

Article

Deep Learning Networks for Automatic Retroperitoneal Sarcoma Segmentation in Computerized Tomography

Giuseppe Salvaggio ¹, Giuseppe Cutaia ^{1,2}, Antonio Greco ¹, Mario Pace ¹, Leonardo Salvaggio ³, Federica Vernuccio ¹, Roberto Cannella ^{1,2}, Laura Algeri ⁴, Lorena Incorvaia ⁵, Alessandro Stefano ^{6,*}, Massimo Galia ¹, Giuseppe Badalamenti ⁴ and Albert Comelli ⁷

- ¹ Section of Radiology, Department of Biomedicine, Neuroscience and Advanced Diagnostics (Bi.N.D.), University of Palermo, 90100 Palermo, Italy; p.salvaggio@libero.it (G.S.); cutaiagiuseppe7@gmail.com (G.C.); antounipa@gmail.com (A.G.); 91mariopace@gmail.com (M.P.); federica.vernuccio@unipa.it (F.V.); rob.cannella89@gmail.com (R.C.); massimo.galia@unipa.it (M.G.)
 - ² Department of Health Promotion, Mother and Child Care, Internal Medicine and Medical Specialties (PROMISE), University of Palermo, 90100 Palermo, Italy
 - ³ Section of Anaesthesia, Department of Surgical, Oncological and Oral Science (Di.Chir.On.S.), University of Palermo, 90100 Palermo, Italy; leonardosalvaggio95@gmail.com
 - ⁴ Section of Medical Oncology, Department of Surgical, Oncological and Oral Science, University of Palermo, 90100 Palermo, Italy; lauraw@hotmail.it (L.A.); giuseppe.badalamenti@unipa.it (G.B.)
 - ⁵ Section of Medical Oncology, Department of Biomedicine, Neuroscience and Advanced Diagnostics (Bi.N.D.), University of Palermo, 90100 Palermo, Italy; lorena.incorvaia@unipa.it
 - ⁶ Institute of Molecular Bioimaging and Physiology, National Research Council (IBFM-CNR), 90015 Cefalù, Italy
 - ⁷ Ri.MED Foundation, Via Bandiera 11, 90133 Palermo, Italy; acomelli@fondazionerimed.com
- * Correspondence: alessandro.stefano@ibfm.cnr.it



Citation: Salvaggio, G.; Cutaia, G.; Greco, A.; Pace, M.; Salvaggio, L.; Vernuccio, F.; Cannella, R.; Algeri, L.; Incorvaia, L.; Stefano, A.; et al. Deep Learning Networks for Automatic Retroperitoneal Sarcoma Segmentation in Computerized Tomography. *Appl. Sci.* **2022**, *12*, 1665. <https://doi.org/10.3390/app12031665>

Academic Editors: Qi-Huang Zheng and Fabio La Foresta

Received: 17 December 2021

Accepted: 2 February 2022

Published: 5 February 2022

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



Copyright: © 2022 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/>).

Featured Application: This study proposes fast and innovative deep learning networks for automatic retroperitoneal sarcoma segmentation in computerized tomography images.

Abstract: The volume estimation of retroperitoneal sarcoma (RPS) is often difficult due to its huge dimensions and irregular shape; thus, it often requires manual segmentation, which is time-consuming and operator-dependent. This study aimed to evaluate two fully automated deep learning networks (ENet and ERFNet) for RPS segmentation. This retrospective study included 20 patients with RPS who received an abdominal computed tomography (CT) examination. Forty-nine CT examinations, with a total of 72 lesions, were included. Manual segmentation was performed by two radiologists in consensus, and automatic segmentation was performed using ENet and ERFNet. Significant differences between manual and automatic segmentation were tested using the analysis of variance (ANOVA). A set of performance indicators for the shape comparison (namely sensitivity), positive predictive value (PPV), dice similarity coefficient (DSC), volume overlap error (VOE), and volumetric differences (VD) were calculated. There were no significant differences found between the RPS volumes obtained using manual segmentation and ENet (p -value = 0.935), manual segmentation and ERFNet (p -value = 0.544), or ENet and ERFNet (p -value = 0.119). The sensitivity, PPV, DSC, VOE, and VD for ENet and ERFNet were 91.54% and 72.21%, 89.85% and 87.00%, 90.52% and 74.85%, 16.87% and 36.85%, and 2.11% and -14.80% , respectively. By using a dedicated GPU, ENet took around 15 s for segmentation versus 13 s for ERFNet. In the case of CPU, ENet took around 2 min versus 1 min for ERFNet. The manual approach required approximately one hour per segmentation. In conclusion, fully automatic deep learning networks are reliable methods for RPS volume assessment. ENet performs better than ERFNet for automatic segmentation, though it requires more time.

Keywords: deep learning; soft tissue sarcoma; volume estimation; segmentation; artificial intelligence

1. Introduction

Soft tissue sarcomas are rare, malignant mesenchymal neoplasms that account for less than 1% of all malignant tumors. Of all sarcomas, the majority occur outside of the retroperitoneum, while around 10% of all sarcomas occur in the retroperitoneum [1], with a mean annual incidence of 2.7 per million [2]. The prognosis for patients with retroperitoneal sarcoma (RPS) is relatively poor, with a 36% to 58% overall 5-year survival rate and a natural history characterized by late recurrence [3]. RPS are frequently underdiagnosed at the early stage, and symptoms appear late, as they are associated with the displacement of adjacent organs and obstructive phenomena. When present, symptoms include abdominal pain, back pain, bowel obstruction, or palpable abdominal mass [1].

A variety of imaging techniques, including computed tomography (CT), positron emission tomography-computed tomography (PET/CT), and magnetic resonance imaging (MRI), may be used to assess RPS. Among them, CT is the most commonly used modality for the identification, localization, and staging of RPS [4]. CT examination allows for tissue components characterization and offers multiplanar reconstructions to easily depict the anatomic site of the origin of a mass, as well as its relationship to adjacent organs and vasculature [5].

RPS is one of the largest tumors of the human body [6]: lesions with a measure of <5 cm are considered rare, while a measure of >20 cm is found in 20 to 50% of masses at the time of resection [7]. Despite RPS's large dimension, the impact of tumor size on the patient's survival remains controversial. Several previous studies have failed to demonstrate any association of tumor size [8–16], while others have found that a size threshold of 10 cm is significant for survival [17,18]. However, these previous studies analyzed just the largest lesion's diameter instead of the volume. RPSs have an irregular shape, and the largest diameter cannot reflect the real tumor volume.

