# Esophagus segmentation from planning CT images using an atlas-based deep learning approach

João Otávio Bandeira Diniz[a,b,*], Jonnison Lima Ferreira[a,c], Pedro Henrique Bandeira Diniz[d], Aristófanes Corrêa Silva[a], Anselmo Cardoso de Paiva[a]

[a]*Federal University of Maranhão, Brazil*
[b]*Federal Institute of Maranhão, Brazil*
[c]*Federal Institute of Amazonas - IFAM, Manaus, AM, Brazil*
[d]*University of Oxford, Oxford, United Kingdom*

## Abstract

**Background and Objective**: One of the main steps in the planning of radiotherapy (RT) is the segmentation of organs at risk (OARs) in Computed Tomography (CT). The esophagus is one of the most difficult OARs to segment. The boundaries between the esophagus and other surrounding tissues are not well-defined, and it is presented in several slices of the CT. Thus, manually segment the esophagus requires a lot of experience and takes time. This difficulty in manual segmentation combined with fatigue due to the number of slices to segment can cause human errors. To address these challenges, computational solutions for analyzing medical images and proposing automated segmentation have been developed and explored in recent years. In this work, we propose a fully automatic method for esophagus segmentation for better planning of radiotherapy in CT.

**Methods**: The proposed method is a fully automated segmentation of the esophagus, consisting of 5 main steps: (a) image acquisition; (b) VOI segmentation; (c) preprocessing; (d) esophagus segmentation; and (e) segmentation refinement.

**Results**: The method was applied in a database of 36 CT acquired from 3 different institutes. It achieved the best results in literature so far: Dice

---

*Corresponding author
*Email address:* `joao.bandeira@ifma.edu.br` (João Otávio Bandeira Diniz)

coefficient value of 82.15%, Jaccard Index of 70.21%, accuracy of 99.69%, sensitivity of 90.61%, specificity of 99.76%, and Hausdorff Distance of 6.1030 *mm*.

**Conclusions**: With the achieved results, we were able to show how promising the method is, and that applying it in large medical centers, where esophagus segmentation is still an arduous and challenging task, can be of great help to the specialists.

*Keywords:* Convolutional neural networks, Computed tomography, Esophagus segmentation, Organs at Risk, Radiotherapy.

## 1. Introduction

Radiotherapy (RT) is a treatment used in about 50% of cancer cases. Because of its efficacy and availability, RT is a widely used treatment, especially when surgery and chemotherapy introduce high risk to the patient's life. About 40% of cured cancer patients have RT as part of their overall treatment [1]. Depending on the type of tumor and the stage of cancer, RT may be the solution to cure it [2–4].

Over the years, more and more studies about radiation therapy have emerged not only for addressing the treatment but also its toxic radiation effects on patients [5, 6]. However, when the RT is performed by a group of professionals aided by computational frameworks, it is possible to reduce these unwanted effects [4, 7, 8]. Reduce them, of course, is very important since there is no point in treating a patient's disease while applying toxic effects on healthy tissues.

Protecting the healthy tissues surrounding the lesions is necessary to eradicate them. These tissues comprise various organs, depending on where the tumor is located. The tissues of healthy organs that need to be protected in the radiotherapy are called as Organs at Risk [9, 10].

A three-dimensional model of the patient's body is generated in the RT planning phase, typically with the use of a computed tomography (CT), known as radiotherapy planning CT. In this stage, there is the segmentation of

2

tumors and OARs. This process is usually done with the support of medical software. The most used is the Varian Eclipse Treatment Planning System [11]. This software presents a Smart Segmentation module that provides an initial segmentation of the organs chosen by the specialist. However, the specialist needs to validate this segmentation slice by slice and, if necessary, adjust the segmentation boundaries.

An extremely important OAR is the esophagus. According to Fechter et al. [12] the esophagus is an important organ and is very difficult to segment. In some slices of the volume, even seasoned specialists have difficulty in defining its boundaries reliably, and they can easily be confused with other structures. In addition, the interpretation of CT slices is very repetitive. These issues make the whole process time-consuming, error-prone, and highly dependable to the variability between observers. Nevertheless, the esophagus is one of the most important OARs in RT due to its radiosensitive mucosa, so a precise segmentation is indispensable. Thus, there is a vital need for automated, reproducible, and consistent approaches for segmenting the esophagus in RT.

Hence, for preparing the CT for radiotherapy, this work proposes an automatic approach to segment the esophagus automatically. The proposed research aims to give a second opinion to the specialist and to assist him in this process.

The method is composed of five steps. Firstly, the materials used by the method are described. Secondly, an initial segmentation is performed, which will generate a volume of interest (VOI) in the planning CT, extracting just a small region containing the esophagus. Third, the VOI is preprocessed to highlight the structures of the esophagus. Fourth, a novel convolutional neural network (CNN) is applied to automatically segment the esophagus. Finally, refinement techniques are performed to generate the final segmentation.

By creating this method, we believe we have achieved the following contributions:

3

1. An automatic and robust method composed of techniques existent in the literature and new techniques, achieving results superior to those published so far for esophagus segmentation in CT.

2. An automatic pre-processing technique that is capable of reducing the volume in a standardized size region containing the esophagus and enhancing its boundaries for better segmentation.

3. A post-processing technique to reduce false-positive and improve the segmentation boundaries.

Thus, we believe that if this method is applied in large medical centers, it can be an important ally for the specialist in the segmentation task of the esophagus.

In addition to this section, the paper presents five more sections. First, the Section 2 presents the related works and explains the advantages of the proposed method over them. In the Section 3, it is detailed the materials and the proposed method for esophagus segmentation. The results achieved by applying the method on the materials are presented in Section 4. In the Section 5, it is discussed the results and presented an overview of the proposed method, highlighting its advances and limitations. Conclusion and future works are presented in Section 6.

## 2. Related works

Studies about esophagus segmentation can be found in the literature in several types of images [13–17]. Also, several approaches propose esophagus segmentation in a semi-automatic fashion. Most authors report that it is difficult to distinguish the esophagus on CT, even for specialists. This happens because of its ill-defined contrast and boundary in CT, turning the segmentation an arduous task. Furthermore, due to its presence in many slices of the volume, the manual segmentation also becomes an exhaustive process. Following, the main works regarding the segmentation of the esophagus in CT will be presented.

4

In Feulner et al. [13], a probabilistic method for esophagus segmentation is proposed. According to the authors, the method consists of some steps. First, a detector trained to learn a discrimination model of appearance is combined with an explicit model of the distribution of respiratory and esophageal air. In the next step, prior knowledge of the previous step is incorporated using a Markov chain model. Then, a "detect and connect" approach is used to obtain the maximum posterior estimate of the approximate shape of the esophagus from the hypotheses about its contour in axial slices. Finally, the surface resulted from this approach is deformed in a non-rigid manner to better adjust to the limits of the organ. The authors' experiments are performed on a private database composed of 144 CTs and achieved a Dice coefficient value of 74%. It is important to note that this method is highly dependent on descriptive models and prior knowledge. Yet, a non-rigid record is necessary, causing deformities in the initial esophageal structure.

The work proposed by Grosgeorge et al. [14] highlights the difficulty in the esophagus segmentation task, since its boundary is not well-defined and has low CT contrast. According to the authors, their method for esophagus segmentation is original. This method achieves its results using a skeleton-shaped model to guide segmentation. Besides, it consists of two steps: a 3D segmentation using a graph cut technique with skeletonization techniques, followed by a 2D propagation. The work is applied to a database composed of six patients and presents a Dice coefficient value equal to 61%. Although it is an automatic method, it depends on the creation of a skeleton model based on the specialist marking.

The work by Larsson et al. [15] proposes a method for the automatic segmentation of various abdominal organs on CT. The first step locates the organs using an atlas model. To define the atlas, the authors use image registration. Next, each voxel is classified by two convolutional neural networks. The authors disregard structures based on a threshold and keep only connected regions. Their experiments are applied in a database of 30 patients, achieving a Dice coefficient value of 66.2% for esophagus segmentation. It is worth

5

mentioning that the work uses the atlas as an initial segmentation, defining the probability of a voxel being a specific organ based on the atlas values. Then, a convolutional neural network decides whether these voxels are the target organ or not. In this type of approach, in addition to not classifying the entire volume, the classification of voxels is completely dependent on the atlas.

Once again, the difficulty in esophagus segmentation is highlighted by Trullo et al. [18]. The authors explain how crucial the segmentation of this organ is for RT. However, they also emphasize that in a recent literature review, there are no contributions to esophagus segmentation [19]. According to the authors, their work presented a fully automatic method consisting of only two steps. In the first one, a convolutional neural network estimates the location of the esophagus (this network was previously published by the same authors in other paper, and presented a Dice of 66% for esophagus segmentation [16]). Then, in the second step, the region estimated in the first step is cropped, and the same network is applied again to it. The method was tested on 30 CTs and presented a Dice coefficient value of 72%. Like most works, this highlights the need for a crop in the esophagus location. This operation is necessary, as the authors themselves point out, because the first network is not able to correctly segment the esophagus, so the crop reduces the scope of the problem for the second network execution. However, when observing the fact that the first network may ignore the location of the esophagus, or mark it in the wrong region, the second network will also not be able to segment it.

