

Article

Automatic Multiorgan Segmentation in Pelvic Region with Convolutional Neural Networks on 0.35 T MR-Linac Images

Emmanouil Koutoulakis ^{1,2,*} , Louis Marage ¹, Emmanouil Markodimitrakis ^{2,3} , Leone Aubignac ¹, Catherine Jenny ⁴ , Igor Bessieres ^{1,*} and Alain Lalande ^{3,5} 

¹ Department of Medical Physics, Centre Georges-Francois Leclerc, 21000 Dijon, France; lmarage@cgfl.fr (L.M.); laubignac@cgfl.fr (L.A.)

² Computation Biomedicine Laboratory (CBML), Foundation for Research and Technology Hellas (FORTH), 70013 Heraklion, Greece; manosmark@ics.forth.gr

³ ICMUB Laboratory, UMR 6302 CNRS, University of Burgundy, 21000 Dijon, France; alain.lalande@u-bourgogne.fr

⁴ Department of Oncologic Radiotherapy, Groupe Hospitalier Pitié-Salpêtrière, Assistance Publique-Hopitaux de Paris, 75013 Paris, France; catherine.jenny@aphp.fr

⁵ Medical Imaging Department, University Hospital of Dijon, 21000 Dijon, France

* Correspondence: manoskout@ics.forth.gr (E.K.); ibessieres@cgfl.fr (I.B.)

Abstract: MR-Linac is a recent device combining a linear accelerator with an MRI scanner. The improved soft tissue contrast of MR images is used for optimum delineation of tumors or organs at risk (OARs) and precise treatment delivery. Automatic segmentation of OARs can contribute to alleviating the time-consuming process for radiation oncologists and improving the accuracy of radiation delivery by providing faster, more consistent, and more accurate delineation of target structures and organs at risk. It can also help reduce inter-observer variability and improve the consistency of contouring while reducing the time required for treatment planning. In this work, state-of-the-art deep learning techniques were evaluated based on 2D and 2.5D training strategies to develop a comprehensive tool for the accurate segmentation of pelvic OARs dedicated to 0.35 T MR-Linac. In total, 103 cases with 0.35 T MR images of the pelvic region were investigated. Experts considered and contoured the bladder, rectum, and femoral heads as OARs and the prostate as the target volume. For the training of the neural network, 85 patients were randomly selected, and 18 were used for testing. Multiple U-Net-based architectures were considered, and the best model was compared using both 2D and 2.5D training strategies. The evaluation of the models was performed based on two metrics: the Dice similarity coefficient (DSC) and the Hausdorff distance (HD). In the 2D training strategy, Residual Attention U-Net (ResAttU-Net) had the highest scores among the other deep neural networks. Due to the additional contextual information, the configured 2.5D ResAttU-Net performed better. The overall DSC were 0.88 ± 0.09 and 0.86 ± 0.10 , and the overall HD was 1.78 ± 3.02 mm and 5.90 ± 7.58 mm for 2.5D and 2D ResAttU-Net, respectively. The 2.5D ResAttU-Net provides accurate segmentation of OARs without affecting the computational cost. The developed end-to-end pipeline will be merged with the treatment planning system for in-time automatic segmentation.

Keywords: automatic segmentation; organ at risk; 2.5D; convolutional neural networks; MRI



Citation: Koutoulakis, E.; Marage, L.; Markodimitrakis, E.; Aubignac, L.; Jenny, C.; Bessieres, I.; Lalande, A. Automatic Multiorgan Segmentation in Pelvic Region with Convolutional Neural Networks on 0.35 T MR-Linac Images. *Algorithms* **2023**, *16*, 521. <https://doi.org/10.3390/a16110521>

Academic Editor: Frank Werner

Received: 6 October 2023

Revised: 6 November 2023

Accepted: 9 November 2023

Published: 15 November 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

External radiation therapy is a widely used curative treatment option for patients with prostate cancer [1]. In radiotherapy, the optimization of the tumor local control and the minimization of the radiation toxicities are highly linked to the accuracy of the definition of the target volume and the OARs [2]. The treatment planning procedure is conventionally based on 3D anatomical information obtained via computed tomography (CT) scans. Nevertheless, because of its superior soft tissue contrast, MR imaging has been used in

radiotherapy treatment planning for several years [3]. The better quality of MR images improves the organs' delineation and reduces the contouring uncertainty [4,5]. Especially for prostate cancer, an MR-only workflow has become more and more common [6–8]. In addition, the recent integration of magnetic resonance imager (MRI) on linear accelerators (LINAC) for magnetic resonance-guided radiotherapy (MRgRT) allows the direct use of MR images for the planning and the guidance of treatment. Indeed, because it is a non-ionizing imaging device, MR images can be acquired continuously during the treatment to monitor the radiation activation (gating process), and the on-board MR image quality enables the online replanning of the treatment by considering the daily position and shape of the organs (adaptive radiotherapy process). MR-Linacs technology represents a higher cost than classical Linacs; nevertheless, it promises improving the clinical outcomes of prostate cancer [9,10]. First, reduced toxicity has been pointed out [11]. In this context, the MRIdian (ViewRay Inc., Oakwood, OH, USA) MR-Linac system has been clinically used in our institution since June 2019. It combines a 6 MV Flattening Filter-Free (FFF) LINAC with a 0.35 T MR imaging system [12]. Nevertheless, with this device, only manual segmentation is possible on the MR planning images. Manual delineation has been described as a tedious process, increasing the workload of radiation oncologists [13]. The segmentations are also subject to intra- and inter-operator variabilities due to different delineation expertise and subjective evaluation from radiation oncologists. Automatic segmentation techniques, on the other hand, have been developed during recent decades in medical imaging for an accurate and efficient analysis of images. It can produce reliable and precise results in a shorter time frame, improving the speed and accuracy of diagnosis and treatment planning [14–17]. These issues have been addressed, and a plethora of automatic segmentation techniques have been proposed in the literature in both CT and MR imaging modalities [13,18,19]. Atlas-based, deep learning, and other image-processing techniques have been developed in the last two decades [20,21]. Atlas-based methods have shown acceptable accuracy for the segmentation of OARs. However, most of the results require a significant amount of time from the radiation oncologist to correct regions that contain small areas, imaging artifacts, deformable organs, or organs near the target volume [22]. Due to the issues with atlas-based methods, researchers focused on multi-atlas segmentation (MAS), where multiple atlases are used as prior knowledge instead of individual atlases [23]. However, even multi-atlas methods have limitations, the most notable being the need for considerable computational resources [24].

Currently, a common trend in healthcare is the development of deep neural networks for automatic detection, classification, and segmentation. Deep learning techniques are capable of learning task-specific features without the need for hand-engineered features. Convolutional neural networks (CNNs) are a branch of deep learning that gain complex information from training samples, which is essential for successfully delineating anatomical structures. Specifically in MR imaging, researchers have shown great interest in alleviating MRgRT workflow [19]. For segmentation, a fully convolutional network (FCN) [25] was proposed to leverage the computational limitation by implementing symmetrical encoding and decoding paths to learn low- and high-level features from the input image. Later, Ronneberger et al. [26] proposed a successful encoder–decoder architecture for medical image segmentation called U-Net. Nowadays, a wide variety of implementations rely on the U-Net architecture [27–30]. Another aspect of neural network models for medical image analysis is the training strategy 2D, 2.5D, or 3D. Each strategy learns features from the training data in a different manner. The most common strategy is 2D, in which a single image (or slice) is passed as an input. Similarly, 2.5D considers stacks of adjacent slices as input, collecting pseudo-3D information using 2D operations [31]. The 3D training strategy is significantly more demanding than the other strategies since it performs 3D operations. Indeed, the 2.5D strategy is regarded as computationally efficient since the memory requirements are reduced [32].