Manual segmentation based on imaging sections could be used for a volume estimation, and it is considered the gold standard for segmentation methods; however it is time-consuming, requires experience, and is strongly operator-dependent. Recently, deep learning methods, especially supervised classification methods based on convolutional neural networks, have been successful in the field of medical imaging for segmenting the anatomy of interest [19,20]. To our knowledge, no prior studies explored deep learning methods for RPS automatic segmentation.

Therefore, the aim of this study is to evaluate fully automated deep learning networks, namely the Efficient Neural Network (ENet) and the Efficient Residual Factorized ConvNet (ERFNet), for RPS segmentation and to compare their results with manual segmentation performed in contrast-enhanced CT examinations of the abdomen.

2. Materials and Methods

The present study is a retrospective study and written informed consent was waived by our ethical committee. All of the patients who underwent CT examination provided written informed consent for the use of their anonymized CT studies for research purposes.

2.1. Population Selection

The tumor registry database of the Department of Oncology of our hospital was queried for patients with RPS between 2013 and 2021. Patients with sarcomas that originated in the gastrointestinal tract (namely gastrointestinal stroma tumors) or in other abdominal visceral organs were excluded from the study due to the different imaging appearances and tumor shapes. The search retrieved 56 patients with histological diagnoses of RPS. From these patients, we identified 20 who underwent contrast-enhanced CT at our hospital. Thirteen of these patients had only CT examinations performed at the time of the diagnosis, while seven patients had CT examinations at the time of diagnosis and after treatments (one patient had three post-treatment CT examinations, two patients had four CTs, two patients had five, one patient had six, and one patient had eight). All CT examinations' images were reviewed by a radiologist (G.S.) with 20 years of experience in

abdominal radiology on a Pictures Archiving and Communication System (PACS—Impax, Agfa-Gevaert, Mortsels, Belgium), confirming the presence of RPS and, eventually, the recurrence in post-treatment CT scan examinations. When recurrence was found, post-treatment CT examinations were also included in the present study. Six post-treatment CT examinations were excluded for the lack of recurrence.

2.2. CT Imaging

All patients included in this study performed a standard protocol CT scan at the radiology department of our hospital. Patients underwent an abdominal contrast-enhanced computed tomography scan on a 16-slice CT scanner (General Electric BrightSpeed, Milwaukee, WI, USA). The scanning parameters were a tube current of 100 mAs, a peak tube voltage of 120 KV, a rotation time of 0.6 s, a detector collimation of 16×0.625 mm, a field of view of $350 \text{ mm} \times 350 \text{ mm}$, and a matrix of 512×512 . Contrast-enhanced CT was performed by injecting about 1.5 mL/kg of iodinated contrast agent (400 mg/mL Iomeprol, Iomeron 400, Bracco Imaging, Milan, Italy; 370 mg/dL Iopromide, Ultravist 370, Bayer Pharma, Berlin, Germany; or 350 mg/dL Iobitidrol, Xenetix 350, Guerbet, Roissy, France, depending on the clinical availability) at a flow of 3 mL/s, followed by the infusion of 20 mL of saline solution with a pump injector (Ulrich CT Plus 150, Ulrich Medical, Ulm, Germany). Images were acquired in the non-enhanced scan and portal-venous phase. Portal venous phase was obtained after 70 s of delay after intravenous injection of the contrast agent.

All CT examinations were performed with the patient in a supine position during a single inspiratory breath-hold whenever possible.

2.3. Manual Segmentation

Manual RPS segmentations were performed in consensus by two radiologists (A.G. and M.P.), both with 3 years of experience in abdominal CT. CT portal-venous phase images were used for manual segmentation. Each CT portal-venous examination was anonymized and imported into an open-source DICOM (Digital Imaging and Communication in Medicine), viewer equipped with Horos (LGPL license at [Horosproject.org](https://horosproject.org) [21]) in order to obtain volume lesion by manual segmentation. Lesion boundaries were manually traced with a contouring tool (pencil), slice by slice. Afterward, the “compute volume” tool was used to obtain the volume rendering of the entire RPS with the volume measurement. The time required by the entire process was dependent on the lesion’s dimension. Manual RPS segmentation volumes were used as a reference standard.

2.4. Automatic and User-Independent Segmentation

Deep learning enables automated identification and delineation of regions of interest in biomedical images, and consequently, it is of great interest in radiology. Nevertheless, it requires high computational power and long training times. In order to overcome this issue, ENet [22] and ERFNet [23] have been proposed as fast and lightweight networks capable of obtaining accurate segmentation with low training time and hardware requirements. Indeed, it was developed for fast inference and high accuracy in augmented reality and automotive scenarios where hardware availability is limited.

Specifically, ENet is based on building blocks of residual networks, with each block consisting of 3 convolutional layers. These are a 1×1 projection that reduces dimensionality, a regular convolutional layer, and a 1×1 expansion with batch normalization. ENet has asymmetric convolutions characterized by separable convolutions with sequences of 5×1 and 1×5 convolutions. The 5×5 convolution has 25 parameters, while the corresponding asymmetric convolution has only 10 parameters to reduce the network size. Finally, ENet uses a single starting block in addition to several variations of the bottleneck layer. The bottleneck layer is used to force the network to learn the most salient features present in the input and, consequently, to ignore the irrelevant parts.

ERFNet is optimized over ENet to improve accuracy and efficiency. The basic building block module is a non-bottleneck 1D layer comprised of two sets of factorized (separable or asymmetric) convolutions of a size of 3×1 followed by the 1×3 with rectified linear unit non-linearity. The input feature map of the main convolution path is added element-wise to the output of the convolution path, which represents the input of the next layer after applying the rectified linear unit non-linearity. The size of input is 512×512 , while the down-sampler block is similar to that of ENet architecture.

Furthermore, both ENet and ERFNet can be used as an alternative to transfer learning, usually used to compensate for the lack of labeled biomedical images. ENET and ERFNet can learn a lot of information, even from small datasets, as demonstrated in several biomedical segmentation studies [24–27], in which they outperformed other state-of-the-art deep learning approaches, such as UNet.

