. [77] proposed the first attempt of using the CNN-based concept to segment OARs in H&N CT scans. Furthermore, the authors smoothed the obtained segmentation results through a Markov random fields algorithm. To address the challenge of low performance on small low-contrast structures in CT, a new loss function, namely, the batch soft Dice loss function, was developed by Kodym et al. [78]. The authors used the new loss function in UNet and achieved better results than the methods using other loss functions. Yan et al. [79] developed a symmetric convolutional neural network to force convolution and deconvolution computation to be symmetric. Table 7 shows the summary of using deep learning-based methods.

Table 7. Summary of deep learning-based methods.

Study	Year	Datasets	No. of Patients	Performance				Image Modalities	Time/Equipment	Category
				Dice (%)	ASD (mm)	HD (mm)	IoU (%)			
Ibragimov [77]	2015	In-house	50	89.50 ± 3.60	—	—	—	CT	4 min/GPU	DL
Kodym [78]	2019	PDDCA	35	94.60 ± 0.70	0.29 ± 0.03	—	—	CT	—	DL
Yan [79]	2018	In-house	93	90.76 ± 2.45	—	—	85.44 ± 3.99	CT	—/GPU	DL
Xue [80]	2021	PDDCA	48	94.00 ± 2.00	0.49 ± 0.18	2.36 ± 0.62	—	CT	—/GPU	DL

Another intuitive method of 3D segmentation is to train a 3D network to process volume data directly. In this way, the 3D segmentation network can extract 3D structure information. AnatomyNet [82] is built in the 3D U-net architecture using residual blocks in encoding layers. Moreover, a new loss function combining the Dice score and focal loss is applied in the training process. Xia et al. [83] presented a methodology to automatically measure the masseter thickness and locate the ideal injection point for botulinum toxin into the masseter from a CT scan, in which a 3D UNet with a Resblock is used for the mandible and masseter segmentation. Willems et al. [84] applied the 3D segmentation network from [116]. The method consists of a 3D CNN and fully connected 3D CRF. Ren et al. [85] introduced multiscale patches as input for representing the center voxel and designed a 3D CNN network for segmentation of OARs. In addition, the authors interleaved the CNNs designated for the individual tissues since the neighboring tissues are often related on the anatomical side. With this strategy, the patch segmentation result

of a specific tissue can be refined by the other neighboring tissues. This study can deal with large tissues, while the small tissues are still not that promising in terms of Dice and 95HD. For contending with the imbalance between classes mainly caused by small organs, He et al. [86] proposed a combined UNet model that consists of 2D UNet, 3D UNet and 3D small UNet. The first model is a 2D UNet model, which has advantages in processing thick slice images. The second model is a 3D UNet model, which can cover most organs with the original resolution in the transverse plane via clipping the scans. The third model is a 3D small UNet model, which focuses on the segmentation of small organs, clipping from the boundary box of the 2D UNet model. Rhee et al. [87] proposed a CNN-based autocontouring tool that can be used to detect the errors in autocontours from a clinically validated atlas-based autocontouring tool. They used the DeepMind model [24] to detect the ill-defined contours from an atlas-based autocontouring tool. Nikolov et al. [24] applied a 3D UNet architecture for OAR segmentation and demonstrated that the 3D UNet approach achieved expert-level performance in delineating multiple organ segmentation in the H&N region. Tong et al. [88] proposed a novel deep learning-based approach based on a generative adversarial network (GAN) with a shape constraint (SC-GAN), where a DenseNet was used to predict the segmentation and a CNN-based discriminator network was applied to correct the predicted errors and image-level inconsistency between the predicted image and ground truth. Moreover, a shape representation loss was applied into the segmentation and adversarial loss function for reducing the false positives. Chan et al. [81] provided a lifelong learning protocol to improve the prediction accuracy of OAR segmentation, in which they used a multitask learning scheme coupled with transfer learning to accomplish this. In this way, the network transfers its shared knowledge to single tasks to help improve the generalizability of the network on limited datasets. To cope with the lack of overall shape and smoothness of OARs, Xue et al. [89] developed the feasibility of learning the signed distance map (SDM) directly from medical scans, in which SDM is usually calculated from object boundary contours and the binary segmentation map. To utilize more information in traditional segmentation training, the authors introduced an approximated Heaviside function to train the model by predicting SDMs and segmentation maps simultaneously. Table 8 lists the summary of literature that used 3D networks for mandible segmentation.

However, 3D networks require clipping or resampling the input data into small 3D patches to reduce GPU memory [92]. In addition, integration of 3D patches in post-processing is a time-consuming workflow. Therefore, to overcome the drawback of 3D networks, using a 2.5D or multiview training strategy instead of 3D is a potential strategy [5,90]. Qiu et al. [5] adopted 2.5D volume of CTs as input in UNet and then combined the resulting 2D segmentations from three orthogonal planes (axial, sagittal and coronal) into a 3D segmentation. This architecture can take into account the spatial information of adjacent slices in order to preserve the connectivity of anatomical structures. Lei et al. [90] first presented a segmental linear function to enhance the intensity of CT images, which makes the organs more separable than the existing simple window width method. Then, the authors proposed a novel 2.5D network for accurate OAR segmentation. Additionally, they applied a novel hardness-aware loss function to make the network give more attention to hard voxels. Liang et al. [91] developed a novel multiview (i.e., axial, coronal, and sagittal view) spatial aggregation framework for joint localization and segmentation of multiple OARs. Additionally, the authors proposed a region-of-interest-based fine-grained representation CNN for predicting the probability maps of OARs for each 2D view of CT images. In this way, the approach unifies the OAR localization and segmentation tasks and trains them in an end-to-end fashion. Qiu et al. [92] integrated a recurrent unit and a vanilla network, which enable the network to learn the continuity of neighborhood slices for the scans in the 2D architecture. Table 9 shows the summary of using 2.5D networks for mandible segmentation.

Table 8. Summary of 3D network strategies of deep learning-based methods.

Study	Year	Datasets	No. of Patients	Performance							Image Modalities	Time/Equipment	Category	
				Dice (%)	ASD (mm)	HD (mm)	95HD (mm)	Rec/Sen (%)	RMSE (mm)	PPV (%)				
Chan [81]	2019	In-house	200	91.00 ± 9.00	—	—	—	—	—	0.66 ± 0.31	—	CT	20 s/GPU	DL
Zhu [82]	2019	PDDCA + TCIA	48 + 223	92.30	—	—	—	—	—	—	—	CT	0.12 s/GPU	DL
Xia [83]	2019	In-house	53	94.60 ± 1.10	0.24 ± 0.034	—	—	95.60 ± 1.20	—	—	—	CT	1 s/GPU	DL
Willems [84]	2018	In-house	70	95.90	0.60	6.48	—	—	—	—	—	CT	1 min/GPU	DL
Ren [85]	2018	PDDCA	48	92.00;	—	—	1.89	—	—	—	—	CT	—/GPU	DL
He [86]	2020	StructSeg2019	50	90.30 (L); 90.80 (R)	—	—	—	—	—	—	—	CT	1 min/GPU	DL
Rhee [87]	2019	In-house + TCIA	1403 + 24	86.80 ± 3.30	—	12.80 ± 9.50	—	—	—	—	—	CT	2 min/GPU	DL
Nikolov [24]	2018	In-house	459	93.10 ± 1.90	—	—	—	—	—	—	—	CT	—/GPU	DL
		TCIA	30	92.90 ± 3.50	—	—	—	—	—	—	—	CT	—/GPU	DL
		PDDCA	15	93.80 ± 1.90	—	—	—	—	—	—	—	CT	—/GPU	DL
Tong [88]	2019	PDDCA	32	93.91 ± 1.30	0.55 ± 0.14	—	2.09 ± 0.63	91.25 ± 2.70	—	96.82 ± 1.70	CT	14 s/GPU	DL	
		In-house	25	81.64 ± 4.44	1.13 ± 0.48	—	2.72 ± 1.31	86.65 ± 6.00	—	78.50 ± 4.30	MRI	14 s/GPU	DL	
Xue [89]	2019	PDDCA	48	95.70 ± 1.80	—	—	0.60 ± 0.49	—	—	—	—	CT	—/GPU	DL

Table 9. Summary of 2.5D networks deep learning-based methods.