One of the most recent works that deal with the segmentation of the esophagus as Organ at Risk (OAR) is proposed in Fechter et al. [12]. The authors highlight the difficulty in esophagus segmentation even by a specialist, being a task that costs time and is prone to human errors. Their method applies a CNN to generate a probabilistic model that serves as input for an active contour model (ACM). Then, both CNN and ACM outputs are applied to a random walker algorithm based on the Hounsfield Units (HU) of the volumes. The method executes on a database composed of 50 patients, reaching a Dice coefficient result of 76%. Although a CNN is part of steps, the final segmentation

6

depends on the Hounsfield scale of the image. This makes the model susceptible to errors when there is a sudden change in HU, especially in cases where the CT acquisition process is not well defined.

In Dong et al. [17], a method for the segmentation of OARs in planning CT is proposed. Their method combines an adversarial network with a fully convolutional network (FCN) to achieve the results. More specifically, as the esophagus comprises only a small region of the CT exam, the method crops the volume based on the location of the lungs. The authors assume the esophagus is located in the centroid of the slice with the largest lung area. Since this specific slice is necessary to crop the volume, the network firstly segments the lungs, and after the crop, it segments the esophagus. The authors present their experiments in a database of 36 patients. However, only 35 were used to develop the method. The results are measured using the leave-one-out technique, averaging the results of the 35 executions. Among the other OARs, the esophagus has a Dice coefficient value of 75%, sensitivity of 73%, and specificity of 99%. Although the work is very promising, it relies on a precise segmentation of the lung for the extraction of the volume of interest where the esophagus may be.

The work proposed by Chen et al. [20] performed a semantic segmentation of the esophagus using a U-Net. Their database contains 15 CTs totalizing more than two thousand slices. Although the segmentation is 2D, the resulting masks are stacked to create a 3D one. The result found is an average Dice value of 79%. Despite being a recent work and using a well-known network, it is only applied to 15 exams.

In turn, Feng et al. [21]'s work presents a segmentation method using deep learning approaches for various organs. The method uses a crop for each organ based on its previous location. In the end, the segmentations of each organ are joined to generate the final multi-organ segmentation. The method is trained and tested on a database composed of 36 CTs. For the esophagus, the method can reach an average Dice value of 72%. In this work, it is observed the importance of the volume crops so that the segmentations are more accurate.

7

However, the method has a low Dice value compared to other studies in the literature.

As can be seen in the literature, most studies highlight three very important facts when it comes to esophagus segmentation. First, it is an organ that is very difficult to segment, even for specialists, and there is no clear distinction between the contrast of the esophagus boundary with the other organs in most slices. Second, many papers show the need to segment this OAR as an important step in the treatment by RT, so there is a lot of interest in automatic methods. Finally, the works often do some localization of the esophagus based either on probabilistic models, on atlases or other organ's location to reduce the region of interest (crop).

Thus, based on the importance of the segmentation of this organ and the need for more robust methods, this work proposes an automatic method for esophagus segmentation. The method takes into account the problems highlighted in the literature, presenting an effective way to locate the esophagus through image registration followed by a crop in the volume of interest using atlases. Then, when observing the need to enhance the esophagus contrast and boundaries against other CT structures, a pre-processing technique for enhancing the structures is proposed. Next, these images are applied to an FCN architecture, showing promising results. Finally, a postprocessing technique is used to keep only the largest connected object within the volume.

### 3. Materials and method

For the segmentation of the esophagus in planning CT, a 5-step method is proposed. The first one is the image acquisition step (Figure 1(a)). The database used in this work is a database for the segmentation of OARs in CT. Esophagus, as many other studies report, is one of the most challenging OARs to segment. Even specialists find it difficult to confidently define the boundaries of the esophagus on CT since its texture is easily confused with other structures around it.

In the second step, volume of interest (VOI) segmentation (Figure 1(b)), the spatial information of an esophagus atlas is used to make a three-dimensional crop in the volume. Then, in the third step, a pre-processing is done to enhance the structures of the esophagus (Figure 1(c)). After the previous steps, in the fourth step (Figure 1(d)), the VOI slices are processed to create the esophagus segmentation. For that, a U-Net with residual blocks use the VOI slices as input and creates a model to suggest an initial segmentation. Finally, a refinement technique is applied to it, resulting in a better-delineated segmentation, i.e., the final segmentation (Figure 1(e)). In the next sections, the details of each step will be explained. Figure 1 describes the steps of the method

### 3.1. Image acquisition

The image database of planning CT used in this work is the database of the AAPM Thoracic Auto-segmentation challenge. The main objective of this challenge is to encourage participants to create algorithms for the segmentation of OARs in CT acquired from patients who will undergo radiotherapy. An example of a volume slice is shown in Figure 2.

The challenge database consists of 36 patients from 3 different institutes (12 patient volumes for each institute). The database also provides the markings of the OARs in files called Radiotherapy Structure Set (RT-STRUCT). Each volume has its respective RT-STRUCT file and each file contains markings of 5 organs (left lung, right lung, heart, esophagus, and spinal cord). The information and results for this challenge are provided in Yang et al. [22].

In addition to the database being composed of volumes from three institutes following different acquisition protocols, another issue is the number of slices of each volume that can range from 134 to 288. This information shows how diverse the database is. Examples of volumes from the 3 institutes are shown in Figure 3.

To make visualization better, and to demonstrate exams in three dimensions, we rotated the image to reflect the entire patient. This can be seen in Figure 3(d), Figura 3(e), and Figura 3(f).
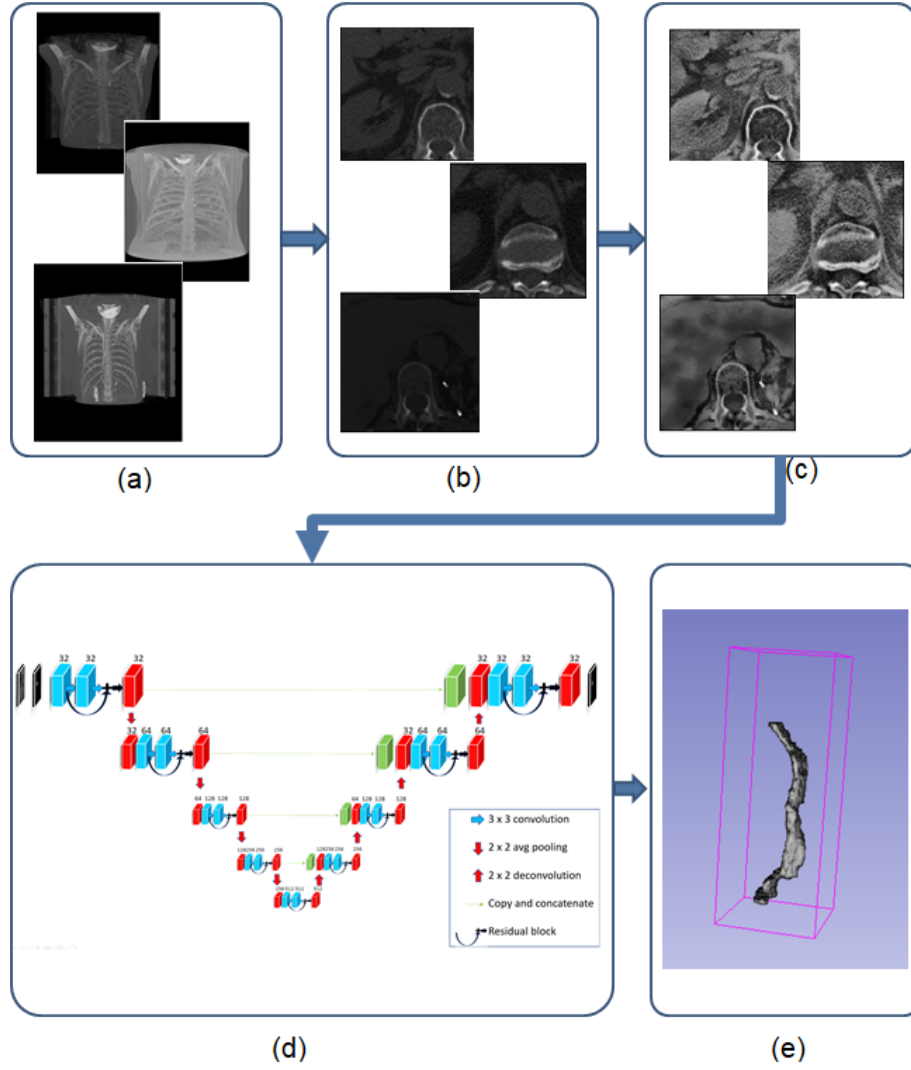
Figure 1: Flowchart of the esophagus method. (a) image acquisition (CT exam rotated for better visualization); (b) VOI segmentation; (c) pre-processing; (d) esophagus segmentation; and (e) segmentation refinement