2.5. Experimental Detail

In the current study, to further overcome the issue related to a limited amount of data, the five-fold cross-validation strategy was adopted by randomly dividing the whole dataset into five folds (9 or 13 studies). Consequently, we trained five models by combining four of the five folds into a training set and by keeping the remaining fold as the validation fold. In our experiments, to avoid including slices from the same patient into the training and validation sets, the patients were firstly split randomly into a training set and validation set and then the slices corresponding to the patients were used to construct the training and validation sets. Therefore, the issue of potential overlap between the training and validation sets does not exist. We used an initial set of 20 lesions to determine the best learning rates experimentally. Specifically, for the training task, the following parameters were used: (i) a learning rate of 0.0001 and 0.00001 with Adam optimizer [28] for ENet and ERFNet, respectively; (ii) a batch size of 8 slices for all studies; (iii) a maximum of 100 epochs with an automatic stopping criterion (if the loss did not decrease for 10 consecutive epochs); and (iv) the Tversky loss function [29] with $\alpha = 0.3$ and $\beta = 0.7$. Data augmentation was obtained by randomly rotating; translating in both the x and y directions; and applying shearing, horizontal flip, and zooming to the input training slices. Consequently, six different types of data augmentation techniques were used to reduce overfitting, while data standardization and normalization were used to help the models converge faster and to avoid numerical instability. A graphics processing unit (GPU), i.e., NVIDIA QUADRO P4000 with 8 GB VRAM and 1792 CUDA Cores, was adopted to train and run inference.

2.6. Evaluation Analyses

Analysis of variance (ANOVA) on the manual and automatic segmentation was used to test the differences (a p -value < 0.05 indicates a significant difference) between the manual and automatic approaches. Specifically, the F-value was calculated to assess if the means between the two populations were significantly different. The F critical value was calculated to compare it with the F-value; if the F-value is larger than the F critical value, the null hypothesis can be rejected. In addition, RPS volumes calculated from the manual and automatic approaches were compared using a correlation graph and Bland Altman plot.

Finally, a set of performance indicators routinely used in the literature for shape comparison [30] were calculated, namely the dice similarity coefficient (DSC), sensitivity, volume overlap error (VOE), volumetric difference (VD), and positive predictive value (PPV):

$$DSC = \frac{2 * TP}{2 * TP + FP + FN} * 100\% \quad (1)$$

$$Sensitivity = \frac{TP}{TN + FN} * 100\% \quad (2)$$

$$VOE = \frac{1 - TP}{TP + FP + FN} * 100\% \quad (3)$$

$$VD = \frac{FN - FP}{2 * TP + FP + FN} * 100\% \quad (4)$$

$$PPV = \frac{TP}{TP + FN} * 100\% \quad (5)$$

where TP , FP , TN , and FN are the number of true positives, false positives, true negatives, and false negatives, respectively.

3. Results

3.1. Population

The final population consisted of 20 patients ($n = 11$ women; $n = 9$ men; mean age, 64 years old; age range, from 40 to 86 years old) with 49 CT examinations and 72 RPS lesions (one lesion in 40 CT examinations, two lesions in eight, three lesions in one, four lesions in one, and nine lesions in one). The most frequent histological subtype of sarcoma was liposarcoma ($n = 14$), while the remaining patients had leiomyosarcoma ($n = 6$). All of the liposarcomas and 4 leiomyosarcomas were poorly differentiated (G3; $n = 18$), while 2 leiomyosarcomas were moderately differentiated (G2).

3.2. ANOVA Analysis

The mean RPS volumes obtained using manual and automatic segmentations were 2697.57 cubic centimeters (cc) (SD = 4075.73 cc) for the manual approach, 2539.88 (SD = ± 3464.10 cc) for ENet, and 1701.48 cc (SD = ± 1996.71 cc) for ERFNet. Table 1 shows the minimum, 25th percentile, median, 75th percentile, and maximum RPS volume obtained using manual and automatic segmentation.

Table 1. Reference values of manually and automatically segmented RPS volumes from CT examinations in cubic centimeters (cc).

	Minimum	25th Percentile	Median	75th Percentile	Maximum
Manual	227.20 cc	515.30 cc	1108.00 cc	2836.00 cc	13,820.00 cc
ENet	205.90 cc	550.00 cc	1173.00 cc	3024.00 cc	11,780.00 cc
ERFNet	163.10 cc	519.50 cc	605.70 cc	1577.00 cc	6196.00 cc

Table 2 reports no significant difference between the RPS volumes obtained using the manual approach and ENet ($p = 0.935$), the manual approach and ERFNet ($p = 0.544$), or ENet and ERFNet ($p = 0.119$).

Table 2. Analysis of variance (ANOVA) on retroperitoneal sarcoma volumes showed no significant difference between manual and automatic segmentations.

	F-Value	F Critic Value	p-Value
ENet vs. Manual	0.0069	4.494	0.935
ERFNet vs. Manual	0.3854	4.494	0.544
ENet vs. ERFNet	2.3263	4.494	0.119

ENet, Efficient Neural Network; ERFNet, efficient residual factorized ConvNet.

The correlation graph showed a high positive correlation between manual and ENet segmentation [$r^2 = 0.99$] and a moderate correlation between manual and ERFNet segmentation [$r^2 = 0.79$], as shown in Figure 1a,c. In the same way, the Bland–Altman plots showed a high consistency between manual and ENet segmentation and a reproducibility coefficient (RPC) of ≤ 1400 cc or $\leq 16\%$ of values; values of 0% or 100% indicate a high or low consistency, respectively, and a low consistency between manual and ERFNet segmentation (RPC of ≤ 5200 cc, or $\leq 100\%$ of values), as shown in Figure 1b,d.