Based on the MR modality, there is considerable literature on semantic segmentation. Maji et al. [33] proposed a modified 2D version of Attention U-Net [26,30,34] with an

attention-guided decoder for semantic segmentation of brain tumors for MR-only treatment planning. Aldoj et al. [35] proposed densely connected blocks instead of plain convolutional blocks [36] in a U-Net-based architecture for the automatic segmentation of prostate parts. Elguindi et al. [37] developed a multiorgan automatic segmentation model for prostate radiotherapy using transfer learning with the fully convolutional network (FCN) architecture [38]. All the proposed architectures used a combination of Dice and cross-entropy loss functions. In the 2.5D strategy, Alkadi et al. [39] implemented a deep FCN based on SegNet for prostate cancer detection and segmentation using T2-weighted MR images. The input with their method was a single slice replicated into three channels. Similarly, Huang et al. [40] proposed a modified version of Fuse U-Net with an attention mechanism to accurately segment multiple OARs (femoral heads, bladder, rectum, and anal canal) through multi-sequence MR images to facilitate MR-only treatment planning preparation for prostate and cervical cancer. Nevertheless, no study has been found focusing on 0.35 T MR Linac images.

In this context, we intend to develop a deep-learning-based automated multiorgan segmentation approach dedicated to the images from a 0.35 T MR-Linac for the pelvic area. A challenge of this study is the evaluation of the 2.5D strategy to mitigate the lack of contextual information in 2D-based networks. State-of-the-art U-Net variations were configured and compared with a 2D training method for multi-class segmentation in the male pelvis. Then, the best-performing model was trained and tested with 2.5D training. The organs of interest are the pelvic OARs: femoral heads, bladder, and rectum. The prostate was considered as the target volume. MRI studies were acquired from the same machine (0.35 T MRIdian, ViewRay Inc., Oakwood, OH, USA) at two institutions to create a large enough dataset. After evaluating deep learning architectures, an automated tool was developed for clinical use.

This work should reduce the radiation oncologists' workload and inter-operator variability and consequently reinforce the use of MR images for treatment planning. To our knowledge, it is the first time that the development of a deep-learning-based automatic segmentation tool has been performed on 0.35 T MR-Linac images.

2. Materials and Methods

2.1. Dataset

Patients in this study were diagnosed with prostate cancer and underwent external radiotherapy between June 2019 and February 2022 on a 0.35 T MR-Linac MRIdian (ViewRay Inc., Oakwood, OH, USA) [9]. In total, 103 cases were collected. More specifically, 93 cases were acquired from the Centre Georges-François Leclerc (CGFL, Dijon, France) and 10 cases from the University Hospital of Pitié Salpêtrière (Paris, France). This study was approved by the human subject ethics board of the Centre Georges-François Leclerc and was conducted in accordance with the Helsinki Declaration of 1975, as revised in 2013. A balanced steady-state free precession pulse sequence (bSSFP) was used for MR image acquisition in 3D with a raw image reconstruction in axial orientation, yielding T2/T1-weighted image contrast. The 0.35 T low magnetic field and the original split gradient coil design of the MRIdian is unusual in comparison to diagnostic MR imaging systems. These aspects make the images acquired with this system unique [12], with possible systemic imaging artifacts [41]. Pixel spacing was $1.5 \times 1.5 \text{ mm}^2$, and the slice thickness was 1.5 mm. In addition to the MR images, a radiation therapy (RT) structure set file was associated with each MRI exam. The RT structure file contains the contextual information of the targets and OARs delineated by an expert with an integrated treatment planning system (TPS) of the MRIdian system. In the case of prostate cancer, the whole prostate or the prostate with the seminal vesicle is generally defined as the target volume that must receive the prescribed dose. All the involved data were acquired without restrictions (i.e., artifacts, malignancy level, etc.). The OARs that the model focused on were the femoral heads, bladder, and rectum; the target volume was the prostate. For the proposed approach, only the axial view

of the images was used in both training strategies, considering the raw image orientation. Figure 1 shows a sample of the dataset.

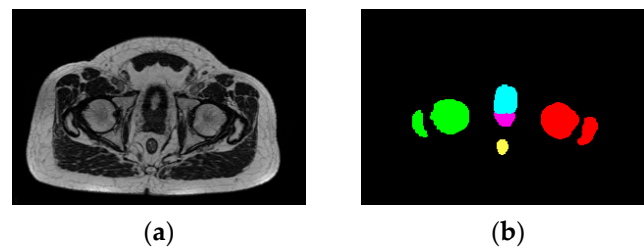


Figure 1. Data sample. (a) MR image in axial orientation, and (b) ground truth segmented OARs and prostate from expert physicians. Each color on (b) corresponds to a different organ.

2.2. Preprocessing

Firstly, image normalization was performed, changing the pixels' value range to [0–1] to stabilize the gradient descent step during training. Secondly, centre cropping or padding was performed to homogenize the shape of the images to 256×256 according to the model's input requirements. The last pre-processing step is the minimum filtering used to discard the potential noise in the background (pixels close to zero). MR images often exhibit regions with pixels that are near or reach values close to zero. The aim is to isolate and preserve these pixels inside the body, as they represent regions with minimal signal and are potentially crucial for maintaining the overall structure of the image. This threshold determination is grounded in a histogram-based examination and utilization of the anatomical information from MRI images. The visualization of the minimum filtering is depicted in Figure 2. Data augmentation was also performed for the training process, thereby contributing to overfitting deterrence by increasing the number of samples. More specifically, the data augmentation was random vertical flipping, random rotation with a limit of $\pm 15^\circ$ and arbitrary scaling with a full scale of 10% of the original image size.

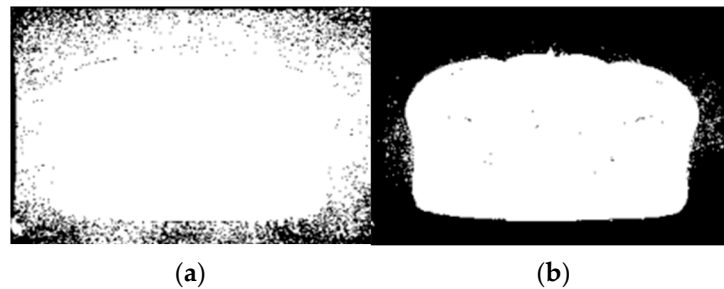


Figure 2. Minimum filtering: (a) the pixels with a normalized gray level >0.1 , and (b) the results after the minimum filtering, removing background noise.

2.3. Residual Attention U-Net Network

The U-Net structure [26] was adopted to consider high-resolution local textures and low-resolution contextual information using encoding and decoding paths. Additional operations were implemented, such as attention gates (AGs) [30] and residual blocks [34]. The integration of the AG into the model provides superior attention to important regions in our case (OARs and prostate) rather than unlabeled regions (background). Also, residual blocks were implemented for in-depth feature information collection in each convolution, increasing the model's performance and reducing the model's parameters. The input layer requires a fixed number of slices to perform the 2.5D training strategy. In our study, this fixed number is defined as 3, where the second slice serves as the primary one, encompassing the segmented OARs. The first and third slices, designated as neighboring slices, are also input into the model to gather contextual information. These neighboring slices provide additional features that help the model in considering the broader context in

which the regions of interest exist. Then, the input shape of the model is $256 \times 256 \times 3$, representing the height, width, and number of slices, respectively.