Study	Year	Datasets	No. of Patients	Performance				Image Modalities	Time/Equipment	Category	
				Dice (%)	ASD (mm)	95HD (mm)	RMSE (mm)				
Qiu [5]	2019	In-house	109	88.10	—	—	0.58	CT	2.5 min/GPU	DL	
		PDDCA	40	93.28 ± 1.44	—	1.43 ± 0.56	—				
Lei [90]	2021	StructSeg2019	50	91.10 ± 2.90 (L); 91.70 ± 1.50 (R)		—	2.81 ± 0.45 (L); 2.70 ± 0.40 (R)		CT	2 min/GPU	DL
		StructSeg2019 + PDDCA + In-house	50 + 48 + 67	90.00 ± 4.20		—	6.54 ± 19.14				
Liang [91]	2020	PDDCA	48	94.10 ± 0.70		0.28 ± 0.14		—	CT	—/GPU	DL
		In-house	96	91.10 ± 1.00 (L); 91.40 ± 2.00 (R)		0.76 ± 0.13 (L); 0.86 ± 0.14 (R)					
Qiu [92]	2020	In-house	109	97.53 ± 1.65		0.21 ± 0.26		—	CT	1.5 min/GPU	DL
		PDDCA	40	95.10 ± 1.21		0.14 ± 0.04					

The attention model has been widely used in various fields of deep learning in recent years [117–119]. The attention mechanism focuses limited attention on key information to save resources and obtain the most effective information quickly [118]. Gou et al. [93] designed a self-channel-and-spatial-attention neural network for H&N OAR segmentation, in which spatial and channelwise attention learning mechanisms can adaptively force the network to emphasize the meaningful features and weaken the irrelevant features at the same time. A novel simple local blockwise self-attention-based segmentation approach was presented by Jiang et al. [94]. This approach can more easily make the network learn the spatial location and interrelation within input images. This study demonstrated that adding the additional attention blocks increases the contextual field and captures focused attention from anatomical structures. Sun et al. [95] presented an end-to-end CNN, called AttentionAnatomy, which was trained with three partially annotated datasets to segment OARs from the whole body. AttentionAnatomy retains the basis structure of UNet and has two branches: a CT region classification path and an OAR segmentation path. The CT region classification path outputs region prediction and attention vectors for OARs, which represent the inference of possible combinations of OARs in the current image. Then, the OAR segmentation path applies these attention vectors to modulate the final output mask. The authors further proposed a recalibration mechanism to solve the partial annotation problem. Moreover, a hybrid loss function composed of batch Dice loss and spatially balanced focal loss is used to address the extreme category imbalance. Liu et al. [96] developed a cross-layer spatial attention map fusion network to combine different spatial attention maps and establish connections for significant features in the feature maps. Furthermore, the authors adopted a top-k exponential logarithmic Dice loss for the imbalanced dataset in their segmentation tasks. Table 10 shows the summary of using attention strategies for mandible segmentation.

Table 10. Summary of attention strategies of deep learning-based methods.

Study	Year	Datasets	No. of Patients	Performance					Image Modalities	Time/Equipment	Category
				Dice (%)	ASD (mm)	95HD (mm)	Rec/Sen (%)	PPV (%)			
Gou [93]	2020	PDDCA	48	94.00 ± 1.00	0.47 ± 0.11	1.40 ± 0.02	93.00 ± 2.00	95.00 ± 2.00	CT	2 s/GPU	DL
Jiang [94]	2019	In-house + PDDCA	48 + 48	93.00 ± 1.00 (trained in in-house dataset)	—	—	—	—	CT	0.1 s/GPU	DL
Sun [95]	2020	In-house	129	94.05 ± 1.08	—	—	—	—	CT	—/GPU	DL
Liu [96]	2020	StructSeg2019	50	90.28 (L); 90.81 (R)	—	—	—	—	CT	—/GPU	DL

There are several two-stage training strategies, such as cascaded CNNs [97–99,105], the combination [101,103] of Faster R-CNN [120] and CNNs. Zhang et al. [97] developed a cascaded CNN architecture for multitasks segmentation in H&N scans, in which a slice classification network is proposed to classify CT slices into the corresponding target categories. Then, the irrelevant slices without the targets were excluded. Next, the slices in the corresponding categories were pushed to a refined 3D segmentation network for target segmentation. In this case, the prediction performance was further improved with the help of the slice classification network. Tappeiner et al. [98] trained two hierarchical 3D neural networks to segment multiple OARs in H&N CT scans. The authors first implemented a coarse network on size-reduced scans for locating the OARs. Then, a subsequent fine network on the original resolution scans was trained for a final accurate segmentation based on the results from the coarse network. Mu et al. [99] applied two cascade networks for location and fine segmentation of organ segmentation, in which the network is modified based on the squeeze and exception module, the residual module and V-Net [121]. There are still some shortcomings in this research. The network input is accomplished by resampling the input images with the same voxel size. However, the difference of the thickness and the resolution within the slices is often too large. Therefore, the operation of resampling in the first network can easily cause localization failure when organs have small volumes. Wang et al. [100] proposed a framework similar to [99]. Tang et al. [32] presented a novel network U_a -Net which contains detection and segmentation stages. The approximate location of the OARs can be identified in the detection stage, and then, fine segmentation can be further performed utilizing the results from the detection stage as a guide in the fine stage. Lei et al. [101,102] adopted a 3D Faster R-CNN to locate the H&N OARs in MRI and CT images and then applied an attention UNet to segment the OARs based on the location results. Liang et al. [103] developed an automatic detection-segmentation network for OAR segmentation in H&N CT scans, in which the authors applied a Faster R-CNN [120] to locate the OARs and a fully convolutional network (FCN) [11] to further predict OAR masks based on the organ bounding boxes from the detection stage. Dijk et al. [104] applied a deep learning-based method from Mirada for OAR segmentation in H&N, in which the method was applied in the AAPM Challenge 2017 [122]. The method first used a CNN network to predict all OARs at a coarse resolution and then applied a series of organ-specific 10-layer networks for each organ segmentation using the predicted coarse results with full resolution images as input. Men et al. [105] developed a cascade CNN for the OAR delineation for radiotherapy in head-and-neck squamous cell carcinoma (HNSCC) of TCIA [33]. The cascaded CNN contains two stages (i.e., a coarse detection stage and a fine segmentation stage). The authors first used a shallow network as coarse detection to find the ROIs. In the fine segmentation stage, they used the coarse prediction as input for fine segmentation. Egger et al. [106] utilized a VGG network to classify the slices where the mandible appears and then employed an FCN network to delineate the mandible. Table 11 shows the summary of using two-stage strategies for mandible segmentation.

A new automatic segmentation algorithm was presented by Xue et al. [80] that employs a novel hybrid neural-like P system to alleviate the challenges in tasks of segmentation of H&N scans. The new P system has the common advantages of cell-like and neural-like P systems so that it can solve more practical problems in parallel. In the new P system, the effective integrated CNNs are implemented at the same time through different initializations to perform the pixel-level segmentation of OARs in order to obtain more effective features and take advantage of the strength of ensemble learning.

Table 11. Summary of two-stage strategies in deep learning-based methods.