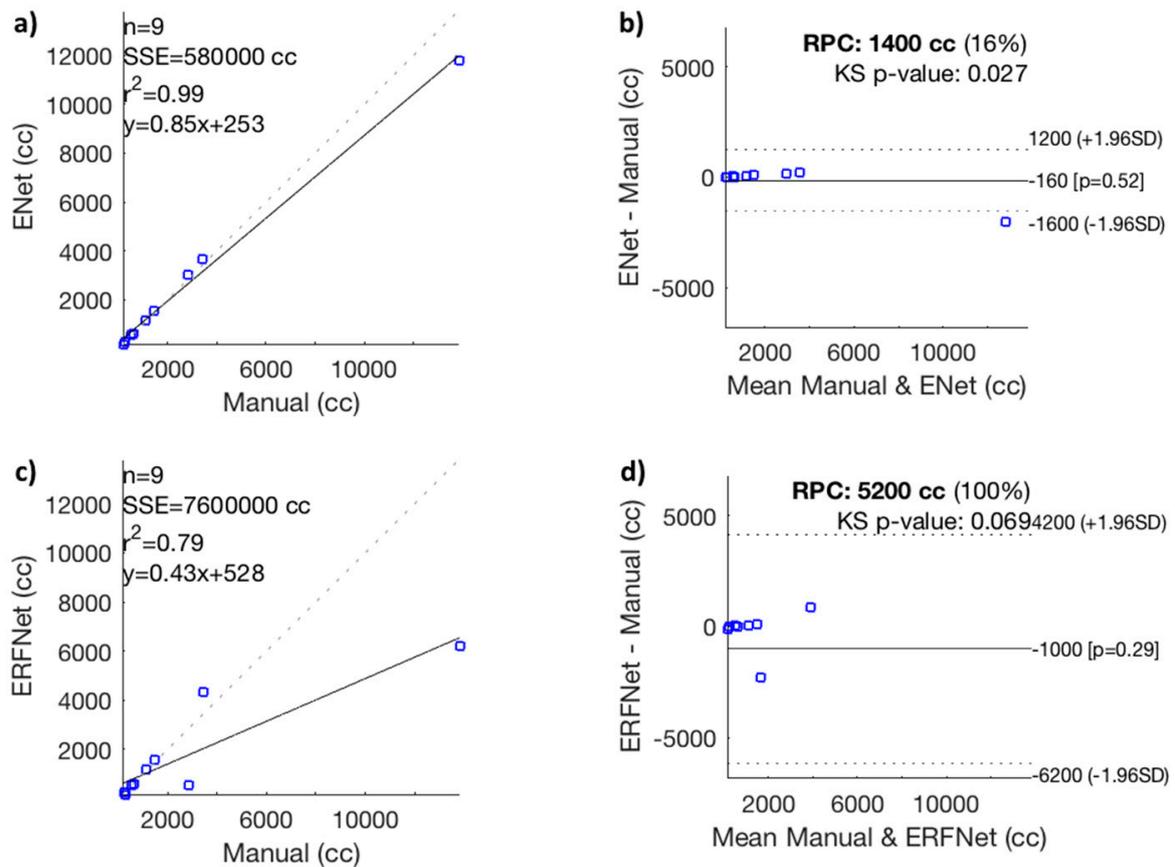


Figure 1. (a) Correlation graph, (b) Bland–Altman plot between manual and Efficient Neural Network (ENet) volumetric segmentation, (c) correlation graph, and (d) Bland–Altman plot between manual and Efficient Residual Factorized ConvNet (ERFNet) segmentation. RPC, reproducibility coefficient; SSE, sum of squared error; n, number of data points; *p*-value, Pearson correlation *p*-value; r^2 , Pearson r-value squared; cc, cubic centimeters; KS, Kolmogorov–Smirnov test; SD, standard deviation.

3.3. Performance Analysis

Table 3 illustrates the performance metrics obtained by comparing the automatic and manual delineations by averaging the results of the five validation folds during the fivefold cross-validation process. A DSC greater than 90% for ENet indicates excellent performances that justify the use of an automatic and independent operator method rather than the manual method, which, although more precise, is very time-consuming, as reported below.

Table 3. Performance results using ENet and ERFNet (fivefold cross-validation strategy).

	Sensitivity	PPV	DSC	VOE	VD
ENet					
Mean ± SD	91.54 ± 7.49%	89.85 ± 5.66%	90.52 ± 5.49%	16.87 ± 8.81%	2.11 ± 8.53%
±CI (95%)	4.89%	3.70%	3.59%	5.76%	5.58%
ERFNet					
Mean ± SD	72.21 ± 26.11%	87.00 ± 8.00%	74.85 ± 19.28%	36.85 ± 21.87%	−14.80 ± 33.32%
±CI (95%)	17.06%	5.23%	12.60%	14.29%	21.77%

ENet, Efficient Neural Network; ERFNet, efficient residual factorized ConvNet; PPV, positive predictive value; DSC, dice similarity coefficient; VOE, volume overlap error; VD, volumetric difference; SD, standard deviation; CI, confidence interval.

Figure 2 shows the training DSC and the loss function plots for one fold: a DSC > 90% was achieved in just ~20 iterations for ENet. ERFNet always turns out to be worse than ENet, except for the segmentation speed, as can be seen in Table 4.

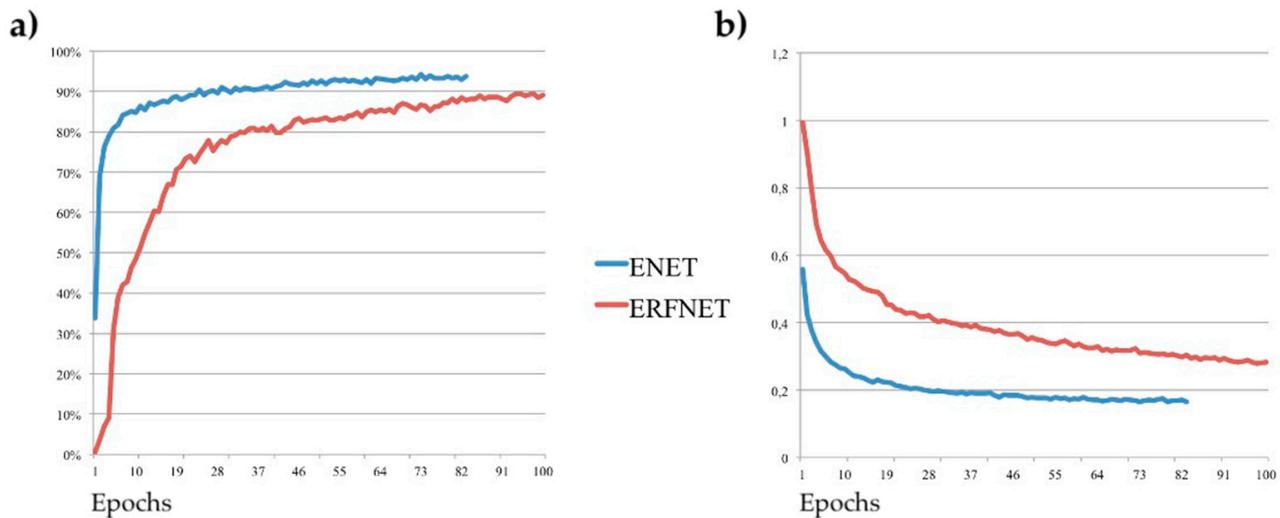


Figure 2. (a) Dice similarity coefficient and (b) loss function; Tversky loss plots for Efficient Neural Network (ENet) in red lines and Efficient Residual Factorized ConvNet (ERFNet) in blue lines during the training process for one fold.