The encoding and decoding paths contain four layers in depth; each layer includes a residual block. Between each layer, both encoding and decoding layers are connected through concatenation connections. In the encoder, each residual block is a pair of convolutional blocks with a fixed sequence of 2D convolutions having a 3×3 filter size, stride one, padding one, batch normalization, and ReLU operations followed by a 2D max-pooling operation with a kernel size of 2×2 and stride 2. After the encoding path, the bottleneck of the model provides a $16 \times 16 \times 1024$ feature map. Then, the attention mechanism is triggered by inputting the feature map from the previous layer and the feature map from the same layer of the encoding path. In the last phase, concatenation is performed between the output of the AG and the up-sampled feature map of the previous layer, followed by the same residual block as in the encoding path. The schematic of Residual Attention U-Net (ResAttU-Net) architecture is presented in Figures 3 and 4.

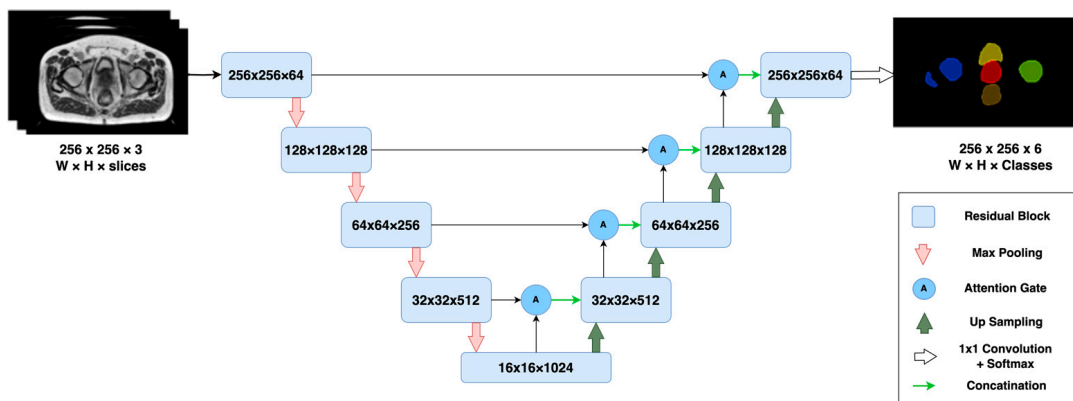


Figure 3. Residual Attention U-Net model. The residual blocks were replaced with the plain convolutional layers. The left side of the model represents the encoding path, and the right side represents the decoding path, respectively.

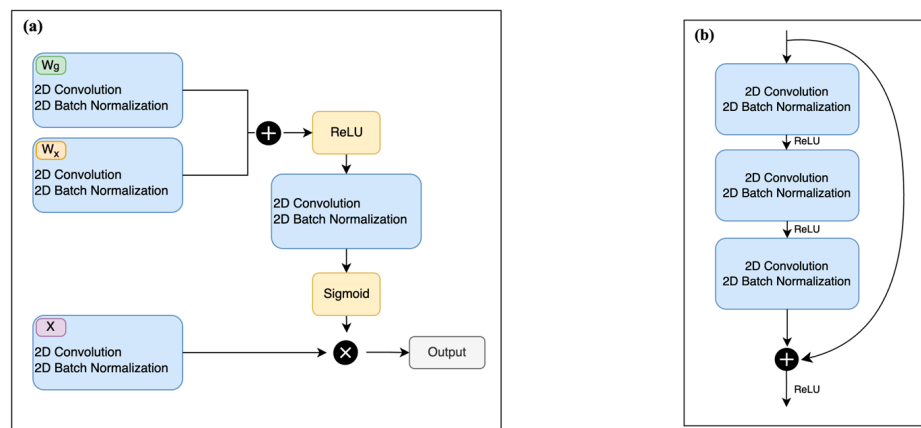


Figure 4. (a) The attention block, where W_g is the feature map derived from the encoding path, and W_x and X are the convolved and normalized output of the previous residual block. (b) The structure of the residual block.

Multiple basic loss functions were tested during the implementation related to the semantic segmentation task. Among them, Focal Loss (FL) [42] is a modified version of the cross-entropy loss, including an additional term to focus on hard misclassified examples,

dealing with the class imbalance issue [43]. Focal loss is defined in Equation (1), where p_t is the estimated probability for the class, and γ is a tunable parameter with $\gamma \in [0,5]$.

$$FL(p_t) = -\alpha_t(1 - p_t)^\gamma \log(p_t) \quad (1)$$

The adaptive moment estimation (ADAM) algorithm was used due to its efficiency [44]. ADAM is an optimization algorithm that updates the network weights iteratively, remaining controlled and unbiased. The learning rate was 0.001, and the batch size was 4 for all the experiments. Additional call-back functions were introduced in the experiments to smooth the learning procedure and decrease the training time. Reduced learning on the plateau was used to reduce the learning rate when there was no update throughout the learning iterations. Additionally, early stopping was implemented to stop the learning process when the model could not learn different features. The patience factor of the call-backs was 15 and 25 epochs, respectively. Apart from the ResAttU-Net, two additional deep learning models were trained to compare the performance of the proposed model using the same hyperparameters: U-Net [26] and Attention U-Net [30].

2.4. Post-Processing

The output of each class in our model is a 2D mask depicting the target's predicted location, or OAR, in the same view as the input image. Sometimes, the predicted classes are not segmented properly, creating gaps inside the segmented regions. Thus, an automatic algorithm with a combination of post-processing techniques and rules was implemented. Firstly, the predicted mask is binarized with a pixels' threshold of 0.5. The optimal threshold was determined after an extensive analysis of binarized predicted outcomes. A hole-filling process then eliminates the faulty gaps inside the segmentation. However, the hole-filling method is not performed for the bladder because neighbouring organs are commonly in the inner area (i.e., the prostate).

A series of rules is also helpful in eliminating misclassified pixels. The rules are set according to each organ's original position and hierarchy. For example, the bladder is localized at the top of the image, followed by the rectum. Similarly, the left femoral head could not be on the right side, and vice versa.

Artifacts are a common issue in medical imaging [45]. The most common artifacts for the acquisitions with the SSFP sequence occur at extreme slices where the magnetic field is less homogeneous.

In whole-volume segmentation, it is a common problem that the model predicts inaccurate segments because of these artifacts. Hence, the classified pixels were eliminated if any organ was shown there according to slice position. In the last step, interpolation proposed by Schenk et al. [46] was used to homogenize the whole volume in three dimensions. Figure 5 illustrates the prediction without post-processing and the results with post-processing.