## 3.2. VOI segmentation

As already mentioned in Section 2, the esophagus is one of the most problematic organs for segmentation in CT. In addition to these image-related problems, there are still other ones related to the anatomical structure of
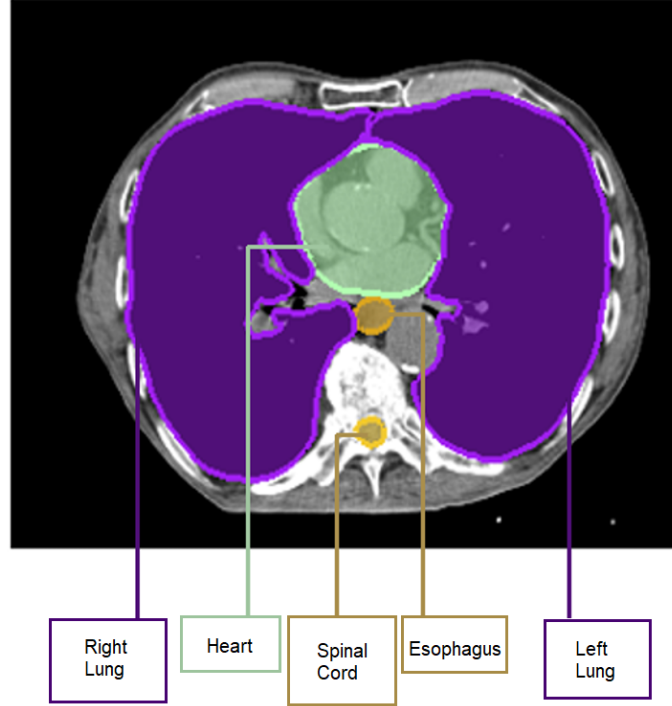
10

Figure 2: Image acquisition: lungs (purple), heart (green), esophagus (yellow on top) and spinal cord (yellow on the bottom).

the esophagus. For example, the esophagus is a small organ, so it occupies a few slices in the planning CT. For this reason, once again the literature shows the importance of delimiting a region of interest as an initial step in the segmentation process [13, 15, 17, 18].

To address the problems of the diversity of the database, difficulties in defining the esophagus structure, and the definition of a small region of interest containing it, a step of VOI segmentation is proposed. This step consists of three substeps, which are the registration of volumes, generation of the atlas, and crop in the esophagus region.

### 3.2.1. Volume registration

As presented in Section 3.1, the database is very diverse, with exams from 3 different institutes. Still, the number of slices varies. Also, some volumes
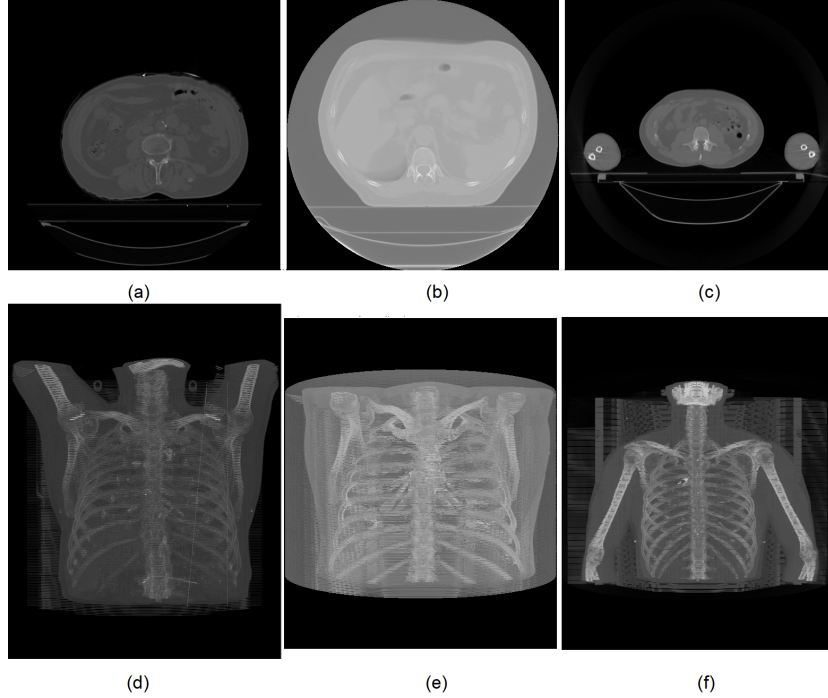
Figure 3: Image acquisition: (a) slice of institute 1; (b) slice of institute 2; (c) slice of institute 3; (d) volume of institute 1; (e) volume of institute 2; and (f) volume of the institute 3.

(like those from institute 3) present slices up to the head region and others do not. One way to standardize the number of slices is by using the registration technique [23, 24].

So, the registration was applied to align all the volumes. One volume was selected as fixed volume and the others were aligned to it. The resulting transformation from the registration was also applied to their respective masks in order to align them too.

In this work, the rigid registration [25] was applied to the entire database, taking as a fixed volume the one with the fewest slices. This criterion was adopted since approximating the slices of a volume with few slices to one with many slices would generate too many estimated voxels that may not be accurate, since some interpolation would be necessary to approximate these volumes.

The image registration was used to align both the patient's volume and its corresponding marking volume. An example of a registered volume can be seen in Figure 4.
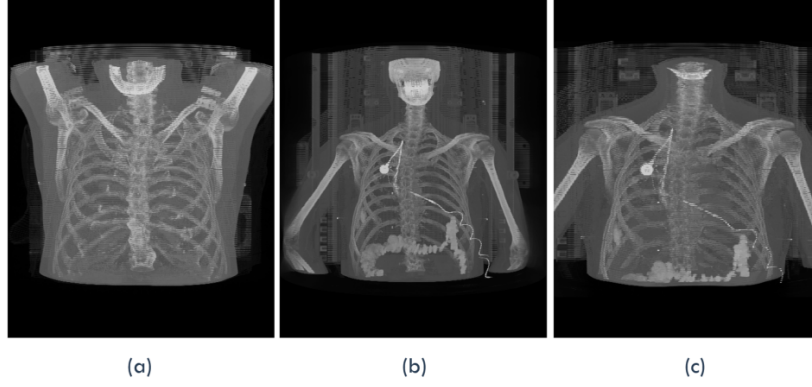


Figure 4: Volume registration: (a) fixed volume with 134 slices, (b) moving volume with 279 slices and (c) moving volume registered with 134 slices (CT exam rotated for better visualization).

With the patient's volumes and markings aligned, the next step is to generate the atlas.

### 3.2.2. Atlas generation

The atlas is a probabilistic volume used to find the probable location of the esophagus. It is generated by combining the information of the marking volumes aligned in the previous step. The atlas voxels values are calculated by averaging the voxels values of all aligned marking volumes. The purpose of the atlas for the VOI segmentation step is to find the region that contains the esophagus in all volumes of the database.

Different from the work of Feulner et al. [13], Larsson et al. [15], where probabilistic and atlas models are proposed to guide OAR segmentation, our method uses atlas only to assist the next step of volume crop.

The atlas corresponds to the spatial distribution of the voxel values belonging to an object of interest. For example, there is a database of CT scans of

13

several patients, and for each patient the specialist makes a specific marking considering the anatomical structure of each patient. When generating the atlas of a specialist's markings, each element represents the sampling probability of the voxels belonging to the object [26]. Figure 5 shows an illustration of how the atlas was generated.
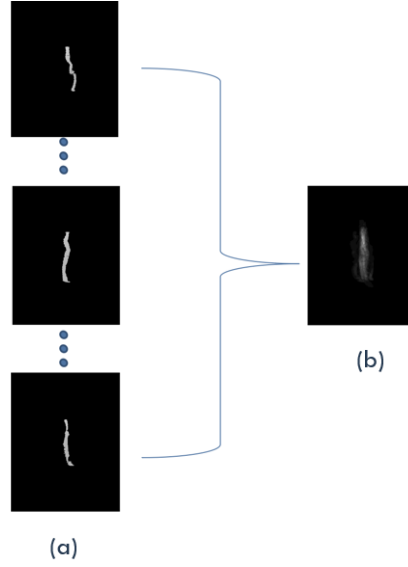


Figure 5: Atlas generation: (a) example of mark volumes and (b) result of atlas generation

It is observed that although the marking of the esophagus in the volumes is still diversified, they are spatially located in a small portion of the volume. With this information, it is possible to crop the volume and generate a VOI.

### 3.2.3. Crop in the esophageal region

From the generation of the atlas, it is possible to obtain spatial information from the esophagus. Thus, a crop is made in the esophagus region in all volumes of the database, following the criterion of the largest region found in the atlas. Thus, all previously registered volumes are now cropped along with their markings, resulting in a region that comprises the possible location of the esophagus. An example can be seen in Figure 6.
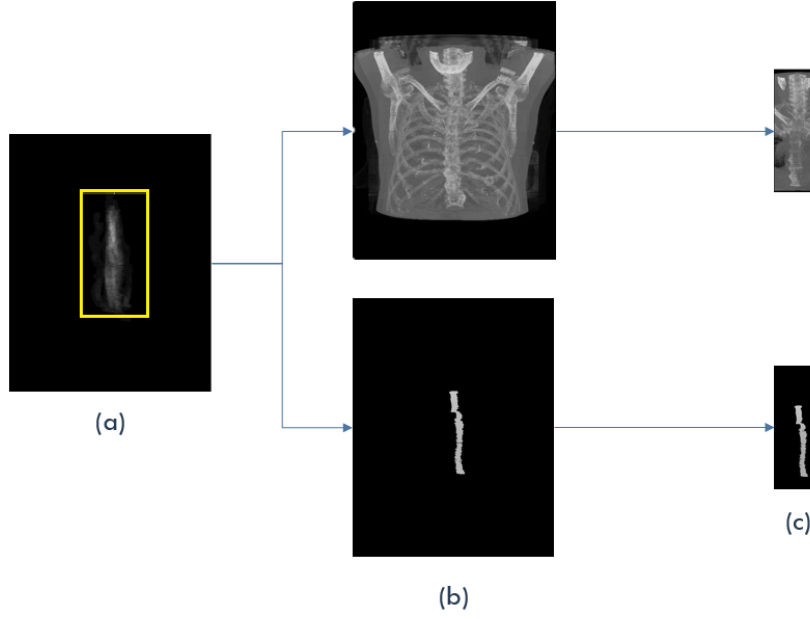
Figure 6: Crop in the esophagus region: (a) atlas, in yellow region corresponding to the volume of the largest object, (b) the patient's original volumes (CT exam rotated for better visualization) and the marking (c) segmentation of the patient's VOI and marking.