Study	Year	Datasets	No. of Patients	Performance			Image Modalities	Time/Equipment	Category
				Dice (%)	HD (mm)	95HD (mm)			
Zhang [97]	2021	In-house	170	89.00 ± 2.00	—	1.66 ± 0.51	CT	40.1 s/GPU	DL
Tappeiner [98]	2019	PDDCA	40	91.00 ± 2.00	—	2.4 ± 0.6	CT	38.3 s/GPU	DL
Mu [99]	2020	In-house	50	89.80 ± 2.70 (L); 90.40 ± 2.00 (R)	—	—	CT	3 s/GPU	DL
Wang [100]	2018	PDDCA	48	93.00 ± 1.90	—	1.26 ± 0.50	CT	6 s/GPU	DL
Tang [32]	2019	In-house	175	93.12 ± 1.41	—	2.48 ± 0.83	CT	2 s/GPU	DL
		HNC+ HNPETCT	35 + 105	89.31 ± 11.59		3.05 ± 2.60			
		PDDCA	48	95.00 ± 0.80		—			
Lei [101]	2020	In-house	15	85.00 ± 4.00	—	—	MRI	—	DL
Lei [102]	2020	In-house	15	88.00 ± 3.00	—	—	CT	—	DL
Liang [103]	2019	In-house	185	91.40 ± 0.04 (L);	—	—	CT	30 s/GPU	DL
				91.20 ± 3.00 (R)					
Dijk [104]	2020	In-house	693	94.00 ± 1.00	—	1.30 ± 0.50	CT	—	DL
Men [105]	2019	HNSCC	100	92.00 ± 2.00	2.40 ± 0.40	—	CT	5.5 min/GPU	DL
Egger [106]	2018	In-house	20	89.64 ± 1.69	—	—	CT	—/CPU	DL

4. Discussion

Numerous studies have been published regarding mandible segmentation. Mandible segmentation methods collected in this paper are classified into seven categories. As shown in Figure 3a, atlas-based (23.4%), CML-based (19.5%) and DL-based (42.9%) methods are frequently used and developed in mandible segmentation. More than three-quarters of the publications (85.8%) utilized these methods. The atlas-based method is the oldest technique in the field of mandible segmentation and has been frequently implemented in commercial software [56–58,60,61]. Low accuracy is reported by using this kind of method due to the image artifacts and low contrast. Atlas construction is a time-consuming task when conducting complex nonrigid registration. An extended approach is to apply level set methods that help the initial curve move towards the boundary of the object [9,63]. The level set function implicitly represents the boundary of a curve or a shape through its zero horizontal tangent plane. The main drawback of using the level set-based methods is the requirement to initialize the contour. Thus, using the contours generated from atlas-based methods as initialization is helpful in this task [9,63]. The SSM-based method is another approach used by various researchers for mandible segmentation, in which the shape information of the object is taken into consideration in ASM. The method performed well compared to the atlas-based method. In addition to considering the shape information, AAM considers the texture information of the image and learns it in an integrated statistical shape model.

The most frequently used image segmentation approach is machine learning, especially deep learning (DL), which produces better results than the other methods. The classical machine learning algorithms demand a considerable human effort in feature extraction. The DL technique has been used frequently since 2018 due to the rapid development of DL techniques, as illustrated in Figure 3b. Deep learning methods have resolved the issue of automatically extracting the features. The major issue of using deep learning methods is the requirement of a substantial quantity of annotated medical data, which limits the use of supervised learning techniques and avoids overfitting [93]. Furthermore, to increase the size of training data, the deep learning-based methods are trained using data augmentation schemes such as scaling, cropping, affine transformation, elastic deformation, rotations, and noise contamination [81,82,93,98]. Moreover, with the evolution of deep learning, mandible segmentation has achieved remarkable performance compared to classical segmentation approaches.

As shown in Figure 3c, most of the publications (86.4%) used the CT imaging modality. Very few publications worked with CBCT (8.6%) and MRI (4.9%). Moreover, most of the publications worked with only one single imaging modality. There are 4 publications that worked with two imaging modalities. As shown in Figure 3c, number of publications using CT modality is much more than that of using CBCT and MRI techniques, which need more effort for manual expert annotation. However, the CBCT and MRI imaging techniques involve very low-dose radiation or no radiation. Mandible segmentation in CBCT and MRI is worthy of further exploration. The ability to derive the exact shape of the mandible from MRI's nonionizing radiation imaging mode provides clinicians with a value-added visualization option to view hard and soft tissues in a visual environment. Especially when high-resolution bone information is not needed, the spatial relationship between soft tissue muscle and the mandibular body can be obtained without additional volume imaging using CBCT. Patients who receive MR imaging but do not require facial CT imaging can benefit from this approach without the need for CBCT or X-ray radiation from CT imaging. This application eliminates the radiation risk to patients [3].

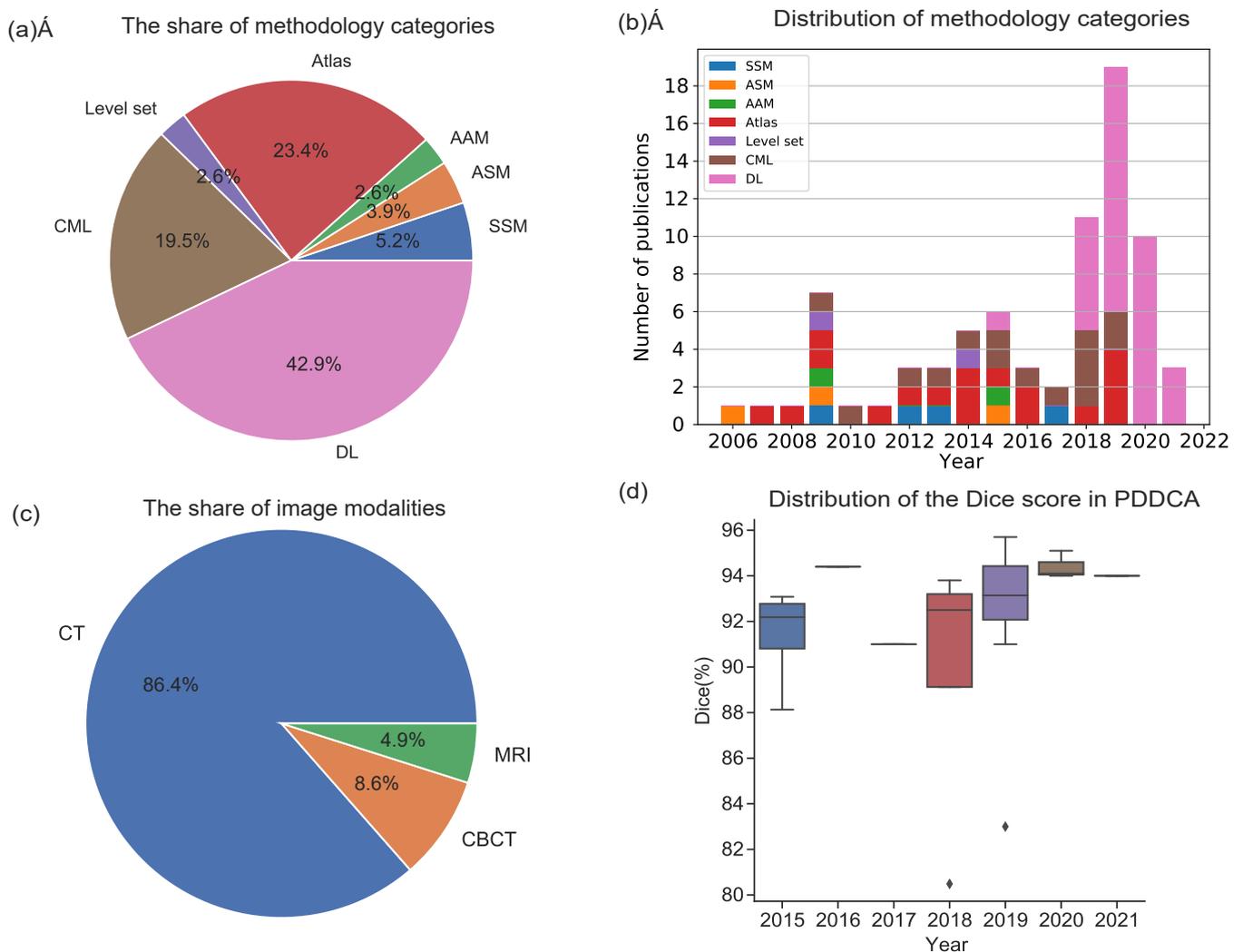


Figure 3. Analysis of publications pertaining to mandible segmentation. (a) The share of methodology categories used in the publications. (b) Distribution of the publications by year. (c) The share of image methodologies used in the publications. (d) Box plot of the distribution of the Dice score of the PDDCA test from the publications. The black lines in the middle of the box indicate the average Dice score.