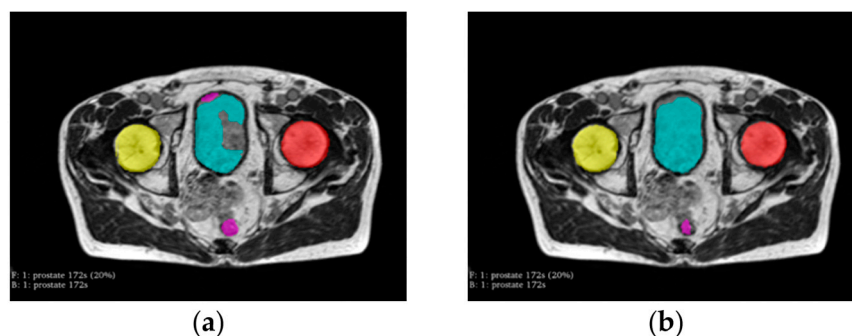


Figure 5. Automatic segmentation and post-processing. (a) The predicted mask which contains false-predicted pixels. (b) Final mask after the post-processing steps, where the wrong segment of the rectum is removed, and the bladder's segmentation morphology was corrected.

2.5. Evaluation of the Segmentation Algorithm

The network was trained and evaluated on our dataset, which has 7866 slices containing at least one class per slice. The dataset was divided into two parts: 80% (83 patients) for training and validation, while the remaining 20% (20 patients) was set aside as a distinct test set. The training process further employed k-fold cross-validation with $k = 5$, enabling a comprehensive evaluation of the model’s performance across different subsets of the training data. Once the cross-validation was completed, all the models were evaluated using the test data. This approach provides an unbiased estimate of the model’s generalization capability. Also, it is worth noting that patient-wise splitting was performed to prevent any slice overlap among the sets. The data division is detailed in Figure 6.

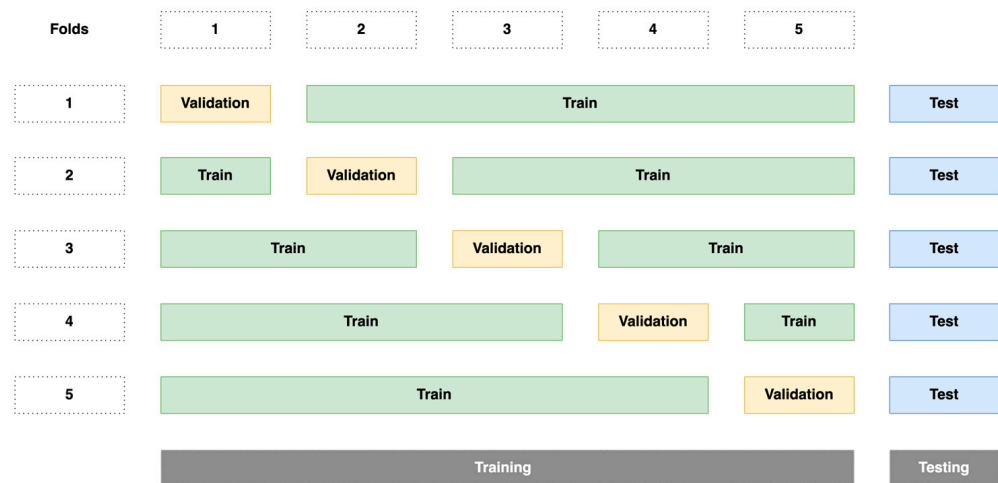


Figure 6. Visualization of the splits regarding the 5-fold cross-validation.

The geometric evaluation of the organs’ delineation was assessed using traditional metrics such as the Dice coefficient (DSC), which is responsible for the global evaluation of the segmentation Equation (2), and the Hausdorff distance (HD), where it is relevant to highlight the outliers Equation (3). These two metrics are then complementary and expressed as the mean \pm standard deviation (STD). The Hausdorff distance (HD) is the maximum distance between two contour point sets, A and B , which are the predicted and the ground truth.

$$H(A, B) = \max \left(\sup_{a \in A} \inf_{b \in B} |a - b|, \sup_{b \in B} \inf_{a \in A} |a - b| \right) \tag{2}$$

where $H(A, B)$ represents the maximum distance between two sets (A and B). The distance between these points is expressed in millimeters (mm). The smaller the HD value, the higher the segmentation accuracy.

The DSC calculates the degree of similarity between the manually segmented and the automatically segmented areas. Both A and B represent the point sets contained in the two regions. The calculation formula is shown in Equation (3).

$$\text{Dice}(A, B) = \frac{2|A \cap B|}{|A| + |B|} \tag{3}$$

where $|A \cap B|$ represents the intersection of A and B . The value range of DSC is [0–1]. The higher the DCS value is, the better the segmentation results are.

The evaluation was conducted through two experiments. The 2D strategy was carried out in the first experiment to investigate the optimal model, comparing the ResAttU-Net with conventional neural networks such as U-Net and Attention U-Net [26,30]. The DSC and the HD were compared between ResAttU-Net and the other networks using the paired

t-test after training. A *p*-value smaller than 0.05 was considered to represent statistical significance. Then, the best model candidate was retrained with a 2.5D training strategy.

2.6. Computing Environment

The computed environment consisted of the following components: Ubuntu 18.04, CPU: Intel(R) Xeon (R) W-2145 CPU @ 3.70 GHz, memory: 32 GB, GPU: TITAN V 12 GB + Quadro P400 2 GB (NVIDIA Corp., Santa Clara, CA, USA). Mask construction was performed using Python 3.9.10, NumPy 1.22.3, OpenCV 4.5.5, PyTorch 1.11.0, and NVIDIA Compute Unified Device Architecture (CUDA) Deep Learning Network library (cuDNN) 10.2.

3. Results

To demonstrate the efficacy of the residual attention network, we compared the result of the modified ResAttU-Net against the plain U-Net and Attention U-Net architectures. These networks were trained on our MRI dataset. The mean slice number per sample was 77 slices containing the prostate and the OARs (femoral heads, bladder, and rectum). All the architectures were evaluated using *k*-fold cross-validation, with approximately 8.5 h of training per fold. Firstly, the U-Net, Attention U-Net and ResAttU-Net were trained with the 2D training strategy. Regarding the evaluation metrics, ResAttU-Net performed slightly better than the other architectures (Table 1). Concerning the DSC, the differences were significant between U-Net and ResAttU-Net ($p = 0.01$) but not between Attention U-Net and ResAttU-Net ($p = 0.10$). On the other hand, for the HD, the differences were always not significant ($p = 0.11$ between U-Net and ResAttU-Net and $p = 0.49$ between Attention U-Net and ResAttU-Net). Even if the differences were not always significant, we decided to keep the model that provided the best results.

Table 1. Evaluation of 3 different architectures in 2D strategy. Best results in bold.

Architectures	Dice \pm STD	HD \pm STD (mm)
U-Net	0.83 \pm 0.14	7.95 \pm 6.03
Attention U-Net	0.84 \pm 0.12	7.50 \pm 6.18
ResAttU-Net	0.85 \pm 0.11	7.49 \pm 6.54

The second task was the evaluation of the ResAttU-Net in the 2.5D strategy. Table 2 shows the results for each organ and the overall performance between the 2D and the 2.5D training strategies. It is worth noting that the bladder showed the best performance, yielding 0.92 DSC. The prostate underperformed the other organs with the lowest DSC value of 0.80. The presence of seminal vesicles is the cause of the low DSC on prostate segmentation. Moreover, the comparison between the different training strategies shows improvements with the 2.5D strategy in all the organs apart from the rectum and the femoral head (right). However, the results are close between the two strategies for these organs. Qualitative results are shown in Figures 7 and 8. Considering the performance, the 2.5D strategy is 0.001 and 0.03 seconds (s) per slice slower than 2D using GPU and CPU, respectively (Table 3).