Table 4. Comparison of computational complexity and performance between the two deep learning models.

	Number of Parameters		Size on Disk	Inference Times (s)/Dataset		Training Times (Days)/Dataset
	Trainable	Non-Trainable		CPU	GPU	GPU
ENet	363,069	8354	5.8 MB	113.10	15.64	5.31
ERFNet	2,056,440	0	25.3 MB	58.08	13.49	4.16

ENet, Efficient Neural Network; ERFNet, efficient residual factorized ConvNet; MB, megabyte; CPU, central processing unit; GPU, graphics processing unit; sec, seconds.

Specifically, using GPU hardware (NVIDIA QUADRO P4000 with 8 GB VRAM and 1792 CUDA Cores), ENet takes about 15 s for a whole segmentation versus 13 s of ERFNet. In the case of a CPU (Intel(R) Xeon(R) W-2125 CPU 4.00GHz processor), ENet takes about 2 min versus 1 min of ERFNet. The manual approach required an average time of 3887 ± 1600 s (approximately one hour per segmentation). Examples of the obtained segmentations are shown in Figure 3 for two patients with abdominal soft tissue sarcomas.

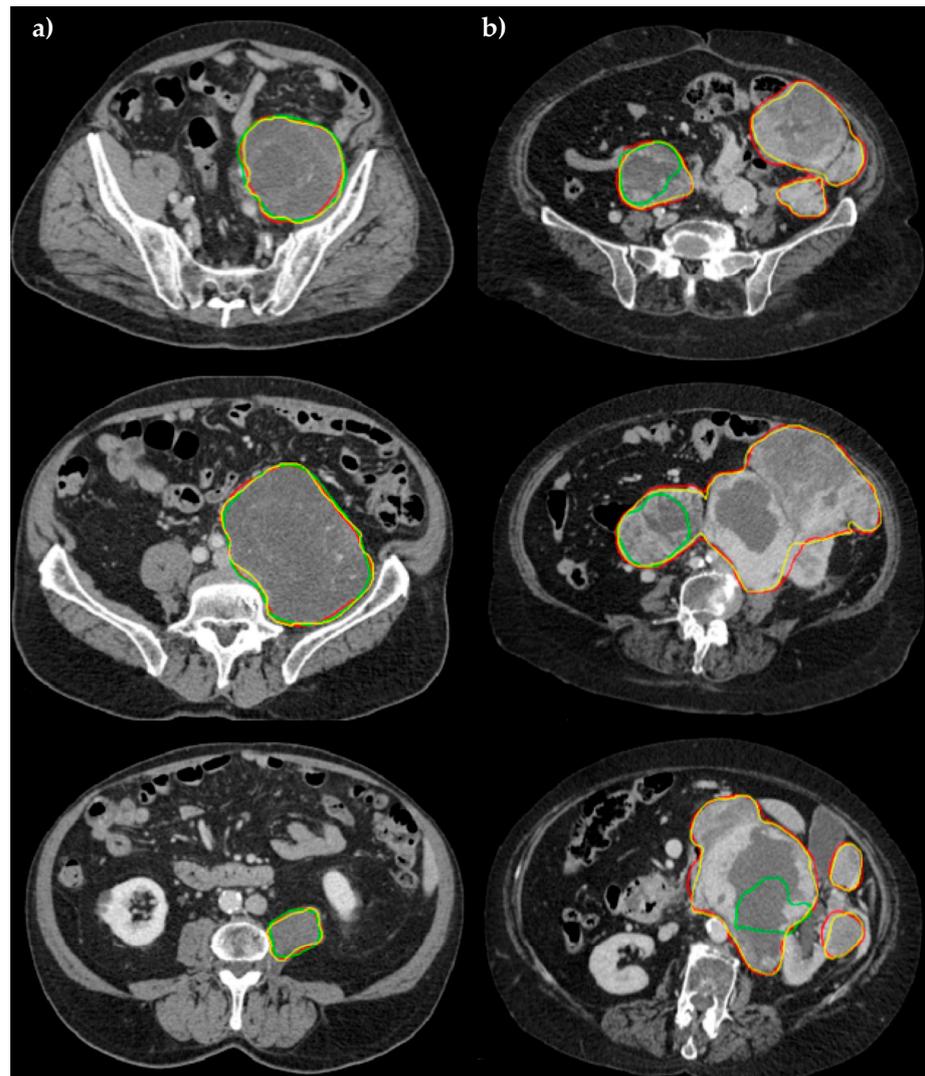


Figure 3. Comparison of segmentation performance in three different slices for (a) patient #005, and (b) patient #033. Manual (yellow), ENet (red), and ERFNet (green) segmentations are superimposed. In the first study, ENET and ERFNet obtained optimal DSCs (96.69% and 93%, respectively). In the second study, ERFNet showed poor DSC (30.55%) while ENET maintained similar performance compared to the first study (95%).

4. Discussion

4.1. Volume Estimation of Retroperitoneal Sarcoma

Due to the nonspecific presentation of their initial symptoms, RPSs are often seen during an initial evaluation on a CT scan as part of a general abdominal survey [31]. A contrast-enhanced CT allows for confirmation of the site and origin of the mass, and often the tissue composition [32], with the additional benefit of wide availability. An MRI is reserved for patients with an allergy to iodinated contrast agents or problem-solving when some finding is equivocal on a CT scan. It has been reported that different histotypes of RPS can differ widely in terms of both histologic features and clinical behavior; these different histopathologic characteristics may influence glucose metabolism and thus the ^{18}F -FDG uptake in PET examination [33]. Moreover, the difference observed in the maximum standardized uptake value (SUV_{max}) of RPS might be explained by the distinctions in cellularity and necrosis percentage in distinctive histotypes [34]. This different behavior may influence segmentation results. For example, Neabauer et al. [35] reported that the same tumor segmentation yields different results in MRI and PET scans: in an MRI scan,

necrosis is considered part of the tumor, but it is not visible on the PET scan, as the necrosis is no longer metabolically active.