The Head and Neck Auto Segmentation MICCAI Challenge (2015) [29], which used the PDDCA dataset in the challenge, has obtained much attention during the past years. There are 25 publications for mandible segmentation in PDDCA [29]. Among all the metrics summaries, the Dice score is the most popular metric used to evaluate the performance of the segmentation methods. The distribution regarding the publication with the PDDCA test is presented in Figure 3d. The Dice score increases with the year. Although it is a common metric used in image segmentation, it may not be the most relevant metric for clinical applications [82]. Identifying a new metric in consultation with the physicians practicing in the field would be an important next step for real clinical applications of the method [82].

5. Conclusions

Mandible segmentation is a challenging task due to the presence of metal artifacts, the low contrast of condyles and the large variation in mandibles among patients. CT modality is the most frequently used imaging technique for OMFS. To date, mandible segmentation methods proposed in the literature have achieved promising results in CT scans. Before the machine learning era, the atlas-based segmentation method was the most common method and performed well. With the advancement of technology, deep learning-based segmentation methods have increased substantially. Deep learning-based methods perform better and much faster than traditional methods, although they require more data and higher computing resources in training. There is still room for improvement since the datasets are limited and cannot fully represent the general patient population in clinic. Furthermore, mandible segmentation in CBCT and MRI is valuable for further exploration due to their low-dose radiation or no radiation. Therefore, mandible segmentation is still an open research area to be improved.

Author Contributions: Conceptualization, all authors; methodology, B.Q. and J.G.; validation, B.Q. and J.G.; formal analysis, B.Q.; writing—original draft preparation, B.Q., H.v.d.W. and J.G.; writing—review and editing, all authors; visualization, B.Q.; supervision, all authors; project administration, J.K., M.J.H.W. and P.M.A.v.O. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All relevant data are included in the study.

Acknowledgments: The author is supported by a joint Ph.D. fellowship from China Scholarship Council.

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

References

1. Shah, J.P.; Gil, Z. Current concepts in management of oral cancer—surgery. *Oral Oncol.* **2009**, *45*, 394–401. [[CrossRef](#)]
2. Ciocca, L.; Mazzoni, S.; Fantini, M.; Persiani, F.; Marchetti, C.; Scotti, R. CAD/CAM guided secondary mandibular reconstruction of a discontinuity defect after ablative cancer surgery. *J. Cranio-Maxillofac. Surg.* **2012**, *40*, e511–e515. [[CrossRef](#)] [[PubMed](#)]
3. Kraeima, J. Three Dimensional Virtual Surgical Planning for Patient Specific Osteosynthesis and Devices in Oral and Maxillofacial Surgery. A New Era. Ph.D. Thesis, University of Groningen, Groningen, The Netherlands, 2019.
4. Kraeima, J.; Dorgelo, B.; Gulbitti, H.; Steenbakkens, R.; Schepman, K.; Roodenburg, J.; Spijkervet, F.; Schepers, R.; Witjes, M. Multi-modality 3D mandibular resection planning in head and neck cancer using CT and MRI data fusion: A clinical series. *Oral Oncol.* **2018**, *81*, 22–28. [[CrossRef](#)]
5. Qiu, B.; Guo, J.; Kraeima, J.; Glas, H.H.; Borra, R.J.; Witjes, M.J.; van Ooijen, P.M. Automatic segmentation of the mandible from computed tomography scans for 3D virtual surgical planning using the convolutional neural network. *Phys. Med. Biol.* **2019**, *64*, 175020. [[CrossRef](#)] [[PubMed](#)]
6. Metzler, P.; Geiger, E.J.; Alcon, A.; Ma, X.; Steinbacher, D.M. Three-dimensional virtual surgery accuracy for free fibula mandibular reconstruction: Planned versus actual results. *J. Oral Maxillofac. Surg.* **2014**, *72*, 2601–2612. [[CrossRef](#)] [[PubMed](#)]
7. Huff, T.J.; Ludwig, P.E.; Zuniga, J.M. The potential for machine learning algorithms to improve and reduce the cost of 3-dimensional printing for surgical planning. *Expert Rev. Med. Devices* **2018**, *15*, 349–356. [[CrossRef](#)]