Table 2. Evaluation of the 2.5D ResAttU-Net.

Organs	2D		2.5D	
	Dice \pm STD	HD \pm STD (mm)	Dice \pm STD	HD \pm STD (mm)
Rectum	0.84 \pm 0.12	5.20 \pm 4.86	0.84 \pm 0.10	5.29 \pm 5.04
Bladder	0.89 \pm 0.12	9.73 \pm 7.78	0.92 \pm 0.09	6.13 \pm 5.46
Femoral head right	0.87 \pm 0.08	6.83 \pm 6.26	0.88 \pm 0.08	7.72 \pm 5.27
Femoral head left	0.88 \pm 0.08	6.64 \pm 5.40	0.88 \pm 0.08	7.02 \pm 6.15
Prostate	0.79 \pm 0.18	9.03 \pm 8.40	0.80 \pm 0.15	7.08 \pm 5.81
Overall Score	0.85 \pm 0.11	7.49 \pm 6.54	0.87 \pm 0.10	6.65 \pm 5.33

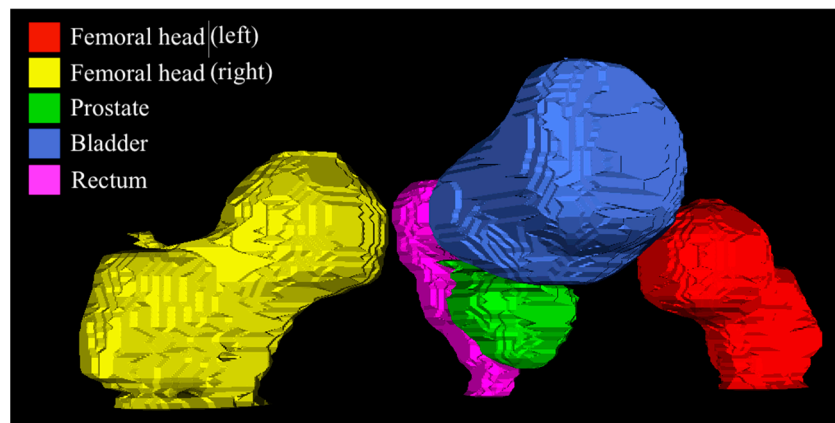


Figure 7. Sample visualization in 3D of a predicted volume.

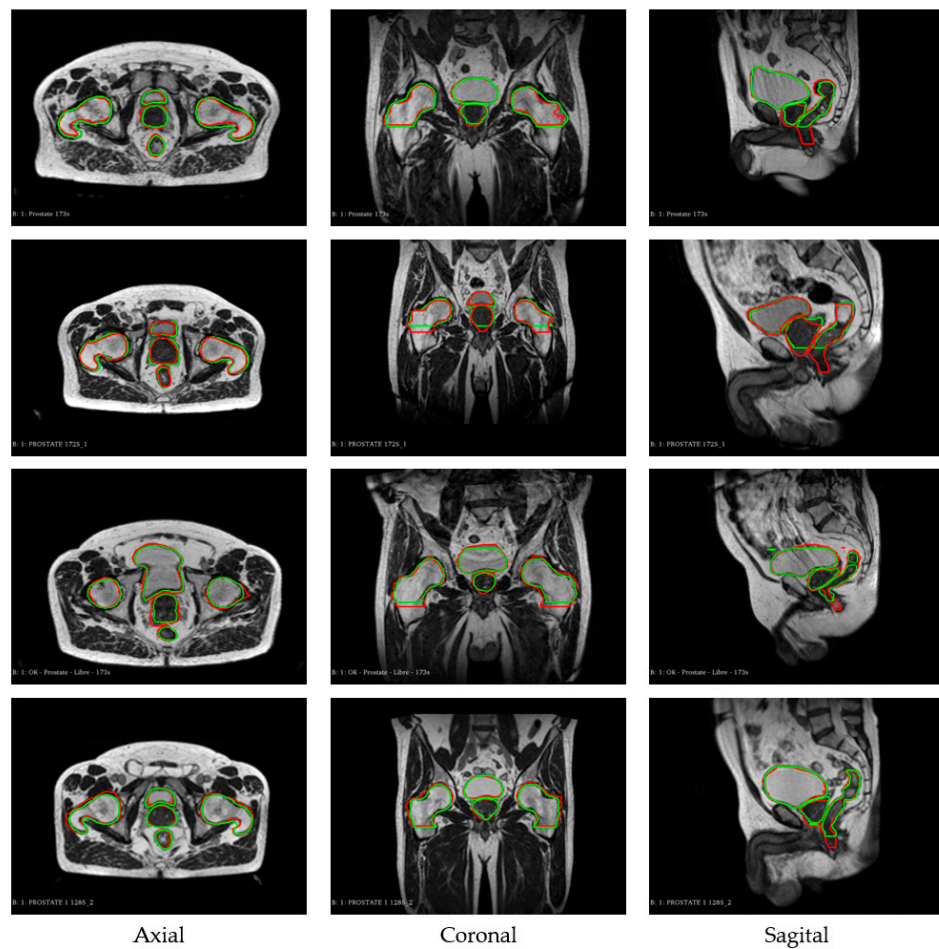


Figure 8. Segmentation using 2.5D ResAttU-Net. The manual delineation is highlighted with red and the predicted of our model with green.

Table 3. Inference time estimation with 2D and 2.5D training strategies. The times shown are seconds per slice.

Strategies	CPU (s)	GPU (s)
2D	1.45	0.113
2.5D	1.48	0.114

4. Discussion

Up to now, the most common screening tool for radiation therapy has been CT imaging [47]. Through the rapid advancements of the MR-Linac radiation therapy systems, MR imaging can screen anatomical structures with superior soft-tissue contrast compared to CT scans. In this context, an automatic multiorgan segmentation tool for the pelvic region dedicated to a 0.35 T MR-Linac system was proposed in our study. A modified version of the U-Net [33] combined with AGs and residual blocks using a 2.5D training strategy was used, utilizing multi-slice input and deep connections. The use of a 2.5D strategy allows for the obtainment of pseudo-3D information of the patient's volume without the excessive computational complexity of 3D-based methods [48–50]. The 3D modeling allows the physicians to access a global view of the region of interest in question, with any kind of view or angle. It has been used in a lot of medical applications [51–53], and this display could complement the advanced imaging capabilities of the MR-Linac device. Minema et al. [32] reported shorter computational times in both strategies than in 3D without significant performance drawbacks. Clearly, the 3D training strategy is not always the optimal solution for robust results [54].