The size of the RPS is an important prognostic criterion included in TNM staging (AJCC). Panda et al. [36] reported that size, as measured for the greatest tumor length, was significant in predicting an early relapse. Moreover, in Cox's proportional hazard model, the time-to-event analysis and large size and weight (together with other variables, such as higher age, male sex, incomplete resection, and high grade) become significant, predicting an early recurrence. There is also increasing evidence for the benefits of radiotherapy for RPS to improve local relapse-free survival. Ecker et al. [37] identified size as the only tumor-related variable associated with the use of neoadjuvant radiotherapy. However, all previous studies based their results on tumor size, as measured for the greatest tumor length. An RPS often has a highly irregular shape; therefore, tumors with the same axis length may have a different volume. Further studies are necessary to evaluate whether an overall volume greater than the maximum diameter changes the prognosis and therapy in the era of precision medicine.

4.2. Deep Learning Network and Volume Estimation

This study shows that automatic segmentations with deep learning networks and using portal-venous CT images are reliable methods for the automatic tumor volumetric measurements of RPS. The best performance for automatic segmentation was reached by ENet, with the VD between the automatic and manual segmentation at 2.11% for ENet and at -14.80% for ERFNet. Furthermore, we observed a VOE value of 16.87% for ENet and of 36.85% for ERFNet. These results indicate a low volumetric overlap error between the segmentation results and the manual segmentation, using ENet. To our knowledge, no other studies investigated the role of a deep learning network for RPS segmentation. Our results are concordant with prior investigations exploring the ENet for automatic segmentation in other organs. In a prior study, Lieman-Sifry et al. [38] presented the FastVentricle, an ENet variation, with skip connections for cardiac segmentation. They compared their results to that of the DeepVentricle, the architecture previously cleared by the FDA for clinical use. Both automatic segmentation methods had a median relative absolute error between 5% and 7%. A study involving 103 patients imaged with a prostate MRI [39] used ENet and UNet for the prostate gland segmentation, in comparison with the manual segmentation, which reported a VD of 6.85% and -3.11%. Comelli et al. [40] tested the ENet, UNet, and ERFNet for prostate gland segmentation, which reported mean VDs of 4.53%, 3.16%, and 5.70%, respectively. They obtained the best VOE value with ENet (16.50%), which is similar to the UNet and ERFNet (17.66% and 22.18%, respectively). The DSC is a widely used measure for evaluating medical image segmentation algorithms. It offers a standardized measure of segmentation accuracy, which has proven useful [41]. In the current study, the obtained DCS was 90.52% for ENet and 74.85% for ERFNet. The higher DSC value obtained in our dataset, using ENet, demonstrates a greater similarity to the manual segmentation contour and high-segmentation accuracy.

4.3. Timing Delineation

Our study showed that automatic and user-independent segmentation was much faster than manual segmentation for RPS volumetric analysis. Manual segmentation required an average time of 3887 s, while automatic segmentation required a few seconds. This would suggest good viability for the use of automatic segmentation in clinical practice, where time constraints may restrict which methods are used. ERFNet had better performance than ENet in terms of training times (4.16 days vs. 5.31 days) and in terms of inference times (58.08 s vs. 113.10 s, using a dedicated GPU, and 13.49 s vs. 14.64 s, using CPU). On the other hand, ENet had fewer order-of-magnitude parameters than ERFNet (363,069 vs. 2,056,440, respectively), requiring lower disk space (5.8 MB vs. 25.3 MB, respectively). However, in our experience, both algorithms can be used in a PC equipped with simple hardware and in portable devices, such as a tablet or smartphone.

4.4. Limitations of the Study

Our study has some limitations that need to be reported. We did not consider different tissue components of lesions. RPS are often highly heterogeneous, with variable tissue components that include cellular tumor, macroscopic fat, necrosis, and cystic change depending on the histopathological tumor subtype. These different components have different density values on CT images, which could influence automatic segmentation performances. Further studies with a larger population should be proposed. Moreover, the inter-observer variability for manual segmentation of CT images was not evaluated, as this was beyond the purpose of the current study. Segmentation from radiologists with different experience levels may provide different performances.

5. Conclusions

This study describes two deep learning automatic segmentation methods for RPS volume assessment without the need for user interaction. Starting from the ANOVA test, the results show that ENet and ERFNet obtained automatic segmentations similar to the manual segmentations, with no significant difference found between the two automatic volume estimations. However, ENet seems to perform much better than ERFNet considering the correlation graph and the Bland–Altman plot (Figure 1), the performance scores in Table 3, and the examples in Figure 3. The lack of significant difference is most likely due to the small sample size and high variance in volume size. Furthermore, although ERFNet was faster in the segmentation process using a CPU, ENet and ERFNet obtained similar segmentation times using a GPU. In the future, it would be desirable for the integration of deep learning networks with PACS systems to obtain fast and accurate RPS volume measurements.

Author Contributions: Conceptualization, G.S.; Data curation, A.C.; Formal analysis, A.C.; Investigation, G.B., L.I. and L.A.; Methodology, A.G., M.P. and L.S.; Project administration, M.G.; Resources, G.C.; Software, A.C.; Supervision, F.V. and R.C.; Validation, A.C.; Visualization, A.C.; Writing—original draft, G.S., G.C. and A.S.; Writing—review and editing, G.S., A.C. and A.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The proposed research has no implication on patient treatment. Review board approval was not sought: the proposed image analysis was performed offline and thus did not change the current treatment protocol.

Informed Consent Statement: The present study is a retrospective study and written informed consent was waived by our Ethical Committee.

Data Availability Statement: Data sharing not applicable.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Choi, J.H.; Ro, J.Y. Retroperitoneal Sarcomas: An Update on the Diagnostic Pathology Approach. *Diagnostics* **2020**, *10*, 642. [[CrossRef](#)]
2. Messiou, C.; Moskovic, E.; Vanel, D.; Morosi, C.; Benchimol, R.; Strauss, D.; Miah, A.; Douis, H.; van Houdt, W.; Bonvalot, S. Primary retroperitoneal soft tissue sarcoma: Imaging appearances, pitfalls and diagnostic algorithm. *Eur. J. Surg. Oncol.* **2017**, *43*, 1191–1198. [[CrossRef](#)]
3. Porter, G.A.; Baxter, N.N.; Pisters, P.W.T. Retroperitoneal sarcoma: A population-based analysis of epidemiology, surgery, and radiotherapy. *Cancer* **2006**, *106*, 1610–1616. [[CrossRef](#)]
4. Varma, D.G. Imaging of soft-tissue sarcomas. *Curr. Oncol. Rep.* **2000**, *2*, 487–490. [[CrossRef](#)]
5. Levy, A.D.; Manning, M.A.; Al-Refaie, W.B.; Miettinen, M.M. Soft-tissue sarcomas of the abdomen and pelvis: Radiologic-pathologic features, part 1—Common sarcomas. *Radiographics* **2017**, *37*, 462–483. [[CrossRef](#)]
6. Liles, J.S.; Tzeng, C.W.D.; Short, J.J.; Kulesza, P.; Heslin, M.J. Retroperitoneal and Intra-Abdominal Sarcoma. *Curr. Probl. Surg.* **2009**, *46*, 445–503. [[CrossRef](#)]