Thus, at the end of this step, we have the VOIs of all patients and all markings. This reduces the scope of working with the entire volume to segment the esophagus. Still, despite the existence of an imbalance between non-esophagus voxels, the proportion decreases significantly. The next step is pre-processing.

*3.3. Pre-processing*

When looking at the literature (Section 2), we did not find studies concerned with using any image improvement to highlight the region of the esophagus. Thus, it is proposed a pre-processing that seeks to enhance the region of the esophagus to improve the accuracy segmentation.

For the pre-processing step, it is proposed to use the sequence of two filters. The first filter is bilateral. It seeks to smooth the images while preserving the edges, through a non-linear combination of close image values. Because the

15

distinction between the esophagus and other surrounding tissues is not clear, this filter helps to highlight the edges of the esophagus [27].

Then, a histogram equalization [28] is used in the images highlighted by the bilateral filter. By creating an image with equally distributed levels of gray, it is believed that it further improves the distinction between the esophagus and other tissues. An example of a VOI slice with the pre-processing step is shown in Figure 7.
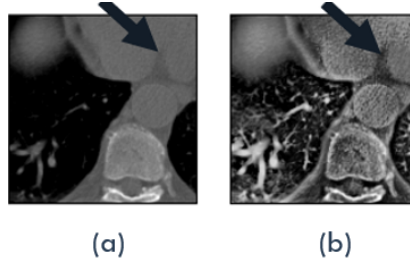


Figure 7: Pre-processing: (a) VOI slice without pre-processing, (b) VOI slice with pre-processing. The arrow indicates the enhancement of the edges of the esophagus.

So, with VOI defined and pre-processed, the next step is to segment the esophagus.

### 3.4. Esophagus segmentation

This step is divided into a few substeps. First, the network input is defined, then training and testing are performed.

### 3.4.1. Generating input images

The result of the VOIs segmentation produces two volumes, one from the patient CT and the other from the marking. The volume of the patient goes through the pre-processing step, to enhance the internal structures and improve the representation of the esophagus. A 2D approach was used, that is, the VOI of each volume was passed slice by slice in the network. However, at the end of the network segmentation, a 3D volume of the correctly classified slices is constructed. The 2D segmented slices of each exam were stacked in 3D volumes. For that, we used the same inter-slice spacing of the original volumes.

We chose to use the slices instead of the volume because, if we consider the number of existing volumes (there are only 36), the network would have few samples in the training phase. As the slices are used, the number of training/validation and test samples is increased.

The next step is the training of U-Net [29] to create a robust model capable of segment new volumes. U-Net is a type of deep convolutional neural network. This network has been used in various medical imaging problems and has reached the state of the art [30–33]. The network used for this task is a U-Net with residual blocks, that we called Residual-U-Net. Figure 8 illustrates the architecture developed to esophagus segmentation.
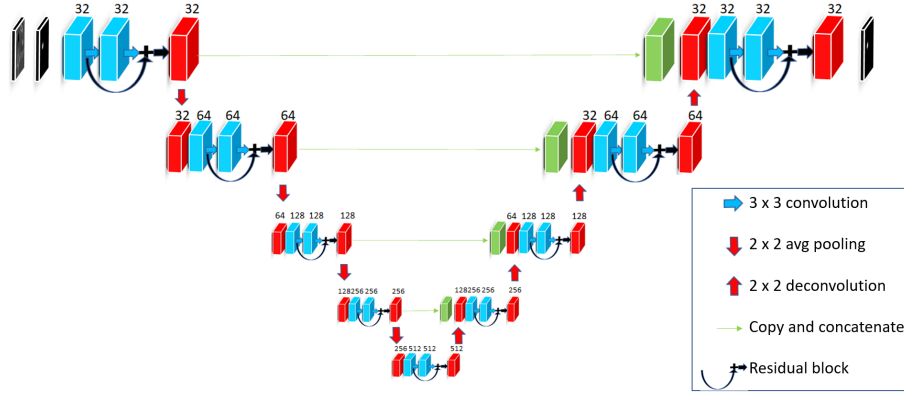


Figure 8: Residual-U-Net: architecture developed to esophagus segmentation

According to He et al. [34], deep neural networks are hard to train. In their work, they explicitly reformulate the layers as learning residual functions regarding the layer inputs, instead of learning unreferenced functions. The results achieved using residual blocks in CNN are superior to several types of approaches. Briefly, the residual block approach inserts shortcut connections between the input and output layers. These shortcut connections simply perform identity mapping and their outputs are added to the outputs of the stacked layers.

Observing the effect of residual blocks in common CNNs and the results achieved using them, we propose to use residual blocks in U-Net. In this way,

it is believed that the network becomes more robust in the task of esophagus segmentation. Inserting the input maps to the output of each layer avoids that the set of pooling operations reduces the necessary information contained in esophagus voxels that comprises only a small region of the image.

By making this change, our network has become more robust. Since the esophagus has a very similar tissue to the structures around them, and corresponds to a small region within the VOI, maintaining its characteristics throughout the network has produced more significant results with the use of Residual-U-Net.

### 3.4.2. Training and testing of Residual-U-Net

In this section, the database is divided into two groups: the first one with the training input slices and the second one with the test input slices. The training database consists of slices from patient's volumes and corresponding slices from marking's volumes. In the training phase, a part of the data is separated for validation. The validation data is used to check the performance of the network with a set of hyperparameters. If a set of hyperparameters creates a model with unsatisfactory results, other ones are chosen and the network is trained again. This process repeats until the performance of the network is good enough. At the end of the training, the best model is selected. Then, this model is applied to the test input slices.

The test phase consists of using the model to classify new images, unknown to the model. To evaluate the test results, the specialist's marking image is compared to the segmentation through the application of evaluation metrics. Thus, they assess whether the model was good enough in segmenting the esophagus as an OAR.

As it is an OAR that does not have a visually defined boundary in the planning CTs, another substep is proposed to refine the segmentation.

### 3.5. Segmentation refinement

The U-Net output represents the probability of each voxel belonging to the target class. These probabilities are thresholded, so if they are greater than 0.5, their pixels are labeled as the target class. In some images, this thresholding results in more than one region.

It was observed that, although Residual-U-Net presents satisfactory results when generating the segmentation predicted by the model, it is possible to identify elements that do not correspond to the esophagus. The most common scenario is the generation of two or more predicted objects in a slice, usually one being the esophagus and the other one(s) being smaller unexpected structure(s). To eliminate these other structures, only the object with the largest area is kept in slices.

As seen in Figure 9a, many structures that are not esophagus are classified as false positives (in blue). The step of keeping the largest object can eliminate these objects (Figure 9b).

To get the largest object, firstly, we use a connected components image filter, so we find all the objects connected in the volume [28]. Then a label is placed for each object. Then, we calculate the area of each label. Finally, all labels are compared and only the one with the most voxels is kept. Some examples can be seen in Figure 9b.

Another point is that after keeping only the largest object (Figure 9b), it is noticed that the region predicted as the esophagus is always larger than the region of the specialist's marking (green marking in Figure 9). This is due to the fact highlighted by the majority of authors who proposed methods for segmenting the esophagus: this organ does not have a noticeable boundary that separates it from the tissues around it. Besides, the $3 \times 3$ convolutional filters applied in the Residual-U-Net convolution steps extract features associated with the neighborhood so they end up considering a larger region in the esophagus. The solution to this problem was the application of erosion in the predicted marking. We use a circle as a structuring element with a radius of $2 \times 2$, the
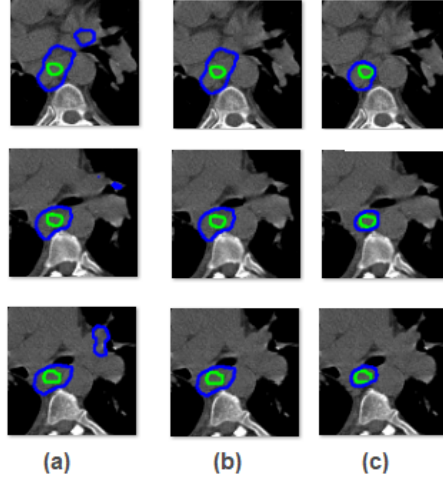
19

Figure 9: Segmentation refinement: (a) examples of specialist markings in green and result of Residual U-Net in blue, (b) specialist markings in green and result of getting largest region in blue, and (c) specialist markings in green and result of erode in blue.

results can be seen in Figure 9c. It was the best parameters that produced coherent results.