For the ResAttU-Net, we conducted experiments on two training strategies as the results of the proposed approach exhibited higher DSC than the U-Net and Attention U-Net in the 2D strategy. Following the optimal network from 2D, the 2.5D ResAttU-Net strategy surpassed the 2D strategy. More specifically, per organ comparison showed a significant increase in the accuracy for the segmentation of the bladder and prostate despite a slight decrease for the rectum. The results show the effectiveness of the 2.5D method, exploiting pseudo-3D information through adjacent segments with insignificant time savings. The segmentation of the bladder obtained the best results among all the contours, having mean DSC: 0.92 ± 0.09 , HD: 6.13 ± 5.46 mm. Both femoral heads obtained stable results between both strategies without affecting the overall accuracy. On the contrary, the segmentation of the prostate showed lower performance with the 2.5D strategy. The observer variability in the prostate segmentation was subject-specific, with target shape and disease staging. This issue is particularly noticeable in cases where patients have intermediate-to-high risk. Similarly, the delineated prostate contains the base of seminal vesicles that can be delineated differently [55]. Additionally, the prostate base is in continuity with seminal vesicles and bladder, which may cause unexpected artifacts throughout the inference. There are not many studies related to multiorgan segmentation in the male pelvis dedicated to 0.35 T MRI. Specifically, the related works are MRI-based but with different sequence types or different magnetic fields. Our work demonstrated better or similar performance compared with the existing studies on T2-weighted images and on the multisequence approach (T1-weighted, T2-weighted, and enhanced Dixon T1-weighted images) [56]. To our knowledge, there is no related work based on the SSFP sequence for multiorgan segmentation in the pelvic region. A comparison shows that our work has comparable results regarding the related approaches (Table 4). However, the prostate remains lower compared to the results of Elguindy et al. The work of Huang et al. has no available prostate results since the work was based on cervical cancer. It is worth noting that it is not possible for the model to work in the same manner with MR imaging systems that have different technical specifications and magnetic field strength due to the differences in image resolution, contrast, and noise levels. Therefore, it is important to train deep learning models on MRIs of the same magnetic strength as those that will be used in clinical practice to ensure optimal performance.

The optimal strategy for the time-efficient segmentation of OARs is both the use of an automatic segmentation tool and the manual correction of the predicted contours from expert physicians if necessary. In addition to the barriers of accurate automatic segmentation, data collection and acquisition is also challenging as fully supervised learning and model performance are dependent on high-quality datasets. In the present study, the model was evaluated using a clinical case provided by an experienced radiologist with over five years of expertise. The automatic segmentation, along with manual modifications by an expert,

required only 10 min, in contrast to the fully manual delineation process that typically takes 20 min or more.

Table 4. Comparison with recent works related to the pelvic region.

Organs	Our Work	Elguindi et al. [37]	Huang et al. [40]
	Dice \pm STD	Dice \pm STD	Dice \pm STD
Rectum	0.84 \pm 0.10	0.82 \pm 0.05	0.78 \pm 0.07
Bladder	0.92 \pm 0.09	0.93 \pm 0.04	0.90 \pm 0.09
Femoral head right	0.88 \pm 0.08	-	0.90 \pm 0.02
Femoral head left	0.88 \pm 0.08	-	0.89 \pm 0.03
Prostate	0.80 \pm 0.15	0.85 \pm 0.07	-
Sequence Protocol	SSFP	T2-weighted	Multisequence

While this study presents promising results, further work is needed to evaluate the approach in a clinical setting. Specifically, a dosimetry analysis will be performed to assess the clinical impact and correlation between automatic segmentation and manual segmentation. Furthermore, we will undertake an extensive examination of model enhancements, with a primary focus on the incorporation of the slices in the 2.5D strategy employing varied input modalities. Notably, introducing additional slices as input to the model holds promise for enhancing segmentation performance. Similarly, the substitution of randomly oriented 2D cross-sections for the existing axial slices as input into the model will be considered since they could exploit further spatial information. On the other hand, even if the 3D deep networks require many training parameters, the use of 3D volume contains important information [56]. An evaluation framework between the 2.5D and 3D strategies will be further investigated.

5. Conclusions

To conclude, the main objective of this study was the automatic segmentation of OARs and prostate in the pelvic region and the construction of a TPS-readable RT structure set. The presented work compared different training strategies with an in-house dataset of 103 cases. The 2.5D ResAttU-Net was the optimal model compared to plain U-Net and Attention U-Net. The mean DSC for each organ in the test dataset was 0.84, 0.92, 0.88, 0.88, and 0.80 for the rectum, bladder, femoral head (right), femoral head (left), and prostate, respectively. Problems arose mainly related to the class imbalance, where the organs were not shown in each slice. The proposed model shows great potential to be adapted to different anatomical regions or additional organs in the pelvic region for radiation therapy planning, which requires a considerable amount of training samples. For future work, dosimetry analysis will be performed. The final tool for the pelvic region is clinically used in our institution and is available for any ViewRay Inc. MRIdian user. The source code for the project is available on GitHub. (https://github.com/manoskout/automatic_segmentation) (accessed on 25 July 2023).

Author Contributions: Conceptualization, E.K., L.M., I.B. and A.L.; methodology, E.K., L.M., I.B. and A.L.; software, E.K. and E.M.; validation, E.K., L.M., I.B. and A.L.; formal analysis, I.B. and A.L.; investigation, E.K. and E.M.; data curation, L.M., L.A., C.J. and I.B.; writing, original draft preparation, E.K.; writing, review and editing, L.M., I.B. and A.L.; funding acquisition, A.L. All authors have read and agreed to the published version of the manuscript.

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

Institutional Review Board Statement: This study was approved by the human subject ethics board of the Centre Georges-François Leclerc and was conducted in accordance with the Helsinki Declaration of 1975, as revised in 2013. Moreover the study was in accordance to the GDPR (General Data Protection Regulation). All participants were given clear information about the study, and their non-opposition was obtained.

Data Availability Statement: Ethical reasons prohibit the public availability of the dataset. However, it can be confidentially communicated to the reviewers and the Journal's Editor if necessary.