7. Matthyssens, L.E.; Creytens, D.; Ceelen, W.P. Retroperitoneal Liposarcoma: Current Insights in Diagnosis and Treatment. *Front. Surg.* **2015**, *2*. [[CrossRef](#)]
8. Singer, S.; Corson, J.M.; Demetri, G.D.; Healey, E.A.; Marcus, K.; Eberlein, T.J. Prognostic factors predictive of survival for truncal and retroperitoneal soft-tissue sarcoma. *Ann. Surg.* **1995**, *221*, 185–195. [[CrossRef](#)]
9. Stoeckle, E.; Coindre, J.M.; Bonvalot, S.; Kantor, G.; Terrier, P.; Bonichon, F.; Bui, B.N. Prognostic factors in retroperitoneal sarcoma: A multivariate analysis of a series of 165 patients of the French Cancer Center Federation Sarcoma Group. *Cancer* **2001**, *92*, 359–368. [[CrossRef](#)]
10. Heslin, M.J.; Lewis, J.J.; Nadler, E.; Newman, E.; Woodruff, J.M.; Casper, E.S.; Leung, D.; Brennan, M.F. Prognostic factors associated with long-term survival for retroperitoneal sarcoma: Implications for management. *J. Clin. Oncol.* **1997**, *15*, 2832–2839. [[CrossRef](#)]
11. Van Dalen, T.; Hennipman, A.; Van Coevorden, F.; Hoekstra, H.J.; Van Geel, B.N.; Slootweg, P.; Lutter, C.F.A.; Brennan, M.F.; Singer, S. Evaluation of a clinically applicable post-surgical classification system for primary retroperitoneal soft-tissue sarcoma. *Ann. Surg. Oncol.* **2004**, *11*, 483–490. [[CrossRef](#)]
12. Gronchi, A.; Casali, P.G.; Fiore, M.; Mariani, L.; Lo Vullo, S.; Bertulli, R.; Colecchia, M.; Lozza, L.; Olmi, P.; Santinami, M.; et al. Retroperitoneal soft tissue sarcomas: Patterns of recurrence in 167 patients treated at a single institution. *Cancer* **2004**, *100*, 2448–2455. [[CrossRef](#)]
13. Perez, E.A.; Gutierrez, J.C.; Moffat, F.L.; Franceschi, D.; Livingstone, A.S.; Spector, S.A.; Levi, J.U.; Sleeman, D.; Koniaris, L.G. Retroperitoneal and truncal sarcomas: Prognosis depends upon type not location. *Ann. Surg. Oncol.* **2007**, *14*, 1114–1122. [[CrossRef](#)]
14. Bonvalot, S.; Rivoire, M.; Castaing, M.; Stoeckle, E.; Le Cesne, A.; Blay, J.Y.; Laplanche, A. Primary retroperitoneal sarcomas: A multivariate analysis of surgical factors associated with local control. *J. Clin. Oncol.* **2009**, *27*, 31–37. [[CrossRef](#)]
15. Lewis, J.J.; Leung, D.; Woodruff, J.M.; Brennan, M.F. Retroperitoneal soft-tissue sarcoma: Analysis of 500 patients treated and followed at a single institution. *Ann. Surg.* **1998**, *228*, 355–365. [[CrossRef](#)]
16. Gronchi, A.; Lo Vullo, S.; Fiore, M.; Mussi, C.; Stacchiotti, S.; Collini, P.; Lozza, L.; Pennacchioli, E.; Mariani, L.; Casali, P.G. Aggressive surgical policies in a retrospectively reviewed single-institution case series of retroperitoneal soft tissue sarcoma patients. *J. Clin. Oncol.* **2009**, *27*, 24–30. [[CrossRef](#)]
17. Haas, R.L.; Baldini, E.H.; Chung, P.W.; van Coevorden, F.; DeLaney, T.F. Radiation therapy in retroperitoneal sarcoma management. *J. Surg. Oncol.* **2018**, *117*, 93–98. [[CrossRef](#)]
18. Nathan, H.; Raut, C.P.; Thornton, K.; Herman, J.M.; Ahuja, N.; Schulick, R.D.; Choti, M.A.; Pawlik, T.M. Predictors of survival after resection of retroperitoneal sarcoma: A population-based analysis and critical appraisal of the AJCC Staging system. *Ann. Surg.* **2009**, *250*, 970–976. [[CrossRef](#)]
19. Cutaia, G.; La Tona, G.; Comelli, A.; Vernuccio, F.; Agnello, F.; Gagliardo, C.; Salvaggio, L.; Quartuccio, N.; Sturiale, L.; Stefano, A.; et al. Radiomics and Prostate MRI: Current Role and Future Applications. *J. Imaging* **2021**, *7*, 34. [[CrossRef](#)]
20. Tian, Z.; Liu, L.; Fei, B. Deep convolutional neural network for prostate MR segmentation. In *Medical Imaging 2017: Image-Guided Procedures, Robotic Interventions, and Modeling*; SPIE: Bellingham, WA, USA, 2017; Volume 10135, p. 101351L.
21. Available online: <https://horosproject.org> (accessed on 4 February 2021).
22. Paszke, A.; Chaurasia, A.; Kim, S.; Culurciello, E. ENet: A Deep Neural Network Architecture for Real-Time Semantic Segmentation. *arXiv* **2016**, arXiv:1606.02147.
23. Romera, E.; Alvarez, J.M.; Bergasa, L.M.; Arroyo, R. ERFNet: Efficient Residual Factorized ConvNet for Real-Time Semantic Segmentation. *IEEE Trans. Intell. Transp. Syst.* **2018**, *19*, 263–272. [[CrossRef](#)]
24. Cuocolo, R.; Comelli, A.; Stefano, A.; Benfante, V.; Dahiya, N.; Stanzione, A.; Castaldo, A.; De Lucia, D.R.; Yezzi, A.; Imbriaco, M. Deep Learning Whole-Gland and Zonal Prostate Segmentation on a Public MRI Dataset. *J. Magn. Reson. Imaging* **2021**, *54*, 452–459.
25. Comelli, A.; Dahiya, N.; Stefano, A.; Benfante, V.; Gentile, G.; Agnese, V.; Raffa, G.M.; Pilato, M.; Yezzi, A.; Petrucci, G.; et al. Deep learning approach for the segmentation of aneurysmal ascending aorta. *Biomed. Eng. Lett.* **2021**, *11*, 15–24. [[CrossRef](#)]
26. Comelli, A.; Coronello, C.; Dahiya, N.; Benfante, V.; Palmucci, S.; Basile, A.; Vancheri, C.; Russo, G.; Yezzi, A.; Stefano, A. Lung Segmentation on High-Resolution Computerized Tomography Images Using Deep Learning: A Preliminary Step for Radiomics Studies. *J. Imaging* **2020**, *6*, 125. [[CrossRef](#)]
27. Stefano, A.; Comelli, A. Customized efficient neural network for covid-19 infected region identification in ct images. *J. Imaging* **2021**, *7*, 131. [[CrossRef](#)]
28. Kingma, D.P.; Ba, J.L. Adam: A method for stochastic optimization. *arXiv* **2015**, arXiv:1412.6980.
29. Salehi, S.S.M.; Erdogmus, D.; Gholipour, A. Tversky loss function for image segmentation using 3D fully convolutional deep networks. In *Proceedings of the Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*; Springer: Cham, Switzerland, 2017.
30. Alongi, P.; Stefano, A.; Comelli, A.; Laudicella, R.; Scalisi, S.; Arnone, G.; Barone, S.; Spada, M.; Purpura, P.; Bartolotta, T.V.; et al. Radiomics analysis of 18F-Choline PET/CT in the prediction of disease outcome in high-risk prostate cancer: An explorative study on machine learning feature classification in 94 patients. *Eur. Radiol.* **2021**, *31*, 4595–4605. [[CrossRef](#)]
31. Francis, I.R.; Cohan, R.H.; Varma, D.G.K.; Sondak, V.K. Retroperitoneal sarcomas. *Cancer Imaging* **2005**, *5*, 89–94. [[CrossRef](#)]