At the end of all steps of esophagus segmentation, the evaluation metrics were calculated - Dice index [35], Jaccard Index, Hausdorff Distance, sensitivity (SEN), specificity (SPE) and accuracy (ACC). The results for the method are presented in the section that follows.

## 4. Results

The results achieved for esophagus segmentation in CT for radiotherapy planning, using the proposed method, will be described in this section. Firstly, we describe the training and test execution environment.

### 4.1. Training/test environment

The training environment for the VOI Segmentation and Pre-processing step was a computer with Intel Core i5-7300HQ at 2.50GHz, 8 gigabytes of DDR3 memory, which has a modest performance. The VOI segmentation

step, consisting of registration and cropping the esophagus region, lasted approximately 30 seconds per volume. In turn, the pre-processing step proved to be more costly, exceeding 1 minute in each volume, mostly because of the bilateral filter. Keras library with the TensorFlow backend was used to design and train the U-Net architectures in a machine with two NVIDIA Titan X graphics cards.

### 4.2. Image acquisition

As already mentioned in Section 3.1, the database consists of volumes from 3 institutes, with 12 volumes per institute, resulting in 36 volumes. To train and test this method, the database was divided into 2 datasets: train and test. This division was done randomly, but guaranteeing the proportion of exams per institutes in each dataset. Thus, we have 10 exams from each of the 3 institutes for training, and 2 from each institute for testing.

In addition, after all the steps using the 30 volumes in training and the 6 volume to test, the evaluation metrics are calculated to show the robustness of the model.

### 4.3. VOI Segmentation and Preprocessing

In the VOI segmentation step, described in Section 3.2 the training and test sets were registered. The fixed volume, used for registration of all other volumes, is the one with the fewest number of slices in the database. For the generation of the atlas, it was calculated the largest region where the specialist's markings are contained, resulting in a VOI size of $128 \times 128 \times 112$ (number of slices). Thus, all the volumes of the image database were cropped to this VOI size. As a result of this operation, the previous volumes with a slice size of $512 \times 512$ are replaced for smaller standardized VOI with fewer structures to be analyzed.

Then, all these VOIs went through the preprocessing step, where the bilateral filter was first applied and followed by histogram equalization. Now, with the pre-processed database, the next step is the final segmentation.

*4.4. Esophagus segmentation and segmentation refinement*

In this step, the U-Net with residual blocks (Residual-U-Net) receives, as input, the results found in the previous steps. First, as highlighted in Section 3.4, the training volume slices are presented to the 2D network. This strategy is adopted to increase the number of sample images in the training phase.

As the VOIs obtained in the pre-processing step have a size of $128 \times 128 \times 112$, the images passed to the network are slices with a size of $128 \times 128$. After several training sessions and hyperparameters tunning, the chosen hyperparameters were: number of epochs equal to 200, size of *batch* equal to 3, Adam optimizer with initial learning rate equal to 0.0001, and 10% of volumes for the validation set. Again, in the validation set, it is guaranteed the proportion of exams per institute, that is, 1 exam from each institute.

After training the network, the test volumes are presented to the model, and the evaluation metrics are calculated. It is worth mentioning that the test set consists of 2 volumes from each institute. As highlighted in Section 3.4, the model produces some false positives, which are disregarded, keeping only the largest object in the slices. Finally, an erosion is applied to reduce the size of the segmentation, since it always exceeds the size of the specialist's marking.

To have an overview of each of the steps, and how much improvement is obtained in each one, the results of Residual-U-Net are presented without pre-processing, with pre-processing, maintaining the largest object and applying erosion in the Table 1.

Table 1: Results of esophagus segmentation steps.

| Steps | Dice(%) | SEN(%) | SPE(%) | ACC(%) |
|---|---|---|---|---|
| No pre-processing | 44,35 | 100 | 98,02 | 98,03 |
| With pre-processing | 65,68 | 100 | 99,19 | 99,18 |
| With pre-processing + Largest Object | 67,05 | 99,89 | 99,22 | 99,23 |
| With pre-processing + Largest Object + Erosion | 82,15 | 90,61 | 99,76 | 99,69 |

Analyzing the Table 1, the pre-processing step showed a significant improvement in the results. It is believed that due to the better definition of the

22

esophagus boundaries, the segmentation results in Dice increased from 44.35% to 65.68%, showing the importance of this step. Although the two results are at 100% sensitivity, the result without pre-processing has a greater number of false positives, so when analyzing the specificity metric it goes up almost 1%.

Observing the step of keeping only the largest object, improvement is guaranteed on the Dice coefficient and specificity, i.e., there are fewer false positives and more precise segmentation. Although there is some loss in the sensitivity, it is worth remembering that the metrics are calculated using the number of voxels and more voxels belong to the class of non-esophagus than the class of esophagus, so any improvement in specificity (correct negative cases) improves the Dice coefficient.

Finally, when using erosion to reduce the size of the object found in the slices, there is a significant improvement in the results. The Dice coefficient value, which was 67% without this technique, rises to 82.15%, and the metrics of specificity and accuracy increase close to 100 %. It is worth mentioning, once again, that the non-esophagus voxels are the majority in the images, so improvements towards this class also benefit the Dice coefficient.

When observing the sensitivity metric with the complete method, it shows a decrease of 10%, contrary to what happens with specificity. This is because there are fewer esophagus voxels in the slices, so any error in this class compromises the sensitivity metric. However, looking at Section 3.4, the results without the erosion step end up classifying a very large area as esophagus class, which despite correctly classifying the esophagus voxels, is of little use because it incorrectly classifies several voxels around it.

To validate the results found by the method concerning the edge surfaces of the specialist marking, the metrics of Hausdorff Distance [36] and Jaccard Index [37] are calculated. In short, Hausdorff Distance is the longest of all distances from one point in one set to the nearest point in the other set. Jaccard Index compares members for two sets to see which members are shared and which are distinct. Table 2, presents the results of the average of these metrics together with the Dice.

Table 2: Results of esophagus segmentation of boundary distance metrics.

| Metrics | Dice(%) | Jaccard(%) | Hausdorff Distance(mm) |
|---------|---------|------------|------------------------|
| Proposed Method | 82,15 | 70.21 | 6.1030 |

Thus, based on the Dice metric, which shows how close the prediction is to the specialist's marking, it is believed that the method for segmenting the esophagus is promising. It is believed that this method could be used in large centers and provide fundamental help in the treatment of radiotherapy.

### 4.5. Case studies

In this section, some case studies are presented. Each case will show a volume from one of the 3 institutes. In addition, the results achieved by the method in each of these cases are discussed.

### 4.5.1. Case study - Institute 1

The first case study is a patient from Institute 1 of the database. In this case, it can be seen how robust the segmentation method of the esophagus was. Figure 10 shows several VOI slices of the same patient segmented by the method.

Figure 10 shows how similar are the segmentation achieved by the proposed method and the specialist marking. The method was able to segment a region very similar to the specialist marking, generating a minimum of false positives. However, it can be seen that there was a loss of the esophagus in one of the slices. This loss, despite only one slice, negatively affects the value of the metrics, especially the Dice. Still, in the slices containing no esophagus voxels, the method did not segment more regions.

Even using a 2D network, a three-dimensional visualization of the results is presented, both as a means of evaluation and for better visualization of the specialist. As explained in Section 3.4, the 2D segmented slices of each exam were stacked in 3D volumes. For that, we used the same inter-slice spacing of the original volumes. This view can be seen in Figure 11.
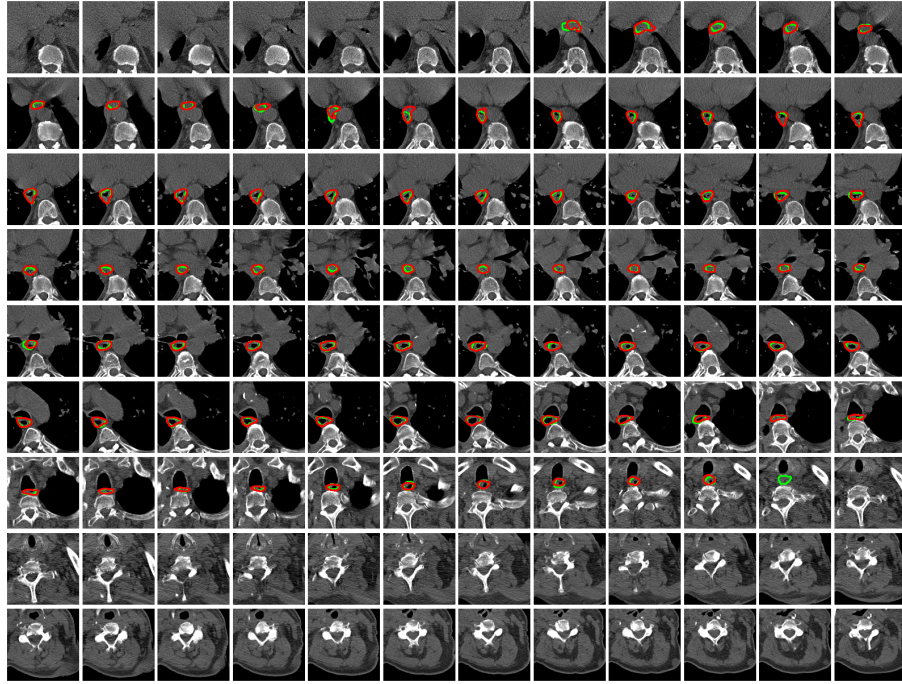
Figure 10: Case study - Institute 1. In green the specialist marking and in red the segmentation predicted by the method.