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

References

1. Boehmer, D.; Maingon, P.; Poortmans, P.; Baron, M.-H.; Miralbell, R.; Remouchamps, V.; Scrase, C.; Bossi, A.; Bolla, M. Guidelines for primary radiotherapy of patients with prostate cancer. *Radiother. Oncol.* **2006**, *79*, 259–269. [[CrossRef](#)] [[PubMed](#)]
2. Segedin, B.; Petric, P. Uncertainties in target volume delineation in radiotherapy—Are they relevant and what can we do about them? *Radiol. Oncol.* **2016**, *50*, 254–262. [[CrossRef](#)] [[PubMed](#)]
3. Gunnlaugsson, A.; Persson, E.; Gustafsson, C.; Kjellén, E.; Ambolt, P.; Engelholm, S.; Nilsson, P.; Olsson, L.E. Target definition in radiotherapy of prostate cancer using magnetic resonance imaging only workflow. *Phys. Imaging Radiat. Oncol.* **2019**, *9*, 89–91. [[CrossRef](#)] [[PubMed](#)]
4. Kupelian, P.; Sonke, J.-J. Magnetic Resonance–Guided Adaptive Radiotherapy: A Solution to the Future. *Semin. Radiat. Oncol.* **2014**, *24*, 227–232. [[CrossRef](#)] [[PubMed](#)]
5. Pollard, J.M.; Wen, Z.; Sadagopan, R.; Wang, J.; Ibbott, G.S. The future of image-guided radiotherapy will be MR guided. *Br. J. Radiol.* **2017**, *90*, 20160667. [[CrossRef](#)]
6. Tyagi, N.; Zelefsky, M.J.; Wibmer, A.; Zakian, K.; Burleson, S.; Happersett, L.; Halkola, A.; Kadbi, M.; Hunt, M. Clinical experience and workflow challenges with magnetic resonance-only radiation therapy simulation and planning for prostate cancer. *Phys. Imaging Radiat. Oncol.* **2020**, *16*, 43–49. [[CrossRef](#)]
7. Greer, P.; Martin, J.; Sidhom, M.; Hunter, P.; Pichler, P.; Choi, J.H.; Best, L.; Smart, J.; Young, T.; Jameson, M.; et al. A multi-center prospective study for implementation of an MRI-only prostate treatment planning work-flow. *Front. Oncol.* **2019**, *9*, 826. [[CrossRef](#)]
8. Ménard, C.; Paulson, E.; Nyholm, T.; McLaughlin, P.; Liney, G.; Dirix, P.; van der Heide, U.A. Role of Prostate MR Imaging in Radiation Oncology. *Radiol. Clin. N. Am.* **2018**, *56*, 319–325. [[CrossRef](#)]
9. Yuan, J.; Poon, D.M.C.; Lo, G.; Wong, O.L.; Cheung, K.Y.; Yu, S.K. A narrative review of MRI acquisition for MR-guided-radiotherapy in prostate cancer. *Quant. Imaging Med. Surg.* **2022**, *12*, 1585–1607. [[CrossRef](#)]
10. Kishan, A.U.; Ma, T.M.; Lamb, J.M.; Casado, M.; Wilhalme, H.; Low, D.A.; Sheng, K.; Sharma, S.; Nickols, N.G.; Pham, J.; et al. Magnetic Resonance Imaging–Guided vs Computed Tomography–Guided Stereotactic Body Radiotherapy for Prostate Cancer. *JAMA Oncol.* **2023**, *9*, 365–373. [[CrossRef](#)]
11. Ma, T.; Neylon, J.; Savjani, R.; Low, D.; Steinberg, M.; Cao, M.; Kishan, A. Treatment Delivery Gating of MRI-Guided Stereotactic Radiotherapy for Prostate Cancer: An Exploratory Analysis of a Phase III Randomized Trial of CT-Vs. MR-Guided Radiotherapy (MIRAGE). *Int. J. Radiat. Oncol.* **2023**, *117*, e692–e693. [[CrossRef](#)]
12. Klüter, S. Technical design and concept of a 0.35 T MR-Linac. *Clin. Transl. Radiat. Oncol.* **2019**, *18*, 98–101. [[CrossRef](#)] [[PubMed](#)]
13. Cardenas, C.E.; Yang, J.; Anderson, B.M.; Court, L.E.; Brock, K.B. Advances in Auto-Segmentation. *Semin. Radiat. Oncol.* **2019**, *29*, 185–197. [[CrossRef](#)] [[PubMed](#)]
14. Kalantar, R.; Lin, G.; Winfield, J.M.; Messiou, C.; Lalondrelle, S.; Blackledge, M.D.; Koh, D.-M. Automatic Segmentation of Pelvic Cancers Using Deep Learning: State-of-the-Art Approaches and Challenges. *Diagnostics* **2021**, *11*, 1964. [[CrossRef](#)] [[PubMed](#)]
15. Almeida, G.; Tavares, J.M.R. Deep Learning in Radiation Oncology Treatment Planning for Prostate Cancer: A Systematic Review. *J. Med. Syst.* **2020**, *44*, 179. [[CrossRef](#)] [[PubMed](#)]
16. Fu, Y.; Lei, Y.; Wang, T.; Curran, W.J.; Liu, T.; Yang, X. A review of deep learning based methods for medical image multi-organ segmentation. *Phys. Med.* **2021**, *85*, 107–122. [[CrossRef](#)]
17. Khan, Z.; Yahya, N.; Alsaih, K.; Al-Hiyali, M.I.; Meriaudeau, F. Recent Automatic Segmentation Algorithms of MRI Prostate Regions: A Review. *IEEE Access* **2021**, *9*, 97878–97905. [[CrossRef](#)]
18. Valentini, V.; Boldrini, L.; Damiani, A.; Muren, L.P. Recommendations on how to establish evidence from auto-segmentation software in radiotherapy. *Radiother. Oncol.* **2014**, *112*, 317–320. [[CrossRef](#)]
19. Cusumano, D.; Boldrini, L.; Dhont, J.; Fiorino, C.; Green, O.; Güngör, G.; Jornet, N.; Klüter, S.; Landry, G.; Mattiucci, G.C.; et al. Artificial Intelligence in magnetic Resonance guided Radiotherapy: Medical and physical considerations on state of art and future perspectives. *Phys. Med.* **2021**, *85*, 175–191. [[CrossRef](#)]
20. Cabezas, M.; Oliver, A.; Lladó, X.; Freixenet, J.; Cuadra, M.B. A review of atlas-based segmentation for magnetic resonance brain images. *Comput. Methods Progr. Biomed.* **2011**, *104*, e158–e177. [[CrossRef](#)]
21. Wang, H.; Yushkevich, P.A. Multi-atlas Segmentation without Registration: A Supervoxel-Based Approach. *Med. Image Comput. Comput.-Assist. Interv.* **2013**, *16*, 535–542.