8. Wallner, J.; Hohegger, K.; Chen, X.; Mischak, I.; Reinbacher, K.; Pau, M.; Zrnc, T.; Schwenzler-Zimmerer, K.; Zemann, W.; Schmalstieg, D.; et al. Clinical evaluation of semi-automatic open-source algorithmic software segmentation of the mandibular bone: Practical feasibility and assessment of a new course of action. *PLoS ONE* **2018**, *13*, e0196378. [[CrossRef](#)] [[PubMed](#)]
9. Wang, L.; Chen, K.C.; Gao, Y.; Shi, F.; Liao, S.; Li, G.; Shen, S.G.; Yan, J.; Lee, P.K.; Chow, B.; et al. Automated bone segmentation from dental CBCT images using patch-based sparse representation and convex optimization. *Med. Phys.* **2014**, *41*, 043503. [[CrossRef](#)]
10. Heimann, T.; Meinzer, H.P. Statistical shape models for 3D medical image segmentation: A review. *Med. Image Anal.* **2009**, *13*, 543–563. [[CrossRef](#)]
11. Long, J.; Shelhamer, E.; Darrell, T. Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 3431–3440.
12. Badrinarayanan, V.; Kendall, A.; Cipolla, R. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 2481–2495. [[CrossRef](#)]
13. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*; Springer: Cham, Switzerland, 2015; pp. 234–241.
14. Yu, F.; Koltun, V. Multi-scale context aggregation by dilated convolutions. *arXiv* **2015**, arXiv:1511.07122.
15. Garcia-Garcia, A.; Orts-Escolano, S.; Oprea, S.; Villena-Martinez, V.; Garcia-Rodriguez, J. A review on deep learning techniques applied to semantic segmentation. *arXiv* **2017**, arXiv:1704.06857.
16. Vaitiekūnas, M.; Jegelevičius, D.; Sakalauskas, A.; Grybauskas, S. Automatic Method for Bone Segmentation in Cone Beam Computed Tomography Data Set. *Appl. Sci.* **2020**, *10*, 236. [[CrossRef](#)]
17. Hirschinger, V.; Hanke, S.; Hirschfelder, U.; Hofmann, E. Artifacts in orthodontic bracket systems in cone-beam computed tomography and multislice computed tomography. *J. Orofac. Orthop.* **2015**, *76*, 152–163. [[CrossRef](#)]
18. Abdolali, F.; Zoroofi, R.A.; Abdolali, M.; Yokota, F.; Otake, Y.; Sato, Y. Automatic segmentation of mandibular canal in cone beam CT images using conditional statistical shape model and fast marching. *Int. J. Comput. Assist. Radiol. Surg.* **2017**, *12*, 581–593. [[CrossRef](#)]
19. Linares, O.C.; Bianchi, J.; Raveli, D.; Neto, J.B.; Hamann, B. Mandible and skull segmentation in cone beam computed tomography using super-voxels and graph clustering. *Vis. Comput.* **2019**, *35*, 1461–1474.
20. Vovk, U.; Pernus, F.; Likar, B. A review of methods for correction of intensity inhomogeneity in MRI. *IEEE Trans. Med. Imaging* **2007**, *26*, 405–421. [[CrossRef](#)]
21. Bag, A.K.; Gaddikeri, S.; Singhal, A.; Hardin, S.; Tran, B.D.; Medina, J.A.; Curé, J.K. Imaging of the temporomandibular joint: An update. *World J. Radiol.* **2014**, *6*, 567. [[CrossRef](#)]
22. Wallner, J.; Mischak, I.; Egger, J. Computed tomography data collection of the complete human mandible and valid clinical ground truth models. *Sci. Data* **2019**, *6*, 190003. [[CrossRef](#)] [[PubMed](#)]
23. Ouzzani, M.; Hammady, H.; Fedorowicz, Z.; Elmagarmid, A. Rayyan—A web and mobile app for systematic reviews. *Syst. Rev.* **2016**, *5*, 210. [[CrossRef](#)] [[PubMed](#)]
24. Nikolov, S.; Blackwell, S.; Mendes, R.; De Fauw, J.; Meyer, C.; Hughes, C.; Askham, H.; Romera-Paredes, B.; Karthikesalingam, A.; Chu, C.; et al. Deep learning to achieve clinically applicable segmentation of head and neck anatomy for radiotherapy. *arXiv* **2018**, arXiv:1809.04430.
25. Fourie, Z.; Damstra, J.; Schepers, R.H.; Gerrits, P.O.; Ren, Y. Segmentation process significantly influences the accuracy of 3D surface models derived from cone beam computed tomography. *Eur. J. Radiol.* **2012**, *81*, e524–e530. [[CrossRef](#)]
26. Branstetter, I.V.B.F.; Blodgett, T.M.; Zimmer, L.A.; Snyderman, C.H.; Johnson, J.T.; Raman, S.; Meltzer, C.C. Head and neck malignancy: Is PET/CT more accurate than PET or CT alone? *Radiology* **2005**, *235*, 580–586. [[CrossRef](#)]
27. Comelli, A. Fully 3D Active Surface with Machine Learning for PET Image Segmentation. *J. Imaging* **2020**, *6*, 113. [[CrossRef](#)]
28. Comelli, A.; Bignardi, S.; Stefano, A.; Russo, G.; Sabini, M.G.; Ippolito, M.; Yezzi, A. Development of a new fully three-dimensional methodology for tumours delineation in functional images. *Comput. Biol. Med.* **2020**, *120*, 103701. [[CrossRef](#)] [[PubMed](#)]
29. Raudaschl, P.F.; Zaffino, P.; Sharp, G.C.; Spadea, M.F.; Chen, A.; Dawant, B.M.; Albrecht, T.; Gass, T.; Langguth, C.; Lüthi, M.; et al. Evaluation of segmentation methods on head and neck CT: Auto-segmentation challenge 2015. *Med. Phys.* **2017**, *44*, 2020–2036. [[CrossRef](#)] [[PubMed](#)]
30. Bosch, W.; Straube, W.; Matthews, J.; Purdy, J. Head-Neck Cetuximab. The Cancer Imaging Archive. 2015. Available online: <https://wiki.cancerimagingarchive.net/display/Public/Head-Neck+Cetuximab> (accessed on 23 February 2021).
31. Clark, K.; Vendt, B.; Smith, K.; Freymann, J.; Kirby, J.; Koppel, P.; Moore, S.; Phillips, S.; Maffitt, D.; Pringle, M.; et al. The Cancer Imaging Archive (TCIA): Maintaining and operating a public information repository. *J. Digit. Imaging* **2013**, *26*, 1045–1057. [[CrossRef](#)] [[PubMed](#)]
32. Tang, H.; Chen, X.; Liu, Y.; Lu, Z.; You, J.; Yang, M.; Yao, S.; Zhao, G.; Xu, Y.; Chen, T.; et al. Clinically applicable deep learning framework for organs at risk delineation in CT images. *Nat. Mach. Intell.* **2019**, *1*, 480–491. [[CrossRef](#)]
33. Vallieres, M.; Kay-Rivest, E.; Perrin, L.J.; Liem, X.; Furstoss, C.; Aerts, H.J.; Khaouam, N.; Nguyen-Tan, P.F.; Wang, C.S.; Sultanem, K.; et al. Radiomics strategies for risk assessment of tumour failure in head-and-neck cancer. *Sci. Rep.* **2017**, *7*, 10117. [[CrossRef](#)]
34. Zuley, M.L.; Jarosz, R.; Kirk, S.; Lee, Y.; Colen, R.; Garcia, K.; Aredes, N. Radiology data from the cancer genome atlas head-neck squamous cell carcinoma [TCGA-HNSC] collection. *Cancer Imaging Arch.* **2016**, *10*, K9.

35. Fan, X.; Yang, J.; Cheng, L. A novel segmentation method for MR brain images based on fuzzy connectedness and FCM. In *International Conference on Fuzzy Systems and Knowledge Discovery*; Springer: Berlin/Heidelberg, Germany, 2005; pp. 505–513.
36. Tam, C.; Yang, X.; Tian, S.; Jiang, X.; Beitler, J.; Li, S. Automated delineation of organs-at-risk in head and neck ct images using multi-output support vector regression. In *Medical Imaging 2018: Biomedical Applications in Molecular, Structural, and Functional Imaging*; International Society for Optics and Photonics: Bellingham, WA, USA, 2018; Volume 10578, p. 1057824.
37. Gollmer, S.T.; Buzug, T.M. Fully automatic shape constrained mandible segmentation from cone-beam CT data. In Proceedings of the 2012 9th IEEE International Symposium on Biomedical Imaging (ISBI), Barcelona, Spain, 2–5 May 2012; pp. 1272–1275.
38. Gollmer, S.T.; Buzug, T.M. Relaxed Statistical Shape Models for 3D Image Segmentation—Application to Mandible Bone in Cone-beam CT Data. *Curr. Med. Imaging* **2013**, *9*, 129–137. [[CrossRef](#)]
39. Kainmueller, D.; Lamecker, H.; Seim, H.; Zachow, S. Multi-object segmentation of head bones. Presented at the Head and Neck Auto-Segmentation Challenge (MICCAI), London, UK, 20–24 September 2009.
40. Lamecker, H.; Zachow, S.; Wittmers, A.; Weber, B.; Hege, H.; Isholtz, B.; Stiller, M. Automatic segmentation of mandibles in low-dose CT-data. *Int. J. Comput. Assist. Radiol. Surg.* **2006**, *1*, 393.
41. Albrecht, T.; Gass, T.; Langguth, C.; Lüthi, M. Multi atlas segmentation with active shape model refinement for multi-organ segmentation in head and neck cancer radiotherapy planning. In Proceedings of the Head Neck Auto-Segmentation Challenge (MICCAI), Munich, Germany, 5–9 October 2015.
42. Kainmueller, D.; Lamecker, H.; Seim, H.; Zinser, M.; Zachow, S. Automatic extraction of mandibular nerve and bone from cone-beam CT data. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*; Springer: Berlin/Heidelberg, Germany, 2009; pp. 76–83.
43. Mannion-Haworth, R.; Bowes, M.; Ashman, A.; Guillard, G.; Brett, A.; Vincent, G. Fully automatic segmentation of head and neck organs using active appearance models. In Proceedings of the Head Neck Auto-Segmentation Challenge (MICCAI), Munich, Germany, 5–9 October 2015.
44. Babalola, K.; Cootes, T. AAM Segmentation of the Mandible and Brainstem. Presented at the Head and Neck Auto-Segmentation Challenge (MICCAI), London, UK, 20–24 September 2009.
45. Chuang, Y.J.; Doherty, B.M.; Adluru, N.; Chung, M.K.; Vorperian, H.K. A novel registration-based semi-automatic mandible segmentation pipeline using computed tomography images to study mandibular development. *J. Comput. Assist. Tomogr.* **2018**, *42*, 306. [[CrossRef](#)] [[PubMed](#)]
46. Mencarelli, A.; Van Beek, S.; Zijp, L.; Rasch, C.; Van Herk, M.; Sonke, J. Automatic detection system for multiple region of interest registration to account for posture changes in head and neck radiotherapy. *Phys. Med. Biol.* **2014**, *59*, 2005. [[CrossRef](#)] [[PubMed](#)]
47. Chen, A.; Dawant, B. A multi-atlas approach for the automatic segmentation of multiple structures in head and neck CT images. In Proceedings of the Head Neck Auto-Segmentation Challenge (MICCAI), Munich, Germany, 5–9 October 2015.
48. Han, X.; Hoogeman, M.S.; Levendag, P.C.; Hibbard, L.S.; Teguh, D.N.; Voet, P.; Cowen, A.C.; Wolf, T.K. Atlas-based auto-segmentation of head and neck CT images. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*; Springer: Berlin/Heidelberg, Germany, 2008; pp. 434–441.
49. Wang, L.; Chen, K.C.; Shi, F.; Liao, S.; Li, G.; Gao, Y.; Shen, S.G.; Yan, J.; Lee, P.K.; Chow, B.; et al. Automated segmentation of CBCT image using spiral CT atlases and convex optimization. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*; Springer: Berlin/Heidelberg, Germany, 2013; pp. 251–258.
50. Zhang, T.; Chi, Y.; Meldolesi, E.; Yan, D. Automatic delineation of on-line head-and-neck computed tomography images: Toward on-line adaptive radiotherapy. *Int. J. Radiat. Oncol. Biol. Phys.* **2007**, *68*, 522–530. [[CrossRef](#)]
51. Qazi, A.A.; Pekar, V.; Kim, J.; Xie, J.; Breen, S.L.; Jaffray, D.A. Auto-segmentation of normal and target structures in head and neck CT images: A feature-driven model-based approach. *Med. Phys.* **2011**, *38*, 6160–6170. [[CrossRef](#)]
52. Gorthi, S.; Duay, V.; Bach Cuadra, M.; Tercier, P.A.; Allal, A.S.; Thiran, J.P. Active Contour-Based Segmentation of Head and Neck with Adaptive Atlas Selection. Presented at the Head and Neck Auto-Segmentation Challenge (MICCAI), London, UK, 20–24 September 2009.
53. Han, X.; Hibbard, L.; O’Connell, N.; Willcut, V. Automatic segmentation of head and neck CT images by GPU-accelerated multi-atlas fusion. Presented at the Head and Neck Auto-Segmentation Challenge (MICCAI), London, UK, 20–24 September 2009.
54. Ayyalusamy, A.; Vellaiyan, S.; Subramanian, S.; Ilamurugu, A.; Satpathy, S.; Nauman, M.; Katta, G.; Madineni, A. Auto-segmentation of head and neck organs at risk in radiotherapy and its dependence on anatomic similarity. *Radiat. Oncol. J.* **2019**, *37*, 134. [[CrossRef](#)]
55. Haq, R.; Berry, S.L.; Deasy, J.O.; Hunt, M.; Veeraraghavan, H. Dynamic multiatlas selection-based consensus segmentation of head and neck structures from CT images. *Med. Phys.* **2019**, *46*, 5612–5622. [[CrossRef](#)]
56. Liu, Q.; Qin, A.; Liang, J.; Yan, D. Evaluation of atlas-based auto-segmentation and deformable propagation of organs-at-risk for head-and-neck adaptive radiotherapy. *Recent Patents Top Imaging* **2016**, *5*, 79–87. [[CrossRef](#)]
57. Zhu, M.; Bzdusek, K.; Brink, C.; Eriksen, J.G.; Hansen, O.; Jensen, H.A.; Gay, H.A.; Thorstad, W.; Widder, J.; Brouwer, C.L.; et al. Multi-institutional quantitative evaluation and clinical validation of Smart Probabilistic Image Contouring Engine (SPICE) autosegmentation of target structures and normal tissues on computer tomography images in the head and neck, thorax, liver, and male pelvis areas. *Int. J. Radiat. Oncol. Biol. Phys.* **2013**, *87*, 809–816.