Figure 11: Case study - Institute 1. (a) volume predicted by the method and (b) volume of the specialist's marking.

Analyzing the Figures 10 and 11, the loss of only one slice did not compromise the three-dimensional visualization of the esophagus segmentation. When comparing it with the specialist's marking, the results prove to be robust and accurate.

Analogous to the first case study, it shows the results of a volume from the Institute 2 of the test database. Again, the method was considered robust in the task of segmenting the esophagus. Figure 12 shows an image of the result of the esophagus segmentation.
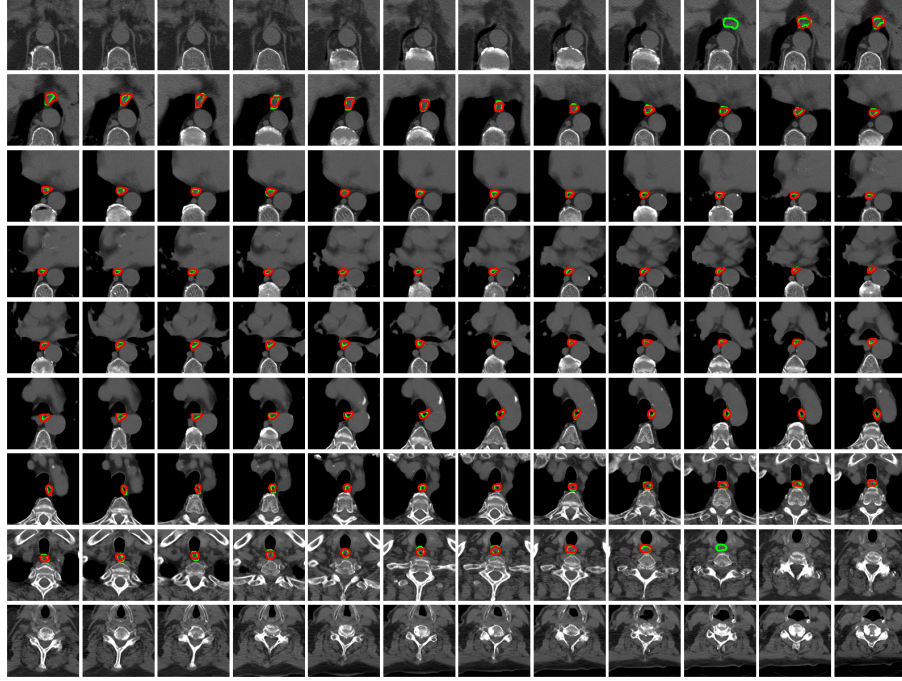


Figure 12: Case study - Institute 2. In green the specialist mark and in red the segmentation predicted by the method.

Once again, the efficiency of the method is shown. Note that in this case there was a loss of two slices marked by the specialist. However, when analyzing the volume as a whole, 112 slices were subjected to segmentation, and only two were lost. Besides, no segmented regions were generated on slices that had no specialist markings. Thus, it is shown how reliable the proposed method is in segmenting the esophagus in planning CT.

As in the case study of Institute 1, a 3D visualization of the result achieved in the volume of Institute 2 is presented. Figure 13 displays a 3D view of the result.
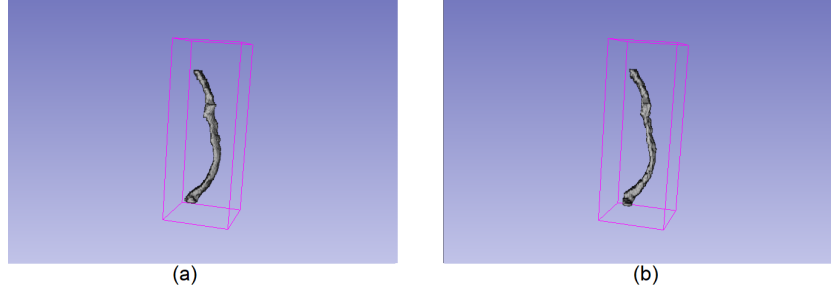


Figure 13: Case study - Institute 2. (a) volume predicted by the method and (b) volume of the specialist's marking.

As in the first case study, there is a high similarity between the specialist marking and the segmentation predicted by the method. This three-dimensional visualization shows a macro view and, consequently, a global analysis of the method result. It is believed that the loss of a few slices, although problematic, does not invalidate the segmentation in a 3D view.

### 4.5.3. Case study - Institute 3

Finally, a case study of a patient from Institute 3 is shown. Analogous to previous cases, Figure 14 shows the segmentation achieved by the method and the specialist's marking.

Again, there was a loss of the esophagus in one slice. However, as already mentioned, the volume has 112 slices. Considering that the other slices were segmented correctly, the method demonstrates to be very efficient in esophagus segmentation.

Analogous to the previous test cases, this also generates a three-dimensional visualization of the method's result compared to the specialist's marking. This view is shown in Figure 15.
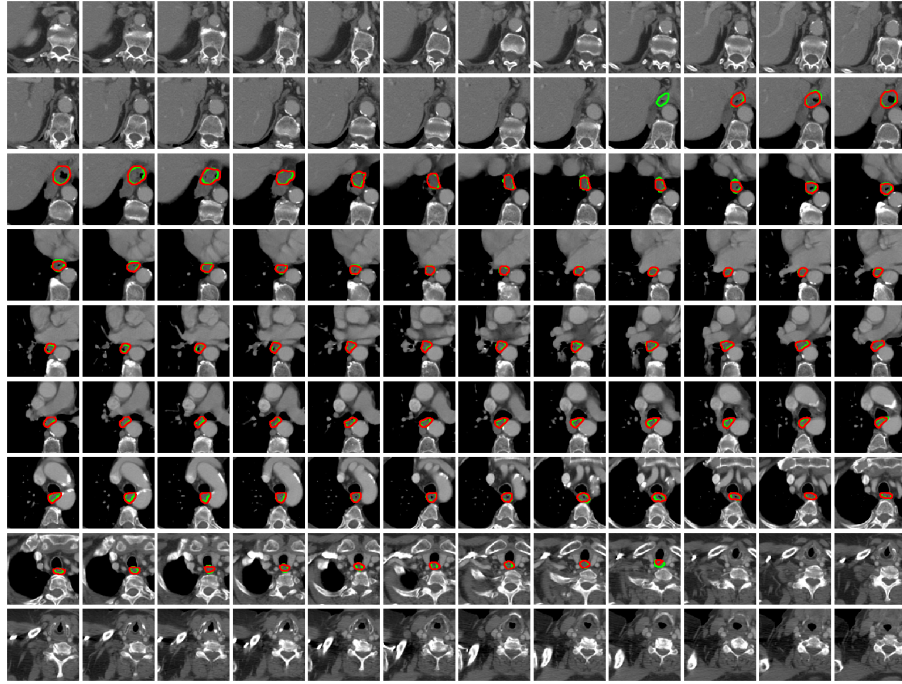
Figure 14: Case study - Institute 3. In green the specialist mark and in red the segmentation predicted by the method.



Figure 15: Case study - Institute 3. (a) volume predicted by the method and (b) volume of the specialist's marking.

Again, it is possible to observe an excellent three-dimensional segmentation when compared to the specialist's marking.

*4.6. Comparison with related works*

The literature provides many approaches for segmenting the esophagus on CT, either as OAR or not. Most authors report the difficulty in distinguishing the esophagus on CT, even when it is done by specialists. It is noteworthy that, because there is no well-defined contrast of this organ in CT, segmenting it has become an arduous task, and consequently, an exhaustive process when performed manually.

In Feulner et al. [13], a probabilistic method for esophagus segmentation is proposed. According to the authors, the method consists of multi-stages. The authors' experiments are performed on a private database with 144 CT and achieved a Dice result of 74%. It is observed that this method is highly dependent on descriptive models and prior knowledge. Yet, non-rigid registration is necessary, causing deformities in the initial esophagus structure. One of the main negative points of the work is the use of deformable registration that ends up modifying the individual anatomical structure of each patient. Besides, segmentation is guided by a probabilistic model. In our work, despite the use of the registration, it is a rigid one and the atlas is only used for the generation of a VOI that will later be presented to Residual-U-Net.

The work proposed by Grosgeorge et al. [14], proposes a method for esophagus segmentation. The method achieves its results using a skeleton-shaped model to guide the segmentation. The method consists of two stages: a 3D segmentation using a graph cut technique, followed by a 2D propagation. This method is applied to a database with 6 patients and presents a Dice equal to 61%. Although the method is said to be automatic, it depends on the creation of a skeleton model based on specialist marking. The method also uses only image processing techniques without applying machine learning classification. Our work uses a very diversified database with a total of 36 patients. Moreover, our work reaches a Dice of 82%.