22. Heckemann, R.A.; Hajnal, J.V.; Aljabar, P.; Rueckert, D.; Hammers, A. Automatic anatomical brain MRI segmentation combining label propagation and decision fusion. *NeuroImage* **2006**, *33*, 115–126. [[CrossRef](#)] [[PubMed](#)]
23. Martínez, F.; Romero, E.; Dréan, G.; Simon, A.; Haigron, P.; de Crevoisier, R.; Acosta, O. Segmentation of pelvic structures for planning CT using a geometrical shape model tuned by a multi-scale edge detector. *Phys. Med. Biol.* **2014**, *59*, 1471–1484. [[CrossRef](#)] [[PubMed](#)]
24. Iglesias, J.E.; Sabuncu, M.R. Multi-atlas segmentation of biomedical images: A survey. *Med. Image Anal.* **2015**, *24*, 205–219. [[CrossRef](#)] [[PubMed](#)]
25. Shelhamer, E.; Long, J.; Darrell, T. Fully convolutional networks for semantic segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 640–651. [[CrossRef](#)]
26. Ronneberger, O.; Fischer, P.; Brox, T. *U-Net: Convolutional Networks for Biomedical Image Segmentation*; Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics); Springer: Cham, Switzerland, 2015; Volume 9351, pp. 234–241. Available online: <https://arxiv.org/abs/1505.04597v1> (accessed on 20 June 2023).
27. Zhou, Z.; Siddiquee, M.M.R.; Tajbakhsh, N.; Liang, J. UNet++: A Nested U-Net Architecture for Medical Image Segmentation. *arXiv* **2018**, arXiv:1807.10165.
28. Cao, Y.; Liu, S.; Peng, Y.; Li, J. DenseUNet: Densely connected UNet for electron microscopy image segmentation. *IET Image Process.* **2020**, *14*, 2682–2689. [[CrossRef](#)]
29. Chen, X.; Yao, L.; Zhang, Y. Residual Attention U-Net for Automated Multi-Class Segmentation of COVID-19 Chest CT Images. *arXiv* **2020**, arXiv:2004.05645.
30. Oktay, O.; Schlemper, J.; Le Folgoc, 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.
31. Ben-Cohen, A.; Diamant, I.; Klang, E.; Amitai, M.; Greenspan, H. *Fully Convolutional Network for Liver Segmentation and Lesions Detection*; Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics); Springer: Berlin/Heidelberg, Germany, 2016; Volume 10008, pp. 77–85. [[CrossRef](#)]
32. Minnema, J.; Wolff, J.; Koivisto, J.; Lucka, F.; Batenburg, K.J.; Forouzanfar, T.; van Eijnatten, M. Comparison of convolutional neural network training strategies for cone-beam CT image segmentation. *Comput. Methods Progr. Biomed.* **2021**, *207*, 106192. [[CrossRef](#)]
33. Maji, D.; Sigedgar, P.; Singh, M. Attention Res-UNet with Guided Decoder for semantic segmentation of brain tumors. *Biomed. Signal Process. Control* **2021**, *71*, 103077. [[CrossRef](#)]
34. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778. [[CrossRef](#)]
35. Aldoj, N.; Biavati, F.; Michallek, F.; Stober, S.; Dewey, M. Automatic prostate and prostate zones segmentation of magnetic resonance images using DenseNet-like U-net. *Sci. Rep.* **2020**, *10*, 14315. [[CrossRef](#)] [[PubMed](#)]
36. Huang, G.; Liu, Z.; van der Maaten, L.; Weinberger, K.Q. Densely Connected Convolutional Networks. *arXiv* **2016**, arXiv:1608.06993.
37. Elguindi, S.; Zelefsky, M.J.; Jiang, J.; Veeraraghavan, H.; Deasy, J.O.; Hunt, M.A.; Tyagi, N. Deep learning-based auto-segmentation of targets and organs-at-risk for magnetic resonance imaging only planning of prostate radiotherapy. *Phys. Imaging Radiat. Oncol.* **2019**, *12*, 80–86. [[CrossRef](#)] [[PubMed](#)]
38. Chen, L.C.; Zhu, Y.; Papandreou, G.; Schroff, F.; Adam, H. *Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation*; Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics); Springer: Berlin/Heidelberg, Germany, 2018; Volume 11211, pp. 833–851. [[CrossRef](#)]
39. Alkadi, R.; El-Baz, A.; Taher, F.; Werghe, N. A 2.5D Deep Learning-Based Approach for Prostate Cancer Detection on T2-Weighted Magnetic Resonance Imaging. In Proceedings of the 2018 European Conference on Computer Vision, Munich, Germany, 8–14 September 2018; pp. 734–739.
40. Huang, S.; Cheng, Z.; Lai, L.; Zheng, W.; He, M.; Li, J.; Zeng, T.; Huang, X.; Yang, X. Integrating multiple MRI sequences for pelvic organs segmentation via the attention mechanism. *Med. Phys.* **2021**, *48*, 7930–7945. [[CrossRef](#)]
41. Marage, L.; Walker, P.-M.; Boudet, J.; Fau, P.; Debuire, P.; Clause, E.; Petitfils, A.; Aubignac, L.; Rapacchi, S.; Bessieres, I. Characterisation of a split gradient coil design induced systemic imaging artefact on 0.35 T MR-linac systems. *Phys. Med. Biol.* **2022**, *68*, 01NT03. [[CrossRef](#)]
42. Lin, T.Y.; Goyal, P.; Girshick, R.; He, K.; Dollar, P. Focal Loss for Dense Object Detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *42*, 318–327. [[CrossRef](#)]
43. Buda, M.; Maki, A.; Mazurowski, M.A. A systematic study of the class imbalance problem in convolutional neural networks. *Neural Netw.* **2018**, *106*, 249–259. [[CrossRef](#)]
44. Kingma, D.P.; Ba, J.L. Adam: A Method for Stochastic Optimization. In Proceedings of the 3rd International Conference on Learning Representations, ICLR 2015-Conference Track Proceedings, San Diego, CA, USA, 7–9 May 2015. [[CrossRef](#)]
45. Lundervold, A.S.; Lundervold, A. An overview of deep learning in medical imaging focusing on MRI. *Z. Med. Phys.* **2018**, *29*, 102–127. [[CrossRef](#)]
46. Schenk, A.; Prause, G.; Peitgen, H.O. *Efficient Semiautomatic Segmentation of 3D Objects in Medical Images*; Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics); Springer: Berlin/Heidelberg, Germany, 2000; Volume 1935, pp. 186–195. [[CrossRef](#)]

47. Samarasinghe, G.; Jameson, M.; Vinod, S.; Field, M.; Dowling, J.; Sowmya, A.; Holloway, L. Deep learning for segmentation in radiation therapy planning: A review. *J. Med. Imaging Radiat. Oncol.* **2021**, *65*, 578–595. [[CrossRef](#)]
48. Zheng, H.; Qian, L.; Qin, Y.; Gu, Y.; Yang, J. Improving the slice interaction of 2.5D CNN for automatic pancreas segmentation. *Med. Phys.* **2020**, *47*, 5543–5554. [[CrossRef](#)] [[PubMed](#)]
49. Li, J.; Liao, G.; Sun, W.; Sun, J.; Sheng, T.; Zhu, K.; von Deneen, K.M.; Zhang, Y. A 2.5D semantic segmentation of the pancreas using attention guided dual context embedded U-Net. *Neurocomputing* **2022**, *480*, 14–26. [[CrossRef](#)]
50. Hu, K.; Liu, C.; Yu, X.; Zhang, J.; He, Y.; Zhu, H. A 2.5D Cancer Segmentation for MRI Images Based on U-Net. In Proceedings of the 2018 5th International Conference on Information Science and Control Engineering (ICISCE), ICISCE 2018, Zhengzhou, China, 20–22 July 2018; pp. 6–10.
51. Battulga, B.; Konishi, T.; Tamura, Y.; Moriguchi, H. The Effectiveness of an Interactive 3-Dimensional Computer Graphics Model for Medical Education. *Interact. J. Med. Res.* **2012**, *1*, e2. [[CrossRef](#)]
52. Sun, L.; Guo, C.; Yao, L.; Zhang, T.; Wang, J.; Wang, L.; Liu, Y.; Wang, K.; Wang, L.; Wu, Q. Quantitative diagnostic advantages of three-dimensional ultrasound volume imaging for fetal posterior fossa anomalies: Preliminary establishment of a prediction model. *Prenat. Diagn.* **2019**, *39*, 1086–1095. [[CrossRef](#)] [[PubMed](#)]
53. Gomes, J.P.P.; Costa, A.L.F.; Chone, C.T.; Altemani, A.M.d.A.M.; Altemani, J.M.C.; Lima, C.S.P. Three-dimensional volumetric analysis of ghost cell odontogenic carcinoma using 3-D reconstruction software: A case report. *Oral Surg. Oral Med. Oral Pathol. Oral Radiol.* **2017**, *123*, e170–e175. [[CrossRef](#)] [[PubMed](#)]
54. Mlynarski, P.; Delingette, H.; Criminisi, A.; Ayache, N. 3D convolutional neural networks for tumor segmentation using long-range 2D context. *Comput. Med. Imaging Graph.* **2019**, *73*, 60–72. [[CrossRef](#)] [[PubMed](#)]
55. Rozet, F.; Mongiat-Artus, P.; Hennequin, C.; Beauval, J.; Beuzebec, P.; Cormier, L.; Fromont-Hankard, G.; Mathieu, R.; Ploussard, G.; Renard-Penna, R.; et al. French ccAFU guidelines—Update 2020–2022: Prostate cancer. *Progrès en Urologie* **2020**, *30*, S136–S251. [[CrossRef](#)]
56. Singh, S.P.; Wang, L.; Gupta, S.; Goli, H.; Padmanabhan, P.; Gulyás, B. 3D Deep Learning on Medical Images: A Review. *Sensors* **2020**, *20*, 5097. [[CrossRef](#)]

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