58. Walker, G.V.; Awan, M.; Tao, R.; Koay, E.J.; Boehling, N.S.; Grant, J.D.; Sittig, D.F.; Gunn, G.B.; Garden, A.S.; Phan, J.; et al. Prospective randomized double-blind study of atlas-based organ-at-risk autosegmentation-assisted radiation planning in head and neck cancer. *Radiother. Oncol.* **2014**, *112*, 321–325. [[CrossRef](#)] [[PubMed](#)]
59. McCarroll, R.E.; Beadle, B.M.; Balter, P.A.; Burger, H.; Cardenas, C.E.; Dalvie, S.; Followill, D.S.; Kisling, K.D.; Mejia, M.; Naidoo, K.; et al. Retrospective validation and clinical implementation of automated contouring of organs at risk in the head and neck: A step toward automated radiation treatment planning for low-and middle-income countries. *J. Glob. Oncol.* **2018**, *4*, 1–11. [[CrossRef](#)]
60. La Macchia, M.; Fellin, F.; Amichetti, M.; Cianchetti, M.; Gianolini, S.; Paola, V.; Lomax, A.J.; Widesott, L. Systematic evaluation of three different commercial software solutions for automatic segmentation for adaptive therapy in head-and-neck, prostate and pleural cancer. *Radiat. Oncol.* **2012**, *7*, 160. [[CrossRef](#)] [[PubMed](#)]
61. Zaffino, P.; Raudaschl, P.; Fritscher, K.; Sharp, G.C.; Spadea, M.F. Plastimatch MABS, an open source tool for automatic image segmentation. *Med. Phys.* **2016**, *43*, 5155–5160. [[CrossRef](#)] [[PubMed](#)]
62. Huang, C.; Badiei, M.; Seo, H.; Ma, M.; Liang, X.; Capaldi, D.; Gensheimer, M.; Xing, L. Atlas Based Segmentations via Semi-Supervised Diffeomorphic Registrations. *arXiv* **2019**, arXiv:1911.10417.
63. Zhang, X.; Tian, J.; Wu, Y.; Zheng, J.; Deng, K. Segmentation of head and neck CT scans using atlas-based level set method. Presented at the Head and Neck Auto-Segmentation Challenge (MICCAI), London, UK, 20–24 September 2009.
64. Ji, D.X.; Foong, K.W.C.; Ong, S.H. A two-stage rule-constrained seedless region growing approach for mandibular body segmentation in MRI. *Int. J. Comput. Assist. Radiol. Surg.* **2013**, *8*, 723–732. [[CrossRef](#)]
65. Wang, L.; Gao, Y.; Shi, F.; Li, G.; Chen, K.C.; Tang, Z.; Xia, J.J.; Shen, D. Automated segmentation of dental CBCT image with prior-guided sequential random forests. *Med. Phys.* **2016**, *43*, 336–346. [[CrossRef](#)]
66. Udupa, J.K.; Odhner, D.; Zhao, L.; Tong, Y.; Matsumoto, M.M.; Ciesielski, K.C.; Falcao, A.X.; Vaideeswaran, P.; Ciesielski, V.; Saboury, B.; et al. Body-wide hierarchical fuzzy modeling, recognition, and delineation of anatomy in medical images. *Med. Image Anal.* **2014**, *18*, 752–771. [[CrossRef](#)]
67. 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 *International Conference on Intelligent Data Engineering and Automated Learning*; Springer: Berlin/Heidelberg, Germany, 2009; pp. 649–655.
68. Orbes-Arteaga, M.; Pea, D.; Dominguez, G. Head and neck auto segmentation challenge based on non-local generative models. In Proceedings of the Head Neck Auto-Segmentation Challenge (MICCAI), Munich, Germany, 5–9 October 2015.
69. Wang, Z.; Wei, L.; Wang, L.; Gao, Y.; Chen, W.; Shen, D. Hierarchical vertex regression-based segmentation of head and neck CT images for radiotherapy planning. *IEEE Trans. Image Process.* **2018**, *27*, 923–937. [[CrossRef](#)]
70. Qazi, A.A.; Kim, J.J.; Jaffray, D.A.; Pekar, V. Probabilistic refinement of model-based segmentation: Application to radiation therapy planning of the head and neck. In *International Workshop on Medical Imaging and Virtual Reality*; Springer: Berlin/Heidelberg, Germany, 2010; pp. 403–410.
71. Torosdagli, N.; Liberton, D.K.; Verma, P.; Sincan, M.; Lee, J.; Pattanaik, S.; Bagci, U. Robust and fully automated segmentation of mandible from CT scans. In Proceedings of the 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017), Melbourne, Australia, 18–21 April 2017; pp. 1209–1212.
72. Wu, X.; Udupa, J.K.; Tong, Y.; Odhner, D.; Pednekar, G.V.; Simone, C.B.; McLaughlin, D.; Apinorasethkul, C.; Lukens, J.; Mihailidis, D.; et al. Auto-contouring via automatic anatomy recognition of organs at risk in head and neck cancer on CT images. In *Medical Imaging 2018: Image-Guided Procedures, Robotic Interventions, and Modeling*; International Society for Optics and Photonics: Bellingham, WA, USA, 2018; Volume 10576, p. 1057617.
73. Tong, Y.; Udupa, J.K.; Wu, X.; Odhner, D.; Pednekar, G.; Simone, C.B.; McLaughlin, D.; Apinorasethkul, C.; Shammo, G.; James, P.; et al. Hierarchical model-based object localization for auto-contouring in head and neck radiation therapy planning. In *Medical Imaging 2018: Biomedical Applications in Molecular, Structural, and Functional Imaging*; International Society for Optics and Photonics: Bellingham, WA, USA, 2018; Volume 10578, p. 1057822.
74. Wu, X.; Udupa, J.K.; Tong, Y.; Odhner, D.; Pednekar, G.V.; Simone, C.B., II; McLaughlin, D.; Apinorasethkul, C.; Apinorasethkul, O.; Lukens, J.; et al. AAR-RT—A system for auto-contouring organs at risk on CT images for radiation therapy planning: Principles, design, and large-scale evaluation on head-and-neck and thoracic cancer cases. *Med. Image Anal.* **2019**, *54*, 45–62.
75. Gacha, S.J.; León, S.A.G. Segmentation of Mandibles in Computer Tomography Volumes of Patients with Foam Cells Carcinoma. In Proceedings of the 2018 IX International Seminar of Biomedical Engineering (SIB), Bogota, Colombia, 16–18 May 2018; pp. 1–7.
76. Spampinato, C.; Pino, C.; Giordano, D.; Leonardi, R. Automatic 3D segmentation of mandible for assessment of facial asymmetry. In Proceedings of the 2012 IEEE International Symposium on Medical Measurements and Applications Proceedings, Budapest, Hungary, 18–19 May 2012; pp. 1–4.
77. Ibragimov, B.; Xing, L. Segmentation of organs-at-risks in head and neck CT images using convolutional neural networks. *Med. Phys.* **2017**, *44*, 547–557. [[CrossRef](#)] [[PubMed](#)]
78. Kodym, O.; Španěl, M.; Herout, A. Segmentation of Head and Neck Organs at Risk Using CNN with Batch Dice Loss. *arXiv* **2018**, arXiv:1812.02427.
79. Yan, M.; Guo, J.; Tian, W.; Yi, Z. Symmetric Convolutional Neural Network for Mandible Segmentation. *Knowl. Based Syst.* **2018**, *159*, 63–71. [[CrossRef](#)]