The work by Larsson et al. [15] proposes a method for the automatic segmentation of various abdominal organs on CT. Among the segmented organs, the esophagus segmentation achieves a Dice of 66.2%. However, the authors use

an atlas to weight the voxels, so that they have a probability of belonging to each organ classes. In our method, we used the atlas only for VOI segmentation, and our Residual-U-Net segmentation achieved better results than the mentioned work.

Once again, the difficulty in segmenting the esophagus is highlighted by Trullo et al. [18]. The authors also highlight how crucial the segmentation of this organ is for RT. According to them, the work presented a fully automatic method. This method consists of only two steps. In the first step, a CNN is used to estimate the location of the esophagus (this CNN was published in other papers by the same authors, and presented a Dice of 66% for the esophagus [16]). So, a crop is made in this location, and the same network is trained to classify only this small region. The method was tested in 30 CT volumes and presented a Dice of 72%. Like most works, it proposes to crop the esophagus location. However, as the authors themselves point out, the first network is not able to correctly locate the esophagus, so cropping is necessary and then the second network classifies the esophagus. Because the first network may not find the location of the esophagus or segment the wrong region, the second network (which has the same architecture of the first) may not be able to segment it properly. In our work, for the VOI segmentation, we first used the atlas, then Residual-U-Net segments the esophagus in the VOI. Thus, our work exceeds the Dice values of both authors' works.

One of the most recent works that deal with the segmentation of the esophagus as OAR is Fechter et al. [12]. The method proposed by the authors uses a CNN to create a probabilistic model that is used as an input to an active contour model (ACM). Then, both segmentations achieved by the CNN model and the one generated by the ACM, are applied to a random walker algorithm based on the Hounsfield units (HU) of the volumes. The method is applied to a database composed of 50 volumes, reaching a Dice result of 76%. Although the method uses CNN as part of its steps, the final segmentation depends on the HU of the image. This makes the model susceptible to errors when there is an abrupt change in HU, in cases where the CT acquisition process is not well

defined. The database used in our work, for example, has different acquisition protocols. Although this method has the highest Dice value in the literature until now, our work was capable to surpass it.

In Dong et al. [17], a method for esophagus segmentation in planning CT is proposed. The authors propose an adversarial network combined with an FCN to achieve their results. This work presents the results in the same database used by us. The method is validated by the leave-one-out technique, ensuring that only one individual is tested at a time, so the training has more samples to be generated. For the esophagus, the method achieves 75% Dice, 73% sensitivity, and 99% specificity. The work is very promising, but it relies on a step of segmentation of the lungs. The lung segmentation is used as a reference for cropping a VOI comprising the esophagus. Thus, the method is dependent on a good segmentation of the lungs, which in our case does not interfere. Also, although this work was carried out in the same database used in our work, the authors used only 35 volumes, but they do not state the reason for this. Our method outperforms this work in all evaluation metrics.

The work proposed by Chen et al. [20] performed a deep learning semantic segmentation to esophagus segmentation. The authors proposed a U-Net and its done slice by slice. And in the end, a 3D model of the result is created to assist the specialist. Despite being a robust method, with recent and promising techniques, the method is applied in 15 CT exams. The results achieved by the method are 79% [20] of Dice. Our method is applied to a larger database, composed of 36 patients from 3 institutes, and reaches an average Dice of 82.15%.

Feng et al. [21] also present a multi-organ segmentation method that targets the esophagus. The method is quite recent and applied to the same database as our work. This method also uses a U-Net approach for organ segmentation. However, the Feng et al. [21] method achieves only 72% of Dice for esophagus segmentation, while our work shows results 10% higher than this one.

Table 3 shows the relationship between these works and the proposed method.

Table 3: Comparison with related works.

| Works | Method | Exams | Acu | Dice |
|---|---|---|---|---|
| Feulner et al. [13] | Probabilistic model | 14 | — | 74% |
| Grosgeorge et al. [14] | Skeleton-shape model | 6 | — | 61% |
| Larsson et al. [15] | Atlases weighted model | 30 | — | 66,2% |
| Trullo et al. [16] | CNN model | 30 | — | 66% |
| Trullo et al. [18] | Two CNN models | 30 | — | 72% |
| Fechter et al. [12] | CNN + ACM | 50 | — | 76% |
| Dong et al. [17] | Adversarial CNN | 35 | — | 75% |
| Chen et al. [20] | U-Net/CNN | 15 | — | 79% |
| Feng et al. [21] | U-Net/CNN | 36 | — | 72% |
| Proposed Method | Atlas-based Residual-U-Net | 36 | 99,69% | 82,15% |

It can be observed that the literature presents several recent works for the segmentation of the esophagus. As it is not a trivial task, not even for a specialist, the highest Dice value found in the literature so far is 79%. All studies show the need for reducing the scope of the patient's volume for better segmentation. This information was valuable when building our method, where registration techniques and atlases were used to segment the VOI.

It was also possible to observe that none of these works used preprocessing, so an important step in our method was to find enhancement techniques that could improve the visualization of the esophagus in the slices. Another important step is the refinement of the Residual-U-Net segmentation. Since the texture of the esophagus is very similar to the other organs and tissues, the segmentation was not accurate enough, so this step was necessary to improve it.

Thus, it is noteworthy that our work presented an automatic method for esophagus segmentation. It is composed of techniques found in the literature and novel techniques to achieve better results. Although some databases used in literature are different from ours, it manages to surpass all the works presented so far. This demonstrates the feasibility of our method for esophagus segmentation.

## 5. Discussion

Esophagus segmentation is not a trivial task. Developing a method capable of circumventing all the adversities of this OAR and achieving a good result is still a challenge. The proposed method for esophagus segmentation proved to be quite robust and presented a series of advantages:

1. A fully automatic method is presented, which does not require prior knowledge to conduct segmentation;

2. The step of VOI segmentation proved to be quite pertinent. Since the esophagus represents a very small portion of the planning CT volume, proposing a method capable of reducing the scope before applying the segmentation impacts the final results, as shown in related works;

3. One of the main challenges of the esophagus, already highlighted in all the related works, is that there is no clear distinction between its tissue and the tissues around it. Following the literature, it was observed that most works relied on locating this OAR from spatial information, either through probabilistic models or scope reduction techniques. Thus, one of the contributions of the proposed work to this task is the pre-processing step;

4. The pre-processing step proved to be crucial to achieve better results. The method achieved an improvement of more than 20% of Dice when using the sequence of the two pre-processing techniques;

5. The strategy of using slices in the training step of the network also proved to be effective. A 3D approach would considerably reduce the number of samples. By using slices, this problem is mitigated;

6. The proposed network for esophagus segmentation was also a crucial step in achieving good results. It is believed that, since the esophagus comprises a small region of the VOI, using the residual blocks between the layers helped to maintain the features, providing better results;

7. Still, we did not find studies in the literature that use residual blocks in the U-Net for the esophagus segmentation, once again presenting advances with the proposed method;

8. Keeping the largest object from the segmentation is also important for the method. It was observed that several tissues around the esophagus present similar features compared to it, so the network ends up predicting some smaller objects other than the esophagus. Therefore, excluding these objects improves the Dice of the method;

9. It is possible to notice that because the tissue around the esophagus is very similar to the texture of the esophagus itself, the network leads to segment a larger region. However, in most cases, there is no loss of the esophagus, resulting in 100% of sensitivity in some tests. Considering this, an erosion step was applied, where it was possible to significantly improve the metrics;

10. The combination of all the techniques employed in this method resulted in a Dice value of 82.15%, sensitivity of 90.61%, specificity of 99.76%, and accuracy of 99.69%. This result surpasses the other works presented so far in the literature;

11. Finally, it is emphasized the problems of the database which, in addition to being diversified, presents the esophagus structure as an OAR that is very difficult to segment. Despite all the adversities, the method managed to reach a state-of-the-art Dice value.

Although the method is very promising, surpassing the works already published, it still has some limitations:

1. The training step has a segmentation network composed of numerous hyperparameters. They are adjusted by evaluating the network performance on the validation dataset, which takes a long time;

2. Another issue is the refinement of segmentation. Although it produces very expressive Dice results, as the erosion technique reduces the object

segmented by Residual-U-Net, there is a certain loss in the sensitivity metric;

3. Finally, there is a great dissimilarity on the data provided by the 3 institutes. Finding a way to standardize their textures could improve the results.

Thus, it is stated that all these advantages and limitations add important value to the proposed method for esophagus segmentation. The combination of the techniques provided promising results, and it is believed that the use of this method in CAD systems could be of fundamental importance to the specialists in the segmentation of this OAR.

## 6. Conclusion

For the esophagus segmentation, we sought to associate the most used techniques in the literature and propose improvements to generate promising results. The VOI segmentation step was crucial for the initial results, with which it was possible to reduce the scope of the volume and thus provide improvements in the results. The pre-processing step proved to be fundamental, with the possibility of enhancing the boundaries of the esophagus and, consequently, passing on valuable information to Residual-U-Net.

The proposed network for esophagus segmentation presented promising results, especially considering that the esophagus corresponds just a small region of the VOI. The subsampling layers lead to further reduce the amount of information. Therefore, combining residual information from the previous layers brought better results to the method.