32. Morosi, C.; Stacchiotti, S.; Marchianò, A.; Bianchi, A.; Radaelli, S.; Sanfilippo, R.; Colombo, C.; Richardson, C.; Collini, P.; Barisella, M.; et al. Correlation between radiological assessment and histopathological diagnosis in retroperitoneal tumors: Analysis of 291 consecutive patients at a tertiary reference sarcoma center. *Eur. J. Surg. Oncol.* **2014**, *40*, 1662–1670. [[CrossRef](#)]
33. Schwarzbach, M.H.M.; Dimitrakopoulou-Strauss, A.; Willeke, F.; Hinz, U.; Strauss, L.G.; Zhang, Y.M.; Mechtersheimer, G.; Attigah, N.; Lehnert, T.; Herfarth, C. Clinical value of [18-F] fluorodeoxyglucose positron emission tomography imaging in soft tissue sarcomas. *Ann. Surg.* **2000**, *231*, 380–386. [[CrossRef](#)]
34. Sambri, A.; Bianchi, G.; Longhi, A.; Righi, A.; Donati, D.M.; Nanni, C.; Fanti, S.; Errani, C. The role of 18F-FDG PET/CT in soft tissue sarcoma. *Nucl. Med. Commun.* **2019**, *40*, 626–631. [[CrossRef](#)]
35. Neubauer, T.; Wimmer, M.; Berg, A.; Major, D.; Lenis, D.; Beyer, T.; Saponjski, J.; Bühler, K. Soft Tissue Sarcoma Co-Segmentation in Combined MRI and PET/CT Data. *Lect. Notes Comput. Sci.* **2020**, *12445*, 97–105.
36. Panda, N.; Das, R.; Banerjee, S.; Chatterjee, S.; Gumta, M.; Bandyopadhyay, S.K. Retroperitoneal Sarcoma. Outcome Analysis in a Teaching Hospital in Eastern India- a Perspective. *Indian J. Surg. Oncol.* **2015**, *6*, 99–105. [[CrossRef](#)] [[PubMed](#)]
37. Ecker, B.L.; Peters, M.G.; McMillan, M.T.; Sinnamon, A.J.; Zhang, P.J.; Fraker, D.L.; Levin, W.P.; Roses, R.E.; Karakousis, G.C. Preoperative radiotherapy in the management of retroperitoneal liposarcoma. *Br. J. Surg.* **2016**, *103*, 1839–1846. [[CrossRef](#)] [[PubMed](#)]
38. Lieman-Sifry, J.; Le, M.; Lau, F.; Sall, S.; Golden, D. Fastventricle: Cardiac segmentation with ENet. *Lect. Notes Comput. Sci.* **2017**, *10263*, 127–138.
39. Salvaggio, G.; Comelli, A.; Portoghese, M.; Cutaia, G.; Cannella, R.; Vernuccio, F.; Stefano, A.; Dispensa, N.; La Tona, G.; Salvaggio, L.; et al. Deep Learning Network for Segmentation of the Prostate Gland With Median Lobe Enlargement in T2-weighted MR Images: Comparison With Manual Segmentation Method. *Curr. Probl. Diagn. Radiol.* **2021**. [[CrossRef](#)] [[PubMed](#)]
40. Comelli, A.; Dahiya, N.; Stefano, A.; Vernuccio, F.; Portoghese, M.; Cutaia, G.; Bruno, A.; Salvaggio, G.; Yezzi, A. Deep learning-based methods for prostate segmentation in magnetic resonance imaging. *Appl. Sci.* **2021**, *11*, 782. [[CrossRef](#)] [[PubMed](#)]
41. Carass, A.; Roy, S.; Gherman, A.; Reinhold, J.C.; Jesson, A.; Arbel, T.; Maier, O.; Handels, H.; Ghafoorian, M.; Platel, B.; et al. Evaluating White Matter Lesion Segmentations with Refined Sørensen-Dice Analysis. *Sci. Rep.* **2020**, *10*, 1–19. [[CrossRef](#)] [[PubMed](#)]