80. Xue, J.; Wang, Y.; Kong, D.; Wu, F.; Yin, A.; Qu, J.; Liu, X. Deep hybrid neural-like P systems for multiorgan segmentation in head and neck CT/MR images. *Expert Syst. Appl.* **2021**, *168*, 114446. [[CrossRef](#)]
81. Chan, J.W.; Kearney, V.; Haaf, S.; Wu, S.; Bogdanov, M.; Reddick, M.; Dixit, N.; Sudhyadhom, A.; Chen, J.; Yom, S.S.; et al. A convolutional neural network algorithm for automatic segmentation of head and neck organs at risk using deep lifelong learning. *Med. Phys.* **2019**, *46*, 2204–2213. [[CrossRef](#)] [[PubMed](#)]
82. Zhu, W.; Huang, Y.; Tang, H.; Qian, Z.; Du, N.; Fan, W.; Xie, X. AnatomyNet: Deep 3D Squeeze-and-excitation U-Nets for fast and fully automated whole-volume anatomical segmentation. *arXiv* **2018**, arXiv:1808.05238.
83. Xia, W.J.; Han, W.Q.; Zhang, Y.; Wang, L.; Chai, G.; Huang, Y.J. Automatic Masseter Thickness Measurement and Ideal Point Localization for Botulinum Toxin Injection. In Proceedings of the 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Berlin, Germany, 23–27 July 2019; pp. 2801–2804.
84. Willems, S.; Crijs, W.; Saint-Estevan, A.L.G.; Van Der Veen, J.; Robben, D.; Depuydt, T.; Nuyts, S.; Haustermans, K.; Maes, F. Clinical implementation of DeepVoxNet for auto-delineation of organs at risk in head and neck cancer patients in radiotherapy. In *OR 2.0 Context-Aware Operating Theaters, Computer Assisted Robotic Endoscopy, Clinical Image-Based Procedures, and Skin Image Analysis*; Springer: Berlin/Heidelberg, Germany, 2018; pp. 223–232.
85. Ren, X.; Xiang, L.; Nie, D.; Shao, Y.; Zhang, H.; Shen, D.; Wang, Q. Interleaved 3D-CNN s for joint segmentation of small-volume structures in head and neck CT images. *Med. Phys.* **2018**, *45*, 2063–2075. [[CrossRef](#)] [[PubMed](#)]
86. He, B.; Jia, F. Head and Neck CT Segmentation Based on a Combined U-Net Model. *J. Integr. Technol.* **2020**, *9*, 17–24.
87. Rhee, D.J.; Cardenas, C.E.; Elhalawani, H.; McCarrroll, R.; Zhang, L.; Yang, J.; Garden, A.S.; Peterson, C.B.; Beadle, B.M.; Court, L.E. Automatic detection of contouring errors using convolutional neural networks. *Med. Phys.* **2019**, *46*, 5086–5097. [[CrossRef](#)]
88. Tong, N.; Gou, S.; Yang, S.; Cao, M.; Sheng, K. Shape constrained fully convolutional DenseNet with adversarial training for multiorgan segmentation on head and neck CT and low-field MR images. *Med. Phys.* **2019**, *46*, 2669–2682. [[CrossRef](#)]
89. Xue, Y.; Tang, H.; Qiao, Z.; Gong, G.; Yin, Y.; Qian, Z.; Huang, C.; Fan, W.; Huang, X. Shape-aware organ segmentation by predicting signed distance maps. In Proceedings of the AAAI Conference on Artificial Intelligence, New York, USA, 7–12 February 2020; Volume 34, pp. 12565–12572.
90. Lei, W.; Mei, H.; Sun, Z.; Ye, S.; Gu, R.; Wang, H.; Huang, R.; Zhang, S.; Zhang, S.; Wang, G. Automatic Segmentation of Organs-at-Risk from Head-and-Neck CT using Separable Convolutional Neural Network with Hard-Region-Weighted Loss. *arXiv* **2021**, arXiv:2102.01897.
91. Liang, S.; Thung, K.H.; Nie, D.; Zhang, Y.; Shen, D. Multi-View Spatial Aggregation Framework for Joint Localization and Segmentation of Organs at Risk in Head and Neck CT Images. *IEEE Trans. Med. Imaging* **2020**, *39*, 2794–2805. [[CrossRef](#)]
92. Qiu, B.; Guo, J.; Kraeima, J.; Glas, H.H.; Borra, R.J.; Witjes, M.J.; Ooijen, P.M.v. Recurrent convolutional neural networks for mandible segmentation from computed tomography. *arXiv* **2020**, arXiv:2003.06486.
93. Gou, S.; Tong, N.; Qi, S.; Yang, S.; Chin, R.; Sheng, K. Self-channel-and-spatial-attention neural network for automated multi-organ segmentation on head and neck CT images. *Phys. Med. Biol.* **2020**, *65*, 245034. [[CrossRef](#)] [[PubMed](#)]
94. Jiang, J.; Sharif, E.; Um, H.; Berry, S.; Veeraraghavan, H. Local block-wise self attention for normal organ segmentation. *arXiv* **2019**, arXiv:1909.05054.
95. Sun, S.; Liu, Y.; Bai, N.; Tang, H.; Chen, X.; Huang, Q.; Liu, Y.; Xie, X. AttentionAnatomy: A unified framework for whole-body organs at risk segmentation using multiple partially annotated datasets. In Proceedings of the 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI), Iowa City, IA, USA, 3–7 April 2020; pp. 1–5.
96. Liu, Z.; Wang, H.; Lei, W.; Wang, G. CSAF-CNN: Cross-Layer Spatial Attention Map Fusion Network for Organ-at-Risk Segmentation in Head and Neck CT Images. In Proceedings of the 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI), Iowa City, IA, USA, 3–7 April 2020; pp. 1522–1525.
97. Zhang, S.; Wang, H.; Tian, S.; Zhang, X.; Li, J.; Lei, R.; Gao, M.; Liu, C.; Yang, L.; Bi, X.; et al. A slice classification model-facilitated 3D encoder–decoder network for segmenting organs at risk in head and neck cancer. *J. Radiat. Res.* **2021**, *62*, 94–103. [[CrossRef](#)]
98. Tappeiner, E.; Pröll, S.; Hönl, M.; Raudaschl, P.F.; Zaffino, P.; Spadea, M.F.; Sharp, G.C.; Schubert, R.; Fritscher, K. Multi-organ segmentation of the head and neck area: An efficient hierarchical neural networks approach. *Int. J. Comput. Assist. Radiol. Surg.* **2019**, *14*, 745–754. [[CrossRef](#)]
99. Mu, G.; Yang, Y.; Gao, Y.; Feng, Q. Multi-scale 3D convolutional neural network-based segmentation of head and neck organs at risk. *J. South. Med. Univ.* **2020**, *40*, 491–498.
100. Wang, Y.; Zhao, L.; Wang, M.; Song, Z. Organ at Risk Segmentation in Head and Neck CT Images Using a Two-Stage Segmentation Framework Based on 3D U-Net. *IEEE Access* **2019**, *7*, 144591–144602. [[CrossRef](#)]
101. Lei, Y.; Zhou, J.; Dong, X.; Wang, T.; Mao, H.; McDonald, M.; Curran, W.J.; Liu, T.; Yang, X. Multi-organ segmentation in head and neck MRI using U-Faster-RCNN. In *Medical Imaging 2020: Image Processing*; International Society for Optics and Photonics: Bellingham, WA, USA, 2020; Volume 11313, p. 113133A.
102. Lei, Y.; Harms, J.; Dong, X.; Wang, T.; Tang, X.; David, S.Y.; Beitler, J.J.; Curran, W.J.; Liu, T.; Yang, X. Organ-at-Risk (OAR) segmentation in head and neck CT using U-RCNN. In *Medical Imaging 2020: Computer-Aided Diagnosis*; International Society for Optics and Photonics: Bellingham, WA, USA, 2020; Volume 11314, p. 1131444.
103. Liang, S.; Tang, F.; Huang, X.; Yang, K.; Zhong, T.; Hu, R.; Liu, S.; Yuan, X.; Zhang, Y. Deep-learning-based detection and segmentation of organs at risk in nasopharyngeal carcinoma computed tomographic images for radiotherapy planning. *Eur. Radiol.* **2019**, *29*, 1961–1967. [[CrossRef](#)]