Another important point was the refinement step. This step is composed of 2 techniques. The first one keeps the largest object and the second one applies an erosion to the segmentation. This step was able to address the problem of similar textures among the esophagus and surrounding tissues. The result achieved by applying these techniques was a Dice coefficient value of 82.15%. It is worth mentioning that this is the best Dice coefficient result in the literature.

Furthermore, this is a very diversified database, which further validates the effectiveness of the proposed method.

However, as already mentioned, the training step has a segmentation network composed of various hyperparameters that takes time to tune. Automating this step could produce even better results. There is a high dissimilarity among the data provided by the three institutes, so standardize their texture could also improve the results.

Thus, it stands out that the presented method is quite promising. We believe that it can be very useful in large radiotherapy centers, where the task of protecting the esophagus is essential. It should be noted that the esophagus is an OAR that is extremely sensitive to radiation. Besides, damaging it can cause several problems. To make things worse, it is an organ that is difficult and time-consuming to segment, even by the specialist. All those issues are enough to prove the importance and need of a precise fully-automated segmentation method, and our method could provide that.

## References

[1] E. Evans, J. Staffurth, Principles of cancer treatment by radiotherapy, Surgery-Oxford International Edition 36 (2018) 111–116.

[2] J. J. Caudell, J. F. Torres-Roca, R. J. Gillies, H. Enderling, S. Kim, A. Rishi, E. G. Moros, L. B. Harrison, The future of personalised radiotherapy for head and neck cancer, The Lancet Oncology 18 (2017) e266–e273.

[3] S. A. Castaneda, L. B. Romak, Radiotherapy for anal cancer: Intensity-modulated radiotherapy and future directions, Surgical Oncology Clinics 26 (2017) 467–475.

[4] J. Gómez-Millán, M. F. Lara, R. C. Generoso, A. Perez-Rozos, Y. Lupiáñez-Pérez, J. A. M. Carmona, Advances in the treatment of prostate cancer

with radiotherapy, Critical reviews in oncology/hematology 95 (2015) 144–153.

[5] M. Button, J. Staffurth, Clinical application of image-guided radiotherapy in bladder and prostate cancer, Clinical Oncology 22 (2010) 698–706.

[6] S. Gwynne, R. Webster, R. Adams, S. Mukherjee, B. Coles, J. Staffurth, Image-guided radiotherapy for rectal cancer—a systematic review, Clinical Oncology 24 (2012) 250–260.

[7] M. Milosevic, M. Gospodarowicz, A. Zietman, F. Abbas, K. Haustermans, L. Moonen, C. Rödel, M. Schoenberg, W. Shipley, Radiotherapy for bladder cancer, Urology 69 (2007) 80–92.

[8] N. McVicar, I. A. Popescu, E. Heath, Techniques for adaptive prostate radiotherapy, Physica Medica: European Journal of Medical Physics 32 (2016) 492–498.

[9] M. J. Zelefsky, Z. Fuks, L. Happersett, H. J. Lee, C. C. Ling, C. M. Burman, M. Hunt, T. Wolfe, E. Venkatraman, A. Jackson, et al., Clinical experience with intensity modulated radiation therapy (imrt) in prostate cancer, Radiotherapy and Oncology 55 (2000) 241–249.

[10] Y. M. Tsang, P. Hoskin, The impact of bladder preparation protocols on post treatment toxicity in radiotherapy for localised prostate cancer patients, Technical Innovations and Patient Support in Radiation Oncology 3 (2017) 37–40.

[11] Varian, Eclipse treatment planning system, 2020. URL: https://www.varian.com/oncology/products/software/treatment-planning/eclipse-treatment-planning-system.

[12] T. Fechter, S. Adebahr, D. Baltas, I. B. Ayed, C. Desrosiers, J. Dolz, Esophagus segmentation in ct via 3d fully convolutional neural network and random walk, Medical physics 44 (2017) 6341–6352.

[13] J. Feulner, S. K. Zhou, M. Hammon, S. Seifert, M. Huber, D. Comaniciu, J. Hornegger, A. Cavallaro, A probabilistic model for automatic segmentation of the esophagus in 3-d ct scans, IEEE transactions on medical imaging 30 (2011) 1252–1264.

[14] D. Grosgeorge, C. Petitjean, B. Dubray, S. Ruan, Esophagus segmentation from 3d ct data using skeleton prior-based graph cut, Computational and mathematical methods in medicine 2013 (2013).

[15] M. Larsson, Y. Zhang, F. Kahl, Robust abdominal organ segmentation using regional convolutional neural networks, in: Scandinavian Conference on Image Analysis, Springer, 2017, pp. 41–52.

[16] R. Trullo, C. Petitjean, S. Ruan, B. Dubray, D. Nie, D. Shen, Segmentation of organs at risk in thoracic ct images using a sharpmask architecture and conditional random fields, in: 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017), IEEE, 2017, pp. 1003–1006.

[17] X. Dong, Y. Lei, T. Wang, M. Thomas, L. Tang, W. J. Curran, T. Liu, X. Yang, Automatic multiorgan segmentation in thorax ct images using u-net-gan, Medical physics 46 (2019) 2157–2168.

[18] R. Trullo, C. Petitjean, D. Nie, D. Shen, S. Ruan, Fully automated esophagus segmentation with a hierarchical deep learning approach, in: 2017 IEEE International Conference on Signal and Image Processing Applications (ICSIPA), IEEE, 2017, pp. 503–506.

[19] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. Van Der Laak, B. Van Ginneken, C. I. Sánchez, A survey on deep learning in medical image analysis, Medical image analysis 42 (2017) 60–88.

[20] S. Chen, H. Yang, J. Fu, W. Mei, S. Ren, Y. Liu, Z. Zhu, L. Liu, H. Li, H. Chen, U-net plus: Deep semantic segmentation for esophagus and

esophageal cancer in computed tomography images, IEEE Access 7 (2019) 82867–82877.

[21] X. Feng, K. Qing, N. J. Tustison, C. H. Meyer, Q. Chen, Deep convolutional neural network for segmentation of thoracic organs-at-risk using cropped 3d images, Medical physics 46 (2019) 2169–2180.

[22] J. Yang, H. Veeraraghavan, S. G. Armato III, K. Farahani, J. S. Kirby, J. Kalpathy-Kramer, W. van Elmpt, A. Dekker, X. Han, X. Feng, et al., Autosegmentation for thoracic radiation treatment planning: A grand challenge at aapm 2017, Medical physics 45 (2018) 4568–4581.

[23] W. R. Crum, T. Hartkens, D. Hill, Non-rigid image registration: theory and practice, The British Journal of Radiology (2014).

[24] L. G. Brown, A survey of image registration techniques, ACM computing surveys (CSUR) 24 (1992) 325–376.

[25] T. S. Yoo, Insight into images: principles and practice for segmentation, registration, and image analysis, AK Peters/CRC Press, 2004.

[26] M. Patel, P. Kelly, M. Smith, A. Plassard, B. Landman, R. G. Abramsom, X. Zhoubing, B. K. Poulose, R. B. Baucom, A. J. Asman, et al., Image segmentation via multi-atlas fusion with context learning, 2019. US Patent 10,402,981.

[27] Y.-T. Wu, J. Wei, L. M. Hadjiiski, B. Sahiner, C. Zhou, J. Ge, J. Shi, Y. Zhang, H.-P. Chan, Bilateral analysis based false positive reduction for computer-aided mass detection, Medical physics 34 (2007) 3334–3344.

[28] R. C. Gonzalez, R. E. Woods, Digital image processing, 2002.

[29] O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation, in: International Conference on Medical image computing and computer-assisted intervention, Springer, 2015, pp. 234–241.

[30] P. H. B. Diniz, T. L. A. Valente, J. O. B. Diniz, A. C. Silva, M. Gattass, N. Ventura, B. C. Muniz, E. L. Gasparetto, Detection of white matter lesion regions in mri using slic0 and convolutional neural network, Computer methods and programs in biomedicine 167 (2018) 49–63.

[31] J. O. B. Diniz, P. H. B. Diniz, T. L. A. Valente, A. C. Silva, A. C. de Paiva, M. Gattass, Detection of mass regions in mammograms by bilateral analysis adapted to breast density using similarity indexes and convolutional neural networks, Computer Methods and Programs in Biomedicine 156 (2018) 191–207.

[32] J. O. B. Diniz, P. H. B. Diniz, T. L. A. Valente, A. C. Silva, A. C. Paiva, Spinal cord detection in planning ct for radiotherapy through adaptive template matching, imslic and convolutional neural networks, Computer methods and programs in biomedicine 170 (2019) 53–67.

[33] J. C. Souza, J. O. B. Diniz, J. L. Ferreira, G. L. F. da Silva, A. C. Silva, A. C. de Paiva, An automatic method for lung segmentation and reconstruction in chest x-ray using deep neural networks, Computer methods and programs in biomedicine 177 (2019) 285–296.

[34] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770–778.

[35] L. R. Dice, Measures of the amount of ecologic association between species, Ecology 26 (1945) 297–302.

[36] D. P. Huttenlocher, G. A. Klanderman, W. J. Rucklidge, Comparing images using the hausdorff distance, IEEE Transactions on pattern analysis and machine intelligence 15 (1993) 850–863.

[37] N. Tustison, J. Gee, Introducing dice, jaccard, and other label overlap measures to itk, Insight J 2 (2009).