104. van Dijk, L.V.; Van den Bosch, L.; Aljabar, P.; Peressutti, D.; Both, S.; Steenbakkers, R.J.; Langendijk, J.A.; Gooding, M.J.; Brouwer, C.L. Improving automatic delineation for head and neck organs at risk by Deep Learning Contouring. *Radiother. Oncol.* **2020**, *142*, 115–123. [[CrossRef](#)]
105. Men, K.; Geng, H.; Cheng, C.; Zhong, H.; Huang, M.; Fan, Y.; Plastaras, J.P.; Lin, A.; Xiao, Y. More accurate and efficient segmentation of organs-at-risk in radiotherapy with convolutional neural networks cascades. *Med. Phys.* **2019**, *46*, 286–292. [[CrossRef](#)] [[PubMed](#)]
106. Egger, J.; Pfarrkirchner, B.; Gsaxner, C.; Lindner, L.; Schmalstieg, D.; Wallner, J. Fully convolutional mandible segmentation on a valid ground-truth dataset. In Proceedings of the 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Honolulu, HI, USA, 18–21 July 2018; pp. 656–660.
107. Krzanowski, W. *Principles of Multivariate Analysis*; OUP Oxford: Oxford, UK, 2000; Volume 23.
108. Wold, S.; Esbensen, K.; Geladi, P. Principal component analysis. *Chemom. Intell. Lab. Syst.* **1987**, *2*, 37–52. [[CrossRef](#)]
109. Cootes, T.; Baldock, E.; Graham, J. An introduction to active shape models. *Image Process. Anal.* **2000**, *328*, 223–248.
110. Cootes, T.F.; Edwards, G.J.; Taylor, C.J. Active appearance models. *IEEE Trans. Pattern Anal. Mach. Intell.* **2001**, *23*, 681–685. [[CrossRef](#)]
111. Cootes, T.F.; Twining, C.J.; Petrovic, V.S.; Schestowitz, R.; Taylor, C.J. Groupwise Construction of Appearance Models using Piece-wise Affine Deformations. *BMVC* **2005**, *5*, 879–888.
112. Pekar, V.; Allaire, S.; Kim, J.; Jaffray, D.A. Head and neck auto-segmentation challenge. *MIDAS J.* **2009**, *11*, 2009.
113. Kass, M.; Witkin, A.; Terzopoulos, D. Snakes: Active contour models. *Int. J. Comput. Vis.* **1988**, *1*, 321–331. [[CrossRef](#)]
114. Otsu, N. A threshold selection method from gray-level histograms. *IEEE Trans. Syst. Man Cybern.* **1979**, *9*, 62–66. [[CrossRef](#)]
115. O'Mahony, N.; Campbell, S.; Carvalho, A.; Harapanahalli, S.; Hernandez, G.V.; Krpalkova, L.; Riordan, D.; Walsh, J. Deep learning vs. traditional computer vision. In *Science and Information Conference*; Springer: Cham, Switzerland, 2019; pp. 128–144.
116. Kamnitsas, K.; Ledig, C.; Newcombe, V.F.; Simpson, J.P.; Kane, A.D.; Menon, D.K.; Rueckert, D.; Glocker, B. Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. *Med. Image Anal.* **2017**, *36*, 61–78. [[CrossRef](#)]
117. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, L.; Polosukhin, I. Attention is all you need. *arXiv* **2017**, arXiv:1706.03762.
118. Oktay, O.; Schlemper, J.; Folgoc, L.L.; Lee, M.; Heinrich, M.; Misawa, K.; Mori, K.; McDonagh, S.; Hammerla, N.Y.; Kainz, B.; et al. Attention U-Net: Learning where to look for the pancreas. *arXiv* **2018**, arXiv:1804.03999.
119. Ji, Y.; Zhang, H.; Zhang, Z.; Liu, M. CNN-based encoder-decoder networks for salient object detection: A comprehensive review and recent advances. *Inf. Sci.* **2021**, *546*, 835–857. [[CrossRef](#)]
120. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2016**, *39*, 1137–1149. [[CrossRef](#)]
121. Milletari, F.; Navab, N.; Ahmadi, S.A. V-net: Fully convolutional neural networks for volumetric medical image segmentation. In Proceedings of the 2016 Fourth International Conference on 3D Vision (3DV), Stanford, CA, USA, 25–28 October 2016; pp. 565–571.
122. Yang, J.; Veeraraghavan, H.; Armato, S.G., III; Farahani, K.; Kirby, J.S.; Kalpathy-Kramer, J.; van Elmpt, W.; Dekker, A.; Han, X.; Feng, X.; et al. Autosegmentation for thoracic radiation treatment planning: A grand challenge at AAPM 2017. *Med. Phys.* **2018**, *45*, 4568–4581. [[CrossRef](#)] [[PubMed](#)]